Extremal and Structural Problems of Graphs

António José Ferra Gomes de Almeida Girão

Sidney Sussex College

and

Department of Pure Mathematics and Mathematical Statistics

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

August 15, 2018

Advisor: Béla Bollobás

Declaration

This document is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. I have not submitted any part of this dissertation for any other qualification.

Chapter 2 is based on joint work with Ross J. Kang.

Chapter 3 is based on joint work with with Katherine Edwards, Jan van den Heuvel, Ross J. Kang, Gregory J. Puleo and Jean-Sébastien Sereni.

Chapter 4 is based on joint work with Shoham Letzter and Julian Sahasrabudhe.

Chapter 5 is based on joint work with David Lewis and Kamil Popielarz.

Chapter 6 is based on joint work with Bhargav Narayanan and Teeradej Kittipassorn.

Chapter 7 is based on joint work with Kamil Popielarz.

Chapter 8 is based on joint work with Teeradej Kittipassorn and Kamil Popielarz.

Chapter 9 is based on joint work with Richard Snyder.

António José Girão

Abstract

In this dissertation, we are interested in studying several parameters of graphs and understanding their extreme values. We begin in Chapter 2 with a question on edge colouring. When can a partial proper edge colouring of a graph of maximum degree Δ be extended to a proper colouring of the entire graph using an 'optimal' set of colours? Albertson and Moore conjectured this is always possible provided no two precoloured edges are within distance 2. The main result of Chapter 2 comes close to proving this conjecture. Moreover, in Chapter 3, we completely answer the previous question for the class of planar graphs.

Next, in Chapter 4, we investigate some Ramsey theoretical problems. We determine exactly what minimum degree a graph G must have to guarantee that, for any two-colouring of $E(G)$, we can partition $V(G)$ into two parts where each part induces a connected monochromatic subgraph. This completely resolves a conjecture of Bal and Debiasio. We also prove a 'covering' version of this result. Finally, we study another variant of these problems which deals with coverings of a graph by monochromatic components of distinct colours.

The following saturation problem proposed by Barrus, Ferrara, Vandenbussche, and Wenger is considered in Chapter 5. Given a graph H and a set of colours $\{1, 2, \ldots, t\}$ (for some integer $t \geq |E(H)|$), we define $\mathrm{sat}_t(n, \mathcal{R}(H))$ to be the minimum number of t-coloured edges in a graph on n vertices which does not contain a rainbow copy of H but the addition of any non-edge in any colour from $\{1, 2, \ldots, t\}$ creates such a copy. We prove several results concerning these extremal numbers. In particular, we determine the correct order of $\mathrm{sat}_t(n, \mathcal{R}(H))$, as a function of n, for every connected graph H of minimum degree greater than 1 and for every integer $t \geq e(H)$.

In Chapter 6, we consider the following question: under what conditions does a Hamiltonian graph on n vertices possess a second cycle of length at least $n - o(n)$? We prove that the 'weak' assumption of a minimum degree greater or equal to 3 guarantees the existence of such a long cycle.

We study the following problem, raised by Caro and Yuster, in Chapter 7. Does every graph G contain a 'large' induced subgraph H which has k vertices of degree exactly $\Delta(H)$? We answer in the affirmative an approximate version of this question. Indeed, we prove that, for every k, there exists $g(k)$ such that any n vertex graph G with maximum degree Δ contains an induced subgraph H with at least $n - g(k)\sqrt{\Delta}$ vertices such that $V(H)$ contains at least k vertices of the same degree $d \geq \Delta(H) - g(k)$. This result is sharp up to the order of $g(k)$.

We solve two problems related to majority colouring in Chapter 8. This topic was recently studied by Kreutzer, Oum, Seymour, van der Zypen and Wood. They raised the problem of determining, for a natural number k, the smallest positive

integer $m = m(k)$ such that every digraph can be coloured with m colours, where each vertex has the same colour as at most a proportion of $\frac{1}{k}$ of its out-neighbours. Our main theorem states that $m(k) \in \{2k-1, 2k\}$.

Finally, in Chapter 9, we move on to examine k-linked tournaments. A tournament T is said to be k-linked if for any two disjoint sets of vertices $\{x_1, \ldots, x_k\}$ and $\{y_1, \ldots, y_k\}$ there are directed vertex disjoint paths $P_1, \ldots, P_k$ such that P_i joins x_i to y_i for $i = 1, \ldots, k$. We prove that any $4k$ strongly-connected tournament with sufficiently large minimum out-degree is k-linked. This result comes close to proving a conjecture of Pokrovskiy.

Acknowledgements

First of all I would like to thank deeply my parents and my supervisor, Béla Bollobás, for everything they have done for me. My parents for their love, constant support, advice and encouragement, and Béla for his great generosity, care and help. I am extremely grateful to Béla for bringing me inside to his wonderful mathematical life, for sharing his passion and his knowledge of mathematics with me and for the fantastic research visits he has organised all over the world. I would also like to thank Béla and his wife, Gabriella Bollobás, for being such pleasant hosts at the innumerous parties I have attended at their house.

Secondly, I cannot thank enough three great friends: Natasha Morrison, Pedro Pamplona and Jon Noel. Without their friendship, wise advice, encouragement and help, this journey would probably not have seen an end.

I would also like to express my gratitude to Mrs Tricia Simmons for being always so helpful.

I am also grateful to all my collaborators whom I have the pleasure to be friends with: Kamil Popielarz, Julian Sahasrabudhe, Richard Snyder, Stephen Smith, Bhargav Narayanan, Gabór Mészáros, David Lewis, Shoham Letzter, Ross J. Kang and Teeradej Kittipassorn.

Finally, I would like thank all my friends who have certainly made my life much happier throughout the last 4 years: (Tia) Paula Bessa, Andrew Thomason, João Pedro Saraiva, Miguel Natário, Leonor Guedes de Oliveira, Aldo Pachiano, Miguel Ribeiro, Frederico Gama, João Saraiva, Nuno Miguel Pires, Margarida Feijó, Inês Sampaio, Filipe Martins, Pedro João Lemos, Houri Tarazi, Alex (the first) Broadhurst, Holly Freeman, Tommaso Cremaschi, Alessandro Masullo, Miguel Cardoso de Brito, Vytautas Gruslys, Joana Ferreira, Joana Nunes and Betina Sousa.

Lastly, I would like to thank my uncle, Tio Rui, whose conversations have always inspired me.

Contents

Chapter 1. Introduction 1

1. Extremal and structural graph theory 1

Chapter 2. Extending partial edge colourings 13

1. Introduction 13
2. Notation and results 16
3. Precolouring a set of far apart edges 19
4. Concluding remarks 23

Chapter 3. Extending partial edge colourings of planar graphs 25

1. Introduction 25
2. $\Delta + 1$ colours are necessary 28
3. Precoloured edges must be at distance at least 2 29
4. Proof of the main result 31
5. Avoiding prescribed colours on a matching 37
6. Concluding remarks 41

Chapter 4. Ramsey colourings 43

1. Introduction 43
2. Notation 46
3. Partitioning into monochromatic connected subgraphs 47
4. Covering with monochromatic components of distinct colours 52
5. Covering a two coloured graph by few monochromatic components 58
6. Concluding remarks 60

Chapter 5. Rainbow saturation colourings 61

1. Introduction 61

2. Organization and notation 66
3. Lower bounds 67
4. Upper bounds for connected graphs 71
5. Upper bounds for disconnected graphs 88
6. Deducing the main results 92
7. Concluding remarks 93

Chapter 6. Long cycles in Hamiltonian graphs 95
1. Introduction 95
2. Organization, notation and preliminaries 97
3. Proof of the main result 99
4. Concluding remarks 110

Chapter 7. Large induced subgraphs with k vertices of almost maximum degree 113
1. Introduction 113
2. Notation and preliminaries 114
3. Proof of the main result 115
4. Concluding remarks 120

Chapter 8. Majority colourings of digraphs 121
1. Introduction and theorems 121
2. Code for the linear program 126

Chapter 9. Highly linked tournaments 127
1. Introduction 127
2. Organization and notation 130
3. Embedding a subdivided complete directed graph 130
4. Proof of the main theorem 135
5. Concluding remarks 145

Bibliography 147

CHAPTER 1

Introduction

In this dissertation, we investigate several graph theoretical problems sharing a common 'extremal' flavour. We shall now describe these topics in more depth and give an overview of the contents of each chapter.

1. Extremal and structural graph theory

Extremal graph theory is a branch of graph theory which is mainly concerned with understanding the relations between various graph parameters. It is fair to say that its systematic study was initiated by Paul Turán who, in 1940, proved the now famous Turán's Theorem [109]. Since then, very much influenced by Erdős numerous results and continued later by Bollobás, a lot of research has been devoted to the study of extremal problems of graphs. We refer the reader to the monograph by Bollobás [21], for a comprehensive survey on the subject.

In the first two chapters, we will be considering some problems regarding *proper edge colourings* of graphs. Given a graph G, G is said to be k*-edge colourable* if there exists a function $\phi : E(G) \to \{1, 2, \ldots, k\}$ satisfying the following property:

(1) For any two edges $e, e' \in E(G)$ sharing an endpoint, $\phi(e) \neq \phi(e')$.

Moreover, given a set of colours $\mathcal{K}$, a colouring $\phi : E(G) \to \mathcal{K}$ is said to be a *proper edge colouring* of G if ϕ satisfies (1). As usual, we denote by $\chi'(G)$ the chromatic index of G, defined as the smallest k for which G is k-edge colourable.

A cornerstone theorem in the area due to Vizing [114] dating to 1964 states that for any graph G of maximum degree $\Delta(G)$, $\chi'(G)$ equals Δ or $\Delta + 1$.

In Chapters 2 and 3, we study the following problem: given a graph G and a palette of colours $\mathcal{K}$, when can a partial proper edge colouring of G using colours from $\mathcal{K}$, be extended to a proper edge colouring of the entire graph G? Note that trivially the size of $\mathcal{K}$ must be at least $\chi'(G)$. Furthermore, it is not too hard to construct, for infinitely many integers Δ, a graph G with $\chi'(G) = \Delta(G)$, containing two edges which must have distinct colours in every proper Δ-edge colouring of G. Moreover, these two edges may be arbitrarily far apart. These examples indicate that $\mathcal{K}$ should have size at least $\Delta + 1$, if we want to answer the above problem without assuming any structural conditions on the graph G. We define the distance between two given edges $e, e' \in E(G)$ as the smallest number of edges in a path between two of their endpoints.

In Chapter 2, based on joint work with Ross J. Kang, our main result states the following. Fix a palette of colours $\mathcal{K}$ with $\Delta + 1$ colours and let G be a graph with maximum degree Δ, also let M be a subset of the edge set of G for which any two edges in M are at distance at least 9. Then, if the edges of M are arbitrarily precoloured from $\mathcal{K}$, i.e. the edges of M induce a (partial) proper edge colouring using colours from $\mathcal{K}$, then there is guaranteed to be a proper edge colouring using only colours from $\mathcal{K}$ that extends the precolouring on M to the entire graph. The main idea of the proof is to use a controlled recolouring process inside a small neighbourhood of the precoloured edges and to do so, we make use of a tool introduced by Vizing which we call a *multifan*. We remark that our main result comes close to proving a conjecture of Albertson and Moore [7]. Finally, using similar methods, we are able to lower the condition on the distance between any two precoloured edges to 5 provided the ground graph G does not contain a cycle on 5 vertices.

In Chapter 3, we continue investigating the above mentioned precolouring problem. Firstly, we show that the size of the palette of colours in the main result of Chapter 2 is optimal. Indeed, we construct for infinitely many integers Δ, *bipartite* graphs of maximum degree Δ containing two edges which which must have different colours in every proper Δ-edge colouring, whose distance may be arbitrarily large. Secondly, we exhibit a collection of graphs with a *non-extendable* (partial) edge colouring of a matching. Note that the edges of a matching have pairwise distance greater or equal than one. This implies that the smallest distance between any two precoloured edges to guarantee an extension to an entire proper edge colouring must be at least 2.

As our main result of the chapter we prove that provided the maximum degree is large enough we can lower the distance condition to 3 for the class of planar graphs. Indeed, we show that if G is a planar graph of maximum degree $\Delta(G) \geq 23$, then using a palette of size $\Delta(G)$, any precolouring whose edges have pairwise distance at least 3 is extendable. It is easy to see this result is best possible regarding the size of $\mathcal{K}$ and the distance between precoloured edges.

Finally, suppose that we wish to obtain a general 'precolouring extension' result when the precoloured edges form a matching. It seems natural to ask the following weaker question: given a graph G, suppose the edges of a matching $M \subset E(G)$ are arbitrarily coloured using colours from $\mathcal{K}$, is there a proper colouring of all edges of G (using colours from $\mathcal{K}$) that *differs* from the given colouring on every edge of M? We answer this question in the affirmative when $\mathcal{K} = [\Delta + 1]$. This chapter is based on a joint work with Katherine Edwards, Jan van den Heuvel, Ross J. Kang, Gregory J. Puleo and Jean-Sébastien Sereni.

In Chapter 4, we will study some problems which lie in the area of Ramsey theory. This topic originated with the classical result of Ramsey [88], from 1930, which states that whenever the edges of a countable infinite complete graph are finitely coloured, one can always find a complete infinite subgraph all of whose edges have the same colour. This theorem has by now been extended in several ways. For a survey of many of these generalisations, we refer the reader the book of Graham, Rothschild and Spencer [54].

Erdős and Rado observed in the 50s that for any graph G, either G or its complement forms a connected graph. In other words, any 2-coloured complete graph contains a spanning connected monochromatic subgraph. This simple remark has been the starting point of extensive research. A natural follow up is the search for a (covering) partition of an r-edge coloured (complete) graph into certain kinds of monochromatic substructures where the number of parts does not depend on the order of the graph. Indeed, an important example appears in a seminal paper by Erdős, Gyárfás and Pyber [47] from 1991, who showed that for any r-colouring of a complete graph, the vertex set can be partitioned into at most $O(r^2 \log r)$ monochromatic cycles and conjectured that r cycles suffice. The first non-trivial case of this conjecture for $r = 2$ was resolved in a strong form by Bessy and Thomassé [101] by showing that any 2-coloured complete graph can be partitioned into at most 2 monochromatic cycles of distinct colours. Unfortunately, this conjecture fails, although not by far, for values of r greater than 3, as shown by Pokrovskiy [85]. Following Erdős and Rado's observation it is natural to pose the following question. What is the smallest integer m such that one can partition any r-coloured complete graph into at most m parts where each part induces a connected monochromatic subgraph? Indeed, Erdős, Gyárfás and Pyber conjectured that m is smaller or equal to $r - 1$ [47] and in 1996, Haxell and Kohayakawa [62] showed r parts are enough.

In Chapter 4, a joint work with Shoham Letzter and Julian Sahasrabudhe, we investigate a similar problem for r-coloured graphs of high minimum degree. Our main result, establishing a conjecture of Bal and DeBiasio [14], shows there exists an integer n_0 such that every 2-coloured graph G on $n \geq n_0$ vertices and with minimum degree at least $\frac{2n-5}{3}$ can be partitioned into two monochromatic connected subgraphs. This result is seen to be sharp both for the number of pieces and for the minimum degree, by a construction of Bal and DeBiasio [14]. We may think of this result as saying that $\frac{2n-5}{3}$ is the minimum degree 'threshold' that guarantees a partition of every 2-coloured graph into two monochromatic connected subgraphs In the proof, we make use of probabilistic tools which allow us to show a stability result, namely that any potential counterexample to Bal and DeBiasio conjecture would have to look like the unique extremal example.

As an easy corollary of a clever observation of Gyárfás [55] from 1997, we shall show that for any positive integer $t \geq 2$, if G is a graph on n vertices with minimum degree greater or equal to $\frac{2n-2t-1}{t+1}$, then in any 2-edge colouring of G, we can cover the vertex set by at most t monochromatic components. We also give constructions, showing this inequality cannot be improved.

In a related topic, Bal and DeBiasio [14] studied the problem of covering the vertex set of an r-coloured graph by monochromatic components of *distinct* colours. Note that any r-coloured *complete* graph has such a covering by taking all monochromatic stars rooted at some vertex. They conjectured that $(1-2^{-r})n$ is the minimum degree 'threshold' to guarantee such a covering for graphs on n vertices and gave examples showing the bound if true is tight. Our final result of the chapter confirms the correctness of this conjecture for 3 colours.

Subsequently in Chapter 5, we shall discuss a coloured saturation problem proposed by Barrus, Ferrara, Vandenbussche, and Wenger [17]. Given a collection of graphs $\mathcal{H}$ and a graph G, G is said to be $\mathcal{H}$-*saturated* if it does not contain any member of $\mathcal{H}$ as a subgraph but the addition of any new edge forms such a copy. In other words, G is a maximal $\mathcal{H}$-free graph. The problem of minimising the number of edges in maximal $\mathcal{H}$-free graphs on n vertices was initially investigated in 1949 by Zykov [117] and independently in 1964 by Erdős, Hajnal and Moon [48]. Indeed, they showed that the minimum number of edges in a K_r-*saturated* graph on n vertices is exactly $\binom{n}{2} - \binom{n-r+2}{2}$. Bollobás later proposed the study of $\mathcal{H}$-saturated graphs for a general family of graphs. In particular, writing $sat(n, \mathcal{H})$ for the minimum number of edges in a maximal $\mathcal{H}$-free graph, he conjectured in 1969 that $sat(n, \mathcal{H}) = O(n)$ for every family of graphs $\mathcal{H}$. This conjecture was confirmed by Kászonyi and Tuza [71] in 1986. Moreover, in independent papers, from 1967, Bollobás [20] and Wessel [116] proved the conjecture of Erdős, Hajnal and Moon concerning the saturation function of a complete bipartite graph in a bipartite ground graph. For further information on this topic we refer the reader to the survey of Faudree, Faudree, and Schmitt [49].

In Chapter 5, based on joint work with David Lewis and Kamil Popielarz, we will investigate the following coloured saturation problem. Given a graph H and a palette of colours $\{1, 2, \ldots, t\}$ (for some integer $t \geq |E(H)|$), we denote $\mathfrak{R}(H)$ to be the collection of all rainbow copies of H using colours from $\{1, 2, \ldots, t\}$. The problem consists in finding the numbers $\mathrm{sat}_t(n, \mathcal{R}(H))$, defined as the minimum number of edges in a t-edge coloured graph on n vertices which does not contain a rainbow copy of H but the addition of any non-edge in any colour from $\{1, 2, \ldots, t\}$ creates such a copy. We shall call $\mathrm{sat}_t(n, \mathcal{R}(H))$ the *t-rainbow saturation number* of H. As in the previous chapter, here, a t-edge colouring does not need to be a *proper* edge colouring. In the first part of Chapter 5, we show that $\mathrm{sat}_t(n, \mathcal{R}(K_r)) = \Theta(n \log n)$, for any r and $t \geq \binom{r}{2}$, thus confirming a conjecture of Barrus, Ferrara, Vandenbussche, and Wenger [17]. In the proof, we reduce the problem to estimating the minimum number of bipartite graphs needed to cover the non-edges of the *saturated* coloured graph. Observe that the growth rates of the rainbow saturation numbers behave very differently from the usual saturation numbers. Recall the result of Kászonyi and Tuza [71] who proved that for any class of graphs $\mathcal{H}$, the $\mathcal{H}$-saturation numbers are always linear in the number of vertices of the graph.

We are also able to prove that stars are the unique graphs H, without isolated vertices, which satisfy $\mathrm{sat}_t(n, \mathcal{R}(H)) = \Theta(n^2)$, for some $t \geq e(H)$, thus answering a question appearing in [17].

Finally, we find, as a corollary of our main theorem, the correct order of $\mathrm{sat}_t(n, \mathcal{R}(H))$, for any connected graph H with minimum degree at least 2 and for any fixed $t \geq e(H)$. Indeed, for any such graph H, we prove that

$$sat_t(n, \mathcal{R}(H)) = \begin{cases} \Theta(n \log(n)) & \text{if every edge of } H \text{ belongs to a triangle,} \\ \Theta(n) & \text{otherwise} \end{cases}$$

for any $t \geq e(H)$.

In Chapter 6, we will be studying a problem related to an important property of graphs called *Hamiltonicity*. A graph is said to be *Hamiltonian* if it contains a cycle which passes through every vertex of the graph. As usual, a rather natural extremal problem concerns estimating the maximum number of edges a graph G on n vertices can have provided G is not Hamiltonian. This problem turns out to have a quite simple and unfortunately rather unsatisfying answer. One may then pose the following much more interesting question: given a non-Hamiltonian graph G, how large can the minimum degree of G be? In 1952, Dirac [43] resolved this question completely; he showed that any graph of minimum degree at least $\frac{n}{2}$ is Hamiltonian. This is easily seen to be sharp by taking two disjoint cliques on $\lceil n/2 \rceil$ and $\lfloor n/2 \rfloor$ vertices. Some years later, in 1960, Ore [84] proved a stronger result which generalized Dirac's Theorem. He showed that any graph on n vertices for which $d(x) + d(v) \geq n$, for every pair of non-adjacent vertices x, y, must be Hamiltonian.

In 1946, Cedric Smith [110] proved that every edge of a cubic graph is contained in an even number of Hamiltonian cycles, implying that any cubic Hamiltonian graph contains a second (actually at least 3) distinct Hamiltonian cycles. This result inspired Sheehan [94] to conjecture, in 1975, that every 4-regular Hamiltonian graph contains a second Hamiltonian cycle. Although this conjecture is still open it has led to many discoveries by several authors. For example, in 1978, Thomason [100] proved, using a beautiful argument, that every d-regular graph (for odd d) contains a second Hamiltonian cycle. Some years later, Thomassen [108] showed the same holds provided d is sufficiently large. In the light of these results one is tempted to ask if regularity is genuinely necessary to force the existence of a second Hamiltonian cycle, or if a weaker condition on the minimum degree might suffice. In particular, one might ask the following natural question: does every Hamiltonian graph G with $\delta(G) \geq 3$ contain a second Hamiltonian cycle? Unfortunately, this is false as shown by a construction of Entringer and Swart [45] dating to 1980.

In Chapter 6, based on a work joint with Teeradej Kittipassorn and Bhargav Narayanan, we show, however, that an asymptotic version of this question holds, in particular confirming Sheehan's conjecture asymptotically, under a much weaker assumption. Indeed, we prove that if an n-vertex graph G of minimum degree at least 3 contains a Hamiltonian cycle, then G must contain another cycle of length at least $n - Cn^{\frac{4}{5}}$, where $C > 0$ is an absolute constant. The proof of this result splits into two regimes: when the graph contains 'many' pairs of *interlacing* chords, a constructive argument allows us to find a long cycle using very few chords, in the other case we rely on a theorem of Thomassen, which is based on the parity-based 'lollipop argument' of Thomason.

In Chapter 7, based on a joint work with Kamil Popielarz, we shall prove an approximate version of a conjecture due to Caro and Yuster [37]. We show that every graph G contains a 'large' induced subgraph $H \subseteq G$ contaning many vertices of the same degree of order 'almost' $\Delta(H)$.

Observe trivially that any graph must contain at least two vertices of the same degree. As there are arbitrarily large graphs which contain exactly two vertices of the same degree, it is natural to ask what is the smallest number of vertices one needs to delete from a graph to ensure that the remaining induced graph is either empty or contains at least k vertices of the same degree. Such question was partially answered by Caro, Shapira and Yuster in [35]. They showed 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.

In the same vein, Caro and Yuster [37] considered the problem of finding the largest induced subgraph H (possibly empty) of a graph G which contains at least k vertices of degree $\Delta(H)$. More precisely, they investigated the size of the smallest number $f_k(n)$ such that given any graph on n vertices, one can delete at most $f_k(n)$ vertices to guarantee the existence of at least k vertices of maximum degree in the induced subgraph. These authors showed $f_2(n) = \Theta(\sqrt{n})$ and conjectured $f_k(n) = O_k(\sqrt{n})$. It is not difficult to construct examples which show $f_k(n) \geq \frac{k}{2}\sqrt{n}$. In Chapter 7, we answer an approximate version of a slightly stronger conjecture posed by Caro, Lauri and Zarb [34]. Indeed, we prove that for every k, there exists $g(k)$ such that in every graph G of maximum degree Δ, there exists an induced subgraph $H \subseteq G$ on at least $n - g(k)\sqrt{\Delta}$ vertices containing at least k vertices of degree at least $\Delta(H) - g(k)$. This result is optimal up to the order of the function $g(k)$.

In Chapter 8, based on a joint work with Teeradej Kittipassorn and Kamil Popielarz, we study a problem proposed by Kreutzer, Oum, Seymour, van der Zypen and Wood in [74]. The problem consists in finding, for every $c \in [0,1]$, the smallest integer $m = m(c)$ such that any digraph D can be partitioned into at most k parts with the property that every vertex $x \in D$ sends at most $c \cdot d^+(x)$ out-edges to its part. The equivalent problem for undirected graphs is easily solvable. Indeed, it is easy to see that for any $c \in [0,1]$, any graph G may be partitioned into at most $\lceil c^{-1} \rceil$ parts with the property that any vertex $x \in V(G)$ contains at most $c \cdot d(x)$ neighbours in its part. Moreover, this bound is tight by taking a complete graph. For directed graphs Kreutzer, Oum, Seymour, van der Zypen and Wood gave some bounds on $m(c)$ and asked whether $m(c) = O(c^{-1})$. In this short chapter we solve completely this question up to an additive constant of 1. We shall show that $m(c) \in \{2\lceil c^{-1} \rceil - 1, 2\lceil c^{-1} \rceil\}$. Note that a $(\lceil c^{-1} \rceil - 1)$-regular tournament on $2\lceil c^{-1} \rceil - 1$ vertices implies the lower bound of $2\lceil c^{-1} \rceil - 1$, since no part can contain more than a vertex. Our proof follows almost immediatelly from a result due to Keith Ball on partitions of matrices. We remark that quite unfortunately we do not even know if $m(\frac{1}{2}) = 3$. We are, however, able to show that we can partition any tournament into three parts such that all but at most 7 vertices have at most a half proportion of its outneighbours in its part. This suggests, at least for tournaments that $m(\frac{1}{k}) = 2k - 1$. Now, note that we could view a partition of the (directed) graph as a colouring of the vertex set. This reformulation suggests a natural variant of the problem, namely when every vertex is instead given a list of colours. Indeed, we show an analogous result holds in this more general setting.

To conclude this dissertation we shall investigate a concept called k-*linkedness*.

Connectivity is probably one of the most important notions in graph theory. Throughout the last century, there has been a lot of research devoted to studying various functions that measure how 'strongly' connected a graph can be. One of them is k-*connectivity*. A graph is said to be k-*connected* if it remains connected after the removal of any set of $(k-1)$-vertices. The first important result in the study of k-connected graphs was Menger's Theorem [82], dating to 1937. It states that a graph is k-connected if and only if there are k-internally disjoint paths joining any pair of vertices. This result provides a nice characterization of the structure of k-connected graphs.

An easy corollary of Menger's Theorem says that for any k-connected graph G and any two disjoint sets of vertices $\{x_1, \dots, x_k\}$ and $\{y_1, \dots, y_k\}$ there are vertex-disjoint paths $P_1, \dots, P_k$ such that P_i goes from x_i to $y_{\sigma(i)}$ for some permutation σ of $[k]$. Notice, however, that one has no control over the permutation σ, i.e. we have no control over the endvertices of the paths P_i. This remark suggests the notion of k-*linkdeness*. A graph is said to be k-*linked* if for any two disjoint sets of vertices $\{x_1, \dots, x_k\}$ and $\{y_1, \dots, y_k\}$ there are vertex disjoint paths $P_1, \dots, P_k$ such that P_i goes from x_i to y_i. Note that a k-linked graph is trivially k-connected. Therefore, it is natural to ask whether any k-connected graph must be t-linked for some $t \leq k$. Indeed, in 1974, Larman and Mani [76], and Jung [66] answered this question by showing that for every k, there is an integer $f(k)$ such that any $f(k)$-connected graph is k-linked. In order to prove their result, they used a theorem of Mader [80] which guarantees a subdivision of a 'large' complete graph in every graph of sufficiently high average degree. The first bounds on $f(k)$ were exponential in k. However, in 1996, in a great breakthrough, Bollobás and Thomason [25] showed that a linear bound on the connectivity suffices. More precisely, they proved that as long as a graph G is $2k$-connected and has average degree at least $22k$ then G is k-linked. A tournament is an oriented complete graph. In Chapter 9, based on a joint work with Richard Snyder, we shall study the notion of k-linkedness in the context of *tournaments*. Observe first that all notions described above have anologous definitions for directed graphs, and indeed Menger's Theorem remains valid for directed graphs.

A directed graph is said to be *k-linked* if for any two disjoint sets of vertices $\{x_1, \ldots, x_k\}$ and $\{y_1, \ldots, y_k\}$ there are vertex disjoint paths $P_1, \ldots, P_k$ such that P_i is a *directed* path from x_i to y_i. In a surprising result, Thomassen [104] constructed an infinite set of digraphs arbitrarily high strongly-connected which are not even 2-linked. So, it seems natural to investigate this problem restricted to the class of tournaments. Indeed, Thomassen was the first to show, in 1984, that there is a constant C such that every $Ck!$-connected tournament is k-linked. Later, Kühn, Lapinskas, Osthus, and Patel [75] improved the bound to $10^4 k \log k$ and finally, in 2015, Pokrovskiy [87] proved that a linear bound is enough. Moreover, in [87], Pokrovskiy conjectured, by analogy with the undirected case, that every $2k$-strongly connected tournament with sufficiently high minimum in-degree and out-degree is k-linked. In Chapter 9, we come close to proving this conjecture. We show that every $4k$-strongly connected tournament with high minimum out-degree is k-linked, thus reducing the bound on the connectivity to within a constant of 2 of the theoretical minimum.

CHAPTER 2

Extending partial edge colourings

1. Introduction

Recall that the classical theorem on proper edge colourings due to Vizing [114] states that the chromatic index $\chi'(G)$ of a simple graph G is either the maximum degree or one larger. This theorem inspired multiple lines of research in the area of edge colourings. I shall mention two of them. We refer the reader to the book by Stiebitz, Scheide, Toft and Favrholdt [98] for detailed references and fuller insights.

The first line is concerned with proper edge colourings of *multigraphs*. A multigraph is a graph where multiple edges are allowed, for our purposes, we exclude loops. As usual we denote by $\mu(G)$ the maximum number of parallel edges sitting on an edge. Observe the trivial lower bound $\Delta \leq \chi'(G)$ holds for any multigraph G.

About a century ago, Kőnig proved that all bipartite multigraphs meet this lower bound with equality. A few decades later, in 1949, Shannon [93] showed that $\chi'(G) \leq \left\lfloor \frac{3}{2}\Delta(G) \right\rfloor$, for any multigraph G. Somewhat later, Vizing proved that $\chi'(G) \leq \Delta(G) + \mu(G)$, for any multigraph G, implying that $\chi'(G) \in \{\Delta(G), \Delta(G)+1\}$, when G is a simple graph. Both Shannon's and Vizing's bounds are tight; this can be seen by taking a triangle with the same number of parallel edges on each edge.

Some years later, an important conjecture arose. The Goldberg–Seymour Conjecture, due to Goldberg [52] and Seymour [92], asserts that $\chi'(G) \leq$

max{ $\Delta(G)+1, \lceil \rho(G) \rceil$ } for any multigraph G where

$$\rho(G) = \max\left\{ \frac{2|E(G[T])|}{|T|-1} : T \subseteq V,\ |T| \geq 3,\ |T| \text{ odd} \right\}.$$

Note that $\rho(G)$ is always a lower bound for $\chi'(G)$. Indeed, since every colour class in a proper edge colouring forms a matching, it contains at most $\lfloor \frac{|T|}{2} \rfloor$ edges inside any subset $T \subset V(G)$. Therefore, one needs at least $\rho(G)$ distinct colours to proper edge colour G. This conjecture remains open and is regarded as one of the most important open problems in the area of edge colourings. Perhaps the most remarkable progress on this problem is due to Kahn [69], in 1996, who established it in an asymptotic form.

Another important concept we would like to mention and which has motivated a lot of research is a generalization of the concept of $\mathcal{K}$-edge colourings, namely list colourings, where each edge is allowed to have its own set of colours. Indeed, given a graph G and a set of lists $\mathcal{L} = \{L_e : e \in E(G)\}$, where each edge of G has its own list L_e, we say G is $\mathcal{L}$-edge colourable if we can find an edge colouring of G where every edge is coloured with an element from its list. Moreover, we say G is *k-edge choosable* if for any family $L = \{L_e : e \in E(G)\}$ with $|L_e| = k$, for every $e \in E(G)$, G is L-edge colourable. Finally, we define the least integer integer k for which G is k-edge choosable to be the *list chromatic index* of G and denote it by χ_l. A famous conjecture, usually called the List Colouring Conjecture, states that for any simple graph G, $\chi' G) = \chi_l$. The List Colouring Conjecture (LCC) was already formulated by Vizing as early as 1975 and has been reformulated several times. We will return in Chapter 3 to the LLC, where we shall explain the bridge to the theory of edge precolourings.

We define the distance between two edges to be the smallest number of edges between two of their endpoints. In this chapter, we are interested in studying the following problem. Let G be a graph of maximum degree Δ. Does there exist an integer d such that, if G has a partial proper edge colouring ϕ using $\Delta+1$ colours on a set of edges whose pairwise distance is at least d, then ϕ extends to a $(\Delta+1)$-proper edge colouring of all of G? Albertson and

Moore [7] conjectured that such a constant d exists and equals 3, noting that the first non-trivial case, when $\Delta(G) = 3$, follows from [67].

Our main goal in the present chapter is to show that such a constant d does indeed exist and it is at most 9. In Chapter 3, we will see that d must be at least 2 and that no such constant exists if we instead only allow $\Delta(G)$ colours for graphs G satisfying $\chi'(G) = \Delta(G)$.

We should point out that our results fit within the broader context of vertex precolouring extensions, which has seen a great deal of activity, (see e.g. [6, 112]). This is an important topic in chromatic graph theory, especially due to Thomassen's ingenious use of a precolouring extension to prove that all planar graphs are 5-choosable [105]. Notably, in a strikingly short answer to a related question of Thomassen, Albertson [4] showed that, given a graph G with chromatic number $\chi(G) = r$ and a set P of vertices with pairwise distance at least 4, any precolouring of P from a palette of $r + 1$ colours extends to a proper $(r + 1)$-colouring of G. Moreover, the distance condition of 4 is best possible.

We should stress that the majority of previous work on (vertex-)precolouring extension allows for one *additional* colour, as in Albertson's seminal result. We highlight two relevant exceptions. Albertson and Moore [7] considered how to extend partial r-colourings of r-chromatic graphs, and proved several sharp results; however, these results only apply to graphs that possess a special r-colouring, namely an r-colouring where no vertex is adjacent to two vertices coloured with r. Later, Axenovich [13] and Albertson, Kostochka and West [5] proved that Δ-precolourings of a set of vertices with minimum distance 8 can be extended to full Δ-colourings for any graph of maximum degree at most Δ ($\Delta \geq 3$), apart from $K_{\Delta+1}$, thereby extending Brooks' Theorem [33].

Finally, we point out that perhaps the most important edge precolouring problem already considered is related to Evans Conjecture. Although Evans Conjecture was originally formulated in terms of latin squares, here, we will rephrase it in terms of precolourings of bipartite graphs. The conjecture states that any (partial) proper edge colouring of at most $n-1$ edges of the complete bipartite graph $K_{n,n}$ can be extended to a proper n-edge colouring of the entire bipartite graph. This conjecture was confirmed by Häggkvist [58], in 1978, for every sufficiently large n, and later by Smetanuik [96] and independently by Anderson and Hilton [10] in full generality.

2. Notation and results

Our notation is standard. We denote $[k] = \{1, 2, \ldots, k\}$. A *multigraph* is a graph where multiple edges are allowed. For our purposes, we exclude the existence of loops. Recall that we have defined the distance between two given edges $e, e' \in E(G)$ as the smallest number of edges in a path between two of their endpoints. Given a multigraph G we denote by $\mu(G)$ the maximum multiplicity of G i.e. the maximum number of parallel edges and by $\Delta(G)$ the maximum degree of G. Given a multigraph G and a proper edge colouring ϕ of G with palette of colours $\mathcal{K}$, we let

$$\phi(z) = \{c \in \mathcal{K} : \text{ there exists an edge } e \text{ incident with } z \text{ with } \phi(e) = c\}$$

and $\overline{\phi}(z) = \mathcal{K} \setminus \phi(z)$, in other words, those colours not seen by the vertex z.

Finally, we define a *multifan* F at a vertex z with respect to an edge $e = (z, v)$ and an edge colouring ϕ as a sequence $F = (f_1, x_1, \ldots, f_p, x_p)$, for some integer $p \geq 1$, where $f_1, \ldots, f_p$ are distinct edges, $f_1 = e$ and f_k has endpoints z and x_k, for all $k \in [p]$, and in addition for every edge $f_k \neq e$, there exists a vertex x_l for $l \in [k-1]$ such that $\phi(f_k) \in \overline{\phi}(x_l)$. We remark that the concept of multifans is crucial in Vizing's proof of Vizing's Theorem. We shall sketch, in an informal way, Vizing's original proof, since some of his ideas will be important for our proofs. We state now Vizing's Theorem [114], for multigraphs.

Theorem 2.1 (Vizing). *Let G be a multigraph with maximum degree $\Delta(G)$ and maximum multiplicity μ. Then G is $(\Delta + \mu)$-edge colourable.*

Proof. (**sketch**)

Suppose the theorem does not hold. Let G be a counterexample with the smallest number of edges. Remove some edge $e = xy \in E(G)$, and let ϕ be a $(\Delta+\mu)$-edge colouring of $G - e$. Now, let $F = (f_1, x_1, \ldots, f_p, x_p)$ be a maximal multifan at vertex y with respect to the edge $e = f_1$, where $x_1 = x$.

Claim 2.2. *$\overline{\phi}(x_i) \cap \overline{\phi}(x_j) = \varnothing$, for every $i \neq j$.*

Suppose not. Then, there exists some colour $c \in \overline{\phi}(x_i) \cap \overline{\phi}(x_j)$, for some $i < j$. Moreover, let $c_1 \in \overline{\phi}(y)$, which must exist. Consider in $G - e$, the connected component Y containing y which is spanned by the edges of colours c and c_1. Clearly, at most one of x_i or x_j belongs to Y; we may assume $x_j \notin Y$. We interchange the colours c and c_1 in the connected component containing x_j spanned by edges of colours c and c_1. In this new edge colouring both x_j and y miss colour c. By the definition of a multifan, there exists a sequence $(e, f_{i_1}, f_{i_2}, \ldots, f_{i_k} = f_j)$ such that $\phi(f_{i_l}) \in \overline{\phi}(x_{i_{l-1}})$, for every $l \in [k]$. We may then 'shift' the colours along those edges and colour the edge f_j with colour c, thus constructing a proper edge colouring of G, which contradicts our assumption.

Now, since F is a maximal multifan, we must have that for every colour $c \in \bigcup_{i=1}^{p} \overline{\phi}(x_i)$, there exists an edge f_{k_i} incident with y of colour c, otherwise a simple recolouring argument by'shifting' colours along the multifan F, would allow us to construct a proper edge colouring of G. To conclude the proof observe that total number of coloured edges between the set $\{x = x_1, x_2, \ldots, x_p\}$, the vertices of F, and y is at most $p \cdot \mu$. Since $|\overline{\phi}(x_i)| \geq \mu$, for every $2 \leq i \leq p$ and $|\overline{\phi}(x)| \geq \mu + 1$, we obtain a contradiction. $\square$

Here is our main result, restated slightly in a more general form, in terms of multigraphs.

Theorem 2.3. *Let G be a multigraph of maximum edge multiplicity μ and maximum degree Δ and let M be a set of edges such that the minimum distance* *between any two edges of M is at least 9. If M is arbitrarily precoloured from the palette $\mathcal{K} = [\Delta + \mu] = \{1, \ldots, \Delta + \mu\}$, then there is a proper edge colouring of G using colours from $\mathcal{K}$ that agrees with the precolouring on M.*

To prove Theorem 2.3,we make use of the following result of Berge and Fournier [18], which is an edge precolouring extension result when all precoloured edges have the same colour.

THEOREM 2.4 (Berge and Fournier). *Let G be a multigraph of maximum edge multiplicity μ and maximum degree Δ and let M be a maximal matching of G. Then there exists a proper $(\Delta + \mu - 1)$-edge colouring of $G \setminus M$.*

To prove our main theorem we will roughly take the following strategy. Let $M' \supseteq M$ be a maximal matching of G. By Theorem 2.4, there is a proper edge colouring of $G \setminus M'$ using only colours from $[\Delta + \mu - 1]$. We would like to colour all edges of $M' \setminus M$ with colour $\Delta + \mu$. Moreover, for every $1 \leq i \leq \Delta + \mu - 1$, if there is an edge of $G \setminus M'$ that is coloured i and incident to some edge precoloured i, then we would like to recolour that edge with the colour $\Delta + \mu$. In the proof, we use a recolouring argument, using multifans, to help us resolve the problems that may arise.

Using the same strategy, we also show how we can afford to relax the distance constraint on the precoloured matching provided we impose a mild structural constraint on the graph.

THEOREM 2.5. *Let G be a multigraph of maximum edge multiplicity μ and maximum degree Δ. Suppose G contains no cycle of length 5 as a subgraph. Let M be a set of edges such that the minimum distance between any two edges is at least 5. If M is arbitrarily precoloured from the palette $\mathcal{K} = [\Delta + \mu] = \{1, \ldots, \Delta + \mu\}$, then there is a proper edge colouring of G using colours from $\mathcal{K}$ that agrees with the precolouring on M.*

3. Precolouring a set of far apart edges

In this section, we shall prove Theorem 2.3 and Theorem 2.5.

Proof of Theorem 2.3. Let $\Phi : M \to \mathcal{K}$ be a precolouring of M. For each $i \in \mathcal{K}$, we write $M_i \subseteq M$ for the set of edges precoloured with colour i. Now, let α be the cardinality of a matching $M' \supseteq M$ of smallest size for which there exists a proper $(\Delta+\mu-1)$-edge colouring of $G \setminus M'$. Note that by Theorem 2.4, α is well defined.

For a matching $M' \supseteq M$ and a proper edge colouring $\varphi : E(G \setminus M') \to [\Delta+\mu-1]$, we define $\Delta+\mu-1$ sets, $A_1^{M',\varphi}, \dots, A_{\Delta+\mu-1}^{M',\varphi}$ as follows. For every $1 \le i \le \Delta+\mu-1$ and each endpoint u of an edge of M_i, we let $A_u^{M',\varphi}$ denote the set of edges which belong to the maximal path $P_u^{M',\varphi}$ beginning at vertex u and which alternates between edges coloured i (by φ) and edges of $M' \setminus M$.

We shall write $P_u^{M',\varphi} = w_0^u e_0^u w_1^u e_1^u w_2^u e_2^u \cdots$ with vertices w_k^u and edges e_k^u, where $w_0^u = u$. Moreover, we take $A_i^{M',\varphi}$ to be the union of all $A_u^{M',\varphi}$, with u an endvertex of an edge of M_i. Hence, $A_i^{M',\varphi} \cup M_i$ induces a disjoint union of paths and cycles for every $i \in [\Delta+\mu-1]$.

We note that if we find a matching M' and a partial edge colouring φ, as defined above, such that *every* set $A_u^{M',\varphi}$ contains at most one edge, then we are done; indeed, giving colour $\Delta+\mu$ to every edge of $A_u^{M',\varphi}$ and to every edge of $M' \setminus M$ and colouring the edges of M with the prescribed colours. (This is the strategy we described informally before the proof.)

More importantly, we would also be done if we could find M' and φ such that $A_i^{M',\varphi}$ induces a subgraph that is disconnected from that of $A_j^{M',\varphi}$, for every $i \neq j \in [\Delta+\mu-1]$ and also disjoint from $M_{\Delta+\mu}$. Indeed, in this case we would be able to recolour the edges in $A_i^{M',\varphi} \setminus M'$ (those edges originally coloured with i) with colour $\Delta+\mu$ and give colour i to every edge in $A_i^{M',\varphi} \cap M'$, for every $1 \le i \le \Delta+\mu-1$.

We now fix a choice of $M' \supseteq M$ and $\varphi : E(G \setminus M') \to [\Delta+\mu-1]$ such that $|M'| = \alpha$. Moreover, we make our choice so that it minimises the number

β of endpoints u of edges in M for which $|A_u^{M',\varphi}| > 1$, and subject to that, it minimises the number γ of edges $e \in M$ with endpoints u and v for which there is even index t for which at least one of the path vertices w_t^u or w_t^v belonging to $P_u^{M',\varphi}$ or $P_v^{M',\varphi}$, respectively, is at distance greater or equal to 3 from e.

The rest of the proof is devoted to showing that under this choice any two subgraphs induced by $A_i^{M',\varphi}$, for $1 \leq i \leq \Delta + \mu - 1$, share no vertex and are also vertex disjoint from $M_{\Delta+\mu}$.

CLAIM 3.1. *For any edge in M with endpoints u and v, either $|A_u^{M',\varphi}| \leq 1$ or $|A_v^{M',\varphi}| \leq 1$.*

Suppose otherwise. We now construct a maximal multifan pivoting on w_2^u, where our aim is to colour $e_1^u \in (M' \setminus M)$ with a colour from $[\Delta + \mu - 1]$ and adjust the partial colouring φ so that the new colouring becomes a proper $(\Delta + \mu - 1)$-edge-colouring of $G \setminus (M' \setminus \{e_1^u\})$. If this succeeds, then e obtain a contradiction to the minimality of α.

Let us then choose $F = (f_1, x_1, \ldots, f_p, x_p)$ to be a maximal multifan at w_2^u with respect to e_1^u and φ. Note that $p \geq 2$ and moreover, there must exist some x_i which is not incident to an edge of M'. Indeed, if this does not hold, following the same argument used in Vizing's proof (see the proof of Theorem 2.1), we would be able to find a $(\Delta + \mu - 1)$-edge colouring of $G \setminus (M' \setminus \{e_1^u\})$, which contradicts the minimality of α.

We may now 'shift' colours along the multifan so that $f_1 = e_1^u$ receives a colour from $[\Delta + \mu - 1]$ and instead we add f_i to M'. This new choice of M' and φ (which still has $|M'| = \alpha$) contradicts the minimality of β. Indeed, observe that for this new choice, $|A_u^{M',\phi}| \leq 1$. Moreover, if $|A_{u'}^{M',\phi}| \leq 1$, then adding the edge f_p to $M' \setminus \{e_1^u\}$ can not change the size of $A_{u'}^{M',\phi}$, since the distance between any two precoloured edges is at least 9. (In fact, here, we just used the fact that any two precoloured edges are at distance at least 5.) This completes the proof of the claim.

CLAIM 3.2. $\gamma = 0$.

Suppose for a contradiction that there is an edge e in M with endpoints u and v for which there is an even index t at which at least one of the path vertices w_t^u or w_t^v is at distance at least 3 from e. Let t be the smallest such index. From Claim 3.1, we may assume, without loss of generality, that this is only the case for w_t^u. Note that w_t^u has distance 3 or 4 from e, and w_s^u is at distance at most 2 from e, for every $s < t-1$.

We know there is no proper $(\Delta+\mu-1)$-edge colouring of $G \setminus (M' \setminus \{e_{t-1}^u\})$, or else there would be a contradiction with the choice of α. We now choose $F = (f_1, x_1, \ldots, f_p, x_p)$ to be a maximal multifan at w_t^u with respect to e_{t-1}^u and φ. As before, by the same argument as in Vizing's proof, we have that $p \geq 2$ and there is some x_j which is not incident to an edge of M'. We then 'shift' the colours along the multifan so that $f_1 = e_{t-1}^u$ receives a colour from $[\Delta+\mu-1]$ and we add instead f_j to M'.

By construction x_j is at distance between 2 and 5 from $e = (u, v)$. Therefore, in the new choice of M' and φ, the edge $f_j \in M'$ can not be appended to the path $P_v^{M',\varphi}$, which implies that under this new choice of M' and φ, there is no even index t at which either of the path vertices w_t^u or w_t^v is at distance at least 3 from e. On the other hand, we could possibly have appended f_j to another path $P_z^{M',\varphi}$, but we are guaranteed by the distance condition on M that in the old choice there would exist already an even index t for which w_t^z was at distance at least 3 from its corresponding edge in M. Therefore, we were able to reduce γ by one in this new choice of M' and new partial colouring φ, contradicting the the minimality of γ. This completes the proof of the claim.

This last Claim implies that for all $i \in [\Delta+\mu-1]$, every edge of $A_i^{M',\varphi}$ is within distance 3 of an edge of M_i. So by the distance condition on M the subgraphs induced by $A_i^{M',\varphi}$, $1 \leq i \leq \Delta+\mu-1$, are disconnected from one another and from $M_{\Delta+\mu}$, and this completes our proof. $\square$

Looking at the argument we easily see that we could slightly relax the condition in Theorem 2.3 on the precoloured edge set. Indeed, we could demand that M is a disjoint union of matchings M' and M'', where any two edges of M' are at distance at least 9 and M' is arbitrarily precoloured from $[\Delta+\mu-1]$, M'' is precoloured with colour $\Delta+\mu$, and the minimum distance between an edge in M' and an edge in M'' is at least 4.

We turn now to the proof of Theorem 2.5. The proof is conceptually the same as that of Theorem 2.3, but perhaps simpler.

Proof of Theorem 2.5. Just as before, let $\Phi : M \to \mathcal{K}$ be the precolouring on M and let α be the cardinality of a smallest size matching $M' \supseteq M$ for which there exists a proper $(\Delta+\mu-1)$-edge colouring of $G \setminus M'$. Theorem 2.4 certifies that α is well defined.

For any matching $M' \supseteq M$ and any proper edge colouring $\varphi : E(G \setminus M') \to [\Delta+\mu-1]$, we say that an edge $e \in M$ is *bad* if there exist two edges e_1, e_2 such that $\varphi(e_1) = \Phi(e)$, $e_2 \in M' \setminus M$ and e_1 is adjacent to both e and e_2. If there are no bad edges, then we may extend Φ to a proper edge colouring of G by colouring the edges $e \notin M'$ with $\varphi(e)$, any edge $e \in M' \setminus M$ with $\Delta+\mu$, and recolouring with colour $\Delta+\mu$ every edge e which is incident with a precoloured edge e', satifying $\varphi(e) = \Phi(e')$.

We now fix a choice of M' and φ as above, with $|M'| = \alpha$ and, subject to this, having the least number β of bad edges. The rest of the proof is devoted to showing $\beta = 0$.

Suppose $e \in M$ is a bad edge and let e_1 and e_2 be two edges certifying its badness. We have that e, e_1, e_2 form a path of length 3. Calling w_2 the endpoint of this path that is incident with e_2, we let $F = (f_1, x_1, \dots, f_p, x_p)$ be a maximal multifan at w_2 with respect to e_2 and φ. As in the previous proof, by following Vizing's argument, $p \geq 2$ and there is some x_i which is not incident with an edge of M'. By 'shifting' the colours along F, we can colour $f_1 = e_2$ from $[\Delta+\mu-1]$ and add f_p to M'.

Under this new choice of φ and M' (which still has $|M'| = \alpha$), any new bad edge would have to be within distance 1 of f_p and thus within distance

4 of e, contradicting the distance requirement on M. Due to the shift, e can no longer be certified bad with the help of e_1 since e_1 is no longer incident to an edge of M'. However, by the choice of β, it must be that e has remained bad with respect to the new choice of φ and M'. So there must exist e'_1 and e'_2 still certifying that e is bad. Therefore $\{e_1, e_2\} \cap \{e'_1, e'_2\} = \varnothing$ and the union of endpoints of e_1 and e_2 is disjoint from that of e'_1 and e'_2. We have furthermore that $e'_2 \neq f_p$ or else e, e_1, e_2, f_p, e'_1 would form a cycle of length 5. Clearly e'_2 and f_p are not incident as both belong to M'.

We can perform another pivot as before but now at the end of the path $ee'_1e'_2$. Calling w'_2 the endpoint of this path, and taking a maximal multifan $F' = (f'_1, x'_1, \ldots, f'_{p'}, x'_{p'})$ at w'_2 with respect to e'_2 and φ, we again must have $p' \geq 2$ and there must exist a vertex x'_j not incident with an edge of M'. In particular, x'_j and f_p are not incident, moreover x'_j is not the common endpoint of e_1 and e_2 or else there would be a cycle of length 5. Again we 'shift' the colours along F' so as to colour $f'_1 = e'_2$ from $[\Delta + \mu - 1]$ and we add $f'_{p'}$ to M'.

Arguing as before, note that under this second new choice of φ and M' (which still has $|M'| = \alpha$), there is no *new* bad edge. Also, observe that we have now modified φ and M' so that neither e_1 nor e'_1 may help to certify that e is bad. Thus e is no longer bad since in any proper partial edge colouring there are at most two edges incident to e coloured with $\Phi(e)$. This is a contradiction to the choice of β.

We may therefore conclude that $\beta = 0$ and this completes the proof. $\square$

Actually, in the previous proof we have just used the fact that no precoloured edge is contained in a C_5.

4. Concluding remarks

We close this chapter by mentioning few remarks. First, note that the result of Albertson [4] implies that for graphs G, where $\chi'(G) = \Delta$, any partial edge colouring using $[\Delta(G) + 1]$ colours is extendable provided any two precoloured edges are at distance at least 3. Indeed, this follows directly by passing to the line graph and noting that if two edges are at distance at least 3, then the corresponding vertices in the line graph are at distance at least 4.

Secondly, we observe that such an edge extension result using an 'optimal' set of colours must impose some conditions on the precoloured edges. As we shall see in the next chapter, Theorem 2.3 does not hold if we allow precoloured edges to be within distance 1. We actually conjecture that 2 should be correct distance requirement; the following conjecture strengthens the Albertson and Moore Conjecture mentioned in Section 1.

CONJECTURE 4.1. *Let G be a multigraph with maximum degree $\Delta(G)$ and maximum multiplicity $\mu(G)$. Using the palette $\mathcal{K} = [\Delta(G) + \mu(G)]$, any precoloured set of edges of pairwise distance at least 2 can be extended to a proper edge colouring of all of G.*

CHAPTER 3

Extending partial edge colourings of planar graphs

1. Introduction

In this chapter, we continue investigating extensions of a partial edge colouring to an entire proper edge colouring using a small palette of colours. Theorem 2.3 in the previous chapter states that such an extension is always possible provided any two precoloured edges are at distance at least 9. Recall that we allowed a palette of size at least $\Delta + 1$, where Δ denotes the maximum degree of the graph.

One might wonder if a strong enough distance requirement on the precoloured set of edges permits us to take a smaller palette, of size $\Delta(G)$, whenever $\chi'(G) = \Delta(G)$. This fails however, even for bipartite graphs, as we shall show in Section 2. These simple examples establish that the size of the palette in Theorem 2.3 is best possible, at least for simple graphs.

In Section 3, we turn to the problem of estimating the smallest positive integer d such that given a simple graph G of maximum degree Δ, one can always extend a partial edge colouring of any set of edges whose pairwise distance is at least d, to a proper edge colouring of the entire graph G, using at most $\Delta + 1$ colours. It is not too hard to construct, for every integer Δ, a graph G of maximum degree Δ with a *non-extendable* partial proper edge colouring of a matching. Thus implying d must lie between 2 and 9.

The *List Colouring Conjecture* (LCC) was formulated by Vizing as early as 1975 and was independently reformulated several times, a brief historical account of which is given by, e.g., Häggkvist and Janssen [59]. For more on the LCC, particularly with respect to the probabilistic method, we refer to the monograph of Molloy and Reed [83]. The results on the LCC most relevant to us also happen to be two of the most striking, both from the mid-1990s. First, Galvin [51] used a beautiful short argument to prove Dinitz's Conjecture, confirming the LCC for bipartite multigraphs. Not long after Galvin's work, Kahn applied powerful probabilistic methods, with inspiration from extremal combinatorics and statistical physics, to asymptotically prove the LCC [69, 70]. An easy observation stated below relates the problem of extending partial edge colourings to a list edge colouring problem, and therefore to the List Colouring Conjecture. Given a non-precoloured edge, we define its *precoloured degree* as the number of adjacent precoloured edges.

OBSERVATION 1.1. Let G be a multigraph with list chromatic index $ch'(G)$. For a positive integer k, take the palette $\mathcal{K} = [ch'(G) + k]$. If G is properly precoloured so that the precoloured degree of any non-precoloured edge is at most k, then the precolouring can be extended to a proper edge colouring of all of G.

In light of this observation, one may deduce trivially from Galvin's Theorem that Conjecture 4.1, stated at the end of the previous chapter, holds for the class of bipartite multigraphs. We state this theorem for completeness.

THEOREM 1.2. *Let G be a bipartite multigraph with maximum degree $\Delta(G)$. Using the palette $\mathcal{K} = [\Delta(G) + 1]$, any precolouring of a set of edges, containing no two edges within distance 1, can be extended to a proper edge colouring of the entire G.*

Moreover, by Observation 1.1 and assuming the LCC, the following weaker form of Conjecture 4.1 holds as well (where we allow one extra colour): Using the palette $\mathcal{K} = [\Delta(G) + \mu(G) + 1]$, any precoloured set of edges whose pairwise distance is greater or equal than 2 extends to a proper edge colouring of the entire G.

It is known that planar graphs G with $\Delta(G) \geq 7$ satisfy $\chi'(G) = \Delta(G)$. This was proved in 1965 by Vizing [115] in the case $\Delta(G) \geq 8$, and much later by Sanders and Zhao [89] when $\Delta(G) = 7$. Vizing conjectured that the same can be said for planar graphs G with $\Delta(G) = 6$, but this long standing question remains open. Vizing also noted that not every planar graph G with $\Delta(G) \in \{4, 5\}$ is $\Delta(G)$-edge colourable.

Regarding list edge colouring, Borodin, Kostochka and Woodall [30] proved the LCC for planar graphs with maximum degree at least 12, i.e., they proved that such graphs have list chromatic index equal to their maximum degree. The LCC remains open for planar graphs with smaller maximum degree, though it is known that if $\Delta(G) \leq 4$ or $\Delta(G) \geq 8$, then $ch'(G) \leq \Delta(G) + 1$ (Juvan, Mohar and Škrekovski [68] for $\Delta(G) \leq 4$; Bonamy [26] for $\Delta(G) = 8$; Borodin [29] for $\Delta(G) \geq 9$). As noted above, it is not true that planar graphs G with $\Delta(G) \in \{4, 5\}$ are always $\Delta(G)$-edge choosable. As there has been significant interest in studying both edge colourings and list edge colourings for the class of *planar* graphs, we thought natural to investigate our precolouring problem when restricted to this class.

In Section 4, we shall prove our main result of the chapter, which is easily seen to be best possible both on the size of the palette and on the distance requirement between any two precoloured edges.

Theorem 1.3. *Let G be a planar graph with maximum degree $\Delta(G) \geq 23$. Using the palette $\mathcal{K} = [\Delta(G)]$, any precoloured set of edges, where any two precoloured edges are at distance at least 3, can be extended to a proper edge colouring of all of G.*

Finally, we obtain the following weak form of a precolouring result when we allow two precoloured edges to be within distance 1.

Given a subset $S \subseteq E(G)$ of edges and an arbitrary (i.e., not necessarily proper) colouring of elements of S using only colours from $\mathcal{K}$, is there a proper colouring of all edges of G (using colours from $\mathcal{K}$) that *differs* from the given colouring on every edge of S? We may view the coloured set S as a set of *forbidden* (coloured) edges, while the full colouring, if it can be produced, is called an *avoidance* of the forbidden edges. We will show the following result, which is *not* directly implied by either the LCC or by other existing precolouring results.

THEOREM 1.4. *Let G be a multigraph with maximum degree $\Delta(G)$ and maximum multiplicity $\mu(G)$. Using the palette $\mathcal{K} = [\Delta(G)+\mu(G)]$, any forbidden matching can be avoided by a proper edge colouring of all of G.*

We use the aforementioned result of Berge and Fournier and a recolouring argument to prove this theorem in Section 5.

2. $\Delta + 1$ colours are necessary

In this short section, we give examples of bipartite graphs G of maximum degree Δ containing two edges (arbitrarily far apart) which must belong to the same colour class in any Δ-edge colouring.

Indeed, for any positive integer m, let D_m denote the bipartite graph on vertex set $\{x\} \cup A^x \cup B \cup A^y \cup \{y\}$, where $|A^x| = |A^y| = m$ and $|B| = 2m - 1$, and whose edge set is the set of all pairs between $\{x\} \cup B$ and A^x and between $\{y\} \cup B$ and A^y. Observe an easy property of the graph D_m: in any proper edge colouring of D_m with colours from $[2m]$, there must be at least one edge of colour 1 incident to x or y. For otherwise, since each vertex in A^x has degree $2m$, there must be m edges of colour 1 between A^x and B; similarly, there must be m edges of colour 1 between A_y and B. But this implies that there are $2m$ distinct edges of colour 1 incident to the $2m - 1$ vertices in B, which means that a vertex of B is incident to two edges of colour 1, a contradiction.

Next, for any positive integers ℓ, m, let $G_{m,\ell}$ be the graph formed by taking ℓ disjoint copies $H_1, \dots, H_\ell$ of D_m with vertex sets labelled $\{x_i\} \cup A_i^x \cup B_i \cup A_i^y \cup \{y_i\}$, identifying y_i with x_{i+1} for all $i = 1, \dots, \ell - 1$, and then adding two new vertices x' and y' and two new edges $x'x_1$ and $y_\ell y'$. See Figure 1 for a depiction of $G_{3,2}$. It is straightforward to check that $G_{m,\ell}$ is bipartite, has maximum degree $2m$, and that the edges $x'x_1$ and $y_\ell y'$ are at distance $4\ell + 1$ in $G_{m,\ell}$.

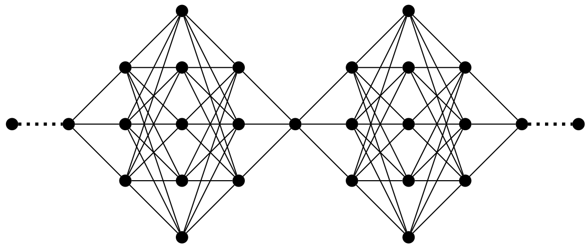

FIGURE 1. A representative G of a class of bipartite graphs, with a non-extendable matching consisting of two edges, using the palette $[\Delta(G)] = [\chi'(G)]$. Dashed lines indicate edges precoloured with colour 1.

Consider a precolouring of $G_{m,\ell}$ from the palette $[2m] = [\Delta(G_{m,\ell})] = [\chi'(G_{m,\ell})]$ in which the edges $x'x_1$ and $y_\ell y'$ are precoloured 1. Suppose, for a contradiction, that there is a proper extension of this precolouring. Then there can be no edge of colour 1 between A_1^y and B_1. By our observation about D_m, there must be an edge of colour 1 between A_1^y and $y_1 = x_2$. It follows by an induction (via copies of D_m) that there is an edge of colour 1 between A_ℓ^y and y_ℓ. Since $y_\ell y'$ is precoloured 1, we have arrived at our desired contradiction.

3. Precoloured edges must be at distance at least 2

In this section, we show that if we omit the distance 2 condition on the precoloured set, then Conjecture 4.1 becomes false whenever $\Delta(G) \geq 4$.

For each $t \geq 3$, we construct a graph G_t of maximum degree $t+1$ with the property that, using the palette $\mathcal{K} = [t+2]$, there is a matching M and a precolouring of M that cannot be extended to a proper edge colouring of all of G_t.

Our construction is based on an observation by Anstee and Griggs [12]. For $t \geq 3$, let H_t be the graph obtained from $K_{t,t}$ by subdividing one edge.

LEMMA 3.1 (Anstee and Griggs). *For every $t \geq 3$, the equality $\chi'(H_t) = \Delta(H_t) + 1 = t + 1$ holds.*

PROOF. Since H_t has $2t+1$ vertices, its largest matching has size t. Since H_t has $t^2 + 1$ edges, we cannot cover all the edges with t matchings. □

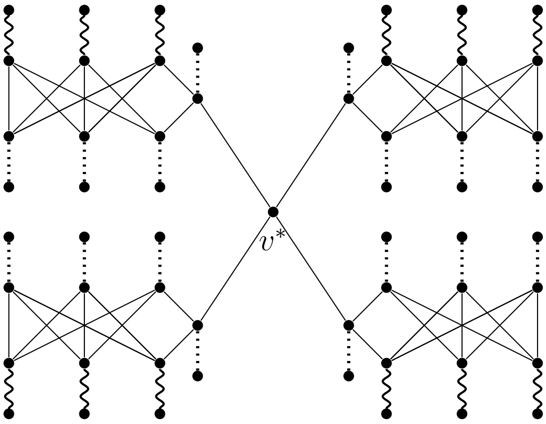

FIGURE 2. The graph G_3 (with maximum degree 4) and a non-extendable precoloured matching using the palette [5]. Wavy edges are precoloured 1, while dotted edges are precoloured 2.

Let $A, B \subseteq V(H_t)$ be the original partite sets of $K_{t,t}$, so that A and B are independent sets of size t in H_t, and the only vertex of H_t not contained in $A \cup B$ is the vertex of degree 2. Let H'_t be the graph obtained from H_t by attaching a pendant edge to each vertex of H_t, and for each $v \in V(H_t)$, let v' be the other endpoint of the pendant edge at v. Finally, set $M_0 = \{ vv' \mid v \in V(H_t) \}$. We precolour the matching M_0 by colouring vv' colour 1 if $v \in A$, and colouring vv' colour 2 otherwise. Now we define the full graph G_t by taking $t + 1$ disjoint copies of H'_t, and adding a new vertex v^* adjacent to the unique vertex of degree 3 in each copy of H'_t. The precoloured matching M in G_t is just the union of each precoloured matching M_0 in each copy of H'_t, with the same precolouring. Figure 2 shows G_3.

THEOREM 3.2. *For every $t \geq 3$, using the palette $\mathcal{K} = [t + 2] = [\Delta(G_t) + \mu(G_t)]$, the precolouring of the matching M as described above cannot be extended to a proper edge colouring of the entire G_t.*

Proof of Theorem 3.2. Suppose for contradiction that G_t has an edge-colouring from $\mathcal{K}$ that extends the precolouring of M. Since every neighbour of v^* has an incident edge precoloured 2, no edge incident to v^* can be coloured 2. Therefore, since $d(v^*) = t + 1$, each of the $t + 1$ colours excluding 2 is used exactly once on the edges incident to v^*. In particular, some edge e incident to v^* has colour 1. Let H be the copy of H_t containing the other endpoint of e. Observe that no edge of H can be coloured 1 or 2: every edge joining A and B has an edge precoloured 1 at one endpoint and an edge precoloured 2 at the other, while the vertex of degree 2 in H is incident to an edge precoloured 2 as well as the edge e coloured 1. Hence all edges of H use only the t remaining colours. Since $\chi'(H_t) = t + 1$ by Lemma 3.1, this is impossible. $\square$

4. Proof of the main result

In this section, we shall prove Theorem 1.3. Note that for brevity, we will write Δ for $\Delta(G)$.

As mentioned earlier, the LCC is known to hold for planar graphs with maximum degree at least 12, [30]. Indeed, given a planar graph G with $\Delta \geq 12$ then $ch'(G) = \Delta$. Combining this result with Observation 1.1 we obtain the following proposition which will be useful in the proof of Theorem 1.3.

Proposition 4.1. *Let G be a planar graph with maximum degree $\Delta(G) \geq 12$. Using the palette $\mathcal{K} = [\Delta(G) + 1]$, any precoloured set where no two edges are within distance 1 can be extended to a proper edge colouring of all of G.*

4.1. Our framework and notation. The notation we use is standard. We shall now outline the general framework and the new terminology we need.

Whenever considering a planar graph G, we fix a drawing of G in the plane. (So we really should talk about a *plane graph.*) Because of this fixed embedding we can talk about the *faces* of the graph. If G is connected, then the boundary of any face f forms a closed walk W_f.

We adopt the following notation to classify vertices of a graph G according to degree and incidence with vertices of degree 1. Let V_i be the set of vertices of degree i. Also, identify by $T_i \subseteq V_i$ the set of those vertices of degree i that are adjacent to a vertex of degree 1, and set $U_i = V_i \setminus T_i$. Write $T = \bigcup_{i \geq 1} T_i$ and $U = V(G) \setminus T$. We also adopt the shorthand notation $V_{[i,j]}$, $U_{[i,j]}$ and $T_{[i,j]}$

to mean, respectively, the sets of vertices in V, U and T with degrees between i and j inclusively.

We move now to the proof of Theorem 1.3.

Proof of Theorem 1.3. Suppose G is not connected, then we extend the edge colouring one component at a time. The precolouring on a component C with $12 \leq \Delta(C) \leq 22$ can be extended using Proposition 4.1. If $\Delta(C) < 12$, a greedy colouring algorithm $(23 \geq 2 \cdot 11)$ easily extends the precolouring to the entire component C.

Next, the statement of Theorem 1.3 is trivially true for graphs with maximum degree 23 and exactly 23 edges. The proof goes now by induction on $E(G)$, and we proceed with the induction step. We may assume that G is connected and has at least 24 vertices, since $\Delta \geq 23$. Let M be the precoloured set.

We first observe that

$$\text{if } uv \in E(G) \setminus M, \text{ then } d(u) + d(v) \geq \Delta + 2. \tag{1}$$

Indeed, suppose that the inequality does not hold for some $uv \notin M$. Then, by induction if $\Delta(G - uv) \geq 23$ and by Proposition 4.1 if $\Delta(G - uv) = 22$, there exists an extension of M to a colouring of $G - uv$ using the palette $\mathcal{K}$. Since at most $\Delta - 1$ colours are used on the edges adjacent to uv, we can easily extend the colouring further to uv. From (1) it follows that every vertex with degree 1 is incident with an edge in M and that if v has degree 2 and $uv \notin M$, then $d(u) = \Delta$. In particular, if a vertex v with degree greater than 1 has a neighbour in T_2, then $d(v) = \Delta$. Moreover, since edges in M are at distance at least 3 in G, a vertex can have at most one neighbour in $V_1 \cup T_2$.

Let V_2' be the set of vertices of degree 2 that are not incident with an edge of M. For a face f, let $V^-(f) = V(f) \setminus (V_1 \cup T_2)$, and let W_f^- be the sequence of vertices on the boundary walk W_f after removing vertices from $V_1 \cup T_2$. For a vertex v, let $v_1, v_2, \ldots, v_{d(v)}$ be the neighbours of v, listed in clockwise order according to the drawing of G. Write f_i for the face incident with v lying between the edges vv_i and vv_{i+1} (taking addition modulo $d(v)$ in $\{1, \ldots, d(v)\}$).

If a vertex v has a (unique) neighbour in $V_1 \cup T_2$, then we always choose v_1 to be this neighbour. In that case $f_{d(v)} = f_1$, and that face is called f_1 again. Note that it is possible for other faces to be the same as well (if v is a cut-vertex), but we will not identify those multiple names of the same face.

Claim 4.2. $|V_\Delta| > |V_2'|$.

Consider the set F of edges in $E(G)$ with one endvertex in V_2' and the other in V_Δ. Note that $F \cap M = \varnothing$ by the definition of V_2'. The subgraph with vertex set $V_2' \cup V_\Delta$ and edge set F is bipartite; we assert it is acyclic. For suppose there exists an (even) cycle C with $E(C) \subseteq F$. By induction if $\Delta(G - E(C)) \geq 23$, by Proposition 4.1 if $\Delta(G - E(C)) \leq 22$, we can extend the precolouring of M to $G - E(C)$ using the palette $\mathcal{K}$. But then we can further extend this colouring to the edges of C, since each one sees only $\Delta - 2$ coloured edges, and even cycles are 2-edge choosable. Since each vertex in V_2' is incident with precisely two edges in F, we have $|V_\Delta| + |V_2'| > |F| = 2|V_2'|$. This finishes the proof of the claim.

We use now a discharging argument to complete the proof. First, let us assign to each vertex v a charge

$\alpha 1$: $\alpha(v) = 3d(v) - 6$,

and to each face f a charge

$\alpha 2$: $\alpha(f) = -6$.

For each vertex v we define $\beta(v)$ as follows.

$\beta 1$: If $v \in V_\Delta$, then $\beta(v) = -2$.

$\beta 2$: If $v \in V_2'$, then $\beta(v) = 2$.

$\beta 3$: In all other cases, $\beta(v) = 0$.

For each edge $e = vu$, we define $\gamma_e(v)$ and $\gamma_e(u)$ as follows.

$\gamma 1$: If $v \in V_1$, then $\gamma_e(v) = -\gamma_e(u) = 3$.

$\gamma 2$: If $v \in T_2$ and $u \in V_\Delta$, then $\gamma_e(v) = -\gamma_e(u) = 3$.

$\gamma 3$: If $v \in U_2 \setminus V_2'$ and $u \in V_\Delta$, then $\gamma_e(v) = -\gamma_e(u) = 2$.

$\gamma 4$: In all other cases, $\gamma_e(v) = \gamma_e(u) = 0$.

Finally, for each face f and vertex $v \in W_f^-$ we define $\delta_f(v)$ and $\delta_v(f)$ as follows.

$\delta 1$: If $v \in U_2$, then $\delta_v(f) = -\delta_f(v) = 1$.

$\delta 2$: If $v \in T$ and $3 \leq d(v) \leq \Delta - 4$, then $\delta_v(f) = -\delta_f(v) = 3 - \dfrac{6}{d(v)-1}$.

$\delta 3$: If $v \in U$ and $3 \leq d(v) \leq \Delta - 4$, then $\delta_v(f) = -\delta_f(v) = 3 - \dfrac{6}{d(v)}$.

$\delta 4$: If $d(v) \geq \Delta - 3$, $|V^-(f)| = 3$, and both neighbours of v in $V^-(f)$ are joined by an edge in M, then $\delta_v(f) = -\delta_f(v) = 4$.

$\delta 5$: If $d(v) \geq \Delta - 3$ and v has a neighbour in $V^-(f) \cap T_3$, then $\delta_v(f) = -\delta_f(v) = 3$.

$\delta 6$: If $d(v) \geq \Delta - 3$ and none of $\delta 4$ and $\delta 5$ applies, then $\delta_v(f) = -\delta_f(v) = \frac{5}{2}$.

For a vertex v, write $\gamma(v)$ for the sum of $\gamma_e(v)$ over all edges e that have v as an endvertex. For a vertex v of degree 1 we set $\delta(v) = 0$. For every other vertex v, write $\delta(v)$ for the sum over the faces f around v of $\delta_f(v)$. Similarly, for a face f, write $\delta(f)$ for the sum over the vertices v on the reduced walk W_f^- around f of the values of $\delta_v(f)$. By the definitions of γ and δ, we have that $\sum_v \gamma(v) + \sum_v \delta(v) + \sum_f \delta(f) = 0$. It follows from Claim 4.2 that $\sum_v \beta(v) < 0$. From Euler's formula we obtain $\sum_v \alpha(v) + \sum_f \alpha(f) < 0$.

Thus, in order to reach a contradiction, it is enough to show that for every vertex v:

$$\alpha(v) + \beta(v) + \delta(v) + \gamma(v) \geq 0, \tag{2}$$

and that for every face f:

$$\alpha(f) + \delta(f) \geq 0. \tag{3}$$

Let f be a face. As G is simple, $|V^-(f)| \geq 3$. Since $\alpha(f) = -6$, it follows that (3) is verified if we can show that $\delta(f) \geq 6$. Let v be a vertex in $V^-(f)$ for which $\delta_v(f)$ is minimum. If $\delta_v(f) \cdot |V^-(f)| \geq 6$, then (3) clearly holds. So, by checking $\delta1-\delta6$, we see we only have to consider the case where $v \in T_{[3,6]} \cup U_{[2,5]}$. (Recall that vertices from $V_1 \cup T_2$ do not appear in W_f^-.)

If $v \in U_2$, then let u and w be the neighbours of v. Consider first the case where both u and w have degree Δ. Then they both belong to $V^-(f)$, so (3) follows, since $\delta_v(f) = 1$ and $\delta_u(f) \geq \frac{5}{2}$, $\delta_w(f) \geq \frac{5}{2}$ by $\delta4-\delta6$. Suppose now that u has degree less than Δ, which implies by (1) that $uv \in M$ and, consequently, $vw \notin M$. In particular, $w \in V^-(f)$ and w has degree Δ. Note also that necessarily $u \in V^-(f)$. If $|V^-(f)| = 3$, then $\delta_w(f) = 4$ by $\delta4$. As $\delta_u(f) \geq \delta_v(f) = 1$, it follows that (3) holds. If $|V^-(f)| \geq 4$, then u has a neighbour u' in $V^-(f) \setminus \{v, w\}$. We assert that $\delta_u(f) + \delta_{u'}(f) \geq \frac{5}{2}$. Indeed, because $uu' \notin M$, we know by (1) and since $\Delta \geq 18$ that (at least) one of u and u' has degree at least 5. Consequently, by $\delta2-\delta6$ we know that $\max\{\delta_f(u), \delta_f(u')\} \geq \frac{3}{2}$. Since $\min\{\delta_f(u), \delta_f(u')\} \geq \delta_f(v) = 1$ and $\delta_f(w) \geq \frac{5}{2}$ by $\delta5$ and $\delta6$, it follows that (3) holds.

If $v \in T_3$, then $\delta_v(f) = 0$, but v has two neighbours in $V^-(f)$ that have degree at least $\Delta - 1$ each. Equation (3) then follows from $\delta5$.

For the remaining cases we always have $\delta_v(f) \geq 1$. Rules $\delta2-\delta6$ ensure that any vertex $u \in V^-(f)$ with $d(u) \geq 13$ satisfies $\delta_u(f) \geq \frac{5}{2}$; hence there can be at most one such vertex and, in particular, a neighbour u of v in $V^-(f)$ must have degree at most 12. As v itself has degree at most 6, by (1) we have $uv \in M$, which also implies that $\{u, v\} \subseteq U$. Hence in particular $v \in U_{[3,5]}$. Let w be the neighbour of v in $V^-(f) \setminus \{u\}$. Since $v \in U_{[3,5]}$ and $vw \notin M$, it necessarily holds that $d(w) \geq \Delta - 3$. If $|V^-(f)| = 3$, then (3) holds by $\delta4$ since $\delta_u(f) \geq \delta_v(f) \geq 1$. If $|V^-(f)| \geq 4$, then u has a neighbour u' in $V^-(f) \setminus \{v, w\}$, which has degree at least $\Delta + 2 - d(u) \geq 10$. Consequently, $\delta_{u'}(f) \geq \frac{12}{5}$ by $\delta3$, $\delta5$ or $\delta6$. We deduce that (3) holds, as $\delta_w(f) \geq \frac{5}{2}$ by $\delta5$ or $\delta6$. This confirms (3) for all faces.

Now let v be a vertex. Recall that $\alpha(v) = 3d(v) - 6$. Furthermore, if v has a neighbour in $V_1 \cup T_2$, then the two consecutive faces incident with that neighbour are counted as one face; all other faces are counted separately. Finally, as noted earlier, a vertex can have at most one neighbour in $V_1 \cup T_2$. If $d(v) = 1$, then $\alpha(v) = -3$ and $\gamma(v) = 3$. Since $\beta(v) = \delta(v) = 0$, we immediately obtain (2).

If $d(v) = 2$, then $\alpha(v) = 0$. If $v \in T_2$, then both $\gamma 1$ and $\gamma 2$ apply; hence $\gamma(v) = 0$. Again one can check that $\beta(v) = \delta(v) = 0$, confirming (2). Otherwise $v \in U_2$, and $\delta 1$ implies that $\delta(v) \geq -2$, as v is incident with at most two faces. If $v \in V_2'$ as well, then $\beta 2$ yields that $\beta(v) = 2$ and $\gamma(v) = 0$. If $v \notin V_2'$, then $\gamma(v) = 2$ while $\beta(v) = 0$. In either case (2) follows.

Next suppose that $3 \leq d(v) \leq \Delta - 4$. Observe that $\beta(v) = 0$. If $v \in T$, then $\gamma(v) = -3$ by $\gamma 1$. Since v has a neighbour with degree one, we know that v is incident with $d(v) - 1$ regions, and so $\delta 2$ yields that $\delta(v) = (d(v) - 1) \cdot \big(-3 + \dfrac{6}{d(v) - 1}\big) = 9 - 3d(v)$. Similarly, if $v \in U$, then $\gamma(v) = 0$, and $\delta 3$ yields that $\delta(v) = 6 - 3d(v)$. This proves (2) for those vertices v.

Suppose now that $d(v) \in \{\, \Delta - 3, \Delta - 2, \Delta - 1 \,\}$. Then $\beta(v) = 0$. If $v \in T$, then $\gamma(v) = -3$ by $\gamma 1$. Since M is distance-3, none of $\delta 4$ and $\delta 5$ applies to v, and v is incident with $d(v) - 1$ faces. From $\delta 6$ we deduce that $\delta(v) = -\frac{5}{2}(d(v) - 1)$. Since $d(v) \geq \Delta - 3 \geq 13$, it follows that $3d(v) - 6 - 3 - \frac{5}{2}(d(v) - 1) = \frac{1}{2}d(v) - \frac{13}{2} \geq 0$, and hence (2) is satisfied again. Next assume that $v \in U$, and so $\gamma(v) = 0$. The fact that M is distance-3 ensures that $\delta 4$ applies to at most one face with respect to v, and $\delta 5$ applies to at most two faces with respect to v. Consequently, $\delta(v) \geq -\big(4 + 6 + \frac{5}{2}(d(v) - 3)\big)$. Combined with the assumption that $\Delta \geq 17$, this is always enough to satisfy (2).

Finally, suppose that $d(v) = \Delta$. In this case $\beta(v) = -2$. If $v \in T$, then the distance condition on M ensures that $\gamma(v) = -3$ and $\delta(v) = -\frac{5}{2}(\Delta - 1)$. Since $\Delta \geq 17$ this confirms (2).

So we are left with the case where $v \in U$. Since any two edge of M are at distance at least 3, it follows that at most one of $\gamma 2$, $\gamma 3$ applies and similarly at most one of $\delta 4$, $\delta 5$ applies. Moreover, if $\gamma 2$ does apply, then $\gamma(v) = -3$ and

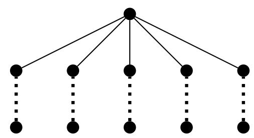

FIGURE 3. A representative G of a class of trees, with a non-extendable precoloured (distance-2) matching, using the palette $[\Delta(G)] = [\chi'(G)]$. Dashed lines indicate edges precoloured with colour 1.

neither $\delta 4$ nor $\delta 5$ applies. This means that the vertex v is incident with Δ faces, and for each of those faces f we have $\delta_f(v) = -\frac{5}{2}$. If $\gamma 2$ does not apply, then $\gamma(v) \geq -2$. The vertex v is incident with Δ faces, and for $\Delta - 1$ of those faces f we have $\delta_f(v) = -\frac{5}{2}$. For the final face f either $\delta 4$ or $\delta 5$ may apply, so $\delta_f(v) \in \{-4, -3, -\frac{5}{2}\}$. Using that $\Delta \geq 23$, we can check that (2) is satisfied in all cases.This confirms (2) for all vertices and completes the proof of the theorem. $\square$

Note that Theorem 1.3 easily becomes false, even for trees, if we replace the the distance condition by 2. For instance, consider stars with each edge subdivided exactly once; see Figure 3.

5. Avoiding prescribed colours on a matching

In this section, we will show the following statement, which directly implies Theorem 1.4.

THEOREM 5.1. *Let G be a multigraph with maximum degree $\Delta(G)$ and maximum multiplicity $\mu(G)$, and let M_1 and M' be two disjoint matchings in G. Suppose that each edge e of G is assigned a list $L(e) \subseteq [\Delta(G)+\mu(G)]$ of colours such that*

- $L(e) = \{1\}$ *if* $e \in M_1$*;*
- $L(e) = \{2, \ldots, \Delta(G) + \mu(G)\}$ *if* $e \in M'$*; and*
- $L(e) = [\Delta(G) + \mu(G)]$ *if* $e \in E(G) \setminus (M_1 \cup M')$*.*

Then there exists a proper edge-colouring ψ of G such that $\psi(e) \in L(e)$ for every $e \in E(G)$.

To establish Theorem 5.1, we use the following theorem, already stated in Chapter 2.

THEOREM 5.2 (Berge and Fournier). *Let G be a multigraph with maximum degree $\Delta(G)$ and maximum multiplicity $\mu(G)$, and let M be a matching in G. Then there exists a proper edge-colouring of G using the palette $[\Delta(G) + \mu(G)]$ such that every edge of M receives the same colour.*

Proof of Theorem 5.1. We may assume without loss of generality that M_1 is a maximal matching in $G \setminus M'$. We set

$$B = \{\, e' \in M' \mid e' \cap e = \varnothing \text{ for all } e \in M_1 \,\}.$$

Let ψ be a *partial* proper edge colouring of G using colours in $[\Delta(G) + \mu(G)]$ such that

(i) $\psi(e) = 1$ for every $e \in M_1$;

(ii) $\psi(e') \neq 1$ for every $e' \in M'$;

(iii) every edge of $E(G) \setminus B$ receives a colour under ψ;

(iv) the number of edges of B that receive a colour under ψ is maximal.

To show that ψ is well defined, we need to prove the existence of a partial proper edge-colouring of $G \setminus B$ using the palette $[\Delta(G) + \mu(G)]$ that satisfies (i) – (iii).

To this end, let $G' = G - B$. By Theorem 5.2, there is a proper edge-colouring ϕ of G' using colours in $[\Delta(G) + \mu(G)]$ such that every edge in M_1 receives colour 1. By the definition of B, each edge in $M' \setminus B$ is incident to at least one edge in M_1. Each edge in M_1 receives colour 1 under ϕ and therefore ϕ does not map any edge of $M' \setminus B$ to colour 1. Thus ϕ ensures that ψ exists.

We now show that every edge of B receives a colour under ψ, which completes the proof. Suppose, on the contrary, that $xy \in B$ is an edge that is not coloured by ψ. We start by making the following observations.

CLAIM 5.3. *For every $e \in E(G)$, we have $\psi(e) = 1$ if and only if $e \in M_1$.*

Indeed, if e is an edge that is coloured 1, then $e \notin M'$ and e is not adjacent to an edge in M_1, since all such edges are also coloured 1. Consequently, $e \in M_1$, as M_1 is a maximal matching of $G - M'$.

Claim 5.3 and the definition of B ensure the following.

CLAIM 5.4. *Neither x nor y is incident with an edge that is coloured* 1.

For each vertex $v \in V(G)$, recall we have defined $\overline{\psi}(v) \subseteq [\Delta(G) + \mu(G)]$ to be the set of colours that do not appear on edges incident to v. Claim 5.4 states that $\overline{\psi}(x)$ and $\overline{\psi}(y)$ both contain the colour 1.

CLAIM 5.5. *If $v \in N_G(x) \setminus \{y\}$, then v is incident to an edge in M_1 and so $\overline{\psi}(v)$ does not contain the colour* 1.

Indeed, for if v is not incident to an edge in M_1, then by Claim 5.4 the edge xv could be added to M_1 to form a larger matching in $G - M'$, thereby contradicting the maximality of M_1.

We know that the edge xy is not yet coloured so both $\overline{\psi}(x)$ and $\overline{\psi}(y)$ must contain some colour different from 1 and we shall from now on redefine $\overline{\psi}(y)$ to be $\overline{\psi}(y) \setminus \{1\}$, which is not empty. We consider the following iterative procedure.

Initially ($t = 0$), we set $D_0 = \{y\}$. At each step $t \geq 1$, we form the set D_t as follows:

$$D_t = \Big\{\, v \in N_G(x) \setminus \bigcup_{i=0}^{t-1} D_i \;\Big|\; \exists\, e \text{ between } v \text{ and } x \text{ with a colour in } \bigcup_{w \in D_{t-1}} \overline{\psi}(w) \,\Big\}.$$

Since $\bigcup_{i\geq 0} D_i \subseteq N_G(x)$ and $D_i \cap D_j = \varnothing$ if $0 \leq i < j$, there exists a least non-negative integer t_0 such that $D_{t_0+1} = \varnothing$. We define $D = \bigcup_{i\leq t_0} D_i$. We consider now two cases.

Case 1. Assume that there exist a vertex $w \in D$ and a colour $c \in \overline{\psi}(w) \cap \overline{\psi}(x)$. Since the subsets $D_0, \dots, D_{t_0}$ are pairwise disjoint, there is precisely one integer t_1 such that $w \in D_{t_1}$. There exists a sequence $y = w_0, w_1, w_2, \dots, w_{t_1} = w$ of vertices such that $w_i \in D_i$ and (at least) one edge e_i between x and w_i has a colour in $\overline{\psi}(w_{i-1})$, whenever $1 \leq i \leq t_1$.

We may then define a partial proper edge colouring ψ' of $E(G)$, using colours in $[\Delta(G) + \mu(G)]$, with

- $\psi'(e) = \psi(e)$ if $e \notin \{\, e_i \mid 1 \leq i \leq t_1 \,\}$;
- $\psi'(e_i) = \psi(e_{i+1})$ for each $i \in \{0, \dots, t_1 - 1\}$; and
- $\psi'(e_{t_1}) = c$.

One can check that ψ' satisfies (i) – (iii) and colours one more edge of B than ψ does, which contradicts the choice of ψ.

For the second case, we need the following two observations.

Claim 5.6. *For every $z \in N_G(x)$, it holds that $\mu(G) \leq |\overline{\psi}(z)|$.*

The only case this is not trivial is when $z = y$, due to our redefinition of $\overline{\psi}(y)$. However, as the edge xy is not coloured, the vertex y sees at most $\Delta(G) - 1$ different colours, which implies the statement.

Now, let H be the bipartite subgraph of G induced by the bipartition $(\{x\}, D)$. (In particular, the edges of G between vertices in D are not in H.) The next statement follows directly from the fact that the number of coloured edges between x and y is at most $\mu(G) - 1$.

Claim 5.7. *The bipartite graph H contains fewer than $|D|\mu(G)$ coloured edges.*

We can now proceed with the second case.

Case 2. For every vertex $w \in D$ and every colour $c \in \overline{\psi}(w)$, there exists an edge e_w between x and a vertex $z \in D$ such that $\psi(e_w) = c$. By Claims 5.6 and 5.7, we know that the number of colours appearing in the bipartite graph H is less than $|D| \cdot \mu(G)$, which is at most $\sum_{w \in D} |\overline{\psi}(w)|$. This implies that there are two distinct vertices v_1 and v_2 in $D \subseteq N_G(x)$ with $\overline{\psi}(v_1) \cap \overline{\psi}(v_2) \neq \varnothing$. Let $c_1 \in \overline{\psi}(v_1) \cap \overline{\psi}(v_2)$ and note that $c_1 \neq 1$ by Claim 5.5. Let $c_2 \in \overline{\psi}(x) \setminus \{1\}$. Then $c_2 \notin \overline{\psi}(v_1) \cup \overline{\psi}(v_2)$ and $c_1 \notin \overline{\psi}(x_1)$. (And hence $c_1 \neq c_2$.)

For $i \in \{1, 2\}$, let P_i be the maximal alternating path with colours c_1 and c_2 beginning at v_i. Note that x cannot belong to both paths. But if x does not belong to P_i, then we may swap c_1 and c_2 along the edges of P_i. This leads us back to Case 1 because then c_2 belongs to $\overline{\psi}(x) \cap \overline{\psi}(v_1)$. (Note that such a swap affects neither the colours of the edges inside H nor those of edges in M_1.)

We have shown that in each case there exists a partial proper edge colouring using colours in $[\Delta(G) + \mu(G)]$ and satisfying (i) – (iii) that assigns colours to more edges of B than ψ does, which is a contradiction. $\square$

6. Concluding remarks

To conclude this chapter, we present a conjecture due to Csóka, Lippner and Pikhurko [40] which is in the same spirit as the problems we have addressed here.

Conjecture 6.1 (Csóka, Lippner and Pikhurko). *Let G be a graph such that every vertex is of degree at most d, except one of degree $d + 1$. Using the palette $\mathcal{K} = [d + 1]$, suppose that at most $d - 1$ pendant edges are precoloured. This precolouring can be extended to a proper edge-colouring of all of G.*

The authors of Conjecture 6.1 proved the weaker statement where $\mathcal{K}$ is replaced by $\mathcal{K} = [d + 9\sqrt{d}]$. Note that even any bound of the form $d + O(1)$ for the palette $\mathcal{K}$ would be extremely interesting to show.

Now, with respect to our precolouring problems, rather than imposing conditions on the set of precoloured edges, we could instead constrain the precolouring. In the light of Theorem 3.2 and the result of Berge and Fournier [18], the following is a natural strengthened version of Conjecture 4.1.

CONJECTURE 6.2. *Let G be a multigraph with maximum degree $\Delta(G)$ and maximum multiplicity $\mu(G)$. Using the palette $\mathcal{K} = [\Delta(G) + \mu(G)]$, any precoloured set such that no two edges precoloured with* ***different*** *colours are within distance 2 can be extended to a proper edge colouring of all of G.*

Unfortunately, we could not even assert Conjecture 6.2 with the constant 2 replaced by any larger fixed integer.

CHAPTER 4

Ramsey colourings

1. Introduction

An old observation by Erdős and Rado says that when the edges of a complete graph are coloured with two colours, there is a spanning monochromatic component. This simple remark has been the starting point of extensive research. A natural example is the search for large monochromatic components in r-edge-coloured complete graphs (see, for example, [55, 56]). Here we focus on a different direction, namely, the search for covers (or partitions) of the vertices into as few as possible monochromatic connected subgraphs.

A classical example appears in a seminal paper by Erdős, Gyárfás and Pyber [47], who showed that for any r-colouring of K_n (the complete graph on n vertices) the vertices can be partitioned into at most $O(r^2 \log r)$ monochromatic cycles. We note that throughout this chapter, when we say that the vertices of a graph are covered (or partitioned) by a collection of subgraphs, we mean that the vertices are covered by the *vertex sets* of these subgraphs.

Gyárfás, Ruszinkó, Sárközy and Szemerédi [57] improved the above result by showing that if the edges of the complete graph are r-coloured then the vertices can be partitioned into $O(r \log r)$ monochromatic cycles. In the other direction, Pokrovskiy [86] showed that one needs strictly more than r cycles, disproving a conjecture of Erdős, Gyárfás and Pyber [47]. Conlon and Stein [39] showed similar results for colourings where every vertex is incident with at most r distinct colours. The question of whether one can partition an r-coloured graph into $O(r)$ monochromatic cycles remains an enticing open problem in this area.

In a slightly different direction, Erdős, Gyárfás and Pyber [47] conjectured that the vertices of an r-coloured complete graph may be partitioned into at most $r-1$ monochromatic connected subgraphs. This conjecture is known to be tight when $r-1$ is a prime power and n is sufficiently large, due to a well-known construction which requires the existence of an affine plane of an appropriate order. Haxell and Kohayakawa [62] proved a slightly weaker result, showing that one can partition an r-coloured complete graph on n vertices into r monochromatic subgraphs, for sufficiently large n.

Interestingly, this problem is closely related to a well-known conjecture of Ryser on packing and covering edges in r-partite, r-uniform hypergraphs. This link was first noted by Gyárfás [55] in 1997 and leads to the following natural formulation of the conjecture of Ryser, appearing in [64], where $\alpha(G)$ is the size of the largest independent set in the graph G.

Conjecture 1.1 (Ryser [64]). *The vertex set of an r-coloured graph G can be covered by at most $(r-1)\alpha(G)$ monochromatic connected subgraphs.*

In this form, it is clear that Ryser's conjecture implies the *covering version* of the aforementioned conjecture of Erdős, Gyárfás and Pyber about monochromatic connected subgraphs. Although not much is known about Ryser's conjecture in general, a few special cases are understood. The case $r=2$ is equivalent to König's classical theorem (see [41], for example), while the case $r=3$ was proved by Aharoni [2] in 2001, who built on the earlier advances of Aharoni and Haxell [3]. The conjecture is also known to hold for $\alpha(G)=1$ (i.e. G is a complete graph) and $r \leq 5$, as was proved by Gyárfás [55] ($r=3$), Duchet [44] and Tuza [111] ($r=4$), and Tuza [111] ($r=5$).

Following Schelp [90] who suggested several variants of Ramsey-type problems (e.g. determining the length of the longest monochromatic path in a 2-coloured graph), we consider variants of the above problems for graphs with large minimum degree. Our first main result proves a conjecture of Bal and

DeBiasio [14] about partitioning the vertices of a 2-coloured graph with large minimum degree; recall that $\delta(G)$ denotes the minimum degree of the graph G.

THEOREM 1.2. *There exists an integer n_0 such that every 2-coloured graph G on $n \geq n_0$ vertices and with minimum degree at least $\frac{2n-5}{3}$ can be partitioned into two monochromatic connected subgraphs.*

We note that this is a generalisation of the result by Haxell and Kohayakawa [62] mentioned above for two colours, where the complete graph is replaced by a graph with large minimum degree. This result is seen to be sharp by a construction of Bal and DeBiasio [14]; in Section 5 we describe a more general family of examples which shows, in particular, the sharpness of the minimum degree condition in this result. One can think of this result as saying that $\frac{2n-5}{3}$ is the minimum degree 'threshold' that guarantees a partition of every 2-colouring into two monochromatic connected subgraphs. It is therefore natural to ask what minimum degree condition on a graph G guarantees a partition into t monochromatic connected subgraphs, no matter how the graph is 2-coloured. We conjecture the following.

CONJECTURE 1.3. *For every t there exists n_0, such that for every 2-colouring of a graph G on $n \geq n_0$ vertices with $\delta(G) \geq \frac{2n-2t-1}{t+1}$ there exists a partition of the vertex set into at most t monochromatic connected subgraphs.*

We support this conjecture by observing an analogous result for *covers* of the vertices by monochromatic components.

PROPOSITION 1.4. *Let t be integer and let G be a 2-coloured graph on n vertices with $\delta(G) \geq \frac{2n-2t-1}{t+1}$. Then the vertices of G can be covered by at most t monochromatic components.*

We also give a construction, showing that the inequality in this proposition (and therefore the conjecture) cannot be improved.

Bal and DeBiasio [14] also considered the problem of covering coloured graphs with monochromatic components of distinct colours. In particular, they conjectured the following.

Conjecture 1.5 (Bal and Debiasio). *Let G be an r-coloured graph on n vertices with $\delta(G) \geq (1 - 1/2^r)n$. Then the vertices can be covered by monochromatic components of distinct colours.*

Again, Bal and DeBiasio provided examples showing that if true, the bound $(1 - 2^{-r})n$ is best possible. We shall prove Conjecture 1.5 for $r = 2, 3$. The case $r = 3$ is the most interesting case but we include a short proof of $r = 2$ for completeness.

Theorem 1.6. *Let G be a 3-coloured graph on n vertices with $\delta(G) \geq 7n/8$. Then the vertices of G can be covered by monochromatic components of distinct colours.*

2. Notation

By an *r-coloured* graph, we mean a graph whose edges are coloured with r colours. When a graph is 2-coloured we call the colours *red* and *blue*; and when it is 3-coloured, we denote the colours by *red*, *blue* and *yellow*.

For a set of vertices W, we denote by $N_r(W)$ the set of vertices in $V(G) \setminus W$ that are adjacent to a vertex in W by a red edge. If $x \in V(G)$ is a vertex, we define $d_r(x) = |N_r(\{x\})|$ which we refer to as the *red degree* of x. We say that y is a *red neighbour* of x if xy is a red edge. By a *red component* of a graph G, we mean the *vertex set* of a component in the graph on vertex set $V(G)$ whose edgse are the red edges of G. We denote the red component that contains x by $C_r(x)$. A *red set* U is a set of vertices that is connected in red, i.e. the red edges induced by U form a connected graph.

All the above definitions and notation, that were defined for red, also works for blue or yellow; e.g. $d_b(x)$ and $d_y(x)$ are the blue and yellow degrees of x, respectively, and a blue set is a set of vertices that is connected in blue.

3. Partitioning into monochromatic connected subgraphs

In this section, we prove Theorem 1.2. We note that the minimum degree condition in this theorem cannot be improved; this can be seen by taking $t = 2$ in the Example 5.1 (see Figure 1. below), which we describe more formally in Section 5.

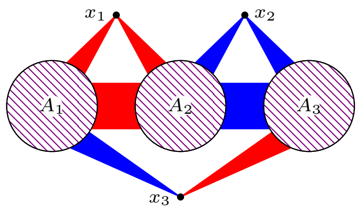

FIGURE 1. A 2-coloured graph on n vertices and with minimum degree $\frac{2n-5}{3}$ which can not be partitioned into 2 monochromatic subgraphs.

Proof of Theorem 1.2. Throughout this proof, we assume that the number of vertices n is sufficiently large. Suppose now, for a contradiction, that the vertices of G cannot be partitioned into two monochromatic sets.

CLAIM 3.1. *There is a blue component of order at most* $(n+2)/6$.

PROOF OF CLAIM. We may assume that there are at least three red components and at least three blue components, as otherwise the vertices may be partitioned into two red sets or two blue sets (recall that a *red set* is defined to be a set of vertices that is connnected in red, and similarly for blue), contradicting our assumption. Let R be a red component of smallest order, so $|R| \leq n/3$.

Let us assume first that $|R| \leq (n-5)/3$. Since every vertex in R sends at least $(2n-5)/3 - (|R|-1) > (n-|R|)/2$ blue edges outside of R, every two vertices in R have a common blue neighbour outside of R. Hence, R is contained in a blue component of order at least $|R| + (2n-5)/3 - (|R|-1) \geq (2n-2)/3$. Since there are at least three blue components, there is a blue component of order at most $(n-(2n-2)/3)/2 = (n+2)/6$.

We now assume that $(n-4)/3 \leq |R| \leq n/3$. If every two vertices in R have a common blue neighbour, then, again, R is contained in a blue component of order at least $(2n-2)/3$ and as before there is a blue component of order at most $(n+2)/6$. Otherwise, there exist two vertices $u, v \in R$ whose blue neighbourhoods do not intersect. But both u and v have at least $(n-5)/3$ blue neighbours outside of R, and therefore every vertex in $R \setminus \{u, v\}$ has a common blue neighbour with either u or w. It follows that there are two blue components (namely, the components $C_b(u)$ and $C_b(w)$) whose union has order at least $|R| + 2(n-5)/3 > n-5$, hence there is a blue component of order at most 4. $\square$

Claim 3.2. *There is a red set U of size at most $27 \log n$ such that $|N_r(U)| \geq 2n/3 - 27 \log n$.*

Proof of claim. By the previous claim, there is a blue component B of order at most $(n+1)/6$. Note that every vertex in B has at least $(2n-5)/3 - |B|$ red neighbours in $V(G) \setminus B$. Fix a vertex $u \in B$ and let N be the set of red neighbours of u outside B. Every $w \in B$ has at least the following number of red neighbours in N.

$$2 \cdot ((2n-5)/3 - |B|) - (n - |B|) = (n-10)/3 - |B| \geq (n-21)/6.$$

Now let U' be a random subset of N where each vertex $w \in N$ belongs to U', independently, with probability $13 \log n/n$. Let I_w be the event that w (where $w \in B$) does not have a red neighbour in U'. We bound

$$\mathbb{P}\Big(\bigcup_{w \in B} I_w\Big) \leq |B| \cdot \mathbb{P}(I_w) \leq n \cdot \left(1 - \frac{13 \log n}{n}\right)^{\frac{n-21}{6}} \leq n \cdot e^{-2 \log n} < 1/2.$$

Note that since $\mathbb{E}(|U'|) \leq 13 \log n$, we have $\mathbb{P}(|U'| \geq 26 \log n) \leq 1/2$, by Markov's inequality. Therefore, there is a choice of $U' \subseteq N$ such that $|U'| \leq 26 \log n$ and every vertex in B is joined by a red edge to some vertex in U'. We choose $U = U' \cup \{u\}$. Note that

$$\begin{aligned} N_r(U' \cup \{u\})| &\geq |N \setminus U'| + |B \setminus \{u\}| \\ &\geq ((2n-5)/3 - |B| - 26\log n) + (|B| - 1) \\ &= 2n/3 - 27\log n. \end{aligned}$$

Hence, the set $U = U' \cup \{u\}$ satisfies the requirements of Claim 3.2. □

Now, let U be a red set as in Claim 3.2 and let $N = N_r(U)$. Now choose a maximal sequence of distinct vertices $x_1, \dots, x_t \in V \setminus (N \cup U)$ so that x_i has at least $\log n$ red neighbours in the set $N \cup \{x_1, \dots, x_{i-1}\}$, for every $i \in [t]$. Then put $\overline{N} = N \cup \{x_1, \dots, x_t\}$ and write $W = V(G) \setminus \big(U \cup \overline{N}\big)$. We may assume $W \neq \varnothing$, otherwise $V(G)$ would form a red component, which is a contradiction to our assumption. Moreover, note that every vertex in W has at most $\log n$ red neighbours in $\overline{N}$.

Claim 3.3. $\big|\overline{N}\big| \leq 2n/3 + 3\log n + 4$.

Proof of claim. For a contradiction, suppose that $\big|\overline{N}\big| > 2n/3 + 3\log n + 4$. We shall deduce that the vertices can be partitioned into a red set and a blue one, a contradiction to our assumption. To define the partition, fix $w \in W$ and let $X = N_b(w) \cap \overline{N}$. Let S be a random subset of X, obtained by taking each vertex of X independently with probability $1/2$. We claim that, with positive probability, $(U \cup \overline{N}) \setminus S$ is red *and* $W \cup S$ is blue. To bound the probability that $W \cup S$ is blue, we consider the probability that every vertex in W is joined by a blue edge to S (an event which would imply that $W \cup S$ is blue). For every $x, y \in V$ we have $|N(x) \cap N(y)| \geq n/3 - 10/3$, hence $|N(x) \cap N(y) \cap \overline{N}| \geq 3\log n$. Since every vertex in W has at most $\log n$ red neighbours in $\overline{N}$, we have $|N_b(x) \cap N_b(y) \cap \overline{N}| \geq \log n$. Therefore the probability that a given $x \in W$ has no blue neighbours in S is at most $2^{-\log n} = 1/n$. Thus, the expected number of vertices in W with no edges to S is smaller than $1/2$ (note that $|W| \leq n/3$). Hence, $\mathcal{P}(W \cup S \textit{ is blue}) > 1/2$. We now estimate the probability that $(U \cup \overline{N}) \setminus S$ is red. First note that as $N = N_r(U)$, we have that

$U \cup N'$ is red for any subset $N' \subseteq N$. So it remains to show that the vertices of $\{x_1, \dots, x_t\} \setminus S$ can be joined, via a red path, to $U \cup (N \setminus S)$, with sufficiently high probability. For $i \in [t]$, let E_i be the event that vertex x_i is joined by a red edge to $(N \cup \{x_1, \dots, x_{i-1}\}) \setminus S$. Note that if the event $E = \bigcap_i^t E_i$ holds, $(U \cup \overline{N}) \setminus S$ is red. Now, to estimate $\mathcal{P}(E_i)$, for $i \in [t]$, note that each vertex x_i has at least $\log n$ forward neighbours, and the probability that one of these vertices is deleted is at most $1/2$. Thus $\mathcal{P}(E_i) \geq 1 - 2^{-\log n} = 1 - 1/n$, therefore $\mathcal{P}\big((U \cup \overline{N}) \setminus S \text{ is red}\big) \geq \mathcal{P}(E) > 1/2$, where the second inequality holds since $t < n/2$. Thus, with positive probability, $W \cup S$ is blue and $(U \cup \overline{N}) \setminus S$ is red. In particular, the vertices can be partitioned into a blue set and a red one, a contradiction. $\square$

Claim 3.4. *There is a vertex of blue degree at most* $90 \log n$.

Proof of Claim. By definition of $\overline{N}$ and since $|\overline{N}| \geq 2n/3 - 27 \log n$, every vertex in W has at least $n/3 - 29 \log n$ blue neighbours in $\overline{N}$.

Fix a vertex $w \in W$. If there is a vertex $v \in W$ with $|N_b(v) \cap N_b(w) \cap \overline{N}| < \log n$, then the blue components containing v and w (at most 2) cover all vertices of W and all but at most $62 \log n$ vertices of $\overline{N}$ (as $|\overline{N}| \leq 2n/3 + 3 \log n + 4$, by the previous claim). Since $|U| \leq 27 \log n$, it follows that these two components cover all but at most $90 \log n$ vertices. Recall that there are at least three blue components, hence there is a component of order at most $90 \log n$, and any vertex in that component has blue degree at most $90 \log n$. Otherwise, every vertex $v \in W$ satisfies $|N_b(v) \cap N_b(w) \cap \overline{N}| \geq \log n$. As in Claim 3.3, let S be an uniformly random subset of $N_b(w) \cap \overline{N}$; we find that, with positive probability, $\big(U \cup \overline{N}\big) \setminus S$ is red and $W \cup S$ is blue, so the vertices can be partitioned into a red set and a blue one, a contradiction to our assumption. $\square$

Let x_r be a vertex of blue degree at most $90 \log n$, which exists by the previous claim. By symmetry, there is a vertex x_b of red degree at most $90 \log n$. Then $d_r(x_r), d_b(x_b) \geq 2n/3 - 90 \log n - 2$. Write $A_1 = N_b(x_b) \setminus N_r(x_r)$,

$A_2 = N_b(x_b) \cap N_r(x_r)$ and $A_3 = N_r(x_r) \setminus N_b(x_b)$. Then $|A_2| \geq n/3 - 180 \log n - 4$ and $|A_1|, |A_3| \leq n/3 + 90 \log n + 2$.

CLAIM 3.5. *There is a vertex with no blue neighbours in A_1, no red neighbours in A_3, and at most $2 \log n$ neighbours in A_2.*

PROOF OF CLAIM. Suppose that the statement does not hold. Let $\{B, R\}$ be a random partition of A_2, obtained by putting vertices in B, independently, with probability $1/2$. Then, with positive probability, every vertex in G has a blue neighbour in $A_1 \cup B \subseteq N_b(x_b)$ or a red neighbour in $A_3 \cup R \subseteq N_r(x_r)$. We thus obtain a partition of the vertices into a red set and a blue one, a contradiction. $\square$

Let x be a vertex with no blue neighbours in A_1, no red neighbours in A_3, and at most $2 \log n$ neighbours in A_2 (its existence is guaranteed by the previous claim). This implies that $|A_2| \leq n/3 + 3 \log n$, so $|A_1|, |A_3| \geq n/3 - 95 \log n$. Furthermore, x has at least $n/3 - 100 \log n$ red neighbours in A_1 and at least $n/3 - 100 \log n$ blue neighbours in A_3. Write $A_1' = A_1 \cap N_r(x)$, $A_2' = A_2 \setminus N(x)$, and $A_3' = A_3 \cap N_b(x)$ (so $|A_1'|, |A_3'| \geq n/3 - 100 \log n$ and $|A_2'| \geq n/3 - 190 \log n$).

CLAIM 3.6. *The vertices x and x_b belong to distinct blue components; similarly, x and x_r belong to distinct red components.*

PROOF OF CLAIM. Suppose for a contradiction that x and x_b are in the same blue component. Then there is a blue path P from $\{x\} \cup A_3'$ to $\{x_b\} \cup A_1' \cup A_2'$. We may assume that the inner vertices of P are outside of $A_1' \cup A_2' \cup A_3' \cup \{x, b\}$. Hence, $|P| \leq 400 \log n$.

Now, let $\{B, R\}$ be a random partition of $(A_2' \cup A_3') \setminus V(P)$, obtained by putting a vertex from $(A_2' \cup A_3') \setminus V(P)$ in B, independently, with probability $1/2$. As every vertex of G has at least $10 \log n$ neighbours in $(A_2' \cup A_3') \setminus V(P)$, we obtain that with positive probability, every vertex in G has a red neighbour in R or a blue neighbour in B, from which it can be deduced that there is a partition of the vertices into a red set and a blue one, which is a contradiction.

Indeed, observe that $P \cup \{x, x_b\} \cup B$ is a blue set and $\{x_r\} \cup R$ is a red set. Thus, we have that x_b and x are in distinct blue components; by symmetry, x_r and x are in different red components. $\square$

We know that $|C_b(x_b)|, |C_r(x_r)| \geq 2n/3 - 91 \log n$ and $|C_b(x)|, |C_r(x)| \geq n/3 - 100 \log n$. Recall also that there are at least three blue components. Hence, there is a vertex w_r which is not in $C_b(x_b)$ or in $C_b(x)$. It follows that $d_b(w_r)$ is at most $191 \log n$, hence it has red degree at least $2n/3 - 192 \log n$, so $w_r \in C_r(x_r)$. Similarly, there is a vertex w_b which is not in $C_r(x_r)$ or in $C_r(x)$, and therefore it must belong to $C_b(x_b)$. We claim that the set $X = \{w_b, w_r, x\}$ is independent. Observe that both edges $w_r x$ or $w_b x$ can not belong to G. Indeed, if the edge $w_r x$ was present then it either has blue colour which contradicts the choice of $w_r \notin C_b(x)$ or colour red in which case it implies x and x_r belong to the same red component, contradicting Claim 3.6. A symmetric argument proves that it can not have red colour. Therefore, $w_r x$ does not belong to G and similarly $w_b x$. If we had $w_r w_b \in E(G)$ and this edge was coloured red then $w_b \in C_r(x_r)$ which is a contradiction, by definition of w_b. If $w_r w_b$ was coloured blue then we arrive at the contradiction $w_r \in C_b(w_b)$. Thus X is independent. Observe that we have actually showed that no two vertices in X can belong to the same monochromatic component. Finally, by the minimum degree condition, there must be a vertex z that is adjacent to all three vertices in X. Indeed, if no such w exists, then the number of edges between X and $V(G) \setminus X$ is at most $2(n-3) < 3(2n-5)/3$, a contradiction. Without loss of generality, w sends two red edges into X, implying that two of these vertices in X belong to the same red component, a contradiction. This completes our proof of Theorem 1.2. $\square$

4. Covering with monochromatic components of distinct colours

In this section, we verify Conjecture 1.5 for $r \in \{2, 3\}$. Most of the difficulty is in the proof for $r = 3$, but we include a short proof for $r = 2$, for completeness. Actually, the $r = 2$ case (for n large) already follows from a difficult result of

Letzter [77], who showed that when $\delta(G) \geq 3n/4$, the vertices can be partitioned into two monochromatic *cycles* of different colours, for every 2-colouring of G. Before turning to the proofs, we mention the following construction of Bal and DeBiasio [14], which shows that the minimum degree condition in Conjecture 1.5 cannot be improved.

EXAMPLE 4.1. Let $n \geq 2^r$; we shall define a graph on vertex set $[n]$ as follows. Partition $[n]$, as equally as possible, into 2^r sets which are indexed by the sequences $s \in \{0,1\}^r$. We write

$$[n] = \bigcup_{s \in \{0,1\}^r} A(s)$$

and define the following, where $\mathbb{1} = (1, \ldots, 1)$.

$$E = [n]^{(2)} \setminus \bigcup_{s \in \{0,1\}^r} \{xy : x \in A(s), y \in A(\mathbb{1} - s)\}.$$

In other words, we include all edges in the graph except for the edges between parts of the partition corresponding to antipodal elements of $\{0,1\}^r$. Now, colour all edges xy, where $x \in A(s), y \in A(s')$, by the first coordinate on which s, s' agree; e.g. the edge between $(0,1,0,0)$ and $(1,0,0,1)$ is coloured 3.

We now show that G cannot be covered by components of distinct colours. Suppose that it can, and note that the i-coloured components are of the form $\bigcup_{s \in S_i} A(s)$ where S_i is a set of elements that agree on their i-th coordinate; denote this coordinate by a_i. It follows that the vertices of $A((1-a_1, \ldots, 1-a_r))$ are not covered by any of these components, a contradiction.

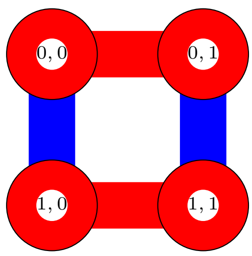

FIGURE 2. an illustration of Example 4.1 for $r = 2$

We now prove Conjecture 1.5 for $r = 2$.

Lemma 4.2. *Let G be a 2-coloured graph with $\delta(G) \geq 3n/4$. Then the vertices of G can be covered by a red component and a blue component.*

Proof of lemma. We first show that there is a monochromatic component of order greater than $n/2$. If G is red connected we are done. Hence, there exists a red component R with $|R| \leq n/2$. Then, any two vertices $u, w \in R$ have a common blue neighbour, as $|N_b(u) \cap N_b(w) \cap \overline{R}| \geq 2 \cdot (3n/4 - (|R| - 1)) - (n - |R|) > 0$. So $R \subseteq C_b(u)$ and $C_b(u)$ is a blue component of order at least $3n/4$, as required.

Without loss of generality, there is a red component R of order larger than $n/2$. Note that there is a vertex x which is not in R (otherwise we are done), and $|N_b(x) \cap R| = |N(x) \cap R| > n/4$, as x does not send red edges to R. In particular, $|C_b(x) \cap R| > n/4$. It follows that every vertex sends at least one edge to $C_b(x) \cap R$ and thus the components R and $C_b(x)$ cover the whole graph. $\square$

We now turn to prove Theorem 1.6, which is the case of three colours in Conjecture 1.5.

Proof of Theorem 1.6. We begin with a series of preparatory claims.

Claim 4.3. *If there are three monochromatic components of distinct colours whose intersection has order at least $n/8$, then the vertices can be covered by monochromatic components of distinct colours.*

Proof of claim. Suppose that R, B and Y are red, blue and yellow components, whose intersection $U = R \cap B \cap Y$ has size at least $n/8$. Then, by the minimum degree condition, every vertex not in U has a neighbour in U, implying that every vertex in the graph belongs to at least one of R, B and Y, as required. $\square$

Claim 4.4. *If there are two monochromatic components of distinct colours whose intersection has order at least $n/4$, then the vertices of G may be covered by monochromatic components of distinct colours.*

Proof of claim. Suppose that R and B are red and blue components whose intersection $U = R \cap B$ has size at least $n/4$. We show that one of the following holds.

(1) $R \cup B = V(G)$;
(2) there is a yellow component whose intersection with $R \cap B$ has size at least $n/8$.

Suppose that the first assertion does not hold. Then there is a vertex $u \notin R \cup B$. By the minimum degree condition, u sends at least $n/8$ edges to $R \cap B$, but these edges cannot be red or blue (because $u \notin R \cup B$), hence they are yellow, so by picking Y to be the yellow component containing u, the second assertion holds. If the first assertion holds, we are done immediately; otherwise, we are done by Claim 4.3. $\square$

Claim 4.5. *If there is a monochromatic component of order at least $n/2$, then the vertices can be covered by three monochromatic components of distinct colours.*

Proof of claim. As in the proof of Claim 4.4, we show that one of the following assertions holds, where R is a red component of order at least $n/2$.

(1) $R = V(G)$;
(2) there are monochromatic components B and Y in colours blue and yellow respectively, such that $R \cup B \cup Y = V(G)$;
(3) there are monochromatic components B and Y in colours blue and yellow respectively, such that $|R \cap B \cap Y| \geq n/8$.

Suppose that $R \neq V(G)$ and let $u \notin R$. Consider the blue and yellow components, B and Y, containing u. By the minimum degree condition, u sends at least $|R| - n/8$ edges to R, none of which are red. So $|(B \cup Y) \cap R| \geq |R| - n/8$.

Suppose that R, B and Y do not cover the whole graph. Let $w \notin R \cup B \cup Y$, and denote the blue and yellow components containing w by B' and Y'. By the same argument as before, $|(B' \cap Y') \cap R| \geq |R| - n/8$, which implies the following.

$$|(B \cup Y) \cap (B' \cup Y') \cap R| \geq |R| - n/4 \geq n/4.$$

Since $B \cap B' = \varnothing$ and $Y \cap Y' = \varnothing$, either $|B \cap Y' \cap R| \geq n/8$ or $|B' \cap Y \cap R| \geq n/8$. This completes the proof that one of the above assertions holds. If one of the first two assertions holds, we are done immediately; and if the third assertion holds, Claim 4.5 follows from Claim 4.3. □

Henceforth, we assume G cannot be covered by monochromatic components of distinct colours.

Claim 4.6. *There are two monochromatic components of distinct colours of order at least* $3n/8$.

Proof of claim. We will show that for every pair of colours there is a monochromatic component of order at least $3n/8$ in one of the two colours; the claim easily follows from this fact. Let the two colours be red and blue. Since G is not connected in yellow, we may find a partition $\{X, Y\}$ of the vertices of G such that no $X - Y$ edges are yellow. Without loss of generality, at least half the edges between X and Y are red; set $H = G_r[X, Y]$, denote $d(x) = d_H(x)$ for any vertex x, and given an edge xy in H, set $s(xy) = d(x) + d(y)$. We will show that there is an edge xy with $s(xy) \geq 3n/8$; note that this would imply the existence of a red component of order at least $3n/8$, as required. Put $e = e(H)$ and, without loss of generality, we assume that $|X| \leq |Y|$. We have

$$\begin{aligned} \frac{1}{e} \sum_{xy \in E(H)} s(xy) &= \frac{1}{e} \sum_{xy \in E(H)} (d(x) + d(y)) \\ &= \frac{1}{e} \left(\sum_{x \in X} d(x)^2 + \sum_{y \in Y} d(y)^2 \right) \end{aligned}$$

$$
\begin{aligned}
&\geq \frac{1}{e}\left(\frac{\left(\sum_{x\in X} d(x)\right)^2}{|X|} + \frac{\left(\sum_{y\in Y} d(y)\right)^2}{|Y|}\right)\\
&= e\left(\frac{1}{|X|} + \frac{1}{|Y|}\right)\\
&\geq \frac{1}{2}|X|(|Y| - n/8)\left(\frac{1}{|X|} + \frac{1}{|Y|}\right)\\
&= \frac{1}{2}\left(|Y| - n/8 + |X| - \frac{|X|\cdot n/8}{n-|X|}\right)\\
&\geq 3n/8.
\end{aligned}
$$

Indeed, the first inequality follows from the Cauchy-Schwarz inequality; the second follows from the minimum degree condition and the assumption that red is the majority colour between X and Y; and the last inequality follows since $|X| + |Y| = n$ and the expression $\frac{|X|}{n-|X|}$ is maximised when $|X| = n/2$ (as we have the constraint $|X| \leq n/2$).

This chain of inequalities shows that the average value of $s(xy)$ is at least $3n/8$; in particular, there is a red component of order at least $3n/8$, as required. $\square$

We remark that the idea of double counting $s(xy)$ as in the proof of the previous claim originated in a paper by Liu, Morris and Prince [79].

By the previous claim, we may assume that R and B are red and blue components of order at least $3n/8$.

Claim 4.7. *Either* $|R \setminus B| < n/4$ *or* $|B \setminus R| < n/4$.

Proof of claim. Assume that $|R \setminus B| \geq n/4$ and $|B \setminus R| \geq n/4$. Note that every edge between the disjoint sets $R \setminus B$ and $B \setminus R$ is yellow. Furthermore, any two vertices in $B \setminus R$ have a common neighbour in $R \setminus B$, and vice versa. Therefore $B \triangle R$ is contained in a yellow component; in particular, there is a yellow component of order at least $n/2$, a contradiction, by Claim 4.5. $\square$

By the previous claim, we may assume that $|B \setminus R| < n/4$. Hence, $|B \cup R| = |R| + |B \setminus R| < n/2 + n/4 = 3n/4$, by Claim 4.5. Therefore, the set $W = V(G) \setminus (R \cup B)$ has size larger than $n/4$. Since all edges between $R \cap B$ and W are yellow, it follows that every two vertices in $R \cap B$ have a common yellow neighbour in W and hence $R \cap B$ is contained in a yellow component. Thus, Claim 4.3 implies that $|R \cap B| < n/8$. It follows that $|B| = |B \cap R| + |B \setminus R| < 3n/8$, in contradiction with the choice of B. This completes the proof of Theorem 1.6. $\square$

5. Covering a two coloured graph by few monochromatic components

In this section, we shall prove Proposition 1.4, which is a weaker version of Conjecture 1.3, where instead of partitioning the vertices, we aim to *cover* the vertices.

Proof of Proposition 1.4. We use the link with König's Theorem first noted by Gyárfás [55]. Let G be a 2-coloured graph with minimum degree at least $\frac{2n-2t-1}{t+1}$. Let $\mathcal{R}$ be the collection of red components (some of which may be singletons, if there are vertices that are not incident with any red edges), and let $\mathcal{B}$ be the collection of blue components. Define an auxiliary bipartite graph $H = (\mathcal{R}, \mathcal{B}, E)$, where for $R \in \mathcal{R}$ and $B \in \mathcal{B}$, we have $RB \in E$ if and only if $R \cap B \neq \varnothing$.

We claim that there is no matching of size larger than t. Indeed, suppose that $\{R_1B_1, \dots, R_{t+1}B_{t+1}\}$ is a matching of size $t+1$. Let $u_i \in R_i \cap B_i$, for $i \in [t+1]$ and $U = \{u_1, \dots, u_{t+1}\}$. Then the vertices of U are in distinct red and blue components. In particular, U is independent, so the number of edges between U and $V(G) \setminus U$ is at least $2n - 2t - 1$. On the other hand, no vertex sends more than one red edge into U (and similarly for blue), so every vertex not in U sends at most two edges into U. It follows that the number of edges between U and $V(G) \setminus U$ is at most $2(n - t - 1) < 2n - 2t - 1$, a contradiction.

By König's theorem, which states that in bipartite graphs, the size of a minimum cover equals the size of a maximum matching, it follows that there is a cover W of size at most t; write $W = \{C_1, \dots, C_t\}$. We claim that $V(G) = C_1 \cup \dots \cup C_t$. Indeed, consider a vertex u and denote its red and blue components by R and B, respectively. Then $R \cap B \neq \varnothing$, hence RB is an edge in H, so either R or B is in W, which implies that $u \in C_1 \cup \dots \cup C_t$, as required. In other words, the vertices of G can be covered by at most t monochromatic components. □

Finally, we note that the restriction on the minimum degree in Proposition 1.4 (and therefore in Conjecture 1.3) cannot be improved. The special case of this example, where $t = 2$, appears in [14] and shows that the minimum degree condition in Theorem 1.2 is best possible.

Example 5.1. Let U be a set of size $n \geq t + 1$, and let $\{X, A_1, \dots, A_{t+1}\}$ be a partition of U, where $|X| = t + 1$ and the sizes of $A_1, A_2, \dots, A_{t+1}$ are as equal as possible; write $X = \{x_1, \dots, x_{t+1}\}$. We define a 2-coloured graph G on vertex set U as follows.

- the sets A_i are cliques, and we colour them arbitrarily;
- we add all possible edges between A_i and A_{i+1}, where $i \in [t]$, and colour them red if i is odd, and blue otherwise;
- we add all edges between x_i and $A_i \cup A_{i+1}$, for $i \in [t + 1]$ (addition is taken modulo $t + 1$). We colour these edges red if i is in $[t]$ and i is odd; and blue if i is in $[t]$ and i is even. Finally, we colour the edges from x_{t+1} to A_1 blue, and colour the edges from x_{t+1} to A_{t+1} red if t is even and blue if t is odd.

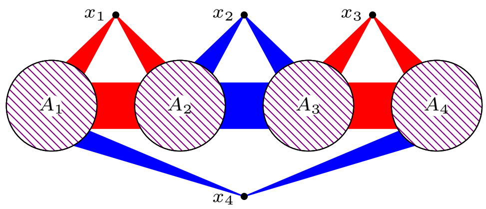

FIGURE 3. an illustration of Example 5.1 for $t = 3$.

An easy calculation shows that G has minimum degree[1] $\lceil (2n - 2t - 1)/(t + 1)\rceil - 1$, and that no two vertices in X belong to the same monochromatic component; in particular, the vertices of G cannot be covered by at most t monochromatic components.

6. Concluding remarks

First, we remind the reader of Conjecture 1.5 by Bal and DeBiasio [14]; In this chapter we proved this conjecture for $r \le 3$.

Another beautiful conjecture stated by Bal and Debiasio [14] concerns the minimum degree needed to ensure that an r-coloured graph can be covered by at most r monochromatic components, whose colours need not to be distinct.

CONJECTURE 6.1 (Bal and Debiasio). *Let G be an r-coloured graph on n vertices with $\delta(G) \ge \frac{r(n-r-1)+1}{r+1}$. Then the vertices of G can be covered by at most r monochromatic components.*

We further recall our Conjecture 1.3. In this conjecture, we attempt to determine the minimum degree condition needed to guarantee the existence of a partition of a 2-coloured graph into t monochromatic connected subgraphs. This is a generalisation of Theorem 1.2 which determines this condition for a partition into two monochromatic connected sets.

[1] In fact, we need to be a bit more careful here. Write $n = a(t + 1) + d$, where a and d are positive integers and $0 \le d \le t$. We consider two cases: $d < \lceil (t + 1)/2\rceil$ and $d \ge \lceil (t + 1)/2\rceil$. In the former case, it is easy to see that $\delta(G) = \lceil (2n - 2t - 1)/(t + 1)\rceil - 1$. In the latter case, note that exactly d of the sets A_i have size a, and the rest have size $a - 1$. Then, again, one can check that $\delta(G) = \lceil (2n - 2t - 1)/(t + 1)\rceil - 1$ if $|A_i| = a$ for every odd $i \in [t + 1]$ (which is possible as $d \ge (t + 1)/2$).

CHAPTER 5

Rainbow saturation colourings

1. Introduction

In Extremal graph theory, over many decades, much attention has been paid to the following two types of question. One is the classical Turán-type problem [109] which asks for the maximum number of edges a graph on n vertices can have provided it does not contain as a subgraph any member of a fixed class of graphs. The other question is concerned with another extreme, namely to determine the minimum number of edges in a graph G on n vertices which is $\mathcal{H}$-free but for which the addition of any edge between two non-adjacent vertices of G creates a copy of H, for some graph H. A maximal (with respect to inclusion) $\mathcal{H}$-free graph G is said to be $\mathcal{H}$-*saturated*. The latter question can then be reformulated: what is the smallest number of edges in a $\mathcal{H}$-saturated graph on n vertices? This number, usually denoted by $\mathrm{sat}(n, \mathcal{H})$, was studied by Zykov [117] and independently by Erdős, Hajnal, and Moon [48] who proved that $\mathrm{sat}(n, K_r) = (r-2)(n-1) - \binom{r-2}{2}$. Soon after, Bollobás [19] showed that $\mathrm{sat}(n, K_s^\ell) = \binom{n}{\ell} - \binom{n-(s-\ell)}{\ell}$, where K_s^ℓ is the complete ℓ-uniform hypergraph on s vertices and he conjectured $\mathrm{sat}(n, H) = O(n)$, for any class of graphs $\mathcal{H}$. Kászonyi and Tuza [71], in 1986 confirmed this conjecture. For more information on saturation numbers we refer the reader to the survey of Faudree, Faudree, and Schmitt [49].

In the present chapter, we will be interested in a variation of the saturation numbers, following the approach of Hanson and Toft [60], who extended this notion to edge coloured graphs. We need introduce some definitions first. We define a t-edge coloured graph to be an ordered pair (G, c), where G is a graph and c is a t-edge colouring of G, i.e., function from the edge set of G to the set $\{1, 2, 3, \ldots, t\}$, whose elements we call colours. An edge coloured subgraph of G is a pair $(H, c|_{E(H)})$, where H is any subgraph of G. Throughout the chapter, we will usually identify the coloured graph (G, c) with the graph G, especially when it is clear from the context which colouring is being used. Note that we do not require edge colourings to be proper. Given an integer t and a family $\mathcal{F}$ of t-edge coloured graphs, we say that a t-edge coloured graph (G, c) is $(\mathcal{F}, t)$-saturated if (G, c) contains no member of $\mathcal{F}$ as an edge coloured subgraph, but the addition of any non-edge in any colour from the set $\{1, 2, \ldots, t\}$ creates a copy of a coloured graph in $\mathcal{F}$. Similarly to the usual saturation problem, one denotes by $\mathrm{sat}_t(n, \mathcal{F})$ the minimum number of edges in a $(\mathcal{F}, t)$-saturated t-edge coloured graph on n vertices. In [60], Hanson and Toft proved that for any sequence of positive integers $2 \leq k_1 \leq k_2 \leq \ldots \leq k_m$,

$$\mathrm{sat}_t(n, \mathcal{M}(K_{k_1}, K_{k_2} \ldots, K_{k_m})) = \begin{cases} \binom{n}{2} & \text{if } n \leq k - 2m \\ \binom{k-2m}{2} + (k-2m)(n-k+2m) & \text{if } n > k - 2m, \end{cases}$$

where $k = \sum_{i=1}^{t} k_i$ and $\mathcal{M}(K_{k_1}, K_{k_2} \ldots, K_{k_m})$ is the collection of coloured graphs consisting of a monochromatic copy of K_{k_i} in colour i, for each $i \in \{1, 2, \ldots, m\}$.

In this chapter, we investigate some problems proposed by Barrus, Ferrara, Vandenbussche, and Wenger [17]. Given a graph H and $t \geq e(H)$, we let $\mathfrak{R}(H)$ to be the collection of all rainbow copies of H, i.e. all t-edge coloured graphs (H, c) where each edge is assigned a different colour from $\{1, 2 \ldots, t\}$. We shall call $\mathrm{sat}_t(n, \mathcal{R}(H))$ the *t-rainbow saturation number of H*, and, if the set of colours is infinite (say the set of natural numbers) we shall simply write $\mathrm{sat}(n, \mathcal{R}(H))$ and call it the *rainbow saturation number of H*. Our goal throughout the chapter is to determine the value of $\mathrm{sat}_t(n, \mathcal{R}(H))$ for a fixed graph H.

The authors of [17] proved several beautiful and surprising results concerning these numbers. In particular, they showed a rather interesting phenomenon, namely that there are graphs whose t-rainbow saturation numbers grow considerably faster as a function of n then the usual saturation numbers. For example, they proved that for every integer r and $t \geq \binom{r}{2}$ there exist two positive constants c_1, c_2 such that

$$c_1 \frac{n \log n}{\log \log n} \leq \mathrm{sat}_t(n, \mathcal{R}(K_r)) \leq c_2 n \log n.$$

In the same paper, the authors determined the t-rainbow saturation number of stars, showing that $\mathrm{sat}_t(n, \mathcal{R}(K_{1,k})) = \Theta(n^2)$ for any positive integers $t \geq k \geq 2$. This result confirms that the growth rates of rainbow saturation numbers behave very differently from the usual saturation numbers. They also state the following conjecture.

Conjecture 1.1 ([17]). *For any integers r and t with $t \geq \binom{r}{2}$, $\mathrm{sat}_t(n, \mathcal{R}(K_r)) = \Theta(n \log n)$.*

One of our aims in this chapter is to prove this lovely conjecture. Moreover, we show that any graph H without isolated vertices satisfying $\mathrm{sat}_t(n, \mathcal{R}(H)) = \Theta(n^2)$, for some $t \geq e(H)$, must be a star. This answers a question posed in [17] asking if stars were the only graphs with quadratic t-rainbow saturation numbers. Observe that the function $\mathrm{sat}_t(n, \mathcal{R}(H))$ is monotonically decreasing in t for every graph H. Therefore, one just needs to show $\mathrm{sat}_t(n, \mathcal{R}(H)) = o(n^2)$ when $t = e(H)$. Indeed, we show the following stronger result.

Theorem 1.2. *Let H be a graph without isolated vertices which is not a star. Then, for any $t \geq e(H)$,*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) = O(n \log n).$$

Observe trivially that the addition of isolated vertices does not change the rainbow saturation numbers for all n sufficiently large.

Given a graph H, we say that a vertex $x \in V(H)$ is *conical* if its degree is $|H| - 1$ and we say an edge is *pendant* if one of its endpoints has degree 1. For any $r \geq 4$, we define K_r *with a rotated edge* to be the graph obtained by taking with a copy of K_r, adding a new vertex, and "rotating" one edge by replacing one of its endpoints with the new vertex, as in Figure 1.

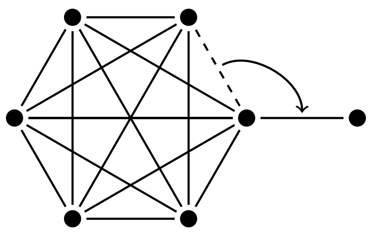

FIGURE 1. K_6 with a rotated edge. The dashed line represents the removed edge.

In the next result, we completely characterize the growth rates of t-rainbow saturation numbers of every connected graph H with no leaves, for every $t \geq e(H)$. Actually, we prove a slightly stronger result.

THEOREM 1.3. *Let H be a connected graph of order at least* 3. *Then, for every $t \geq e(H)$,* $\operatorname{sat}_t(n, \mathcal{R}(H))$ *equals:*

(1) $\Theta(n^2)$, *if H is a star.*
(2) $\Theta(n \log n)$, *if H has a conical vertex but is not a star.*
(3) $\Theta(n \log n)$, *if every edge of H is in a triangle.*
(4) $\Theta(n)$, *if H contains a non-pendant edge which does not belong to a triangle.*
(5) $\Theta(n)$, *if H is a K_r with a rotated edge, for some even $r \geq 4$.*

We note that if H is connected with no pendant edges, then, for any $t \geq e(H)$, $\operatorname{sat}_t(n, \mathcal{R}(H)) = \Theta(n \log n)$ if every edge belongs to a triangle (by 3) and $\operatorname{sat}_t(n, \mathcal{R}(H)) = \Theta(n)$ otherwise (by 4).

It is easy to check that all graphs excluded from the classification of Theorem 1.3 can be constructed by starting with a connected graph in which every edge lies in a triangle and adding pendant edges to the graph. Observe that not all graphs constructed in this way are excluded, as the class of such graphs includes all cliques with a rotated edge and some graphs with a conical vertex. For simplicity, we denote by $\mathcal{B}$ the class of all connected graphs excluded from the classification of Theorem 1.3.

Although we have not determined the correct order of magnitude of the t-rainbow saturation numbers of any graph H in $\mathcal{B}$ for all $t \geq e(H)$, in almost all cases, we were able to determine the order of magnitude of $\mathrm{sat}_t(n, \mathcal{R}(H))$ for all sufficiently large values of t. The authors of [17] also showed that if H is a graph on at least five vertices with a leaf whose neighbour is not a conical vertex and the rest of the vertices do not induce a clique then for any $t \geq \binom{|H|-1}{2}$ we have $\mathrm{sat}_t(n, \mathcal{R}(H)) = \Theta(n)$. Our next result covers almost all the remaining graphs containing a pendant edge. We show that for every H in $\mathcal{B}$ (with the exception of K_r with a rotated edge, r odd), the t-rainbow saturation number of H is linear in n, for all t sufficiently large.

Theorem 1.4. *Let H be a connected graph with no conical vertex and containing at least one pendant edge. Moreover, suppose H is not a copy of K_r with a rotated edge for odd $r \geq 5$. Then, for every $t \geq |H|^2$,*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) = \Theta(n).$$

In all results discussed above, we assumed that the number of available colours, t, is always fixed and does not grow with n. In Theorem 4.13 we scratch the surface of the case when $t = t(n)$ grows with n and prove that for any $r \geq 3$ there exists a constant $c_r > 0$ such that, for any $t \geq \binom{r}{2}$, we have

$$\mathrm{sat}_t(n, \mathcal{R}(K_r)) \leq \max\left\{\frac{c_r}{\log t} n \log n, 2(r-2)n\right\}.$$

In particular, this shows (by taking $t(n)$ to be at least linear in n) that $\mathrm{sat}(n, \mathcal{R}(K_r)) = \Theta(n)$, for any $r \geq 3$.

Finally, we shall remark that we did not rule out the existence of a 'sharp threshold' for some connected graph H, i.e., a $t \geq e(H)$ such that $\mathrm{sat}_{t+1}(n, \mathcal{R}(H)) = o(\mathrm{sat}_t(n, \mathcal{R}(H)))$ as a $n \to \infty$. However, if such graph exists it must belong to $\mathcal{B}$, by Theorem 1.3. Note also that the set of connected graphs for which we have *not* determined the correct growth rate of their t-rainbow saturation numbers for large enough t consists exactly of the aforementioned K_r's with a rotated edge for odd $r \geq 5$.

2. Organization and notation

In section 3, we prove lower bounds for the t-rainbow saturation number of two classes of graphs, namely graphs where every edge belongs to a triangle and graphs which contain a conical vertex, allowing us to establish the correctness of Conjecture 1.1. In Section 4, we shall prove Theorem 1.2 when restricted to the class of connected graphs, as well as the main parts of the proof of Theorem 1.3 and Theorem 1.4. We split the argument in the following way. First, in Subsection 4.1, we show item 4 of Theorem 1.3 and in Subsection 4.2, we prove Theorem 1.2 assuming the graph is connected. Secondly, in Subsection 4.3, we establish item 5. In Subsection 4.4, we shall give upper bounds (depending on t), for the t-rainbow saturation numbers of complete graphs. We also show that, when the palette of colours is infinite, the rainbow saturation numbers of complete graphs are linear. In Section 5, we complete the proof of Theorem 1.2, showing it also holds for disconnected graphs without isolated vertices. In Section 6, we deduce from the results proved in previous Sections Theorem 1.3 and Theorem 1.4. Finally, in Section 7 we make some remarks and propose some conjectures and questions that we would like to be investigated.

The notation we use is mostly standard. For a graph G we define $e(G)$ to be the number of edges in G. For $S \subseteq V(G)$, we denote by $e(S)$ the number of edges with both endpoints in S, and, for $S, T \subseteq V(G)$, we denote by $e(S,T)$ the number of edges with one endpoint in S and the other in T. A *non-edge* of G is an edge of $\overline{G}$. Moreover, we say a non-edge in a graph G is $\mathfrak{R}(H)$-saturated if adding e in any colour from the palette of colours understood by the context creates a rainbow copy of H. Also, if v is a vertex in an edge coloured graph, we say informally that v *sees* a given colour if it is incident with an edge of that colour. For any positive integer k, we define the *k-star* to be the graph $K_{1,k}$. All logarithms are base 2.

3. Lower bounds

In this Section, we show that if a graph possesses certain properties then its t-rainbow saturation numbers grow at least as fast as $n \log n$. Before doing so, we will need the following trivial lower bound for the rainbow saturation numbers of a connected graph on at least three vertices.

LEMMA 3.1. *If H is a connected graph on at least three vertices then* $\mathrm{sat}(n, \mathcal{R}(H)) \geq \frac{n-1}{2}$.

PROOF. It is easy to check that if G is an $\mathcal{R}(H)$-saturated graph then it has at most one isolated vertex, hence $e(G) \geq \frac{n-1}{2}$. Indeed, observe first that, since H is connected and has at least three vertices, every edge in H has an endpoint with degree at least 2. Therefore, if there are two isolated vertices in G, say x and y, then adding the edge xy to G with any colour must create a copy of H, hence either x or y must have degree at least 1, which gives a contradiction. □

The following theorem improves a result appearing in [17] and confirms Conjecture 1.1.

THEOREM 3.2. *Let H be a graph in which every edge lies in a triangle, then if $t \geq e(H)$,*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) \geq \left(\frac{1}{4t} + o(1)\right) n \log n.$$

Proof of Theorem 3.2. For each positive integer n, let $(G, c) = (G_n, c_n)$ be a $\mathfrak{R}(H)$-saturated t-edge coloured graph on n vertices and $m = m(n)$ edges. Note that, by Lemma 3.1, $m \geq \frac{n-1}{2}$. Moreover, we must have $d(v) \geq 2$ for all $v \in V(G)$.

For every colour $i \in \{1, 2, \ldots, t\}$ and every vertex v, let $d_i(v)$ be the degree of v in the subgraph spanned by the i-coloured edges and m_i be the total number of i-coloured edges. Now, pick a colour, say 1, and, for each vertex v and each pair $i < j$ of colours different from 1, consider the complete bipartite graph $B_v^{i,j}$ with parts S_v^i and S_v^j, where, for any colour k, $S_v^k = \{u \in V(G) : uv \text{ is a } k\text{-coloured edge in } G\}$. Since the addition of a new edge to G in colour 1 must create a rainbow triangle, every non-edge of G must belong to at least one of these bipartite graphs. Let

$$\left\{X_v^{i,j} \sim \text{Bernoulli}\left(\frac{1}{2}\right) : v \in V(G), i < j, i, j \neq 1\right\}$$

be an independent set of random variables and, for each $v \in V(G)$ and every pair of colours $i < j, i, j \neq 1$, set

$$T_v^{i,j} = \begin{cases} S_v^i & \text{if } X_v^{i,j} = 0 \\ S_v^j & \text{if } X_v^{i,j} = 1. \end{cases}$$

Now let $U = V(G) \setminus \bigcup\{T_v^{i,j} : v \in V(G), i, j \in [t], 1 \notin \{i, j\}\}$. Notice that, if uw is a non-edge, then at least one of u and w is not in U. U is therefore a clique, so as G has m edges,

$$|U| \leq \sqrt{2m + \frac{1}{4}} + \frac{1}{2} \leq \sqrt{2m} + 1.$$

We also have the lower bound

$$\mathbb{E}[|U|] = \sum_{v \in V(G)} 2^{-(t-2)(d(v) - d_1(v))} \geq n \cdot 2^{-2(t-2)\frac{(m-m_1)}{n}}.$$

Combining these inequalities, we have that

$$n \cdot 2^{-2(t-2)\frac{(m-m_1)}{n}} \leq \sqrt{2m} + 1.$$

Since this holds for every colour, by taking the average over all colours, we obtain

$$n \cdot 2^{-2\frac{(t-1)(t-2)}{tn}m} \leq \sqrt{2m} + 1.$$

Let γ be a constant for which $m < (\gamma + o(1))n \log n$, then $m = n^{1+o(1)}$ and

$$\begin{aligned}\sqrt{2m} + 1 = m^{\frac{1}{2}+o(1)} \geq n \cdot 2^{-2\frac{(t-1)(t-2)}{tn}m} \geq n^{1-2\gamma\frac{(t-1)(t-2)}{t}+o(1)} \implies \\ n^{1-2\gamma\frac{(t-1)(t-2)}{t}+o(1)} \leq m^{\frac{1}{2}+o(1)} = n^{\frac{1}{2}+o(1)} \implies \\ 1 - 2\gamma\frac{(t-1)(t-2)}{t} \leq \frac{1}{2} \implies \gamma \geq \frac{t}{4(t-1)(t-2)} \geq \frac{1}{4t}.\end{aligned}$$

□

Using a similar argument we can show that every graph with a conical vertex also has large t-rainbow saturation numbers.

THEOREM 3.3. *If H is a graph with a conical vertex and $|H| \geq 3$, then, for any $t \geq e(H)$,*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) \geq \left(\frac{1}{4t^2} + o(1)\right) n \log n.$$

Proof of Theorem 3.3. Let H be a graph which is not a star containing a conical vertex v. For every positive integer n, let $(G, c) = (G_n, c_n)$ be an $\mathfrak{R}(H)$-saturated t-edge coloured graph. As G has at most one isolated vertex, we can find a set $S \subset V(G)$ of size at least $\frac{n-1}{t}$ such that every vertex in S sees the same colour, say colour 1. Now, we claim that for every non-edge xy, with $x, y \in S$, there must exist a rainbow path of length 2 between x and y using colours in $\{2, 3, \ldots, t\}$. Suppose, for a contradiction, this is not the case. When $e = xy$ is added and coloured 1, we must create a copy H' of H, which implies one of the endpoints of e (say x) must play the role of v and the other (say y) plays the role of a leaf in H, the latter must hold by the assumption that there is no rainbow path of length 2 between x and y. However, in this case, there would already exist a rainbow copy of H in G, namely $H' \setminus \{y\} \cup \{z\}$, where z is a neighbour of x with the edge xz coloured 1. We may now apply the same technique used in the proof of Theorem 3.2. Let m be the number of edges of G.

As before, for each vertex $x \in G$ and each pair $i < j$ of colours other than 1, we consider the complete bipartite graph $B_x^{i,j}$ with parts S_x^i and S_x^j, where, for any colour k, $S_v^k = \{u \in S : uv \text{ is a } k\text{-coloured edge in } G\}$. Since every non-edge between vertices of S is joined by a rainbow path in colours other than 1, each of them is covered by at least one of these bipartite graphs. Let $\left\{X_x^{i,j} \sim \text{Bernoulli}\left(\frac{1}{2}\right) : x \in V(G), i < j, i, j \neq 1\right\}$ be an independent set of random variables and, for each $x \in V(G)$ and each pair of colours $i < j, i, j \neq 1$, set

$$T_v^{i,j} = \begin{cases} S_v^i & \text{if } X_v^{i,j} = 0 \\ S_v^j & \text{if } X_v^{i,j} = 1. \end{cases}$$

Let $S \setminus \bigcup \{T_v^{i,j} | v \in V(G), i, j \in [t], i, j \neq 1\}$. If uw is a non-edge, then at least one of u and w is not in U. Hence U is a clique, so $|U| \leq m^{\frac{1}{2}+o(1)}$. We also have

$$\mathbb{E}[|U|] = \sum_{u \in S} 2^{-(t-2)(d(u)-d_1(u))} \geq \sum_{u \in S} 2^{-(t-2)(d(u)-1)} \geq$$
$$|S| \cdot 2^{-(t-2)\frac{2e(S)+e(S,V(G)\setminus S)}{|S|}-1} \geq \frac{n-1}{t} \cdot 2^{-2t(t-2)\frac{m}{n-1}-1}.$$

Where the second inequality holds by convexity of 2^{-x}. Suppose γ is a constant for which $m < (\gamma + o(1))(n-1)\log(n-1)$, then

$$(n-1)^{\frac{1}{2}+o(1)} = m^{\frac{1}{2}+o(1)} \geq \frac{n-1}{t} \cdot 2^{-2t(t-2)\frac{m}{n-1}-1} \geq$$

$$\frac{n-1}{2t} \cdot 2^{-2t(t-2)\gamma \log(n-1)} = (n-1)^{1-2t(t-2)\gamma+o(1)},$$

which implies that $\gamma \geq \frac{1}{4t(t-2)}$. Therefore

$$m \geq \left(\frac{1}{4t(t-2)} + o(1)\right)(n-1)\log(n-1) \geq \left(\frac{1}{4t^2} + o(1)\right) n \log n.$$

$\square$

4. Upper bounds for connected graphs

Throughout this section we will assume all graphs are connected and have at least three vertices. The aim of this section is to provide constructions of rainbow saturated graphs which, in some cases, are optimal up to multiplicative constants.

First, we show that if H has a cycle then $\mathrm{sat}_t(n, \mathcal{R}(H)) \leq O(n \log n)$, for any $t \geq e(H)$. Next, for any graph H with a non-pendant edge not contained in any triangle, we give constructions of t-coloured graphs on n vertices and with $\Theta(n)$ edges which are $\mathcal{R}(H)$-saturated. Observe that if H is not a star then either H contains a cycle or H is a tree which has a non-pendant edge, hence by the aforementioned results $\mathrm{sat}_t(n, \mathcal{R}(H)) \leq O(n \log n)$ for any $t \geq e(H)$. This answers a question from [17] for connected graphs, namely that stars are the only connected graphs with quadratic rainbow saturation numbers. We also provide constructions of $\mathcal{R}(K_r)$-saturated graphs on t colours, when t is a function of n.

4.1. Graphs with a non-pendant edge not in a triangle. In this subsection, we show that if H is a graph with a non-pendant edge not contained in any triangle then for any integers $t \geq e(H)$, $n \geq 1$ we have $\mathrm{sat}_t(n, \mathcal{R}(H)) \leq c_H n$, where c_H depends only on H.

Let H be a connected graph on $p \geq 3$ vertices and m edges and slet $e = xy \in E(H)$ be an edge which is not contained in a triangle. For $n \geq |H| \cdot e(H)$, we shall construct a graph $G = G_{n,H,e}$ on n vertices together with an edge colouring

$c = c_{n,H,e} : E(G) \to [m]$ such that the vast majority of the non-edges of (G, c) are $\mathcal{R}(H)$-saturated and, if H satisfies some additional conditions, (G, c) is $\mathcal{R}(H)$-free. Observe that our coloured graph (G, c) uses exactly $m = e(G)$ colours, therefore any rainbow copy of H in G must use all these colours.

First, let $\{e_1, \dots, e_m = e\}$ and $\{v_1, \dots, v_{p-1} = x, v_p = y\}$ be enumerations of the edges and vertices of H, respectively. For every $i \in [m]$, let H_i be a copy of $H \setminus \{x, y\}$ with the vertex set $V_i = \left\{v_1^i, \dots, v_{p-2}^i\right\}$, where v_j^i in H_i corresponds to v_j in H.

Now, define a graph $G = K \cup L$ where $G[K] = H_1 \cup \dots \cup H_m$ is a disjoint union of H_i's and L is an independent set of size $n - |K|$. Moreover, for every $u \in L$, u is joined with $v_j^i \in K$ if and only if either xv_j or yv_j is an edge in H.

Having defined G, let us define an edge colouring c of G. Let w_1w_2 be an edge in G. Since L is independent we may assume that $w_1 = v_j^i$, for some $i \in [m]$ and $j \in [p-2]$. Consider now two cases depending on which part w_2 belongs:

(1) if $w_2 \in K$, then $w_2 = v_k^i$ for some $k \in [p-2]$, we let s be such that $e_s = v_jv_k$;

(2) if $w_2 \in L$, we let s be such that $e_s = xv_j$ or $e_s = yv_j$.

It follows from the fact that e is not in a triangle that s is well defined. We then define $c(w_1w_2) = s$ if $s \neq i$ and $c(w_1w_2) = m$ otherwise.

First, we shall show that every non-edge in L is $\mathcal{R}(H)$-saturated.

Proposition 4.1. *Every non-edge in L is $\mathcal{R}(H)$-saturated.*

Proof. Take any non-edge w_1w_2 in L and any colour $i \in [m]$. It is easy to check that adding the i-coloured edge w_1w_2 to the graph creates a rainbow copy of H in $\{w_1, w_2, H_i\}$. □

Now we shall describe the properties H must have if there exists a rainbow copy of H in (G, c).

Lemma 4.2. *Let W be a rainbow copy of H in (G, c). Then, all the following must hold.*

(1) If $v_i v_j$ is an edge of H, for some $i, j \in [p-2]$, then there is k such that $v_i^k v_j^k$ is an edge in W.

(2) There is exactly one $i \in [p-2]$ such that there exist distinct k, k' with $v_i^k, v_i^{k'} \in W$ (we shall say that i is not unique *in W).*

(3) There is exactly one vertex in W, say z, such that $z \in L$.

(4) If $v_i^k \in W$ and v_i is adjacent to x or y in H then v_i^k is adjacent to z in W.

(5) $d_W(z) = d_H(x) + d_H(y) - 1$.

(6) If $v_i^k v_j^k \in E(W)$ and $v_i^{k'} v_j^{k'} \in E(W)$ then $k = k'$.

Proof. For every $k \in [m]$, we let $f_k \in E[W]$ be the edge of W of colour k. Observe, that for every $k \in [m-1]$, the only k-coloured edges in (G, c) are exaclty those edges which are 'copies' of e_k, in other words,

(a) if $e_k = v_i v_j$, for $i, j \in [p-2]$ then $f_k = v_i^{k'} v_j^{k'}$ for some $k' \neq k$;

(b) and if $e_k = v_i v_j$, for $i \in [p-2]$, $j \in \{p-1, p\}$, then $f_k = v_i^{k'} z$, for some $z \in L$, $k' \neq k$.

Note that since H is connected and W must intersect at least two distinct H_i's, it follows that $|W \cap L| \geq 1$. Moreover, it follows from (a) and (b) that for every $i \in [p-2]$, there exists some $k' \in [m]$ such that $v_i^{k'} \in W$. Hence, (1) holds.

To see (2) and (3), observe first that if there are two different indices $i \neq i' \in [p-2]$ for which there exists two copies of v_i, v_j in W then $|W| \geq (p-2) + 2 + 1 = p + 1$, which is a contradiction. Therefore, there is at most one index which is not unique.

To finish the proof of (2) and (3), it is enough to show that $|W \cap K| \geq p - 1$. Let us consider where the edge f_m, of colour m appears in W. If $f_m \in G[K]$, then $f_m = v_i^k v_j^k$ for some i, j, k such that $v_i v_j = e_k$. Since we know by (a), that $f_k = v_i^{k'} v_j^{k'}$ for some $k' \neq k$ we have that both i and j are not unique in

W, which cannot happen as we have seen. Therefore, we may assume that $f_m = zv_i^k$ for some $z \in L$ and i, k. By construction v_i is adjacent to either x or y. Without loss of generality, we can assume that v_i is adjacent to x, and again by construction, $e_k = v_i x$. Since $f_k = wv_i^{k'}$, for some $w \in L$ and $k' \neq k$, we have that i is not unique in W. Hence, $|W \cap K| = p - 1$ and $|W \cap L| = 1$ and $w = z$.

Now, to prove (4), suppose $v_i^k \in W$. Notice that we already showed that if i is not unique in W then z is adjacent to v_i^k in W. Therefore, we may assume that i is unique in W. Since v_i is adjacent to either x or y, without loss of generality, we may assume that v_i adjacent to x, and therefore we have that $v_i x = e_\ell$ for some ℓ. Hence, as observed before, $f_\ell = wv_i^{k'}$ for some $w \in L$ and $k' \in [m]$. Since there is only one vertex in L, namely z, and i is unique in W we have that $w = z$ and $k' = k$ hence $f_\ell = zv_i^k$ is an edge in W.

Next, to show (5), note that since z is the only vertex in $W \cap L$, it must be incident with f_m and $d_H(x) - 1 + d_H(y) - 1$ edges of other colours. Hence, $d_W(z) = d_H(x) + d_H(y) - 1$.

Finally, if (6) does not hold then both i and j are not unique in W, which contradicts (2). $\square$

PROPOSITION 4.3. *Suppose H has an edge e which is in a cycle but not in a triangle then there is no rainbow copy of H in $(G, c) = (G_{n,H,e}, c_{n,H,e})$.*

PROOF. Suppose for contradiction that W is a rainbow copy of H in (G, c). Let g be the length of a longest cycle in H which uses e. We shall show that there is a natural correspondence between the g-cycles in W and the g-cycles in W not using the edge e, thus yielding a contradiction, since the number of g-cycles in W is then strictly smaller than the number of g-cycles in H.

Let C be a g-cycle in W. We shall find a corresponding g-cycle K_C in H. If C does not use vertices from L, i.e., $C = v_{k_1}^i \dots v_{k_g}^i v_{k_1}^i$, with $k_1, \dots, k_g \leq p - 2$, then let $K_C = v_{k_1} \dots v_{k_g} v_{k_1}$. Note that by construction K_C is a g-cycle in H.

Otherwise, by (3) in Lemma 4.2, C uses exactly one vertex from L, i.e., $C = uv^i_{k_1} \dots v^i_{k_{g-1}}u$ with $u \in L$ and $k_1, \dots, k_{g-1} \leq p-2$. In that case let $K_C = wv_{k_1} \dots v_{k_{g-1}}w$, where $w = x$ if v_{k_1} is a neighbour of x in H, or $w = y$ otherwise.

We claim that K_C is a g-cycle in H. Indeed, observe first that by construction $v_{k_1} \dots v_{k_{g-1}}$ is a path in H. Note also that v_{k_1} and $v_{k_{g-1}}$ both have exactly one neighbour in $\{x, y\}$. Therefore, if v_{k_1} and $v_{k_{g-1}}$ are both adjacent to the same vertex $w \in \{x, y\}$ then K_C is indeed a g-cycle. We can therefore assume, without loss of generality, that k_1 adjacent to x and k_{g-1} is adjacent to y. We note that $k_1, \dots, k_{g_1}, x, y$ is then a $(g+1)$-cycle in H using the edge $e = xy$, which contradicts the assumption that g is the size of a longest cycle in H using the edge e.

It is easy to check now that if $K_C = K_{C'}$ then we obtain a contradiction to (6) of Lemma 4.2. Finally, there is no g-cycle C in W such that K_C is a g-cycle in H using the edge e, thus we obtain a contradiction. $\square$

Recall that that an edge is a *bridge* if its removal increases the number of connected components.

PROPOSITION 4.4. *If H has a non-pendant bridge then there is an edge $e \in H$ such that there is no rainbow copy of H in $(G_{n,H,e}, c_{n,H,e})$.*

PROOF. If there is an edge e' in H which is in a cycle but not in a triangle then the result follows from Proposition 4.3, by taking $(G_{n,H,e'}, c_{n,H,e'})$. Hence, we may assume that every edge in H which is not in a triangle is a bridge. Let $e = xy$, with $d(x) \geq d(y)$, be a non-pendant bridge in H for which $d(x)$ is maximized. By the assumption e is well defined.

Suppose for contradiction that W is a rainbow copy of H in (G, c). We will show that the number of non-pendant bridges in W is strictly smaller than the number of non-pendant bridges in H, thus obtaining a contradiction.

Observe first that we cannot have a non-pendant bridge in W incident with any vertex $z \in L$ as then, by (5) of Lemma 4.2, we have $d(z) \geq d(x)+d(y)-1 \geq$

$d(x)+1$ which would contradict the maximality of $d(x)$. Therefore, if there is a non-pendant bridge in W then it must be within K.

Let $b = v_i^k v_j^k$ be a non-pendant bridge in W, for some i, j, k. We shall show that $e_b = v_i v_j$ must be a non-pendant bridge in H. By assumption every edge in H which is not in a triangle is a bridge hence $v_i v_j$ is contained in a triangle, say in $v_i v_j v_\ell$ for some $\ell \in [p] \setminus \{i, j\}$.

Observe that if $v_i v_j x$ or $v_i v_j y$ is a triangle in H then, by (4) of Lemma 4.2, $v_i^k v_j^k z$ is a triangle in W; this contradicts the assumption that $v_i^k v_j^k$ is a bridge. Therefore we can assume that $v_i v_j v_\ell$ is a triangle with $\ell \le p-2$. Since $v_i^k v_j^k$ is a bridge in W it follows that the edge cannot belong to any triangle in W. Therefore either $v_i^k v_\ell^k$ or $v_j^k v_\ell^k$ is not an edge in W. Without loss of generality we can assume that $v_i^k v_\ell^k$ is not an edge in W. Hence we must have by (1) of Lemma 4.2 that, for some $k' \neq k$, $v_i^{k'} v_\ell^{k'}$ is an edge in W. By the same lemma, there also must exist k'' such that $v_j^{k''} v_\ell^{k''}$ is an edge in W. But then there are two indices i and ℓ which are not unique in W contradicting (2) of Lemma 4.2. Therefore, we have that $e_b = v_i v_j$ is indeed a bridge in H.

Note that by (6) of Lemma 4.2 we have that $e_b \neq e_{b'}$ for distinct non-pendant bridges b, b' in W. Hence we found a correspondence between the non-pendant bridges in W and the non-pendant bridges in $H \setminus \{e\}$, which gives a contradiction as then the number of non-pendant bridges in W is strictly smaller than the number of non-pendant bridges in H. $\square$

THEOREM 4.5. *If H has a non-pendant edge not contained in a triangle then for any integers $t \ge e(H)$ and n we have*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) \le c_H \cdot n,$$

where $c_H = e(H) \cdot (|H| - 2)$.

PROOF. When $n \le e(H) \cdot (|H|-2)$, the result follows easily by considering a monochromatic K_n. We may then assume that $n > e(H) \cdot (|H|-2)$. Consider an edge in H as in the statement of Proposition 4.3 or 4.4. Then there is no rainbow copy of H in $(G = G_{n,H,e}, c_{n,H,e})$ and every non-edge in L is $\mathcal{R}(H)$-saturated. If there are non-edges in G which are not $\mathcal{R}(H)$-saturated for some colour i, we can simply add those edges to G and colour them with an appropriate colour, obtaining (G', c'). Note that $e(G') \leq |L||K| + \binom{|K|}{2} \leq (n - |K|)|K| + |K|^2 = n|K| \leq n \cdot e(H) \cdot (|H| - 2)$. $\square$

4.2. Graphs with a cycle. The construction presented in this subsection will be very similar to the one in Subsection 4.1. Let H be a graph on p vertices with a cycle. Observe that if H is triangle-free then there is an edge in H which in a cycle but not in a triangle hence by a result from previous section we have that $\mathrm{sat}_t(n, \mathcal{R}(H)) = O(n)$. Therefore, we can assume that H has a triangle. Let $e = xy$ be an edge of H which is contained in a triangle.

As before, for n large enough we shall construct a graph $G = G^r_{n,H,e}$ on n vertices together with an edge colouring $c = c^r_{n,H,e} : E(G) \to [t]$ such that the vast majority of the non-edges of (G, c) is $\mathcal{R}(H)$-saturated and (G, c) is $\mathcal{R}(H)$-free.

Let $\{e_1, \dots, e_m = e\}$ and $\{v_1, \dots, v_{p-1} = x, v_p = y\}$ be enumerations of the edges and vertices of H, respectively. For all $i \in [m]$ and $j \in [h]$, where $h = \lceil \log(n^2 m + 1) \rceil$, let $H_{i,j}$ be a copy of $H \setminus \{x, y\}$ with the vertex set $V_{i,j} = \left\{v^{i,j}_1, \dots, v^{i,j}_{p-2}\right\}$, where $v^{i,j}_l$ in $H_{i,j}$ corresponds to v_l in H.

Now we define a graph $G = K \cup L$, where $G[K] = \bigcup_{i,j} H_{i,j}$ is a disjoint union of $H_{i,j}$'s and L is an independent set of size $n - |K|$. Moreover, for every $u \in L$ and $H_{i,j}$ we shall toss a coin and based on the result decide how to join the vertices in $H_{i,j}$ with u. More precisely, for $u \in L$, $i \in [m]$ and $j \in [h]$, let $X_{u,i,j}$ be a random variable such that $\mathbb{P}\{X_{u,i,j} = x\} = \mathbb{P}\{X_{u,i,j} = y\} = \frac{1}{2}$, and let all the $X_{u,i,j}$'s be independent. Now join u with $v^{i,j}_k \in H_{i,j}$ if and only if $v_k X_{u,i,j} \in E(H)$.

Having defined G let us define the edge colouring c. Let $w_1 w_2$ be an edge in G. Since there are no edges in L we can assume that $w_1 = v^{i,j}_k$, for some i, j, k. Consider two cases depending on w_2:

(1) if $w_2 \in K$ and $w_2 = v^i_{k'}$ for some k', then let s be such that $e_s = v_k v_{k'}$;

(2) if $w_2 \in L$ then let s be such that $e_s = v_k X_{w_2,i,j}$.

Now $c(w_1w_2) = s$ if $s \neq i$ and $c(w_1w_2) = m$ otherwise.

Proposition 4.6. *With positive probability every non-edge in L is $\mathcal{R}(H)$-saturated.*

Proof. Let $f = uv$ be a non-edge in L and $i \in [m]$ some colour. Notice, that if f is i-coloured and there is some j for which $X_{u,i,j} \neq X_{v,i,j}$, then we can find a rainbow copy of H in $\{u, v, H_{i,j}\}$. Call the pair (uv, i) *bad* if $X_{u,i,j} = X_{v,i,j}$ for every $j \in [h]$. The probability that (uv, i) is bad is equal to

$$\mathbb{P}\{X_{u,i,j} = X_{v,i,j} \text{ for every } j\} = 2^{-h}.$$

Since we have $\binom{|L|}{2} \leq n^2$ non-edges in L and m colours the expected number of bad pairs is

$$\mathbb{E}[\#\text{bad pairs}] \leq 2^{-h}n^2m \leq \frac{n^2m}{n^2m+1} < 1,$$

therefore with positive probability there is no bad pair, hence with positive probability every non-edge in L is $\mathcal{R}(H)$-saturated. $\square$

Proposition 4.7. *There is no rainbow copy of H in (G, c).*

Proof. Suppose W is a rainbow copy of H in (G, c). We shall show that there is a natural correspondence between the triangles in W and the triangles in H not using the edge xy, thus obtaining a contradiction, since the number of triangles in W is then strictly smaller than the number of triangles in H.

Let T be a triangle in W. We shall find a corresponding triangle K_T in H. If T does not uses vertices from L, i.e., $T = \left\{v^{i,j}_{k_1}, v^{i,j}_{k_2}, v^{i,j}_{k_3}\right\}$, with $k_1, k_2, k_3 \leq p-2$, then let $K_T = \{v_{k_1}, v_{k_2}, v_{k_3}\}$. Note that by construction K_T is a triangle in H.

Otherwise, since L is independent, T uses exactly one vertex from L, i.e., $T = \left\{v^{i,j}_{k_1}, v^{i,j}_{k_2}, u\right\}$ with $u \in L$ and $k_1, k_2 \leq p-2$. In that case let $K_T = \{v_{k_1}, v_{k_2}, X_{u,i,j}\}$. Again, by construction K_T is a triangle in H.

It is easy to check now that if $K_T = K_{T'}$ for some distinct triangles T and T' in W then at least one colour appears twice in $E(T) \cup E(T')$, which is a contradiction. Finally, there is no triangle T in W such that K_T is a triangle in H using the edge xy. This proves that there is no rainbow copy of H in G. $\square$

Using those two propositions we are ready to prove the main theorem of this subsection.

THEOREM 4.8. *If H contains a cycle then*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) \leq (1 + o_H(1))c_H \cdot n \log n,$$

where $c_H = 2e(H)(|H| - 2)$.

PROOF. By Theorem 4.5 from the previous subsection we may assume that H contains a triangle. Let e be an edge in H contained in a triangle. For n large enough, it follows from Propositions 4.6 and 4.7, that there is (G, c), with vertex partition $K \cup L$ (where $|K| = e(H) \cdot |H| \cdot h$), such that every non-edge in L is $\mathcal{R}(H)$-saturated and there is no rainbow copy of H in (G, c). If there are any non-edges which are not $\mathcal{R}(H)$-saturated we can just add those edges with appropriate colours to G obtaining (G', c'). Therefore (G', c') is $\mathcal{R}(H)$-saturated and the number of edges in G' is at most $(n - |K|) \cdot |K| + |K|^2 = n \cdot |K| = (1 + o_H(1))c_H \cdot n \log n$, where $c_H = 2e(H)(|H| - 2)$. $\square$

Theorem 1.2 restricted to the class of connected graphs follows easily as a corollary of the previous theorem and Theorem 4.5.

COROLLARY 4.9. *Let H be a connected graph on at least three vertices which is not a star. Then, for every $t \geq e(H)$,*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) = O(n \log n).$$

PROOF. If H contains a cycle then we are done by Theorem 4.8. If not, then H is a tree containing a non-pendant edge and the result follows from Theorem 4.5. $\square$

4.3. Graphs with leaves. In this subsection we are concerned with connected graphs which contain a leaf. In [17] Barrus et al showed that, with few exceptions, if a connected graph H has a leaf, then for $t \geq \binom{|H|-1}{2}$, $\mathrm{sat}_t(n, \mathcal{R}(H)) = \Theta(n)$.

THEOREM 4.10 (Barrus, Ferrara, Vandenbussche, and Wenger [17]). *Let H be a graph on at least five vertices with a leaf whose neighbour is not a conical vertex and such that the rest of the vertices do not induce a clique. Then, for any $t \geq \binom{|H|-1}{2}$, we have $\mathrm{sat}_t(n, \mathcal{R}(H)) = \Theta(n)$.*

To prove similar bounds for the remaining connected graphs containing a leaf we shall introduce some terminology. We let $H_{k,\ell}$ to be the graph obtained by taking a K_k (for some $k \geq 3$) and adding two new vertices x and y, where x adjacent to some ℓ vertices of the clique and yx is a pendant edge. We shall call x the *middle* vertex and y the *leaf* vertex. Note that all such graphs are isomorphic however we choose the ℓ neighbours of x in K_k. Also, observe that the graph K_k with a *rotated* edge is just $H_{k-1,k-2}$.

The following proposition shows that for any $\ell \leq k-2$, the t-rainbow saturation number of $H_{k,\ell}$ is linear in n when the number of colours is sufficiently large.

THEOREM 4.11. *For any $2 \leq \ell \leq k-2$ and $t \geq k(k-1)$ we have that $\mathrm{sat}_t(n, \mathcal{R}(H_{k,\ell})) = O(n)$.*

PROOF. Let $G = K \cup L$ where $G[K]$ is a disjoint union of two cliques of size k, say C_1, C_2, and L is independent set on $n - 2k$ vertices. Now, fix $\ell + 1$ vertices C_1 and $\ell + 1$ vertices of C_2 and join each vertex in L to all of those vertices.

Let $A, B \subseteq [k(k-1)]$, with $|A| = |B| = \frac{k(k-1)}{2}$ be a disjoint union of colours and $A' \subseteq A$, $B' \subseteq B$ be any subsets of size $\ell+1$. We shall describe the colouring of the edges of G. First, colour the edges of C_1 using distinct colours from A, and colour the edges of C_2 using distinct colours from B. Now, for every vertex $v \in L$ colour the edges incident with v with distinct colours from B' if the edges are incident with C_1 and distinct colours from A' if the edges are incident with C_2. Note that in this colouring each vertex in L is incident with $2(\ell+1)$ edge of different colours.

We claim that there is no rainbow copy of $H_{k,\ell}$ in G. Suppose for contradiction that W is a rainbow copy of $H_{k,\ell}$ in G. First let us find a copy of k-clique C in W. Up to symmetry there are two cases: either C uses all the vertices from C_1 or it uses $k-1$ vertices from C_1 and one vertex from L. In the former case the middle vertex must be in L and the leaf vertex must be in C_2. Which is a contradiction since C uses all colours of A and the edge between the middle and leaf vertices uses a colour from $A' \subset A$, therefore W is not rainbow. In the other case, when C uses a vertex from L, say z, note that $\ell = k-2$ and therefore the edges between z and the rest of the clique C use all of the colours from B'. Observe now that the middle vertex cannot be in C_2 since it has to be adjacent to at least two vertices of the clique C (we assumed that $\ell \geq 2$). Also, the middle vertex cannot be in L since it has to be adjacent to at least one vertex from $C_1 \cap C$, hence must be incident with an edge of colour from B' but all the colours of B' have already been used by the edges incident with z. Therefore, the middle vertex z must belong to $C_1 \setminus C$ and the leaf must be in L. This is impossible as z is not joined to any vertex of L, which is a contradiction.

Now we shall show that every non-edge in L is $\mathcal{R}(H)$-saturated for any colour $i \in [t]$. By symmetry, we can assume that $i \in B$. (If $i \notin A \cup B$ the same argument holds). It is easy to check now that adding the edge xy, with $x, y \in L$, and colouring it with colour i, we create a rainbow copy of $H_{k,\ell}$ using all the vertices from C_1 and x, y, where x and y play the roles of the middle and leaf vertices, respectively. $\square$

The following theorem shows that, when $r \geq 4$ is even, the t-rainbow saturation of K_r with a rotated edge is linear.

THEOREM 4.12. *Let $r \geq 4$ be even and H be K_r with a rotated edge. Then, for any $t \geq \binom{r}{2}$, $\mathrm{sat}_t(n, \mathcal{R}(H)) = O(n)$.*

PROOF. Assume $t = \binom{r}{2}$. We first define a graph Γ with vertex set $[r]^{\frac{r}{2}}$ and an edge between each pair of vertices that differ in exactly one component. Now we will define an edge colouring of Γ.

We identify the elements of $[t]$ with the edges of K_r (with vertex set $[r]$). It is well known that K_r has a proper edge colouring with $r-1$ colours if r is even. Fix one such colouring c. The edges of any given colour i form a matching with $\frac{r}{2}$ edges, and every vertex is incident with exactly one edge of colour i. For each $i \in [r-1]$, choose an arbitrary bijection g_i from $[\frac{r}{2}]$ to the set of edges of colour i. For each vertex x of Γ, let $S(x)$ be the sum of the components of x modulo r. We define the edge colouring of of Γ as follows: If x and y are two vertices of Γ that differ in the k^{th} component, colour the edge xy by $g_{c(e)}\Big(k + g_{c(e)}^{-1}(e)\Big)$, where $e = \{S(x), S(y)\}$. We claim that every clique in Γ is rainbow and that every vertex is incident with exactly one edge of each colour. For the first claim, observe that the restriction of S to a maximal clique is a bijection from the vertices of that clique to those of our K_r, and the function $e \mapsto g_{c(e)}\Big(k + g_{c(e)}^{-1}(e)\Big)$, where k is the component on which all the elements of the clique differ, permutes the edges of K_r. For the second claim, let f be any edge of our K_r and let $i = c(f)$ be its colour. Given a vertex x of Γ, let v be the unique vertex of K_r such that $\{v, S(x)\}$ is coloured i. Notice that x has exactly $\frac{r}{2}$ neighbours y such that $S(y) = v$, and each of these neighbours differs from x in a different component, hence each edge xy is a coloured with a different i-coloured edge of K_r, hence x sees the colour f. Therefore every vertex of Γ sees every colour. But every vertex of Γ has degree $\binom{r}{2}$, so it must be incident with exactly one edge of each colour.

To show that Γ is $\mathfrak{R}(H)$-free, we first observe that every clique in Γ is a subset of a maximal clique. Hence if there is a rainbow copy of H in Γ, the "missing" edge of this copy must have the same colour as the pendant edge, contradicting the fact that the colouring of Γ is proper.

Now, for any n, let G be a graph on n vertices consisting of the disjoint union of $\left\lfloor \frac{n}{r^{\frac{r}{2}}} \right\rfloor$ copies of Γ and a monochromatic clique on the leftover vertices. G is $\mathfrak{R}(H)$-free because each of its components is. Suppose we add to G a new edge e in any colour i. One endpoint x of this new edge must be in a copy of Γ. Since x is incident with an edge of colour i and this edge is in a rainbow copy of K_r, removing this edge and adding e creates a rainbow H. G is therefore an $\mathfrak{R}(H)$-saturated graph with at most $\frac{1}{2}\binom{r}{2}r^{\frac{r}{2}}\left\lfloor \frac{n}{r^{\frac{r}{2}}} \right\rfloor + \binom{r^{\frac{r}{2}}-1}{2}$ edges. $\square$

4.4. Complete graphs.

THEOREM 4.13. *For any $r \geq 3$ there exists a positive constant c_r (depending only on r) such that the following holds. For any n and $t = t(n) \geq \binom{r}{2}$,*

$$\mathrm{sat}_t(n, \mathcal{R}(K_r)) \leq \max\left\{\frac{c_r}{\log t} n \log n, 2(r-2)n\right\}.$$

Proof of Theorem 4.13. First, it is clear we may assume n is sufficiently large, by taking c_r large enough. Note that if $t \leq r^7$, by Theorem 4.8 we have

$$\mathrm{sat}_t(n, \mathcal{R}(K_r)) \leq 2\binom{r}{2} rn \log n \leq \frac{r^3 \log r^7}{\log t} n \log n = \frac{7r^3 \log r}{\log t} n \log n,$$

for n large enough, depending only on r. We may then assume that $t \geq r^7$. Let ℓ be a positive integer (to be specified later) and G be the union of 2ℓ disjoint $(r-2)$-cliques together with an independent set M of size $n - 2(r-2)\ell$, where each edge with one endpoint in M and the other in one of the cliques is present, and there are no edges between two distinct cliques. Observe that G does not contain any copies of K_r, because any such copy would need to use at least two vertices from M.

Let A, B an equipartition of the integers $\{1, 2, \ldots, t\}$ (thus, A, B partition $[t]$ and $||A| - |B|| \leq 1$). Now, we shall arbitrarily colour the edges of the first ℓ $(r-2)$-cliques with the colours from A and the edges of the remaining ℓ $(r-2)$-cliques with the colours from B, such that in each clique no colour appears twice. For each $(r-2)$-clique K, let C_K be the set colours used by the edges of K. Moreover, for each vertex $x \in M$ and each clique K, we shall take a subset $B_{x,K} \subseteq A \setminus C_K$, if $C_K \subseteq A$, or $B_{x,K} \subseteq B \setminus C_K$ otherwise, of size $r-2$ uniformly at random (and independently for every choice of x and K) and colour each edge from x to K with a different element of $B_{x,K}$. Our aim is to prove that with positive probability the addition of any coloured edge between two vertices in M will form a rainbow copy of K_r. To do so, let us compute the probability that some edge $e = xy$, with both endpoints in M, coloured c creates a rainbow copy of K_r. By symmetry, we can assume that $c \in B$. Let $t' = |A| - \binom{r-2}{2}$. Suppose K' is a rainbow copy of K_{r-2} such that $C_{K'} \subseteq A$.

First, we need the following easy claim.

CLAIM 4.14. *For positive integers s, u with $s \geq 2u-1$ the following holds*

$$\frac{\binom{s-u}{u}}{\binom{s}{u}} \geq 1 - \frac{u^2}{s-u+1}.$$

PROOF. Note first, that since $s \geq 2u-1$ we have $\frac{u}{s-u+1} \leq 1$. Hence

$$\begin{aligned}
\frac{\binom{s-u}{u}}{\binom{s}{u}} &= \frac{(s-u)!(s-u)!}{s!(s-2u)!} = \frac{(s-2u+1)\cdot(s-2u+2)\cdots(s-u)}{(s-u+1)\cdot(s-u+2)\cdots s} \\
&= \frac{s-2u+1}{s-u+1}\cdot\frac{s-2u+2}{s-u+2}\cdots\frac{s-u}{s} = \left(1-\frac{u}{s-u+1}\right)\left(1-\frac{u}{s-u+2}\right)\cdots\left(1-\frac{u}{s}\right) \\
&\geq \left(1-\frac{u}{s-u+1}\right)^u \geq 1-\frac{u^2}{s-u+1},
\end{aligned}$$

where the last inequality follows from Bernoulli's inequality: $(1-x)^p \geq 1-px$ for $p \geq 1$ and $x \in [0,1]$. $\square$

Observe that by construction $c \notin (C_{K'} \cup B_{x,K'} \cup B_{y,K'})$ hence as long as $B_{x,K'}$ and $B_{y,K'}$ are disjoint we are done, i.e., there is a rainbow copy of K_r in $\{x,y\} \cup K'$. Let us bound the probability that $B_{x,K'}$ and $B_{y,K'}$ are disjoint. To do that, we apply Claim 4.14 with $s = t'$ and $u = r-2$:

$$\mathbb{P}\{B_{x,K'} \cap B_{y,K'} = \varnothing\} = \frac{\binom{t'-(r-2)}{r-2}}{\binom{t'}{r}} \geq 1 - \frac{(r-2)^2}{t'-r+3}.$$

Hence, since $t' \geq t/2 - 1$ and $t \geq r^7$, we have

$$\mathbb{P}\{\{x,y\} \cup K' \text{ is not rainbow } K_r\} = 1 - \mathbb{P}\{B_{x,K'} \cap B_{y,K'} = \varnothing\} \leq \frac{(r-2)^2}{t' - r + 3} \leq \frac{1}{\sqrt{t}}.$$

Note, there are ℓ rainbow copies of K_{r-2} which only use colours from A, so we deduce that

$$\mathbb{P}\{e \text{ in colour c does not create a rainbow } K_r\} \leq t^{-\ell/2}.$$

Therefore, the probability that some edge with both endpoints in M is *bad*, i.e. the addition of e in some colour does not form a rainbow copy of K_r is at most

$$e(G) \cdot t^{-\ell/2}.$$

This holds because if we colour e in some colour not appearing in the edges of the graph, then we clearly form a rainbow copy of K_r. Hence, taking $\ell = \max\left\{\left\lceil \frac{10 \log n}{\log(t)} \right\rceil, 1\right\}$, we get

$$\mathbb{P}\{ \text{ some edge is } \textit{bad}\} \leq e(G)\binom{M}{2} t^{-\ell/2} \leq n^4 2^{-5 \log n} \leq \frac{1}{n} < 1.$$

We have thus proved there exists a colouring of G for which no edge with both endpoints in M is *bad*. If there are still some unsaturated non-edges in G, just keep adding them with appropriate colours to G. Let $N = V(G) \setminus M$. We are done as

$$\begin{aligned} e(G) &\leq |N|(n - |N|) + \binom{|N|}{2} \leq |N|n - |N|^2 + |N|^2 \leq |N|n \\ &\leq 2\ell(r-2)n. \end{aligned}$$

So if $\ell = 1$ then $e(G) \leq 2(r-2)n$ and if $\ell > 1$ then $e(G) \leq \frac{20(r-2)}{\log t} n \log n$. In order for the graph to be well-defined we must take n big enough (depending on r only) so that $2(r-2)\ell \leq n$. □

Observe that as long as $t(n) \geq \Omega(n)$ we have $\mathrm{sat}_t(n, \mathcal{R}(K_r)) = \Theta(n)$.

Corollary 4.15. *For any $r \geq 3$ we have*

$$\mathrm{sat}(n, \mathcal{R}(K_r)) \leq 2(r-2)n.$$

Proof. When $n \leq 2(r-2)$ then the results follows trivially by considering a monochromatic K_n. We can therefore assume that $n \geq 2(r-2)$. Observe that when there is not a restriction on the number of colours then in our construction we can assign each edge a different colour. In that case we can take $\ell = 1$, which corresponds to a disjoint union of an independent set A and two $(r-2)$-cliques B and C, such that all the edges between A and $B \cup C$ are present, and possibly some edges between B and C. The number of edges is then at most $2(r-2)n$. $\square$

We conjecture that this bound is best possible up to an additive constant.

The following construction gives a better upper bound for the rainbow saturation numbers of a triangle, at least when t is not too large compared to n.

Theorem 4.16. *For any $t \geq 3$ with $t \equiv 1$ or $3 \pmod 6$, then*

$$\mathrm{sat}_t(n, \mathcal{R}(K_3)) \leq \frac{3}{\log \binom{t}{2}} n \log n + 3n.$$

In particular, $\mathrm{sat}_3(n, \mathcal{R}(K_3)) \leq \frac{3}{\log 3} n \log n + 3n$.

Proof. Let S be a Steiner triple system of order t, i.e., a set of three-element subsets of $[t]$ such that every pair of elements of $[t]$ is contained in exactly one element of S. We call the elements of t *points* and the elements of S *lines*. It can be shown[1] that such a system exists if and only if $t \equiv 1$ or $3 \pmod 6$ and that any such system has exactly $\frac{t(t-1)}{6}$ lines. We define a binary operation $\star : [t]^2 \to [t]$ as follows:

[1]An easy divisibility argument shows that, if a Steiner triple systems of order t exists, then $t \equiv 1$ or $3 \pmod 6$. Conversely, Bose [31] describes a simple construction for $t \equiv 3 \pmod 6$ and Skolem [95] describes one for $t \equiv 1 \pmod 6$. Bollobás [22] gives another construction for systems of prime order. Other examples of Steiner triple systems include the projective spaces over $\mathbb{F}_2$ and affine spaces over $\mathbb{F}_3$.

$$a \star b = \begin{cases} a & \text{if } a = b \\ c, \text{ where } c \text{ is the unique point such that } abc \text{ is a line} & \text{if } a \neq b. \end{cases}$$

This operation has the property that, for every fixed a in $[t]$, the map $b \mapsto a \star b$ permutes the elements of each line containing a. We also let $F = \{(\ell, p) : p \in \ell \in S\}$ and call the elements of F the *flags* of S. The number of flags is $3|S| = \binom{t}{2}$. For each line ℓ, we choose an arbitrary ordering of the points on ℓ and, for any $i \in [3]$, we let $\ell^{(i)}$ denote the i^{th} point of ℓ.

Given n, let k be the smallest natural number such that $\binom{t}{2}^k + 3k \geq n$. Clearly, $k \leq \frac{1}{\log \binom{t}{2}} \log n + 1$. Let G be the complete bipartite graph with parts $V \subseteq F^k$ and $K = [k] \times [3]$, with $|V| = n - 3k$. We define a colouring c of the edges of G as follows: for each $f \in V$ and $(i, j) \in K$, let $c(\{f, (i, j)\}) = p \star \ell^{(j)}$, where (ℓ, p) is the i^{th} component of f. To show that adding an edge between two vertices in V creates a rainbow triangle, it suffices to show that every pair of such vertices is joined by either two disjoint rainbow paths of length two using disjoint sets of colours or three such paths that each use a different pair of colours from a set of three. Suppose f and f' are k-tuples of flags that differ in the i^{th} component, say $f_i = (\ell, p)$ and $f'_i = (\ell', p')$. First, consider the possibility that $\ell = \ell'$ and $p \neq p'$. In this case, for every $j \in [3]$, $p \star \ell^{(j)} \neq p' \star \ell^{(j)}$, and neither is equal to $(p \star p') \star \ell^{(j)}$. Thus each path f–(i, j)–f' is a rainbow path of length two using a distinct pair of colours from ℓ. Next, if $\ell \neq \ell'$, then each edge $\{f, (i, j)\}$ is coloured with a different point from ℓ and each edge $f', (i, j)$ is coloured with a different point from ℓ' for $j \in [3]$. Since ℓ and ℓ' have at most one point in common, at most one path f–(i, j)–f' is monochromatic. If this is the case, then the other two such paths are rainbow with disjoint sets of colours. Otherwise, all such paths are rainbow, and at most one pair of them have a colour in common, so there is a pair that uses disjoint sets of colours.

It is possible that adding an edge between two vertices in K in some colour does not create a rainbow triangle; there are at most $\binom{|K|}{2}$ such edges. We can add these coloured edges to (G,c) to form an $\mathfrak{R}(K_3)$-saturated t-edge coloured graph (G',c') with at most

$$|V||K| + \binom{|K|}{2} \leq (n-|K|)|K| + |K|^2 = n|K| \leq \frac{3}{\log\binom{t}{2}} n\log n + 3n$$

edges. $\square$

When $t=3$, the coefficient of the $n\log n$ term in the upper bound is $\frac{3}{\log 3}$, while for large values of t it is approximately $\frac{1.5}{\log t}$. Note that, for values of t that aren't congruent to 1 or 3 modulo 6, we can obtain similar bounds with slightly better coefficients t using maximum partial Steiner systems, as defined and constructed in [97][2].

5. Upper bounds for disconnected graphs

In this section, we shall show that the rainbow saturation number of a disconnected graph can be bounded above by the rainbow saturation number of one of its connected components, up to additive $O(n)$ term. Moreover, we shall show that if H is a disconnected graph with no isolated vertices, then the t-rainbow saturation number of H is at most $O(n\log n)$ answering a question from [17] for disconnected graphs. Throughout the section, we assume, for simplicity of exposition, that H has no isolated vertices.

For a sequence of graphs $H_1,\dots,H_k$ we say that H_i is *maximal*, for some $i\in[k]$, if H_i is not isomorphic to any proper subgraph of H_j for any $j\in[k]$. Observe that every sequence has a maximal element; for example, we can take one with the largest total number of vertices and edges.

Proposition 5.1. *Let H be a graph with connected components $H_1,\dots,H_k$ and let H_i be a maximal component. Then, for every $t\geq e(H)$, we have*

$$\mathrm{sat}_t(n,\mathcal{R}(H)) \leq \mathrm{sat}_t(n,\mathcal{R}(H_i)) + O(n).$$

[2]The maximum number of lines in a partial Steiner triple system of order t is $\left\lfloor\frac{t}{3}\left\lfloor\frac{t-1}{2}\right\rfloor\right\rfloor - 1$ if $t\equiv 5 \pmod 6$ and $\left\lfloor\frac{t}{3}\left\lfloor\frac{t-1}{2}\right\rfloor\right\rfloor$ otherwise.

PROOF. Without loss of generality we may assume that $i = 1$ and $H_1 \cong H_2 \cong \ldots \cong H_\ell$ (for some $\ell \in [k]$), and that no other component is isomorphic to H_1. Let $H' = H_{\ell+1} \cup H_{\ell+2} \cup \ldots \cup H_k$.

Let $t' = e(H) \leq t$ and consider the following graph G on n vertices. First add vertex disjoint copies of all possible rainbow copies of H' for every subset of size $|e(H')|$ in $[t']$. Write V_1 for the set of vertices spanned by these copies. Second, consider the following coloured graph $H_1^\star$: for every set A of colours of size $e(H_1)$ inside $[t']$, we add a rainbow of copy of H_1 with colours in A, where all rainbow copies share exactly one vertex. Now we add $\ell - 1$ vertex disjoint copies of $H_1^\star$ to G and define V_2 to be the set of vertices spanned by these copies. In the set $V(G) \setminus (V_1 \cup V_2)$, consisting of the remaining vertices, we add a $\mathcal{R}(H_1)$-saturated graph on t colours. It is easy to check that every non-edge in $V(G) \setminus (V_1 \cup V_2)$ is $\mathcal{R}(H)$-saturated. Finally, if there are any non-edges which are not $\mathcal{R}(H)$-saturated, we add those edges to G in some colour that does not create a rainbow H. Clearly, there are at most $O(n)$ such edges.

Let us show G does not contain a rainbow copy of H. Suppose for contradiction that it does. We shall obtain a contradiction by showing that the number of vertex disjoint rainbow copies of H_1 in G is strictly smaller ℓ. Note that H_1 cannot be a subgraph of $G[V_1]$ as, by construction, H_1 is not isomorphic to any connected component of $G[V_1]$ and, by maximality, H_1 cannot be a subgraph of any connected component of $G[V_1]$. Observe as well that each copy of $H_1^\star$ contains at most one rainbow copy of H_1. Finally, by construction, $V(G) \setminus V_1 \cup V_2$ does not contain a rainbow copy of H_1. Therefore there are at most $\ell - 1$ vertex disjoint rainbow copies of H_1.

Let $p = |V_1 \cup V_2|$. Observe that $p = \Theta(1)$ as n goes to infinity. Therefore the number of edges in G is at most $\binom{p}{2} + p(n-p) + \mathrm{sat}_t(n-p, \mathcal{R}(H_1)) \leq pn + \mathrm{sat}_t(n, \mathcal{R}(H_1)) = \mathrm{sat}_t(n, \mathcal{R}(H_1)) + O(n)$. $\square$

We have the following immediate corollary.

Corollary 5.2. *Let H be a graph containing at least one component which is not a star and let H' be a maximal component among the components of H which are not stars. Then, for every $t \geq e(H)$, we have*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) \leq \mathrm{sat}_t(n, \mathcal{R}(H')) + O(n) \leq O(n \log n).$$

Proof. Observe that H' cannot be a subgraph of a star, hence by Proposition 5.1 and Corollary 4.9, we have that

$$\mathrm{sat}_t(n, \mathcal{R}(H)) \leq \mathrm{sat}_t(n, \mathcal{R}(H')) + O(n) \leq O(n \log n).$$

□

We showed that if a disconnected graph contains a component which is not a star then its rainbow saturation number is subquadratic. Since stars have rainbow saturation number which is quadratic in n, one might suspect that the same should hold for disconnected graphs where each component is a star. The following proposition shows that this is not the case.

Proposition 5.3. *Let $H = S_1 \cup S_2 \cup \cdots \cup S_k$ be a graph with more than one component, each of which is a star. Then for every $t \geq e(H)$ we have*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) \leq O(n).$$

Proof. Suppose $|S_1| \leq |S_2| \leq \cdots \leq |S_k|$. First we shall show the case when $k = 2$. Let $a = |S_1| - 1$ and $b = |S_2| - 1$. Let $G = K \cup L$ where $G[K]$ is a complete graph of size $a + b - 1$ and L is an independent set of size $n - |K|$. Let $K = \{x_1, \ldots, x_{a+b-1}\}$. First we join every vertex $x_i \in K$ and with every vertex $y \in L$ and give the edge colour i. Next we shall describe the colouring of the edges inside K. Let $x_i, x_j \in K$ where $i \leq j$. If $i \leq a$ and $j \geq a$ then assign $a + b$ as the colour of $x_i x_j$, otherwise assign j as the colour of $x_i x_j$.

We claim that there is no rainbow copy of $S_1 \cup S_2$ in G. To see that, observe first that every rainbow copy of S_i in G uses at least $|S_i| - 1$ vertices of K. Indeed, suppose for contradiction that it is not the case and that there is a rainbow copy of S_i which uses fewer than $|S_i|-1$ vertices of K. Then it must use at least two vertices, say x, y, of L. It follows from independence of L that the center z of that rainbow copy must be in K. We obtain a contradiction by noticing that zx and zy have the same colour. Therefore if there is a rainbow copy of $S_1 \cup S_2$ then it has to use at least $a+b$ vertices of K, which is a contradiction since there are only $a+b-1$ such vertices.

Next we shall show that every non-edge is $\mathcal{R}(H)$-saturated. Consider any non-edge xy in L and any colour $c \in [t]$.

If $c \leq a$ then we find a copy of S_1 in $\{x, y, x_1, \ldots, x_a\} \setminus \{x_c\}$ with x being the center and a copy of S_2 in $\{x_c, x_{a+1}, \ldots, x_{a+b-1}, z\}$ with x_{a+1} as the center, for any $z \in L \setminus \{x, y\}$. Observe that those two copies are vertex disjoint and the copy of S_1 uses only colours from $[a]$ and the copy of S_2 uses colours from $[a+1, a+b]$. Hence we have a rainbow copy of $S_1 \cup S_2$.

If $c \in [a+1, a+b-1]$ then we find a copy of S_2 in $\{x, y, x_a, \ldots, x_{a+b-1}\} \setminus \{x_c\}$ with x being the center and a copy of S_1 in $\{x_1, \cdots, x_{a-1}, x_c, z\}$ with x_1 as the center, for any $z \in L \setminus \{x, y\}$. Observe that those two copies are vertex disjoint and the copy of S_1 uses only colours from $[a-1] \cup \{a+b\}$ and the copy of S_2 uses colours from $[a, a+b-1]$. Hence we have a rainbow copy of $S_1 \cup S_2$.

In the remaining case when $c \geq a+b$, it is easy to check that we can find a rainbow copy of $S_1 \cup S_2$ where both of the centers are in L.

Observe that we have $e(G) \leq (n-|K|)|K| + |K|^2 = |K|n = (a+b-1)n = (|S_1|+|S_2|-3)n$.

Now, suppose $k \geq 3$. We let $t^\star = e(H)$. Moreover, let $G = G' \cup G''$ where G' is an $\mathcal{R}(S_1 \cup S_2)$-saturated graph on $n' = n-(k-2)(t^\star+1)$ vertices with $\mathrm{sat}_{t^\star}(n', \mathcal{R}(S_1 \cup S_2))$ edges and G'' is the vertex-disjoint union of $k-2$ rainbow copies of $t^\star$-stars. It is easy to check that there is no rainbow copy of H in G. Indeed, by assumption there can not be two vertex-disjoint rainbow copies of distinct components of H appearing in G'. Note as well that there can only be at most $k-2$ vertex-disjoint stars in G'', hence in total there are at most $k-1$ disjoint rainbow components of H in G. Finally, it is clear that the addition of any coloured non-edge inside G' creates a rainbow copy of H. Now, we keep adding edges to G (with both endpoints in G'' or with one endpoint in G' and one in G'') until G is saturated. The case $k=2$ shows that $\mathrm{sat}_{t^\star}(n, \mathcal{R}((S_1 \cup S_2))) \leq (|S_1|+|S_2|-3)n$, hence, the number of edges in G is at most $(n-|G''|)|G''|+|G''|^2+e(G') \leq n|G''|+(|S_1|+|S_2|-3)n \leq O(n)$. $\square$

We have the following corollary from Propositions 5.1 and 5.3.

Corollary 5.4. *Let H be a disconnected graph. Then for every $t \geq e(H)$ we have*

$$\mathrm{sat}_t(n, \mathcal{R}(H)) \leq O(n \log n) \leq o(n^2).$$

6. Deducing the main results

We are now ready to deduce Theorems 1.3 and 1.4.

Proof of Theorem 1.3. First, note that item 1 is a result appearing in [17] and item 5 is just a restatement of Theorem 4.12. Now, the lower bounds in items 2, 3 follow by Theorems 3.2 and 3.3, respectively, and the upper is a consequence of Theorem 4.8 since in both cases H must contain a cycle.

In item 4 the lower bound follows from Lemma 3.1 and the upper bound follows from Theorem 4.5. $\square$

Proof of Theorem 1.4. Observe first that if H is a connected graph on at most four vertices which contains a leaf and no conical vertex, then H must be a path on four vertices, hence by Theorem 1.34 its t-rainbow saturation number is linear. We may therefore assume that $|H| \geq 5$. Let xy be a pendant edge of H. If $H \setminus \{x, y\}$ is not a clique then we are done by Theorem 4.10. Hence, we may then assume $H = H_{k,\ell}$ for some $k \geq 3$ and $\ell \leq k-1$. Suppose $\ell \leq k-2$, then result follows by Theorem 4.11. Hence, we may assume $\ell = k-1$ in which case k must be odd, by assumption, and therefore H is a K_{k+1} with a rotated edge, so we are done by Theorem 4.12. $\square$

7. Concluding remarks

We have shown that for any $t \geq \binom{r}{2}$, $\mathrm{sat}_t(n, \mathcal{R}(K_r)) = \Theta(n \log n)$ when $n \to \infty$, i.e., there exist constants $c_1 = c_1(t,r)$ and $c_2 = c_2(t,r)$ such that $c_1 n \log n \leq \mathrm{sat}_t(n, \mathcal{R}(K_r)) \leq c_2 n \log n$. There is still an enormous gap between our lower and upper bounds. However, recently, in [73], Korándi showed the following.

Theorem 7.1 (Korándi). *For every $r \geq 3$, and any $t \geq \binom{r}{2}$,*

$$\mathrm{sat}_t(n, \mathcal{R}(K_r)) \geq \frac{t(1+o(1))}{(t-r+2)\log(t-r+2)} n \log n$$

as $n \to \infty$, with equality for $r = 3$.

This theorem, together with Theorem 4.13 gives:

$$\mathrm{sat}_t(n, \mathcal{R}(K_r)) = \Theta_r\left(\frac{n \log n}{\log t}\right)$$

.

Now, when H is an even clique with a rotated edge, we know that $\mathrm{sat}_t(n, \mathcal{R}(H))$ is always $\Theta(n)$ for $t \geq e(H)$. However, for odd cliques with rotated edges, we do not even know the asymptotic behaviour of $\mathrm{sat}_t(n, \mathcal{R}(H))$ for large values of t.

Question 7.2. If H is a copy of K_r with a rotated edge (as shown in Figure 2) for some *odd* $r \geq 5$ and $t \geq \binom{r}{2}$, what is the asymptotic growth rate of $\mathrm{sat}_t(n, \mathcal{R}(H))$?

The following conjecture together with Theorem 1.3 and Question 7.2 would completely classify the possible rates of growth of $\mathrm{sat}_t(n, \mathcal{R}(H))$ for all connected graphs H and every constant $t \geq e(H)$.

Conjecture 7.3. *Let H be a connected graph (other than an odd clique with a rotated edge) with an edge not in a triangle and no conical vertex. Then, for every $t \geq e(H)$, $\mathrm{sat}_t(n, \mathcal{R}(H)) = O(n)$.*

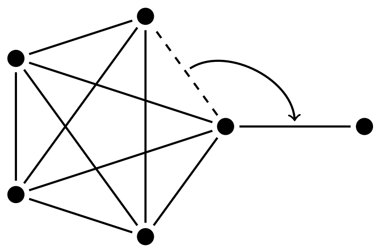

FIGURE 2. K_5 with a rotated edge. The dashed line represents the removed edge.

Note that we can confirm this conjecture when the number of available colours is at least $\binom{|H|-1}{2}$. Indeed, either H is in one of the classes defined in Theorem 1.3, in which case we are done, or H has a leaf and is not a clique with a rotated edge, hence by Theorem 1.4 we have $\mathrm{sat}_t(n, \mathcal{R}(H)) = \Theta(n)$.

One different direction would be to allow the palette of colours to be infinite. We have only considered this question for complete graphs and showed that $\mathrm{sat}(n, \mathcal{R}(K_r)) \leq 2(r-2)n$ for any $r \geq 3$.

Recall that the construction in Corollary 4.15 is a disjoint union of an independent set A and two $(r-2)$-cliques B and C, such that all the edges between A and $B \cup C$ are present and all the edges in B, C and between A and $B \cup C$ receive different colours. We conjecture that, for $n \geq 2(r-2)$, the above construction is best possible up to the configuration of the edges between B and C.

CONJECTURE 7.4. *For any integer $r \geq 3$, there exists a constant C_r depending only on r such that, for any $n \geq 2(r-2)$,*

$$\mathrm{sat}(n, \mathcal{R}(K_r)) = 2(r-2)n + C_r.$$

Finally, we conjecture that, like the ordinary saturation numbers, the rainbow saturation numbers of any graph are at most linear in n.

CONJECTURE 7.5. *For any graph H, $\mathrm{sat}(n, \mathcal{R}(H)) = O(n)$.*

CHAPTER 6

Long cycles in Hamiltonian graphs

1. Introduction

A *Hamiltonian cycle* in a graph G is a cycle spanning the vertex set of G, and a graph is said to be *Hamiltonian* if it contains a Hamiltonian cycle. Over the last seventy years, the following problem has received a great deal of attention: under what conditions does a graph G with a Hamiltonian cycle $\mathcal{C}$ contain another long cycle distinct from $\mathcal{C}$? Of course, for this question to be interesting, one needs to ensure that G contains additional edges (not already in $\mathcal{C}$); a moment's thought further reveals that additional edges are not enough in and of themselves, but rather, one requires additional edges that are 'equidistributed' over the vertex set of G. This problem, namely understanding when the presence of additional edges in a Hamiltonian graph forces the existence of another long (possibly Hamiltonian) cycle, has a storied history; see the surveys of Gould [53] and Bondy [27] for an overview.

The main contribution of this chapter is to show that perhaps the weakest possible condition promising some form of 'equidistribution of additional edges' in a graph with a Hamiltonian cycle is sufficient to guarantee the existence of another long cycle; writing $\delta(G)$ for the minimum degree of a graph G, we prove the following.

Theorem 1.1. *For all $n \in \mathbb{N}$, if an n-vertex graph G with $\delta(G) \geq 3$ contains a Hamiltonian cycle, then G contains another cycle of length at least $n - cn^{4/5}$, where $c > 0$ is an absolute constant.*

At first glance, it is tempting to conclude that Theorem 1.1 must hold since a Hamiltonian graph with minimum degree at least 3 should, necessarily,

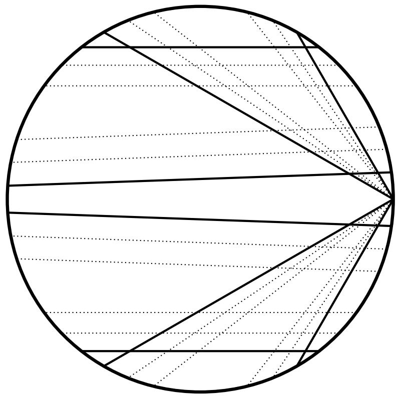

FIGURE 1. Chord patterns of bounded complexity (i.e. using a bounded number of chords) are insufficient to find long cycles.

contain 'short chords'; however, it is not difficult (see Figure 1) to construct Hamiltonian graphs with minimum degree at least 3 that do not contain any such chords, or for that matter, any 'chord pattern' of bounded complexity that gives rise to a long second cycle. Indeed, perhaps the most interesting aspect of Theorem 1.1 is the fact that its proof is based on a combination of constructive and non-constructive arguments: to prove our main result, we use poset-based techniques and parity-based arguments in conjunction with each other, so our methods might be of independent interest.

To provide some context for Theorem 1.1, we remind the reader of the most famous open problem in the area; the following long outstanding conjecture is due to Sheehan [94].

CONJECTURE 1.2. *For each integer $d \geq 3$, every d-regular Hamiltonian graph contains a second Hamiltonian cycle.*

Conjecture 1.2 was proposed as an extension of the classical result of Smith, see [110], that establishes the above conjecture in the case where $d = 3$. Sheehan's conjecture was subsequently shown to hold for all odd $d \geq 3$ by Thomason [100] using a beautiful, non-constructive, parity-based argument, and for all $d \geq 300$ by Thomassen [106, 108] using an ingenious combination of Thomason's argument and the Lovász local lemma. We refer the reader to the paper of Haxell, Seamone and Verstraëte [63] for both the current state of the art as well as a discussion of why existing methods are unlikely to settle Conjecture 1.2 in its full generality.

In light of Sheehan's conjecture, it is natural to ask if regularity is genuinely necessary to force the existence of a second Hamiltonian cycle, or if a weaker condition on the minimum degree, say, might suffice instead. In particular, the following question suggests itself: does every Hamiltonian graph G with $\delta(G) \geq 3$ contain a second Hamiltonian cycle? Entringer and Swart [45] answered this question negatively by constructing infinitely many Hamiltonian graphs without a second Hamiltoninan cycle, all with minimum degree 3. While the Hamiltonian graphs with minimum degree 3 constructed by Entringer and Swart only contain a single Hamiltoninan cycle each, these graphs do contain other long cycles that almost span the entire vertex set; it is therefore natural to ask if such a situation is unavoidable in general.

Problem 1.3. *If an n-vertex graph G with $\delta(G) \geq 3$ contains a Hamiltonian cycle, then must G contain another cycle of length $n - o(n)$?*

Of course, Problem 1.3 is closely related to Conjecture 1.2 since an affirmative answer to the above question would assert precisely that an asymptotic form of Sheehan's conjecture holds under significantly milder degree conditions than the regularity restrictions prescribed in Conjecture 1.2; our main result furnishes, in a quantitative form, precisely such an affirmative answer.

2. Organization, notation and preliminaries

This chapter is organised as follows. We first introduce some notation and collect together the tools that we need for the proof of our main result in Section 2. We then prove Theorem 1.1 in Section 3. Finally, we conclude in Section 4 with a discussion of some open problems.

It will be convenient to begin by establishing some notation for dealing with Hamiltonian graphs.

Given a graph G with a designated Hamiltonian cycle $\mathcal{C}$, we shall always fix one of the two possible cyclic orderings of $V(G)$ obtained by traversing $\mathcal{C}$ to be canonical. Therefore, when we speak, for example, about following $\mathcal{C}$ from x to y for $x, y \in V(G)$, we mean this with respect to the canonical ordering. We use $\prec$ to specify relative positions with respect to the canonical ordering, so for instance, given $x, y, z \in V(G)$, we write $x \prec y \prec z$ (or equivalently either $y \prec z \prec x$ or $z \prec x \prec y$) to mean that we encounter x, y and z in that order around $\mathcal{C}$. Finally, for $x, y \in V(G)$, we write $d_{\mathcal{C}}(x, y)$ for the length of the path from x to y around $\mathcal{C}$ following the canonical ordering, noting that $d_{\mathcal{C}}(x, y) \neq d_{\mathcal{C}}(y, x)$ in general.

Let G be a graph with a designated Hamiltonian cycle $\mathcal{C}$. Any cycle of G distinct from $\mathcal{C}$ is said to be *nontrivial*. We call any edge of G not in $\mathcal{C}$ a *chord*. Observe that there exist two subsets of the vertex set of G corresponding to each chord e of G, namely the vertex sets of the two paths traversing $\mathcal{C}$ between the endpoints of e; we call these two sets of vertices the two *domains* of e, and note that the domains of e intersect precisely in the endpoints of e. We say that a chord e is *minimal* if at least one of its domains induces no chords of G other than e itself, and we call the corresponding domain of e its *minimal domain*; here, if both domains of e induce no chords, then we arbitrarily choose one these domains to be the minimal domain of e. We say that a pair of chords *interlace* if their endpoints are all distinct and appear in alternating order around $\mathcal{C}$ (in the canonical ordering of the vertex set, say); otherwise, we say that they are *parallel*. Also, we say that a set of chords is *independent* if no two of the chords in the set share an endpoint. Finally, we say that two vertices $x, y \in V(G)$ are *chord-adjacent* if they are connected by a chord of G.

Next, we collect together some tools that we shall require for the proof of our main result.

To handle the constructive half of our argument, we shall require a well-known consequence of a classical result of Dilworth [42]. Recall that in a partially ordered set (or poset for short), a *chain* is a subset in which each pair of elements is comparable (which makes a chain a linearly ordered set), and an *antichain* is a subset in which no two elements are comparable; we need the following fact.

PROPOSITION 2.1. *For $r, s \in \mathbb{N}$, every poset of size rs contains either a chain of size r or an antichain of size s.* $\square$

The non-constructive half of our argument depends on the following convenient formulation, due to Thomassen [107], of the parity-based 'lollipop argument' of Thomason [100].

Recall that a set X of vertices *dominates* another set Y of vertices and edges in a graph if each vertex in Y is adjacent to some vertex in X and if each edge in Y is incident to some vertex in X.

PROPOSITION 2.2 (Thomassen). *Let G be a graph with a designated Hamiltonian cycle $\mathcal{C}$. If there exists a set $X \subset V(G)$ such that*

(1) X is independent in the graph $G' = (V(G), E(\mathcal{C}))$, and

(2) X dominates $V(G) \setminus X$ in the graph $G'' = (V(G), E(G) \setminus E(\mathcal{C}))$,

then G contains a nontrivial Hamiltonian cycle. $\square$

Finally, we use standard asymptotic notation throughout to suppress absolute constants, and for the sake of clarity of presentation, we systematically omit floor and ceiling signs whenever they are not crucial.

3. Proof of the main result

We begin with the following lemma that allows us to handle Hamiltonian graphs with many interlacing chords.

LEMMA 3.1. *Let G be an n-vertex graph with a designated Hamiltonian cycle $\mathcal{C}$. If G contains a set I of $2m$ independent chords made up of m interlacing pairs for some $m \geq 1$, then G contains a nontrivial cycle missing $O(n/m^{1/3})$ vertices.*

Proof of lemma. Note that if G has at least one chord, then G contains a nontrivial cycle. In what follows, we therefore suppose, as we may, that m is sufficiently large. We shall show, assuming m is suitably large, that it is possible to construct a cycle of the required length using at most 4 chords of G and the edges of $\mathcal{C}$.

We begin by constructing two posets on any set S of independent chords in G as follows. We fix some edge f of $\mathcal{C}$, and for a chord e of G, we call the domain of e containing the endpoints of f the *interior* of e, and the other domain the *exterior* of e. We then define a partial order $\mathcal{P}_S$ on S by saying $e_1 < e_2$ for $e_1, e_2 \in S$ if the interior of e_1 is contained in the interior of e_2. Next, we fix a linear order $\mathcal{L}$ of the vertices of G by starting at one of the endpoints of f and following $\mathcal{C}$ to the other endpoint of f, and then define another poset $\mathcal{Q}_S$ on S by saying that $e_1 < e_2$ for $e_1, e_2 \in S$ if both the endpoints of e_1 precede both the endpoints of e_2 in $\mathcal{L}$.

The following observation guarantees the existence of a large set of chords with useful structural properties.

Claim 3.2. *For any $K > 0$, given a set S of independent chords in G of size Km, we may find either*

(1) a chain in $\mathcal{P}_S$ of size $Km^{1/3}$,
(2) a chain in $\mathcal{Q}_S$ of size $m^{1/3}$, or
(3) an antichain in both $\mathcal{P}_S$ and $\mathcal{Q}_S$ of size $m^{1/3}$.

Moreover, in either of the latter two cases, we may find a nontrivial cycle of length at least $n - n/m^{1/3}$ in G.

Proof of claim. By Proposition 2.1, we see that $\mathcal{P}_S$ contains either a chain of size $Km^{1/3}$ or an antichain of size $m^{2/3}$. Applying Proposition 2.1 again to such an antichain if it exists, we see that either $\mathcal{Q}_S$ contains a chain of size $m^{1/3}$, or there exists an antichain in both $\mathcal{P}_S$ and $\mathcal{Q}_S$ of size $m^{1/3}$.

If $\mathcal{Q}_S$ contains a chain of size $m^{1/3}$, then it is easy to see that this chain contains a chord whose exterior contains at most $n/m^{1/3}$ vertices, in which case we are done.

If there exists an antichain in both $\mathcal{P}_S$ and $\mathcal{Q}_S$ of size $m^{1/3}$, then it is clear that this antichain consists of pairwise interlacing chords. We may then find, using the pigeonhole principle, chords uv and xy in this antichain with $u \prec x \prec v \prec y$ such that $d_{\mathcal{C}}(u,x) + d_{\mathcal{C}}(v,y) \le n/m^{1/3}$, in which case we are again done. $\square$

For the rest of the proof, we restrict our attention to the set I and the poset $\mathcal{P} = \mathcal{P}_I$; in what follows, any ordering of chords in I will implicitly mean their ordering in $\mathcal{P}$. Furthermore, we may assume going forwards that in any set $S \subset I$ of size at least $m/8$, there exists a chain in $\mathcal{P}$ of size at least $m^{1/3}/8$; indeed, we are done by Claim 3.2 if this is not the case.

We say that a triple $\{u_1v_1 < u_2v_2 < u_3v_3\}$ of independent chords in I with $u_1 \prec u_2 \prec u_3 \prec v_3 \prec v_2 \prec v_1$ is *tight* if

$$d_{\mathcal{C}}(u_1,u_3) + d_{\mathcal{C}}(v_3,v_1) \le 24n/m^{1/3}.$$

This definition of a tight triple is motivated by the following observation.

Claim 3.3. *If G contains two tight triples whose middle chords interlace, then G contains a nontrivial cycle of length at least $n - 48n/m^{1/3}$.*

Proof of claim. This claim follows from a somewhat tedious analysis of a few different cases; this analysis requires us to establish some notation first. For a tight triple $U = \{u_1v_1 < u_2v_2 < u_3v_3\}$ with $u_1 \prec u_2 \prec u_3 \prec v_3 \prec v_2 \prec v_1$, we say that a vertex *lies inside the strip* of U if it lies either on the path $P(u_1,u_3)$ between u_1 and u_3 in $\mathcal{C}$ containing u_2, or on the path $P(v_3,v_1)$ between v_3 and v_1 in $\mathcal{C}$ containing v_2.

Suppose that $T_1 = \{u_1v_1 < u_2v_2 < u_3v_3\}$ with $u_1 \prec u_2 \prec u_3 \prec v_3 \prec v_2 \prec v_1$ and $T_2 = \{x_1y_1 < x_2y_2 < x_3y_3\}$ with $x_1 \prec x_2 \prec x_3 \prec y_3 \prec y_2 \prec y_1$ are two tight triples whose middle chords u_2v_2 and x_2y_2 interlace.

Assume first that T_1 and T_2 are not disjoint, and say $u_1v_1 = x_1y_1$ with $u_1 = x_1$ and $v_1 = y_1$. Suppose, as we may, that $u_1 \prec x_2 \prec u_2$; we then obtain a cycle using the chords u_2v_2 and x_2y_2 missing at most

$$d_{\mathcal{C}}(u_1, u_3) + d_{\mathcal{C}}(y_3, y_1) \leq 48n/m^{1/3}$$

vertices of G, as required.

Therefore, we may suppose that T_1 and T_2 are disjoint. Suppose first that x_2 and y_2 lie inside the strip of T_1. If both x_2 and y_2 lie on $P(u_1, u_3)$, then we obtain a cycle using just the chord x_2y_2 missing at most $d_{\mathcal{C}}(u_1, u_3) \leq 24n/m^{1/3}$ vertices. If x_2 lies on $P(u_1, u_3)$ and y_2 lies on $P(v_3, v_1)$ on the other hand, then we obtain a cycle using the chords u_2v_2 and x_2y_2 missing at most

$$d_{\mathcal{C}}(u_1, u_3) + d_{\mathcal{C}}(v_3, v_1) \leq 24n/m^{1/3}$$

vertices of G.

Therefore, suppose that x_2 lies outside the strip of T_1 and that u_2 lies outside the strip of T_2. Suppose without any loss of generality that $u_2 \prec u_3 \prec x_2 \prec v_3 \prec v_2$ and $y_2 \prec y_1 \prec u_2 \prec x_1 \prec x_2$, so either $u_2 \prec u_3 \prec x_1 \prec x_2$ or $u_2 \prec x_1 \prec u_3 \prec x_2$.

First, suppose that $u_2 \prec u_3 \prec x_1 \prec x_2$, in which case, both u_2v_2 and u_3v_3 interlace with both x_1y_1 and x_2y_2. We may then obtain a cycle using the chords u_2v_2, u_3v_3, x_1y_1 and x_2y_2 missing at most

$$d_{\mathcal{C}}(u_1, u_3) + d_{\mathcal{C}}(v_3, v_1) + d_{\mathcal{C}}(x_1, x_3) + d_{\mathcal{C}}(y_3, y_1) \leq 48n/m^{1/3}$$

vertices of G.

Now, suppose that $u_2 \prec x_1 \prec u_3 \prec x_2$. If $u_3 \prec v_3 \prec x_3$, then we obtain a cycle using the chord u_3v_3 missing at most $d_{\mathcal{C}}(x_1, x_3) \leq 24n/m^{1/3}$ vertices. Therefore, suppose that $u_3 \prec x_3 \prec v_3$. If $y_3 \prec v_2 \prec y_2$, then we obtain a cycle using the chords u_2v_2 and x_2y_2 missing at most

$$d_{\mathcal{C}}(u_2, x_2) + d_{\mathcal{C}}(v_2, y_2) \leq d_{\mathcal{C}}(u_1, u_3) + d_{\mathcal{C}}(x_1, x_3) + d_{\mathcal{C}}(y_3, y_1) \leq 48n/m^{1/3}$$

vertices. Hence, suppose that $v_2 \prec y_3 \prec y_2$, so that both u_2v_2 and u_3v_3 interlace with both x_2y_2 and x_3y_3. In this case, we obtain a cycle using the chords u_2v_2, u_3v_3, x_2y_2 and x_3y_3 missing at most

$$d_{\mathcal{C}}(u_1, u_3) + d_{\mathcal{C}}(v_3, v_1) + d_{\mathcal{C}}(x_1, x_3) + d_{\mathcal{C}}(y_3, y_1) \leq 48n/m^{1/3}$$

vertices of G. □

Continuing the proof of Lemma 3.1, recall our assumption that in any set $S \subset I$ of size at least $m/8$, there exists a chain in $\mathcal{P}$ of size at least $m^{1/3}/8$. This assumption implies that there are many pairwise disjoint tight triples in I, as we demonstrate below.

Claim 3.4. *For $K \geq 1/2$, any set $S \subset I$ of size Km contains $Km/4$ pairwise disjoint tight triples.*

Proof of claim. We shall show that given any collection T of at most $Km/4$ pairwise disjoint tight triples from S, we may find a tight triple from the remaining chords in S which is pairwise disjoint from each of the tight triples in T. We know that S contains a subset S' of at least $Km - 3Km/4 \geq Km/4 \geq m/8$ chords none of which appear in any of the triples in T. By our assumption, we know that S' contains a chain $u_1v_1 < u_2v_2 < \cdots < u_kv_k$ of size $k = m^{1/3}/8 \geq 6$ in $\mathcal{P}$ with $u_1 \prec u_2 \prec \ldots \prec u_k \prec v_k \prec v_{k-1} \prec \ldots \prec v_1$. By considering a partition of $\mathcal{C}$ into paths with endpoints in $\{u_1, u_3, \ldots, v_1, v_3, \ldots\}$, we have

$$\sum_{i=1}^{\lceil k/2 \rceil - 1} \left(d_{\mathcal{C}}(u_{2i-1}, u_{2i+1}) + d_{\mathcal{C}}(v_{2i+1}, v_{2i-1})\right) \leq n,$$

so there exists an index $1 \leq i \leq \lceil k/2 \rceil - 1$ such that

$$d_{\mathcal{C}}(u_{2i-1}, u_{2i+1}) + d_{\mathcal{C}}(v_{2i+1}, v_{2i-1}) \leq \frac{n}{k/2 - 1} \leq \frac{3n}{k} = \frac{24n}{m^{1/3}};$$

this implies that the triple $\{u_{2i-1}v_{2i-1} < u_{2i}v_{2i} < u_{2i+1}v_{2i+1}\}$ is tight, proving the claim. □

We may now finish the proof of Lemma 3.1 as follows. By Claims 3.3 and 3.4, we see that I contains $m/2$ pairwise disjoint tight triples whose middle chords are all parallel and independent. Applying Claim 3.4 again to the $m/2$ interlacing partners of the middle chords of the triples above, we obtain $m/8$ new pairwise disjoint tight triples; in particular, there exist two tight triples whose middle chords interlace, so we are done by Claim 3.3. $\square$

In order to handle Hamiltonian graphs with many parallel chords, we shall rely on the non-constructive argument implicit in Lemma 2.2. In order to apply this lemma in the proof of our main result, we shall require a fair bit of preparation; this is accomplished in the somewhat technical lemma that follows below.

LEMMA 3.5. *Let G be an n-vertex graph with a designated Hamiltonian cycle $\mathcal{C}$ with the property that no two chords of G interlace. Suppose that no vertex of G is chord-adjacent to two consecutive vertices of $\mathcal{C}$, and that no two vertices of G of degree greater than 3 are chord-adjacent. Also, assume that there are subsets R and B of $V(G)$ (whose elements we shall call red and blue respectively) such that*

(1) every vertex in $R \cup B$ has degree 3, and
(2) no two vertices in $R \cup B$ are chord-adjacent.

Then, writing $M \geq 2$ for the number of minimal chords in G and setting $r = |R|$, there exists a set $S \subseteq V(G)$ of vertices such that

(1) S dominates the chords of G,
(2) S contains no red vertices, and
(3) S contains at most $r + M - 2$ pairs of consecutive vertices of $\mathcal{C}$, and none of these pairs contains a blue vertex.

PROOF OF LEMMA. We prove this lemma by induction on the number of minimal chords as follows.

First, we prove the base case. Suppose that G has exactly two minimal chords. Let $e = xy$ and $f = uv$ be the two minimal chords, and since e and f cannot interlace by assumption, we may assume that $x \prec u \prec v \prec y$. We say that a vertex is *upstairs* if it lies between x and u on $\mathcal{C}$, and *downstairs* if it lies between v and y on $\mathcal{C}$; we write U and D for the sets of vertices upstairs and downstairs respectively. Note that $E(G) \setminus E(\mathcal{C})$ is a collection of stars, each of which is such that its centre is upstairs and all of its leaves are downstairs, or vice versa; let these stars be $S_1, S_2, \ldots, S_k$. Note that the centres of these stars are necessarily uncoloured; we adopt the convention that the centre of a trivial star consisting of a single edge is one of its uncoloured vertices. Furthermore, these stars come with a natural ordering: for $i < j$, all the vertices of S_i upstairs are closer to x than all the vertices of S_j upstairs, and all the vertices of S_i downstairs are closer to y than all the vertices of S_j downstairs. To ensure that S dominates the chords of G, we shall construct S by choosing, for each $1 \le i \le k$, either to add all the vertices of S_i that are upstairs to S, or to add all the vertices of S_i that are downstairs to S. Since no pair of leaves of any of these stars are consecutive vertices of $\mathcal{C}$, S can contain a pair of consecutive vertices of $\mathcal{C}$ only if the pair spans two stars. We may assume that there are r stars containing a red vertex; we denote these stars by $S_{i_1}, S_{i_2}, \ldots, S_{i_r}$. We partition the set of all stars into $r+1$ blocks as

$$\{S_{i_0}, \ldots, S_{i_1 - 1}\} \cup \{S_{i_1}, \ldots, S_{i_2 - 1}\} \cup \cdots \cup \{S_{i_{r-1}}, \ldots, S_{i_r - 1}\} \cup \{S_{i_r}, \ldots, S_{i_{r+1} - 1}\},$$

where $i_0 = 1$ and $i_{r+1} = k+1$. For each $0 \le j \le r-1$, we shall pick vertices in the block $\{S_{i_j}, \ldots, S_{i_{j+1}-1}\}$ ensuring that the last vertex picked is not blue, and that we pick at most one pair of consecutive vertices of $\mathcal{C}$ from $\{S_{i_j}, \ldots, S_{i_{j+1}}\}$. In the case where $j = r$, we shall ensure that we create no pair of consecutive vertices of $\mathcal{C}$ from the last block.

For $0 \le j \le r$, we handle the corresponding block of stars as follows. Without loss of generality, suppose that there is a red vertex downstairs in S_{i_j}, and consider the sequence

$$S_{i_j} \cap U, S_{i_j+1} \cap D, S_{i_j+2} \cap U, \ldots$$

of candidates for addition to S, where the sequence above goes up to the star with the index $i_{j+1} - 1$. We enlarge S using the block under consideration as follows. If $j = r$, then we add all the vertices in the sequence above. If $j < r$ and the last element in the sequence above containing vertices of $S_{i_{j+1}-1}$ is on the same side (upstairs or downstairs) as a red vertex of $S_{i_{j+1}}$, then we again add all the vertices in the sequence above. Suppose now that $j < r$ and that the last element in the sequence containing vertices of $S_{i_{j+1}-1}$ is on the opposite side as a red vertex of $S_{i_{j+1}}$. Let $i_j + t$ denote the index of the last set in the above sequence that does not contain a blue vertex, and note that $t \geq 0$. In this case, we add all the vertices in the sequence above up to the index $i_j + t$, and then add all the vertices in the complementary sequence (obtained by selecting vertices on the opposite side) from the index $i_j + t + 1$ to the index $i_{j+1} - 1$. It is clear from the properties that G is assumed to have that this selection procedure generates at most one pair of consecutive vertices of $\mathcal{C}$ (possibly between S_{i_j+t} and S_{i_j+t+1}) from this block, and it is also clear that the last vertex added to S from this block is not blue. Note that in the case where $j = 0$, if the corresponding block is nonempty, then there are no red vertices in this block; therefore, we can ensure that when considering the first nonempty block (which corresponds to either $j = 0$ or $j = 1$), the first set in the sequence above contains the centre but not the leaves of the first star in the block; we shall need this additional property later in the induction step.

It is easy to check that the above procedure applied to each of the $r + 1$ blocks of stars produces a set S as required, proving the base case of the induction.

Next, suppose that $M \geq 3$. Pick a minimal chord f. Among all chords whose domain inducing f induces no other chords (except the chord in question itself), pick a chord $e = xy$ which is maximal with respect to the order of its domain inducing f; denote the domains of e by A and B, where A is the domain of e inducing f. Clearly, both $G[A]$ and $G[B]$ are Hamiltonian graphs satisfying the conditions of the lemma; moreover, $G[A]$ has at most 2 minimal chords, and by our maximal choice of e, it is also clear that $G[B]$ has exactly $M-1$ minimal chords.

We now apply the inductive hypothesis to the graphs G_A and G_B that we now define. First, G_A is obtained from $G[A]$ by adding a new uncoloured vertex z and joining it to x and y. It is clear that G_A has at most two minimal chords; say G_A contains r_1 red vertices, and set $r_2 = r - r_1$. Next, we obtain G_B from $G[B]$ by recolouring some vertices as follows. Without loss of generality, we may assume that y is the uncoloured centre of the star containing e in $E(G) \setminus E(\mathcal{C})$. Let w be the neighbour of y in $\mathcal{C}$ that belongs to $G[B]$. We make w red in G_B if it was coloured blue in G (and do not alter its colour otherwise), and if x was red or blue in G, then we make x an uncoloured vertex in G_B. Clearly, G_B has $M-1$ minimal chords, and either at most r_2+1 or at most r_2 red vertices depending on whether or not the colour of w was altered in G_B.

Let S_A and S_B be the sets obtained inductively in G_A and G_B respectively. First, $e = xy$ is a minimal chord in G_A, and G_A has at most two minimal chords, so we can ask for S_A to contain y but not x by arguing as in the base case earlier. Next, note that S_B either contains at most $(r_2+1)+(M-1)-2$ pairs of consecutive vertices of $\mathcal{C}$, or at most $r_2+(M-1)-2$ pairs of consecutive vertices of $\mathcal{C}$, depending on whether or not we had to alter the colour of w in G_B. Also, observe that x has degree 2 in G_B, so we may assume that S_B does not contain x.

We now claim that $S = S_A \cup S_B$ is sufficient for our purposes. It is clear that S dominates $E(G) \setminus E(\mathcal{C})$ and contains no red vertices of G. It is also clear, by induction, that S does not contain a consecutive pair of $\mathcal{C}$ in which one of the vertices is coloured blue in G. Next, if the colour of w was altered in G_B, then S does not contain any consecutive pairs of $\mathcal{C}$ spanning S_A and S_B since $x \notin S_A \cup S_B$ and $w \notin S_B$, and if not, then S contains at most one such pair (namely, the edge yw); it follows that the number of pairs of consecutive vertices of $\mathcal{C}$ in S is at most $(r_2+1)+(M-1)-2+r_1 = r+M-2$ in the former case, and at most $r_2+(M-1)-2+r_1+1 = r+M-2$ in the latter case, thereby completing the proof. □

Armed with Lemmas 3.1 and 3.5, we are now in a position to prove our main result.

Proof of Theorem 1.1. Let G be an n-vertex graph with a designated Hamiltonian cycle $\mathcal{C}$. We assume, without loss of generality, that G is minimal in the sense that no two vertices with degree greater than 3 in G are chord-adjacent.

Let $2m$ be the maximum size of a set I of independent chords in G which may be partitioned into m interlacing pairs. If $m \geq n^{3/5}$, then the result follows from Lemma 3.1, so we may suppose that $m \leq n^{3/5}$.

Let P denote the set of $4m$ endpoints of the chords in I, and consider the graph G' on the same vertex set as G obtained by deleting every chord of G incident to some vertex in P; of course, G' is also an n-vertex graph in which $\mathcal{C}$ is the designated Hamiltonian cycle, and from the maximality of I, we see that no two chords of G' interlace. We now transform G' as follows: if x and y are consecutive vertices of $\mathcal{C}$ that are both chord-adjacent to some vertex of G', then we contract the edge xy of $\mathcal{C}$, and repeat this operation until it is no longer possible to do so. Let H be the resulting graph, and let $\mathcal{D}$ be its designated Hamiltonian cycle obtained from $\mathcal{C}$ after these contractions; note that our contractions ensure that no vertex of H is chord-adjacent to two consecutive vertices of $\mathcal{D}$.

Now, the set of minimal chords of G' with respect to $\mathcal{C}$ is the same (up to the obvious identification) as the set of minimal chords of H with respect to $\mathcal{D}$, and furthermore, the size of the minimal domains of these minimal chords are identical in both G' and H. Moreover, it is easy to see that H does not contain a pair of interlacing chords. We call any vertex of H that corresponds to one or more contracted edges of G' a *contracted vertex*, and we colour a contracted vertex red in H if it is the image of $n^{1/5}$ or more contracted edges, and blue otherwise. By the minimality of G assumed above, we see that each contracted vertex of H is the image under contractions of some set of vertices all of which have degree 3 in G; hence, no contracted vertex is chord-adjacent in G to any vertex in P, and no two contracted vertices are chord-adjacent.

Write M for the number of minimal chords of H, and let r denote the number of red vertices in H. Note that, by definition, we have $r \leq n^{4/5}$ since each red vertex corresponds to a set of at least $n^{1/5}$ vertices of G, and these sets are all pairwise disjoint. Next, since H does not contain any interlacing pairs of chords, the minimal domains of the minimal chords of H are all pairwise disjoint, so if $M \geq n^{1/2}$, then one of these minimal domains contains at most $n^{1/2}$ vertices in H, and therefore in G' and G as well, in which case we are done. Therefore, we may suppose that $M \leq n^{1/2}$.

We now apply Lemma 3.5 to H with $\mathcal{D}$ as its designated Hamiltonian cycle to get a set S of vertices such that S dominates $E(H) \setminus E(\mathcal{D})$, contains no red vertices, and contains at most $r + M - 2$ pairs of consecutive vertices of $\mathcal{D}$ with none of these pairs containing a blue vertex. Let us now add back to H the chords that we deleted earlier, namely, those chords incident to some vertex in P; we call the resulting graph H'. Note that $X = P \cup S$ dominates the $V(H') \setminus X$ in the graph spanned by the chords of H' since every vertex of degree 2 in H is chord-adjacent to some vertex in P; furthermore, X contains at most $8m + r + M - 2$ consecutive pairs of vertices of $\mathcal{D}$.

We would like to apply Lemma 2.2 to H'; to do so, we need to ensure that X is independent in the graph spanned by the edges of $\mathcal{D}$. To ensure this, we shall contract every edge of $\mathcal{D}$ between two vertices of X; we call the resulting graph F and let $\mathcal{E}$ be its designated Hamiltonian cycle obtained from $\mathcal{D}$ after these contractions. Clearly, the image of X in F is a set that satisfies all the conditions of Lemma 2.2 with respect to F and $\mathcal{E}$; therefore, it follows from Lemma 2.2 that F contains another Hamiltonian cycle $\mathcal{F}$. Note that we have not contracted any edge incident to some red vertex in H' in constructing F; moreover, we have contracted at most $8m$ blue vertices of H' in constructing F.

Now, this cycle $\mathcal{F}$ in F gives rise to a cycle $\mathcal{D}'$ in H' missing at most $8m + r + M - 2$ vertices of H'; indeed, at most $8m$ of the missing vertices are blue, no red vertex is missed, and the remaining missing vertices are non-contracted vertices of G. Now, we lift this cycle $\mathcal{D}'$ in H' to a cycle $\mathcal{C}'$ in G by replacing each red or blue vertex in $\mathcal{D}'$ with an appropriate path of the original vertices of G; we can always choose this path to contain all the pre-images of the coloured vertex in question since, as mentioned earlier, all such vertices have degree 3 in H'. It then follows that $\mathcal{C}'$ misses at most $8mn^{1/5} + r + M - 2$ vertices of G. Also, note that $\mathcal{C}' \neq \mathcal{C}$ since $\mathcal{F}$ contains at least one chord of F (and also G), and this chord is present in $\mathcal{C}'$.

It is now clear that $\mathcal{C}'$ is a nontrivial cycle of G, and that the length of $\mathcal{C}'$ is at least

$$n - (8mn^{1/5} + r + M - 2);$$

the result follows since we know that $m \leq n^{3/5}$, $r \leq n^{4/5}$ and $M \leq n^{1/2}$. $\square$

4. Concluding remarks

Our results raise a number of questions. Perhaps the most fundamental of these concerns the nature of the error term in Theorem 1.1. We expect that it should be possible to improve the exponent of 4/5 in the error term in our main result using the methods developed here, possibly up to an exponent of 1/2; however, we chose to keep the presentation simple because we expect much more to be true.

CONJECTURE 4.1. *If an n-vertex graph G with $\delta(G) \geq 3$ contains a Hamiltonian cycle, then G contains another cycle of length at least $n - K$, where $K > 0$ is an absolute constant.*

We remark that it is not impossible that Conjecture 4.1 holds even with $K = 2$;

Next, while a minimum degree of 3 is not sufficient, as discussed earlier, to guarantee a second Hamiltonian cycle in a Hamiltonian graph, we remind the reader that it is still unknown if a minimum degree of 100, say, suffices instead; see [65, 28, 50] for more details.

To close the chapter, let us mention a conjecture due to Verstraëte [113] that seems closely related to the problem we have addressed here.

CONJECTURE 4.2 (Verstraëte). *If an n-vertex graph G with $\delta(G) \geq 3$ contains a Hamiltonian cycle, then G contains cycles of $\Omega(n)$ distinct lengths.*

It is easy to deduce a lower bound of the form $\Omega(\sqrt{n})$ for the above problem using the poset-based arguments developed here.

CHAPTER 7

Large induced subgraphs with k vertices of almost maximum degree

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., [15, 24, 36, 38, 46]), 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 $\alpha(G) = (1 + o(1))n/k$ ([24]). 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 [35]. 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 ([37]) 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 [34].

CONJECTURE 1.1. *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, the authors proved that $g(2, \Delta) = \left\lceil \frac{3+\sqrt{8\Delta+1}}{2} \right\rceil$ 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 [34] 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.

In this chapter, our main aim is to prove the following approximate version of Conjecture 1.1.

THEOREM 1.2. *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 $\Delta(H) - g_2(k)$.*

2. Notation and preliminaries

Our notation is mostly standard. We need to introduce the following defintions. 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 $\mathcal{A}$ 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 $|\{A \in \mathcal{A} : j \in A\}| = r_i$. Note that in a multiset we allow repetitions.

We call an $(r_1, r_2, \ldots, r_t)$-uniform cover $\mathcal{A}$ 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 $\mathcal{B} \subset \mathcal{A}$, for some strictly decreasing sequence of non-negative integers $r'_1 > r'_2 > \ldots > r'_t$.

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

As usual, we write $R(k)$ (see e.g. [23]) for the *two coloured Ramsey number*, the least integer n such that in any two colouring of the edges of the complete graph on n vertices, there is a monochromatic K_k.

Finally, in order to prove our main theorem, we make use of the following Theorem of Caro, Shapira and Yuster, appearing in [35] from 2013.

THEOREM 2.1. *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.*

Indeed, Caro et al proof of Theorem 2.1 is inspired by a similar idea used by Alon and Berman in [9]. The idea relies on a clever application of the following beautiful theorem due to Sevast'janov [91].

THEOREM 2.2 (Sevast'janov). *Let V be any d-dimensional space normed space. Suppose $v_1, \ldots, v_n \in V$ where $\|v_i\| \leq 1$ and $\sum_{i=1}^{n} v_i = 0$. Then there is a permutation α on $\{1, \ldots, n\}$ such that for all $j = 1, \ldots, n$,*

$$\left\| \sum_{i=1}^{j} v_{\alpha(i)} \right\| \leq d.$$

3. Proof of the main result

In this section, we shall prove Theorem 1.2. To do so, we use two lemmas.

LEMMA 3.1. *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 $\mathcal{A}$ of $\{1, 2, \ldots, n\}$ contains a proper $(r'_1, r'_2, \ldots, r'_t)$-uniform subcover $\mathcal{B} \subset \mathcal{A}$ with $r'_1 \leq f(n)$.*

PROOF. We will show 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 $w_{B_{l_{i_1}}}(A)\le w_{B_{l_{i_2}}}(A)\le\ldots$. 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) $\mathcal{A},\mathcal{B}$ for which $w_{\mathcal{A}}(F)\le w_{\mathcal{B}}(F)$ for every $F\subseteq\{1,2,\ldots,n\}$. This implies $\mathcal{A}\subseteq\mathcal{B}$, which is a contradiction. Take $f(n)$ to be the maximum r_1 over all *irreducible* uniform covers of $\{1,2,\ldots,n\}$. $\square$

LEMMA 3.2. *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 3.1. 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 neighbour of w in G. Note that $\mathcal{A}=\{A_w : w\in B\}$ is an $(r_1,r_2,\ldots,r_t)$-uniform cover of $\{1,2,\ldots,n\}$. Applying now Lemma 3.1, we can find a $(r'_1,r'_2,\ldots,r'_t)$-uniform sub-cover $\mathcal{B}\subseteq\mathcal{A}$ with $r'_1\le f(n)$. Let $W=\{w\in B : A_w\in\mathcal{B}\}$ and $G'=G[A,(B\setminus W)]$. It is easy to see that

$|W| \leq n \cdot f(n)$ and that the property is satisfied by the definition of uniform cover. $\square$

Given a positive integer k, and a graph G with the vertex set $\{x_1, \ldots, x_n\}$ such that $d(x_1) \geq \cdots \geq d(x_n)$, we 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 3.3. *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 $r_k(H) \leq h(k) \cdot k$.*

Proof of Theorem 3.3. The proof consists of two parts. Firstly, we show that we can remove at most $k\sqrt{\Delta}$ vertices from G so that in the remaining graph H' we have $r_k(H') \leq \sqrt{\Delta}$. Then we iteratively apply Lemma 3.2 (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 $r_k(H) \leq h(k) \cdot k$. We may take $h(k)$ to be $f'(k)$ from Lemma 3.2.

We start with the first part of the proof.

CLAIM 3.4. *There is an induced subgraph H' of G on at least $n - k\sqrt{\Delta}$ vertices such that $r_k(H') \leq \sqrt{\Delta}$.*

The idea 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 3.2. 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 $H_0, \dots, H_i$ have already been defined. If $r_k(H_i) \le k \cdot h(k)$ then we are done, so we may assume that $r_k(H_i) > k \cdot h(k)$. Let $A = \{x_1, \dots, x_k\}$ 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 $d_{H_i}(x_1) \ge \dots \ge d_{H_i}(x_k)$. Since $r_k(H_i) \ge k \cdot h(k)$ there must exist $l \in \{2, \dots, k\}$ 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 3.2, 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

$$\text{for any } x, y \in A, \text{ if } d_K(x) < d_K(y) \text{ then } d_K(x) - d_{K'}(x) < d_K(y) - d_{K'}(y) \quad (4)$$

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

Claim 3.5. *$r_k(H_{i+1}) < r_k(H_i)$.*

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 $t \ge l$ and $d_{H_{i+1}}(w) \ge d_{H_{i+1}}(x_s)$ for some $s < l$. First, notice that $d_{H_i}(x_t) - d_{H_i}(x_s) \le d_{H_i}(x_1) - d_{H_i}(x_k) = r_k(H_i)$. Hence, $r_k(H_{i+1}) = d_{H_{i+1}}(z) - d_{H_{i+1}}(w) \le d_{H_{i+1}}(x_t) - d_{H_{i+1}}(x_s) < d_{H_i}(x_t) - d_{H_i}(x_s) \le r_k(H_i)$, where the strict inequality follows from (4) 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 $r_k(H) \le k \cdot h(k)$ and $|H| \ge |H'| - h(k)\sqrt{\Delta} \ge n - (h(k) + k)\sqrt{\Delta}$. $\square$

We are now ready to conclude the proof of our main Theorem.

Proof of Theorem 1.2. Firstly, we apply Theorem 3.3 with $k = R(k)$ to find a large induced subgraph $G' \subset G$ of order at least $n' \ge n - (h(R(k)) +$

$R(k))\sqrt{\Delta}$ and with vertex set $\{x_1, \ldots, x_{n'}\}$ where $d(x_1) \geq d(x_2) \geq \cdots \geq d(x_{n'})$ and $d(x_1) - d(x_{R(k)}) \leq h(R(k)) \cdot R(k) = M$. We should point out from now on, our approach is the same as in the proof of Theorem 1.1 in [35].

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,

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

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 2.1, 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. Concluding 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 1.2 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 8

Majority colourings of digraphs

1. Introduction and theorems

For a natural number $k \geq 2$, a $\frac{1}{k}$-*majority colouring* of a digraph is a colouring of the vertices such that each vertex receives the same colour as at most $1/k$ proportion of its out-neighbours. We say that a digraph D is $\frac{1}{k}$-*majority* m-*colourable* if there exists a $\frac{1}{k}$-majority colouring of D using m colours. The following natural question was recently raised by Kreutzer, Oum, Seymour, van der Zypen and Wood [74].

QUESTION 1.1. 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 colouring of this tournament must be a proper vertex-colouring, and hence it needs at least $2k-1$ colours. Conversely, we prove that $m(k) \leq 2k$.

THEOREM 1.2. *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 [32]) about partitions of matrices. We shall use a slightly more general version proved by Alon [8].

LEMMA 1.3. *Let* $A = (a_{ij})$ *be an* $n \times n$ *real matrix where* $a_{ii} = 0$ *for all* i, *and* $a_{ij} \geq 0$ *for all* $i \neq j$, *and* $\sum_j a_{ij} \leq 1$, *for all* $i \in \{1, 2, \ldots, n\}$. *Then, for every positive integer* t *and all positive reals* $c_1, \ldots, c_t$ *whose sum is* 1, *there is* *a partition of* $\{1, 2, \ldots, n\}$ *into* t *pairwise disjoint sets* $S_1, S_2, \ldots, S_t$*, such that for every* r*,* $1 \leq r \leq t$ *and every* $i \in S_r$*,* $\sum_{j \in S_r} a_{ij} \leq 2c_r$*.*

Proof of Theorem 1.2. 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 1.3 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, $1 \leq r \leq 2k$ and every $i \in S_r$, $\sum_{j \in S_r} a_{ij} \leq \frac{1}{k}$. Equivalently, the number of out-neighbours of v_i that have the same colour as v_i is at most $\frac{d^+(v_i)}{k}$ where the colouring of D is defined by the partition $S_1 \cup S_2 \cup \cdots \cup S_{2k}$. $\square$

Question 1.1 has now been reduced to whether $m(k)$ is $2k - 1$ or $2k$.

Question 1.4. 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 [74] gave an elegant argument showing that every digraph is $\frac{1}{2}$-majority 4-colourable and they conjectured that $m(2) = 3$.

Conjecture 1.5. *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 1.6. *Every tournament can be* 3*-coloured in such a way that all but at most* 205 *vertices receive the same colour as at most half of their out-neighbours.*

Proof of Theorem 1.6 . 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 colours to each vertex independently with probability 1/3. Given a vertex x, let B_x be the number of out-neighbours of x which receive the same colour 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 $x \in S_i$,

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

Notice that if $i \geq 10$ then $\mathbb{P}(\text{x is bad}) \leq 2^{-(2i-7)}$. Let X denote the total number of bad vertices. Since the vertices of out-degree 0 cannot be bad,

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

Hence, there is a 3-colouring such that all but at most 205 vertices receive the same colour as at most half of their out-neighbours.

□

Observe also that the exact same argument proves the following special case of Conjecture 1.5.

THEOREM 1.7. *Every tournament with minimum out-degree at least* 2^{10} *is* $\frac{1}{2}$*-majority* 3*-colourable.*

We remark that Theorem 1.6 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 strictly less than 7.75. To do so, let V_i be the set of vertices of out-degree i for $i \in \{1, \ldots, 1024\}$ and note that the expectation of the number of *bad* vertices of degree at most 1024 is $f(v_1, \ldots, v_{1024}) = \sum_{i=1}^{1024} v_i p_i$ 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 less or equal than i is at most $2i+1$, therefore $\sum_{j=1}^{i} v_i \leq 2i+1$.

$$\text{Maximize: } f(v_1, \ldots, v_{1024})$$

$$\text{Subject to: } \sum_{j=1}^{i} v_j \leq 2i+1, \text{ for } i \in \{1, \ldots, 1024\}$$

$$\text{Subject to: } v_i \geq 0, \text{ for } i \in \{1, \ldots, 1024\}$$

See 2 for the source code. Similarly, we can replace 2^{10} in Theorem 1.7 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 colours to the vertices, there exists a $\frac{1}{k}$-majority colouring where each vertex gets a colour 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 [74] asked whether there exists a finite number m such that every digraph is $\frac{1}{2}$-majority m-choosable. Anholcer, Bosek and Grytczuk [11] showed that the statement holds with $m = 4$. We generalise their result as follows.

THEOREM 1.8. *Every digraph is $\frac{1}{k}$-majority $2k$-choosable for all $k \geq 2$.*

Theorem 1.8 was independently proved by Fiachra Knox and Robert Šámal [72]. We prove Theorem 1.8 using a slight modification of Lemma 1.3.

LEMMA 1.9. *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 OF LEMMA. 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 minimises 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 1.8. The proof is the same as that of Theorem 1.2, using Lemma 1.9 instead of Lemma 1.3. □

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

COROLLARY 1.10. *Every digraph is $\frac{2}{m}$-majority m-choosable for all $m \geq 2$.*

This statement generalises a result of Anholcer, Bosek and Grytczuk [11] where they prove the case $m = 3$ which says that, given a digraph with colour lists of size three assigned to the vertices, there is a colouring from these lists such that each vertex has the same colour as at most two thirds of its out-neighbours.

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 1.4.

QUESTION 1.11. Is every digraph $\frac{1}{k}$-majority $(2k-1)$-choosable?

2. Code for the linear program

We used the toolkit [1] to solve the linear program, using 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 9

Highly linked tournaments

1. Introduction

Given a positive integer k, a graph is said to be *k-linked* if for any two disjoint sets of vertices $\{x_1, \dots, x_k\}$ and $\{y_1, \dots, y_k\}$ there are vertex disjoint paths $P_1, \dots, P_k$ such that P_i joins x_i to y_i for $i = 1, \dots, k$. Clearly, k-linkedness is a stronger notion than k-connectivity for graphs with at least $2k$ vertices. But how much stronger is it? Larman and Mani [76] and Jung [66] showed that there is an $f(k)$ such that any $f(k)$-connected graph is k-linked. They based their result on a theorem of Mader [80], which implies that for any k, any sufficiently connected graph contains a subdivision of a complete graph on $3k$ vertices, and noticed that any $2k$-connected graph containing such a subdivision must be k-linked. Their proofs show that $f(k)$ can be taken to be exponential in k. Later, Bollobás and Thomason [25] proved that $f(k) = 22k$ will do.

The definitions of k-connectivity and k-linkedness carry over to directed graphs. A directed graph is *strongly connected* if for any pair of distinct vertices x and y there is a directed path from x to y, and is strongly k-connected if it remains connected upon removal of any set of at most $k-1$ vertices. In what follows, we shall omit the use of the word 'strongly' with the understanding that we always mean strong connectivity. A directed graph D is *k-linked* if for any two disjoint sets of vertices $\{x_1, \dots, x_k\}$ and $\{y_1, \dots, y_k\}$ there are pairwise vertex disjoint directed paths $P_1, \dots, P_k$ such that P_i has initial vertex x_i and terminal vertex y_i for every $i \in [k]$. Thus, D is 1-linked if and only if it is connected.

Menger's Theorem carries over in the directed case as well and asserts that a directed graph is k-connected if and only if for any two distinct vertices x and y there are k internally vertex disjoint *directed* paths from x to y. And finally, the notion of k-linkedness is the same for directed graphs with the condition that all paths must be directed.

Directed graphs exhibit quite different behaviour from undirected graphs with respect to the relations they bear between connectivity and linkedness. Indeed, Thomassen [104] constructed directed graphs with arbitrarily large connectivity which are not even 2-linked. Since large connectivity does not necessarily imply linkedness for general directed graphs, it is natural to consider the situation for a restricted class of directed graphs, namely, *tournaments*. A tournament is a complete graph where every edge has a unique direction. Thomassen [103] proved that there is a $g(k)$ such that every $g(k)$-connected tournament is k-linked, where $g(k)$ can be taken to be $Ck!$, for some absolute constant C. Greatly improving Thomassen's bound on $g(k)$, Kühn, Lapinskas, Osthus, and Patel [75] showed that one may take $g(k) = 10^4 k \log k$ and still ensure k-linkedness. They went on to conjecture that $g(k)$ may be taken to be linear in k. Pokrovskiy [87] resolved this conjecture by showing that any $452k$-connected tournament is k-linked. Except for small values of k, an optimal bound for $g(k)$ is not known. Bang-Jensen [16] showed that any 5-connected tournament is 2-linked, and there exists a family of 4-connected tournaments which are not 2-linked. Moreover, it is easy to construct $(2k-2)$-connected tournaments with arbitrarily large out and in-degree which are not k-linked: consider the blow up of a directed triangle with vertex sets A, B, C such that $|C| = 2k-2$ and A and B have size at least $2k$.

Going back to undirected graphs for a moment, if some density conditions are assumed on the graph, then Bollobás and Thomason's $22k$ can be taken all the way down to $2k$, since Mader [80] proved that a graph with sufficiently large average degree contains a subdivision of a complete graph of order $3k$. Note that $2k$ is close to the theoretical minimum connectivity in any k-linked graph (a k-linked graph is necessarily $(2k-1)$-connected). Recently, Thomas and Wollan [99] showed that any $2k$-connected graph with average degree at least $10k$ is k-linked, greatly reducing the bound on the required average degree. Motivated by this result, Pokrovskiy [87] conjectured that a similar phenomenon should occur for tournaments with a natural 'density' condition: high minimum out-degree and in-degree. In particular, he conjectured that there is a function $f : \mathbb{N} \to \mathbb{N}$ such that any $2k$-connected tournament with minimum out and in-degree at least $f(k)$ is k-linked. Here is our main result, which solves Pokrovskiy's conjecture within a factor of two on the connectivity bound.

Theorem 1.1. *For every positive integer k there exists $f(k)$ such that every $4k$-connected tournament T with $\delta^+(T) \geq f(k)$ is k-linked.*

We remark that we do not assume any lower bound on the minimum in-degree.

Recall that the complete directed graph $\overrightarrow{K}_k$ is the directed graph on k vertices where, for every pair x, y of distinct vertices, both xy and yx are present. In order to prove Theorem 1.1 we shall show that large minimum out-degree allows us to embed subdivisions of the complete directed graph $\overrightarrow{K}_k$. As we mentioned earlier, Mader [80] showed that for any positive integer k there is $g(k)$ such that any graph with average degree at least $g(k)$ contains a subdivision of K_k. The following theorem can be viewed as an analogue of Mader's result for tournaments, replacing 'average degree' with 'minimum out-degree', and may be of independent interest.

Theorem 1.2. *For any positive integer k there exists a $d(k)$ such that the following holds. If T is a tournament with $\delta^+(T) \geq d(k)$, then T contains a subdivision of $\overrightarrow{K}_k$.*

We remark that, as shown by Mader [81], this theorem does not hold if we replace T by a general digraph. This fact also follows from a result of

Thomassen [102], who showed that for every integer n there exist digraphs on n vertices with minimum out-degree at least $\frac{1}{2}\log n$ which do not contain a directed cycle of even length. But since any subdivision of $\overrightarrow{K}_3$ must contain an even directed cycle, these digraphs do not contain any subdivision of a complete directed graph on at least 3 vertices.

In order to prove Theorem 1.1, we shall need a little more than Theorem 1.2. Roughly speaking, we shall first embed in T a subdivided $\overrightarrow{K}_k$, and then attach a few additional paths to it (see Section 3).

2. Organization and notation

The remainder of this chapter is organized as follows. In Section 3, we prove Theorem 1.2 which allows us to embed subdivisions of a complete directed graph and related structures in tournaments with high minimum out-degree. In Section 4, we shall prove one preparatory lemma and then finish our proof of Theorem 1.1. Our final section concludes with some open problems.

Our notation is standard. Thus, for a directed graph D we use $N^+(x), N^-(x)$, and $d^+(x), d^-(x)$ to denote the out-neighbourhood, in-neighbourhood and out-degree, in-degree of a vertex x, respectively. We use $\delta^+(D)$ to denote the minimum out-degree of D. A directed path $P = x_1 \dots x_\ell$ in D is a sequence of distinct vertices such that $x_i x_{i+1}$ is an edge for every $i = 1, \dots, \ell - 1$. We call x_1 the *initial vertex* and x_ℓ the *terminal vertex* of P. The length of P is the number of its directed edges. We say that P is *internally disjoint* from some subset $X \subset V(D)$ if $\ell \geq 3$ and $\{x_2, \dots, x_{\ell-1}\} \cap X = \varnothing$. If A and B are subsets of $V(D)$, then we shall write $A \to B$ if every edge with one endpoint in A and the other endpoint in B is directed from A to B. Lastly, if $\mathcal{P}$ is a family of directed paths in a digraph, then we use $\bigcup \mathcal{P}$ to denote the set $\bigcup_{P \in \mathcal{P}} V(P)$.

3. Embedding a subdivided complete directed graph

The first proof of the result that graphs with sufficiently large connectivity are k-linked use a result of Mader, which allows one to embed a subdivision of a complete graph in a graph with sufficiently large average degree. Our proof of Theorem 1.1 follows a similar strategy. In order to proceed, we need a directed analogue of Mader's result for tournaments: we prove this in the present section. We shall use the following simple lemma of Lichiardopol [78] (independently rediscovered by Havet and Lidický [61]). We include a short proof for convenience of the reader.

LEMMA 3.1. *Every tournament with minimum out-degree at least k has a subtournament with minimum out-degree k and order at most $3k^2$.*

PROOF. Let T be a tournament with minimum out-degree at least k, and let T' be a vertex-minimal subtournament of T such that $\delta^+(T') \geq k$. Denote by L the collection of vertices in T' with out-degree k in T', and let $|T'| = t$ and $|L| = \ell$. By minimality, for every vertex $v \in T'$ we have $\delta^+(T' \setminus \{v\}) \leq k-1$. Hence, every vertex in $T' \setminus L$ has an in-neighbour in L, and so there are at least $t - \ell$ edges from L to $T' \setminus L$. On the other hand, the number of such edges is exactly

$$\ell k - \binom{\ell}{2},$$

and so $t - \ell \leq \ell k - \ell^2/2 + \ell/2$. It follows that

$$\ell^2 - \ell(2k+3) + 2t \leq 0,$$

implying the bound $(2k+3)^2 - 8t \geq 0$. In other words, $t \leq \frac{1}{8}(2k+3)^2$, so since t must be an integer we get $t \leq \frac{1}{8}((2k+3)^2 - 1) = k^2/2 + 3k/2 + 1 \leq 3k^2$, as required.

□

We are now ready to prove Theorem 1.2. In the following, for a positive integer k and nonnegative integer $m \leq 2\binom{k}{2}$, an *m-partial* $\overrightarrow{K}_k$ is any spanning subdigraph of $\overrightarrow{K}_k$ with precisely m directed edges present. Our proof shows that we can find a subdivision of $\overrightarrow{K}_k$ by inductively finding subdivisions of m-partial $\overrightarrow{K}_k$'s for each $m \leq 2\binom{k}{2}$.

Proof of Theorem 1.2. For a positive integer k and nonnegative integer $m \le 2\binom{k}{2}$, let $d(k,m)$ denote the smallest positive integer such that any tournament with $\delta^+(T) \ge d(k,m)$ contains a subdivision of an m-partial complete directed graph on k vertices. We shall show that if $m < 2\binom{k}{2}$, then $d(k,m+1) \le 7d(k,m)^2$. We use induction on k, and for each fixed k, induction on m. For $k=1$ there is nothing to show and we can take $d(1,0)=1$. So let us assume $k \ge 2$ is given and that we can embed a subdivision of an m-partial $\overrightarrow{K}_k$ in any tournament with minimum out-degree at least $d(k,m)$, and let T be a tournament with $\delta^+(T) \ge 7d(k,m)^2$.

Claim 3.2. *We may assume that there is a subdivision of an m-partial $\overrightarrow{K}_k$ contained in the out-neighbourhood of some vertex of T, and which spans at most $3d(k,m)^2$ vertices.*

Since certainly we have $\delta^+(T) \ge d(k,m)$, by Theorem 3.1 we may find a subtournament T' of size at most $3d(k,m)^2$ and with minimum out-degree at least $d(k,m)$. By induction we may embed in T' a subdivision of an m-partial $\overrightarrow{K}_k$. Denote this subdivision by K. We wish to add a missing directed edge, say xy. In other words, we must find a directed path from x to y in T such this path is internally disjoint from $V(K)$. Let $T'' = T \setminus T'$ and partition it into strongly connected subtournaments $T'' = S_1 \cup \cdots \cup S_\ell$ such that $S_i \to S_j$ for all $1 \le i < j \le \ell$ (unless, of course, T'' itself is strongly connected). Observe that since $d^+(x) \ge 7d(k,m)^2$ and $|T'| \le 3d(k,m)^2$, we have that x has an out-neighbour in T''. Therefore, if some vertex of S_ℓ is joined to y we are done, as we can find a directed path from x to y outside of T'. So we may assume that $S_\ell \subseteq N^+(y)$. Now, as $|T'| \le 3d(k,m)^2$ and no vertex of S_ℓ is joined to any vertex of S_i for $i < \ell$, we have that

$$\delta^+(S_\ell) \ge 7d(k,m)^2 - 3d(k,m)^2 \ge d(k,m),$$

Applying Theorem 3.1 to S_ℓ, we find a subtournament $S \subseteq S_\ell$ such that $\delta^+(S) \ge d(k,m)$ and with size at most $3d(k,m)^2$. It follows by induction that we may embed a subdivision of an m-partial $\overrightarrow{K}_k$ in S. But since $S \subseteq S_\ell \subseteq N^+(y)$ and $|S| \leq 3d(k,m)^2$, the claim holds.

By Claim 3.2, choose a vertex z with the smallest possible minimum out-degree satisfying the property that there is a subdivision of an m-partial $\overrightarrow{K}_k$ contained in $N^+(z)$ spanning at most $3d(k,m)^2$ vertices. Denote by N the out-neighbourhood of z and K_z the subdivision with $K_z \subseteq N$. We wish to add one more directed edge to this subdivision, say uv with $u, v \in K_z$. From N remove all vertices of K_z except for u and v and call this set N'. If $T[N']$ is strongly connected then we are done; otherwise, partition $T[N']$ into strongly connected subtournaments, say $T[N'] = S'_1 \cup \cdots \cup S'_t$ where $S'_i \to S'_j$ for all $1 \leq i < j \leq t$. Suppose that some vertex $w \in S'_t$ is joined to a vertex $w' \in N^-(z)$. Then since there is a directed path P from u to w in $T[N']$ we have that $uPww'zv$ is a directed path from u to v which avoids $K_z \setminus \{u, v\}$. Hence we may assume that every vertex of $N^-(z)$ dominates S'_t. But then, since $|K_z| \leq 3d(k,m)^2$ and there are no edges from S'_t to S'_i for $i < t$, one has that $\delta^+(S'_t) \geq 7d(k,m)^2 - 3d(k,m)^2 = 4d(k,m)^2$. So we can repeat the argument in Claim 3.2 to S'_t with minimum out-degree $4d(k,m)^2$ instead of $7d(k,m)^2$ (observe that we need $4d(k,m)^2 - 3d(k,m)^2 \geq d(k,m)$ to hold, which is clearly true). Accordingly, there is a vertex $q \in S'_t$ such that $N^+(q)$ contains a subdivision of an m-partial $\overrightarrow{K}_k$ spanning at most $3d(k,m)^2$ vertices. However, since $\bigcup_{i<t} S'_i \neq \varnothing$ (as $T[N']$ is not strongly connected), and q is not joined to any vertex of $\bigcup_{i<t} S'_i \cup N^-(z)$, we have $d^+(q) < d^+(z)$, a contradiction to the minimality of z. This completes the proof of Theorem 1.2, as we may take $d(k) = d(k, 2\binom{k}{2})$. $\square$

In fact, we need to embed a slightly more complicated structure in T. In particular, we shall need to attach a few special paths to our subdivided complete directed graph. Say a subdivision $\mathcal{S}$ is *minimal* in a tournament T if all of its paths have minimal length. This implies that every path in $\mathcal{S}$ is *backwards transitive*, i.e if $x_1 \ldots x_t$ is a path in $\mathcal{S}$ between branch vertices, then $x_i x_j \notin E(T)$ whenever $i \in [t-2]$ and $i + 1 < j$. Let $\mathcal{K}_r^{\min}$ denote a minimal subdivision of a $\overrightarrow{K}_r$. Since any subdivision of $\overrightarrow{K}_r$ contains a minimal subdivision, Theorem 1.2 allows us to find a $\mathcal{K}_r^{\min}$ in tournaments with sufficiently large out-degree. If U denotes the set of branch vertices of this subdivision, then for every $u, v \in U$, $\mathcal{K}_r^{\min}$ consists of directed paths P_{uv}, P_{vu} going from u to v and from v to u, respectively. Since T is a tournament and $\mathcal{K}_r^{\min}$ is minimal, precisely one of these paths is a directed edge.

Now we define our augmented subdivision, denoted by $\mathcal{K}_r^*$, as follows. Let $\mathcal{K}$ denote a copy of $\mathcal{K}_r^{\min}$ in T. The *branch vertices* of $\mathcal{K}_r^*$ are precisely the branch vertices of $\mathcal{K}$; denote this set by U. We form $\mathcal{K}_r^*$ by adding a collection $\mathcal{L}$ of special 'loop' paths in the following manner. For each pair $u, v \in U$, if, say, P_{uv} is the path between u and v in $\mathcal{K}$ of length at least two, then each of u and v has an associated directed path from $\mathcal{L}$: one directed path L_{uv}^u going from the second vertex of P_{uv} to u, and another directed path L_{uv}^v going from v to the penultimate vertex of P_{uv}; we require that these paths are internally disjoint from $V(\mathcal{K})$. We also impose that the paths in $\mathcal{L}$ are minimal and hence backwards transitive. For $u \in U$, we let $\mathcal{L}_u$ denote the collection of paths in $\mathcal{L}$ which contain u. Note that $\mathcal{K}_r^*$ and $\mathcal{K}_r^{\min}$ really denote *families* of subdigraphs which depend on the underlying tournament T. When we speak of 'a $\mathcal{K}_r^*$' we really mean 'a member of $\mathcal{K}_r^*$ in T'; we hope this usage of notation does not cause confusion, but we think that it is simpler. Now the proof of the existence of a $\mathcal{K}_r^*$ follows exactly in the same way as the proof of Theorem 1.2, namely by induction on the number of 'loops'. We state it as a corollary and provide only a sketch of the proof.

Corollary 3.3. *For any positive integer k there exists a $d^*(k)$ such that the following holds. If T is a tournament with $\delta^+(T) \geq d^*(k)$, then T contains a $\mathcal{K}_k^*$.*

(Sketch). Similarly as in Theorem 1.2, for a positive integer k and non-negative integer $m \leq 2\binom{k}{2}$, an *m-partial* $\mathcal{K}_k^*$ is any minimal subdivision of $\overrightarrow{K}_k$ with precisely m loop paths present. Let $d^*(k, m)$ denote the smallest positive integer such that any tournament with $\delta^+(T) \geq d^*(k,m)$ contains a subdivision of an m-partial $\mathcal{K}_k^*$. We show, as before, that if $m < 2\binom{k}{2}$, then $d^*(k,m+1) \leq 7d^*(k,m)^2$. For $k=1$ there is nothing to show and we can take $d^*(1,0)=1$. So assume $k \geq 2$ is given. Then $d^*(2,0)$ exists by Theorem 1.2 (i.e., we can embed a subdivision of $\overrightarrow{K}_2$ which contains a minimal such subdivision). Thus let $m \geq 1$ and suppose we can embed an m-partial $\mathcal{K}_k^*$ in any tournament with minimum out-degree at least $d^*(k,m)$. Let T be a tournament with $\delta^+(T) \geq 7d^*(k,m)^2$. Then the same proof used to show Theorem 1.2 gives that we may attach one more loop path, which we may assume has minimal length. Therefore we can embed an $(m+1)$-partial $\mathcal{K}_k^*$ in T, as claimed. $\square$

4. Proof of the main theorem

In this section, we finish the proof of Theorem 1.1. The structure of the proof is as follows. First, assuming the minimum degree of our tournament is sufficiently large, we shall embed in T a copy $\mathcal{S}$ of $\mathcal{K}_r^*$ where $r=r(k)$ is sufficiently large. If $x_1,\dots,x_k$, $y_1,\dots,y_k$ are the vertices we want to link, then we shall show that there exists a collection of k directed paths going from the x_i's to the branch vertices of $\mathcal{S}$, and a collection of k directed paths going from the branch vertices of $\mathcal{S}$ to the y_i's, all of these paths being pairwise vertex disjoint. Here we only use the assumption that T is $4k$-connected (see Lemma 4.1 below). Finally, we show that, provided one chooses these paths appropriately, one can link each x_i to y_i by rerouting the paths through $\mathcal{S}$. The rerouting step is slightly more complicated than one might expect, and we shall see that we do need the richer structure $\mathcal{K}_r^*$ rather than just a subdivided complete directed graph.

We need a small bit of terminology first before proceeding. If X and Y are two disjoint sets of vertices in a directed graph, then we say that there is an *out-matching* (resp., *in-matching*) of X to Y if there is a matching from X into Y such that all matching edges are directed from X to Y (resp., directed from Y to X).

Lemma 4.1. *Let T be a $4k$-connected tournament. Suppose $A, B \subset V(T)$ are two disjoint subsets of size k, and let $L \subset V(T)$ be a set of $4k$ vertices disjoint from $A \cup B$. Then there are k directed paths from A to L, and k directed paths from L to B, all these paths pairwise vertex disjoint and internally disjoint from L.*

Proof. Choose two disjoint subsets W_A, W_B disjoint from $A \cup B \cup L$ with maximum size subject to the following properties:

- Every vertex in W_A has at least $2k$ out-neighbours in L, and every vertex in W_B has at least $2k$ in-neighbours in L.
- There is an in-matching $\mathcal{M}_A$ from W_A to A, and an out-matching $\mathcal{M}_B$ from W_B to B.

We shall assume, without loss of generality, that $|W_A| \leq |W_B|$. Let A' denote the set of $|W_A|$ vertices in A that are incident with an edge of $\mathcal{M}_A$, and let $A'' = A \setminus A'$. Let B', B'' denote the analogous sets of vertices in B. As T is $4k$-connected, we can find pairwise vertex disjoint directed paths from some $k - |W_B|$ vertices of L to B'' avoiding $A \cup W_A \cup B' \cup W_B$. Choose a collection of such paths $\mathcal{P}$ which minimizes $|\bigcup \mathcal{P}|$, and subject to that, maximizes the number of paths whose second vertex has at least $2k$ in-neighbours in L. Partition $\mathcal{P}$ into sets $\mathcal{P}', \mathcal{P}''$ where the former denotes the collection of paths in $\mathcal{P}$ whose second vertex has at least $2k$ in-neighbours in L, and the latter denotes the collection of remaining paths. Denote by X' the set of all second and third vertices on paths in $\mathcal{P}'$, and denote by X'' the set of all first and second vertices on paths in $\mathcal{P}''$. Consider the set $Y := A' \cup W_A \cup X' \cup X'' \cup B \cup W_B$ and note that we can bound the size of Y as

$$|Y| \leq 2|W_A| + 3(k - |W_B|) + 2|W_B|.$$

We now find $k - |W_A|$ disjoint directed paths from the vertices in A'' to some subset of L, avoiding Y. This is possible since T is $4k$-connected and

$$4k - |Y| \geq 4k - (2|W_A| + 3(k - |W_B|) + 2|W_B|)$$

$$= k - 2|W_A| + |W_B| \geq k - |W_A|,$$

where the last inequality holds since we are assuming that $|W_A| \leq |W_B|$. Therefore, choose a collection $\mathcal{Q}$ of pairwise disjoint directed paths from A'' to L avoiding Y with $|\bigcup \mathcal{Q}|$ as small as possible. We claim that these new paths do not intersect any path from $\mathcal{P}$:

CLAIM 4.2. *No path from $\mathcal{Q}$ intersects a path from $\mathcal{P}$.*

Suppose that some path $Q \in \mathcal{Q}$ intersects a path $P \in \mathcal{P}$. Let $P = x_1 \dots x_s$ and $Q = y_1 \dots y_t$, and let $L_A = (\bigcup \mathcal{Q}) \cap L$ and similarly $L_B = (\bigcup \mathcal{P}) \cap L$. We consider two cases, according to whether $P \in \mathcal{P}'$ or $P \in \mathcal{P}''$. Suppose first the former holds, and let y_i ($i \geq 2$) be the first vertex of Q that intersects P. We may assume that $y_i \neq x_1$; indeed, if $y_i = x_1$, then $|L_A \cup L_B| \leq 2k - 1$, and since $P \in \mathcal{P}'$, we have that x_2 has at least $2k$ in-neighbours in L. Therefore, we may choose some in-neighbour x' disjoint from $L_A \cup L_B$ and replace P with $P' := x'x_2 \dots x_s$. Moreover, since the paths in $\mathcal{Q}$ avoid $\{x_2, x_3\}$ we may assume that $y_i = x_4$. Consider y_{i-1} and pick any vertex $z \in L \setminus (L_A \cup L_B)$. If $y_{i-1}z \in E(T)$, then we may replace Q with the shorter directed path $y_1 \dots y_{i-1}z$, contradicting the minimality of $|\bigcup \mathcal{Q}|$. So we have $zy_{i-1} \in E(T)$. But then as long as $i \geq 3$ we may replace P with the shorter path $zy_{i-1}x_4 \dots x_s$, contradicting the initial minimal choice of $|\bigcup \mathcal{P}|$. It remains to consider when $i = 2$. In this case, $zy_2 \notin E(T)$ for every $z \in L \setminus (L_A \cup L_B)$, since otherwise we can replace P with a shorter directed path. Thus y_2 has at least $2k$ out-neighbours in L, and we can add y_1y_2 to the matching $\mathcal{M}_A$, a contradiction to the maximality of this matching. It follows that $P \cap Q = \varnothing$ for $P \in \mathcal{P}'$.

So let us assume that $P \in \mathcal{P}''$. Since the paths in $\mathcal{Q}$ avoid $\{x_1, x_2\}$, we may assume in this case that $y_i = x_3$. The same argument as in the previous paragraph shows that we may assume $i \geq 3$ (otherwise, we obtain a larger matching than $\mathcal{M}_A$). Also, as before, if $z \in L \setminus (L_A \cup L_B)$, then $y_{i-1}z \notin E(T)$; otherwise we can replace Q with the shorter path $y_1 \dots y_{i-1}z$. Hence y_{i-1} has at least $|L| - |L_A \cup L_B| \geq 2k$ in-neighbours in L. Choose one of these in-neighbours

u (disjoint from $L_A \cup L_B$) and consider the path $P^* := uy_{i-1}x_3 \dots x_s$. Then P^* has the same length as P and its second vertex has at least $2k$ in-neighbours in L, so we could replace P with P^*, contradicting the maximality of $\mathcal{P}'$. Therefore, we must have $P \cap Q = \varnothing$, and the proof of Claim 4.2 is complete.

Armed with Claim 4.2, the proof of Lemma 4.1 is essentially complete. Indeed, every vertex in W_A has at least $2k$ out-neighbours in L, and so each of these vertices has at least

$$2k - |L_A \cup L_B| = |W_A| + |W_B|,$$

out-neighbours in $L \setminus (L_A \cup L_B)$. So for each vertex in W_A we may select a distinct out-neighbour in $L \setminus (L_A \cup L_B)$. Then every vertex in W_B has at least $|W_B|$ in-neighbours from the remaining vertices of L, so we can pick a distinct in-neighbour for every vertex of W_B. The paths of length 2 using vertices of $W_A \cup W_B$ together with $\mathcal{P}$ and $\mathcal{Q}$ form the required collection of paths. $\square$

We can now finish the proof of Theorem 1.1.

Proof of Theorem 1.1. Let $k \geq 2$ be an integer and let $f(k) := d^*(12k^2) + 2k$, where $d^* : \mathbb{N} \to \mathbb{N}$ is the function provided by Corollary 3.3. Suppose that T is a $4k$-connected tournament with minimum out-degree at least $f(k)$, and let $X = \{x_1, \dots, x_k\}$, $Y = \{y_1, \dots, y_k\}$ be two disjoint k-sets of vertices. We wish to find pairwise vertex disjoint directed paths going from x_i to y_i for each $i \in [k]$. Remove $X \cup Y$ from T; the tournament induced on $V(T) \setminus (X \cup Y)$ has minimum out-degree at least $d^*(12k^2)$, so by Corollary 3.3 we may embed in T a $\mathcal{K}^*_{12k^2}$ disjoint from $X \cup Y$. Denote this subdivision by $\mathcal{S}$. We shall use the same notation as in Section 3, namely, U denotes the branch vertices of $\mathcal{S}$, $\mathcal{K}$ denotes the underlying minimal subdivision of $\overrightarrow{K}_{12k^2}$ composed of minimal paths P_{uv}, P_{vu} for every pair of branch vertices $u, v \in U$, and $\mathcal{L}$ denotes the collection of minimal paths attached to $\mathcal{K}$. We call a *path of $\mathcal{S}$* any path P_{uv} between branch vertices of length at least 2, and any member of $\mathcal{L}$. We consider the following edges to belong to the structure $\mathcal{S}$:

- The edges belonging to paths in $\mathcal{K}$, except the paths of length one.
- The edges belonging to paths in $\mathcal{L}$.
- For every pair $u, v \in U$, every edge in T between $\{u, v\}$ and $V(P_{uv}) \cup V(P_{vu})$.
- For every $u \in U$, every edge in T between u and $\bigcup \mathcal{L}_u$.

We denote the set of edges of $\mathcal{S}$ by $E(\mathcal{S})$. For example, whenever we speak of distances in $\mathcal{S}$, we insist that they are computed using only these directed edges. Let $\mathcal{P}$ and $\mathcal{Q}$ be any two collections of pairwise disjoint directed paths such that every path in $\mathcal{P}$ goes from U to Y, every path in $\mathcal{Q}$ goes from X to U, and all of these paths are internally vertex disjoint from U; by Lemma 4.1, such collections exist. We say that a pair $(u, x) \in U \times V(\mathcal{S})$ is at *in-distance d in $\mathcal{S}$* if d is the smallest integer such that there is a directed path P' of length d using only edges of $\mathcal{S}$, and such that P' goes from u to x. We shall also sometimes say that x has in-distance d in $\mathcal{S}$ from u. Similarly, we say that $(u, x) \in U \times V(\mathcal{S})$ is at *out-distance d in $\mathcal{S}$* if d is the smallest integer such that there is a directed path Q' of length d using only edges of $\mathcal{S}$, and such that Q' goes from x to u in $\mathcal{S}$; we shall also sometimes say that x has out-distance d in $\mathcal{S}$ from u. We denote in-distance by $d^{\text{in}}(u, x)$ and out-distance by $d^{\text{out}}(u, x)$ (where we have suppressed the dependence on $\mathcal{S}$).

Observation 4.3. Let $x \in V(\mathcal{S}) \setminus U$. Then x is at in-distance (or out-distance) at least 3 from every vertex of U, except possibly the branch vertex (or vertices) belonging to the path of $\mathcal{S}$ containing x.

Proof. If $x \in V(\mathcal{S}) \setminus U$, then either $x \in P_{uv}$ for some $u, v \in U$ or $x \in L^u_{uv} \in \mathcal{L}_u$ (or possibly both). Let $w \in U \setminus \{u, v\}$. In order to get from w to x using only edges of $\mathcal{S}$, we must first reach either u or v. However, recall that the single edge paths in $\mathcal{K}$ are *not* edges of $\mathcal{S}$, so the path from w to u or v in $\mathcal{S}$ has length at least 2. Therefore, x has in-distance at least 3 from w, as required. A symmetric argument shows that the observation remains true with 'out-distance' instead of 'in-distance'. $\square$

In the following, we shall always assume that any family $\mathcal{F}$ of directed paths in T between $X \cup Y$ and U are internally disjoint from U. We also denote by $U_{\mathcal{F}}$ the set $U \cap (\bigcup \mathcal{F})$. Our first claim asserts that we may assume the paths in one of the collections $\mathcal{P}$, $\mathcal{Q}$ contains few vertices which are 'close' in $\mathcal{S}$ to a vertex in U.

LEMMA 4.4. *We may choose either $\mathcal{P}$ or $\mathcal{Q}$ such that there are at most $8k^2+4k$ vertices $u \in U \setminus U_{\mathcal{P}}$ (resp., $U \setminus U_{\mathcal{Q}}$) with $d^{in}(u,x) \leq 2$ (resp., $d^{out}(u,x) \leq 2$) for some $x \in \bigcup \mathcal{P} \setminus U_{\mathcal{P}}$ (resp., for some $x \in \bigcup \mathcal{Q} \setminus U_{\mathcal{Q}}$).*

PROOF. Apply Lemma 4.1 with $A = X$, $B = Y$, and $L = U$. Using the proof and notation of Lemma 4.1, assume that $|W_X| \leq |W_Y|$. Then recall that we may choose the paths from U to Y first minimally (with respect to the number of vertices used) upon the removal of $W_X \cup W_Y$, a set of at most $2k$ vertices. Recall also that each such path which uses a vertex of $W_X \cup W_Y$ has length two. Suppose there is a set $U' \subset U \setminus U_{\mathcal{P}}$ of more than $8k^2 + 4k$ vertices such that for every $u \in U'$ there is $x \in \bigcup \mathcal{P} \setminus U_{\mathcal{P}}$ with $d^{\text{in}}(u,x) \leq 2$. We claim that this contradicts minimality. Indeed, by pigeonhole there is a set $U'_0 \subset U'$ of size more than $8k+4$, and a path $P \in \mathcal{P}$ such that for each $u \in U'_0$ there is some $x \in P$ with $d^{\text{in}}(u,x) \leq 2$. From Observation 4.3, it follows that for each interior vertex v of P there are at most two vertices of U'_0 that are at in-distance 2 from v. Therefore P must have more than two edges so does not intersect $W_X \cup W_Y$. For each vertex $u \in U'_0$, pick some vertex $v_u \in P$ at in-distance exactly 2 from u, and denote by D the set containing all such vertices v_u. Note that P contains at most one vertex at in-distance 1 from a vertex in $U \setminus U_{\mathcal{P}}$, as otherwise we may reroute P and obtain a shorter path avoiding $W_X \cup W_Y$. Using Observation 4.3 again, there is a set D' of at least $\frac{1}{2}(8k+4) = 4k+2$ vertices in D corresponding to distinct vertices of U'_0. Let $P = p_0 \dots p_\ell$, where $p_0 \in U$ and $p_\ell \in X$, $F := D' \setminus \{p_1, p_2\}$. For each $p_j \in F$, we may choose vertex disjoint directed paths $u_j m_j p_j$ of length 2 in $\mathcal{S}$, where $u_j \in U'_0$. Accordingly, there are at least $4k$ 'middle vertices' m_j, at least $2k$ of which are disjoint from $W_X \cup W_Y$; let M denote the set of middle vertices disjoint from $W_X \cup W_Y$.

Now, suppose some $m_j \in M$ does not intersect any path in $\mathcal{P}$. Then we may replace P with the path $u_j m_j p_j P$, which is shorter and still avoids $W_X \cup W_Y$, a contradiction. Thus, each middle vertex in M belongs to some member of $\mathcal{P}$ and so by pigeonhole there is a path P' which contains at least two vertices of M. But both of these vertices are at in-distance 1 from a vertex in $U \setminus U_{\mathcal{P}}$, which, as noted before, is a contradiction. Hence at most $8k^2 + 4k$ vertices in $U \setminus U_{\mathcal{P}}$ have the stated property, as claimed. A symmetric argument shows that we may choose $\mathcal{Q}$ with the stated property in the event that $|W_Y| \leq |W_X|$. This completes the proof of the lemma. $\square$

Suppose $\mathcal{F}$ is a collection of pairwise disjoint directed paths from U to Y (internally disjoint from U), and let $P = p_0 \dots p_t$ be any path in $\mathcal{F}$. We call the pairs (p_0, p_1) and (p_0, p_2) *trivial* if they have in-distance at most 2 in $\mathcal{S}$; any other pair with in-distance at most 2 is *nontrivial*. For a subset $U' \subseteq U$ we shall say that $\mathcal{F}$ is *U'-good* if no nontrivial pair of vertices from $U' \times (\bigcup \mathcal{F} \setminus U_{\mathcal{F}})$ is at in-distance at most 2 in $\mathcal{S}$. In particular, each path $P \in \mathcal{F}$ intersects U' in at most one vertex, namely its initial vertex. Suppose that $\mathcal{F}$ satisfies the property stated in Lemma 4.4. Then we have the following:

Claim 4.5. *There exists a subset $U' \subset U \setminus U_{\mathcal{F}}$ of size at least $2k$ such that $\mathcal{F}$ is U'-good.*

This follows immediately from the previous lemma. Indeed, remove from U every vertex in $U_{\mathcal{F}}$ and every vertex in $U \setminus U_{\mathcal{F}}$ at in-distance at most 2 in $\mathcal{S}$ from some vertex of $\bigcup \mathcal{F} \setminus U_{\mathcal{F}}$; let U' denote the remaining set of vertices. By Lemma 4.4, we have removed at most $8k^2 + 5k$ vertices. As $|U| = 12k^2$ we have $|U'| \geq 12k^2 - (8k^2 + 5k) \geq 2k$, since $k \geq 2$. Clearly $\mathcal{F}$ is U'-good.

We shall assume without loss of generality that we may choose the paths from U to Y with the property stated in Lemma 4.4. So the previous two claims show that we may find collections of vertex disjoint directed paths $\mathcal{P}, \mathcal{Q}$ which are internally disjoint from U and such that the paths in $\mathcal{P}$ go from U to Y, the paths in $\mathcal{Q}$ go from X to U, and $\mathcal{P}$ is U'-good for some $U' \subset U \setminus U_{\mathcal{P}}$ with

$|U'| \geq 2k$. Conditioned on this, we assume that $\mathcal{P} \cup \mathcal{Q}$ minimizes the number of edges outside of $\mathcal{S}$, and again conditioned on this, we take such a pair with $|\bigcup \mathcal{P}| + |\bigcup \mathcal{Q}|$ as small as possible. Let $U'' = U' \setminus U_{\mathcal{Q}}$ so that $|U''| \geq k$ and it is disjoint from $U_{\mathcal{P}} \cup U_{\mathcal{Q}}$; we may assume that $U'' = \{u_1, \ldots, u_k\}$ has precisely k elements. We now show that one can reroute the paths in $\mathcal{P} \cup \mathcal{Q}$ through U'' in order to create the desired paths linking x_i to y_i for each $i \in [k]$. Let $U_{\mathcal{P}} = \{z_1, \ldots, z_k\}$ and $U_{\mathcal{Q}} = \{w_1, \ldots, w_k\}$ so that z_i is the initial vertex in U of the path $P_i \in \mathcal{P}$ with terminal vertex $y_i \in Y$, and w_i is the terminal vertex in U of the path $Q_i \in \mathcal{Q}$ with initial vertex $x_i \in X$. Recall that for every pair of branch vertices $u, v \in U$, P_{uv} and P_{vu} denotes the path in $\mathcal{K}$ from u to v, and from v to u, respectively. The following sequence of claims show that we can control intersections of paths in $\mathcal{P} \cup \mathcal{Q}$ with appropriate paths in $\mathcal{S}$ in order to link each x_i to y_i.

CLAIM 4.6. *Suppose some path $Q \in \mathcal{Q}$ intersects $L^{u_i}_{w_i u_i} \in \mathcal{L}_{u_i}$, for some $i \in [k]$. Let z be the first vertex of $L^{u_i}_{w_i u_i}$ in the intersection. Then one of the following holds: z is the terminal vertex of $L^{u_i}_{w_i u_i}$ and $z \in Q_i$, or z is the second vertex of $L^{u_i}_{w_i u_i}$.*

Suppose z is not the second vertex of $L^{u_i}_{w_i u_i}$. If z is an interior point of $L^{u_i}_{w_i u_i}$, then $z u_i \in E(T)$ by minimality of the path $L^{u_i}_{w_i u_i}$. Note that if Q has an edge which is not in $E(\mathcal{S})$ after z then we have a contradiction: indeed replacing Q with $Q z u_i$ yields a collection of paths with fewer edges outside of $E(\mathcal{S})$. Otherwise, $Q = Q_i$ and it must use at least 2 edges after z, so we obtain a contradiction to the minimality of $|\bigcup \mathcal{P}| + |\bigcup \mathcal{Q}|$ by rerouting the path as before. Therefore, z must be the terminal vertex of $L^{u_i}_{w_i u_i}$. Finally, z must belong to Q_i, otherwise we may similarly reroute Q through u_i, decreasing the number of edges used outside $E(\mathcal{S})$.

CLAIM 4.7. *No path in $\mathcal{P}$ intersects $P_{w_i u_i}$. Moreover, if q_i denotes the last vertex in $P_{w_i u_i}$ which occurs as the intersection of some path in $\mathcal{Q}$, then $q_i \in Q_i$.*

No path in $\mathcal{P}$ intersects $\{u_i, w_i\}$, so it suffices to show that no such path intersects the interior of $P_{w_iu_i}$. Therefore, we may assume that $P_{w_iu_i}$ has length at least 2. Suppose first that some $P \in \mathcal{P}$ contains a vertex v in the interior. Note that v must be the penultimate vertex of $P_{w_iu_i}$. Otherwise, $u_iv \in E(T) \cap E(\mathcal{S})$ by the minimality of the subdivision $\mathcal{K}$, and this contradicts the fact that $\mathcal{P}$ is U'-good. Consider the loop path $L = L^{u_i}_{w_iu_i} \in \mathcal{L}_{u_i}$ at u_i ending at v, and recall that the edges of L are edges of $\mathcal{S}$. Let z be the first vertex in $L^{u_i}_{w_iu_i}$ belonging to some path $P' \in \mathcal{P}$: such a vertex and path exist since we may take $z = v$ and $P' = P$. Let L' be the initial segment of the path $L^{u_i}_{w_iu_i}$ ending at z.

Suppose first that no path in $Q \in \mathcal{Q}$ intersects L', and replace P with $P'' = u_iL'zP'$. Since P' cannot intersect u_i or w_i it must have an edge which is not in $E(\mathcal{S})$ before z. It follows that P'' has fewer edges outside of $\mathcal{S}$. This is a contradiction to our choice of $\mathcal{P} \cup \mathcal{Q}$, provided $\mathcal{P}'' := (\mathcal{P} \setminus \{P'\}) \cup \{P''\}$ is U'-good. To see this, observe that any vertex of $L \setminus \{v\}$ is at in-distance at least 3 from w_i. Moreover, if $w_i \in U'$, and $z = v$ (and hence $P' = P$), then z is also at in-distance at least 3 from w_i. Accordingly, if $w_i \in U'$, then every vertex of P'' is still at in-distance at least 3 from w_i. By the minimality of L, every vertex in the interior of L (except the second) is directed towards u_i; thus, the only vertices at in-distance at most 2 from u_i are the second and third vertices of L, say x and y, respectively. But the pairs (u_i, x) and (u_i, y) are trivial pairs, and thus do not contradict U'-goodness. Lastly, by Observation 4.3 every vertex of P'' (except possibly u_i) is at in-distance at least 3 from every vertex of $U' \setminus \{u_i, w_i\}$. It follows that $\mathcal{P}''$ is U'-good, which is a contradiction to our choice of $\mathcal{P} \cup \mathcal{Q}$.

On the other hand, if some path $Q' \in \mathcal{Q}$ intersects L' in some vertex r, then by Claim 4.6 r must the second vertex of $L^{u_i}_{w_iu_i}$. Note that by U'-goodness, no path in $\mathcal{P}$ contains the third vertex r_1 of $L^{u_i}_{w_iu_i}$, hence we can replace Q' by $Q'rr_1u_i$ thus decreasing the number of edges outside $E(\mathcal{S})$. Therefore we conclude that no path in $\mathcal{P}$ can intersect $P_{w_iu_i}$. Let us now show the second part of the claim. Suppose that $q_i \in Q_j$ for some $j \neq i$. Since Q_j must avoid $\{u_i, w_i\}$ it contains an edge which is not in $E(\mathcal{S})$ after q_i. Replace Q_j with $Q' = Q_j v P_{w_i u_i}$. Then by the previous paragraph, no path in $\mathcal{P}$ intersects Q' and the resulting collection of paths has fewer edges outside of $\mathcal{S}$, a contradiction. This completes the proof of the claim.

It remains to establish the analogous claims for the path $P_{u_i z_i}$, namely that intersections of paths in $\mathcal{P} \cup \mathcal{Q}$ with $P_{u_i z_i}$ and $L^{u_i}_{u_i z_i}$ behave as one expects. The arguments are similar to those in the previous two claims. Theorem 1.1 will then be an immediate consequence.

Claim 4.8. *For every $i \in [k]$, no path in $\mathcal{P}$ intersects $L^{u_i}_{u_i z_i} \in \mathcal{L}_{u_i}$.*

Suppose some $P \in \mathcal{P}$ intersects $L^{u_i}_{u_i z_i}$ in a vertex z. Then z cannot be the first vertex of $L^{u_i}_{u_i z_i}$, as this would contradict the fact that $\mathcal{P}$ is U'-good. Therefore, if z' denotes the vertex preceding z in $L^{u_i}_{u_i z_i}$, then by the minimality of paths in $\mathcal{L}$, we have $u_i z' \in E(T) \cap E(\mathcal{S})$. But then z is at in-distance 2 from u_i, contradicting U'-goodness.

Claim 4.9. *Let p_i denote the first vertex in $P_{u_i z_i}$ which occurs as the intersection of some path in $\mathcal{P}$. Then no path in $\mathcal{Q}$ intersects $P_{u_i z_i}$ and $p_i \in P_i$.*

As before, it suffices to show that no path in $\mathcal{Q}$ intersects the interior of $P_{u_i z_i}$, so we may assume that $P_{u_i z_i}$ has length at least 2. Suppose some $Q \in \mathcal{Q}$ intersects the interior of $P_{u_i z_i}$ at v. Note that since Q does not meet $\{u_i, z_i\}$, it must leave $\mathcal{S}$ at some time after v. If v is not the second vertex of $P_{u_i z_i}$, then $v u_i \in E(T) \cap E(\mathcal{S})$, and so we may replace Q with $Q v u_i$. This path has fewer edges outside of $\mathcal{S}$ than Q, and this contradicts our minimal choice of $\mathcal{P} \cup \mathcal{Q}$. If v is the second vertex, then let $L = L^{u_i}_{u_i z_i} \in \mathcal{L}_{u_i}$ be the loop path at u_i directed from v to u_i. Let z be the last vertex of L which occurs as the intersection of some path $Q' \in \mathcal{Q}$ (z and Q' exist since we may take $z = v$ and $Q' = Q$), and let L' be the subpath of L from z to u_i. By Claim 4.8, no path in $\mathcal{P}$ intersects L', so replace Q' with $Q' z L' u_i$. Again, the edges of L' are in $E(\mathcal{S})$ so this path has fewer edges outside $\mathcal{S}$ than Q', a contradiction. It follows that no path in $\mathcal{Q}$ intersects $P_{u_iz_i}$ as claimed. For the second part of the claim, suppose that $p_i \in P_j$ for some $j \neq i$. Then P_j avoids $\{u_i, z_i\}$ and therefore leaves $\mathcal{S}$ at some time before p_i. Now, no path in $\mathcal{P} \cup \mathcal{Q}$ intersects the interior of the subpath $P_{u_iz_i}p_i$ so replace P_j with $P' = P_{u_iz_i}p_iP_j$. This path has fewer edges outside of $\mathcal{S}$. We claim that $\mathcal{P}' = (\mathcal{P} \setminus \{P_j\}) \cup \{P'\}$ is U'-good. Indeed, note that since $\mathcal{P}$ is U'-good, the subpath $P_{u_iz_i}p_i$ has length at least 3. Also, for every $v \in P_{u_iz_i}$ we have that $vu_i \in E(T)$ by the minimality of $\mathcal{K}$. So the only pairs at in-distance at most 2 in $U' \times (\bigcup \mathcal{P}' \setminus U_{\mathcal{P}'})$ are the trivial pairs (u_i, x) and (u_i, y), where x, y are the second and third vertices, respectively, of $P_{u_iz_i}$. But these pairs, by definition, do not contradict U'-goodness. It follows that $j = i$, and the claim is proved.

By Claims 4.7 and 4.9, the directed paths $Q_iq_iP_{w_iu_i}u_iP_{u_iw_i}p_iP_i$, for each $i \in [k]$, are pairwise vertex disjoint and link x_i to y_i. This completes the proof of Theorem 1.1. $\square$

5. Concluding remarks

The most obvious open problem is to reduce our bound of $4k$ on the connectivity in Theorem 1.1. We remark that an improvement on the connectivity bound in Lemma 4.1 translates directly into a better bound in Theorem 1.1. Unfortunately, we could not go beyond $4k$. Furthermore, our Lemma 4.1 does *not* hold if we replace $4k$ with anything smaller than $3k$. The following construction, of a $(3k-1)$-connected tournament T where Lemma 4.1 fails, was communicated to us by Kamil Popielarz. Suppose $V(T) = [n]$ and partition $V(T)$ into disjoint sets A, S, B, L, where $L = V(T) \setminus (A \cup S \cup B)$, and $|A| = |B| = k$, $|S| = 2k-1$. Direct the edges from L to A; from B to L; from A to S and from S to B; and from A to B. Inside L we place a balanced blow-up of a directed triangle. That is, equitably partition L into sets L_1, L_2, L_3 with directed edges $L_1 \to L_2$, $L_2 \to L_3$, $L_3 \to L_1$, and inside each of the L_i's we orient the edges arbitrarily. Now, join every vertex in S to all of L_1 and join every vertex of L_2 to all of S. Finally, orient the edges between S and L_3, and the edges inside A, B, and S, arbitrarily.

Provided n is sufficiently large (depending on k), it is not hard to show that T is $(3k-1)$-connected. Observe that we cannot get from A to L (disjointly from B) without using vertices of S. Similarly, we cannot get from L to B (disjointly from A) without using vertices of S. As $|S| = 2k-1$, any path system as in Lemma 4.1 will not be pairwise disjoint. Accordingly, Lemma 4.1 fails for this tournament. We remark that a slight modification of this construction yields a tournament which additionally has large minimum in and out-degree.

Aside from improving our bound of $4k$ on the connectivity and resolving completely Pokrovskiy's conjecture, there are a few other open problems of interest. For example, what is the smallest function $d(k)$ such that Theorem 1.2 holds?

PROBLEM 5.1. *Determine the smallest function $d : \mathbb{N} \to \mathbb{N}$ such that any tournament T with $\delta^+(T) \geq d(k)$ contains a subdivision of the complete directed graph $\overrightarrow{K}_k$.*

Note that our proof gives a doubly exponential bound on $d(k)$. Indeed, it is easy to check that $d(k) \leq 2^{2^{Ck^2}}$. Finally, while the conclusion of Theorem 1.2 does not hold if we replace T with a general digraph, can we embed subdivisions of *acyclic* digraphs in digraphs of large minimum out-degree? We end by recalling the following beautiful conjecture of Mader [81] from 1985.

CONJECTURE 5.2. *For every positive integer k, there exists a function $f(k)$ such that every digraph with minimum out-degree at least $f(k)$ contains a subdivision of the transitive tournament of order k.*

Of course, since every acyclic digraph is contained in the transitive tournament of the same order, this conjecture (if true) would give an affirmative answer to the preceding question.

Bibliography

1. *GLPK (GNU Linear Programming Kit)*, https://www.gnu.org/software/glpk/.
2. Ron Aharoni, *Ryser's conjecture for tripartite 3-graphs*, Combinatorica **21** (2001), no. 1, 1–4.
3. Ron Aharoni and Penny Haxell, *Hall's theorem for hypergraphs*, J. Graph Theory **35** (2000), no. 2, 83–88.
4. Michael O. Albertson, *You can't paint yourself into a corner*, J. Combin. Theory Ser. B **73** (1998), no. 2, 189–194.
5. Michael O. Albertson, Alexandr V. Kostochka, and Douglas B. West, *Precoloring extensions of Brooks' theorem*, SIAM J. Discrete Math. **18** (2004/05), no. 3, 542–553.
6. Michael O. Albertson and Emily H. Moore, *Extending graph colorings*, J. Combin. Theory Ser. B **77** (1999), no. 1, 83–95.
7. ———, *Extending graph colorings using no extra colors*, Discrete Math. **234** (2001), no. 1-3, 125–132.
8. Noga Alon, *Splitting digraphs*, Combin. Probab. Comput. **15** (2006), no. 6, 933–937.
9. Noga Alon and Kenneth A. Berman, *Regular hypergraphs, Gordon's lemma, Steinitz' lemma and invariant theory*, J. Combin. Theory Ser. A **43** (1986), no. 1, 91–97.
10. Lars Døvling Andersen and Anthony J. W. Hilton, *Thank evans!*, Proc. London Math. Soc. **3** (1983), no. 47, 507–522.
11. Marcin Anholcer, Bartłomiej Bosek, and Jarosław Grytczuk, *Majority choosability of digraphs*, Electron. J. Combin. **24** (2017), no. 3, #P3.57.

12. Richard P. Anstee and Jerrold R. Griggs, *An application of matching theory of edge-colourings*, Discrete Math. **156** (1996), no. 1-3, 253–256.
13. Maria Axenovich, *A note on graph coloring extensions and list-colorings*, Electron. J. Combin. **10** (2003), 5 pp.
14. Deepak Bal and Louis DeBiasio, *Partitioning random graphs into monochromatic components*, Electron. J. Combin. **24** (2017), #P1.18.
15. Paul Balister, Béla Bollobás, Jenő Lehel, and Michał Morayne, *Repeated degrees in random uniform hypergraphs*, SIAM J. Discrete Math. **27** (2013), no. 1, 145–154.
16. Jørgen Bang-Jensen, *On the 2-linkage problem for semicomplete digraphs*, Ann. Discrete Math. **41** (1989), 23–38.
17. Michael D. Barrus, Michael Ferrara, Jennifer Vandenbussche, and Paul S. Wenger, *Colored saturation parameters for rainbow subgraphs*, J. Graph Theory **86** (2017), no. 4, 375–386.
18. Claude Berge and Jean-Claude Fournier, *A short proof for a generalization of Vizing's theorem*, J. Graph Theory **15** (1991), no. 3, 333–336.
19. Béla Bollobás, *Determination of extremal graphs by using weights*, Wiss. Z. Techn. Hochsch. Ilmenau **13** (1967), 419–421.
20. Béla Bollobás, *On a conjecture of Erdős, Hajnal and Moon*, Amer. Math. Monthly **74** (1967), 178–179.
21. ______, *Extremal graph theory*, London Mathematical Society Monographs, no. 11, Academic Press, Inc., London-New York, 1978.
22. Béla Bollobás, *Combinatorics: Set systems, hypergraphs, families of vectors and combinatorial probability*, Cambridge University Press, Cambridge, 1986.
23. Béla Bollobás, *Modern graph theory*, Graduate Texts in Mathematics, 184, Springer-Verlag, New-York, 1998.
24. Béla Bollobás and Alexander Scott, *Independent sets and repeated degrees*, Discrete Math. **170** (1997), no. 1, 41–49.

25. Béla Bollobás and Andrew Thomason, *Highly linked graphs*, Combinatorica **16** (1996), no. 3, 313–320.
26. Marthe Bonamy, *Planar graphs with* $\Delta \geq 8$ *are* $(\Delta+1)$*-edge-choosable*, SIAM J. Discrete Math. **29** (2015), no. 3, 1735–1763.
27. Adrian Bondy, *Beautiful conjectures in graph theory*, European J. Comb. **37** (2014), 4–23.
28. John A. Bondy and Bill Jackson, *Vertices of small degree in uniquely Hamiltonian graphs*, J. Combin. Theory Ser. B **74** (1998), no. 2, 265–275.
29. Oleg V. Borodin, *A generalization of Kotzig's theorem and prescribed edge coloring of planar graphs*, Math. Notes **48** (1990), no. 5-6, 1186–1190.
30. Oleg V. Borodin, Alexandr V. Kostochka, and Douglas R. Woodall, *List edge and list total colourings of multigraphs*, J. Combin. Theory Ser. B **71** (1997), no. 2, 184–204.
31. Raj C. Bose, *On the construction of balanced incomplete block designs*, Annals of Eugenics **9** (1939), 353–399.
32. Jean Bourgain and Lior Tzafriri, *Restricted invertibility of matrices and applications*, Analysis at Urbana, Vol. II (Urbana, IL, 1986–1987), London Math. Soc. Lecture Note Ser., vol. 138, Cambridge Univ. Press, Cambridge, 1989, pp. 61–107. MR 1009186
33. Roland L. Brooks, *On colouring the nodes of a network*, Math. Proc. of Camb. Phil. Soc. **37** (1941), 194–197.
34. Yair Caro, Josef Lauri, and Christina Zarb, *Equating two maximum degrees*, ArXiv e-prints (2017).
35. Yair Caro, Asaf Shapira, and Raphael Yuster, *Forcing k-repetitions in degree sequences*, Electron. J. Combin. **21** (2014), no. 1, #P1.24.
36. Yair Caro and Douglas B. West, *Repetition number of graphs*, Electron. J. Combin. **16** (2009), no. 1, #R7.
37. Yair Caro and Raphael Yuster, *Large induced subgraphs with equated maximum degree*, Discrete Math. **310** (2010), no. 4, 742–747.

38. Guantao Chen and Richard H. Schelp, *Ramsey problems with bounded degree spread*, Combin. Probab. Comput. **2** (1993), no. 3, 263–269.
39. David Conlon and Maya Stein, *Monochromatic cycle partitions in local edge colourings*, J. Graph Theory **81** (2016), no. 2, 134–145.
40. Endre Csóka, Gábor Lippner, and Oleg Pikhurko, *Kőnig's line coloring and Vizing's theorems for graphings*, Forum Math. Sigma **4** (2016), e27, 40 pp.
41. Reinhard Diestel, *Graph theory*, Graduate Texts in Mathematics, 173, Springer, Berlin, 2017.
42. Robert P. Dilworth, *A decomposition theorem for partially ordered sets*, Ann. of Math. **51** (1950), 161–166.
43. Gabriel A. Dirac, *Some theorems on abstract graphs*, Proc. London Math. Soc. **2** (1952), 69–81.
44. P. Duchet, *Représentations, noyaux en théorie des graphes at hypergraphes*, (1979).
45. Roger C. Entringer and Hendrika Swart, *Spanning cycles of nearly cubic graphs*, J. Combin. Theory Ser. B **29** (1980), no. 3, 303–309.
46. Paul Erdős, Guantao Chen, Cecil Rousseau, and Richard Schelp, *Ramsey problems involving degrees in edge coloured complete graphs of vertices belonging to monochromatic subgraphs*, European J. Combin. **14** (1993), no. 3, 183–189.
47. Paul Erdős, András Gyárfás, and László Pyber, *Vertex coverings by monochromatic cycles and trees*, J. Combin. Theory Ser. B **51** (1991), no. 1, 90–95.
48. Paul Erdős, András. Hajnal, and John W. Moon, *A problem in graph theory*, Amer. Math. Monthly **71** (1964), 1107–1110.
49. Jill R. Faudree, Ralph J. Faudree, and J. R. Schmitt, *A survey of minimum saturated graphs*, Electron. J. Combin. **18** (2011), 36, Dynamic Survey 19.
50. Herbert Fleischner, *Uniquely Hamiltonian graphs of minimum degree 4*, J. Graph Theory **75** (2014), no. 2, 167–177.

51. Fred Galvin, *The list chromatic index of a bipartite multigraph*, J. Combin. Theory Ser. B **63** (1995), no. 1, 153–158.
52. Mark K. Goldberg, *Multigraphs with a chromatic index that is nearly maximal (in Russian)*, Diskret. Analiz **23** (1973), 3–7, 72.
53. Ron J. Gould, *Advances on the Hamiltonian problem – a survey*, Graphs Comb. **19** (2003), 7–52.
54. Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer, *Ramsey Theory*, Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., NJ, 1990.
55. András Gyárfás, *Partition covers and blocking sets in hypergraphs (in Hungarian)*, MTA SZTAKI Tanulmanyok **71** (1977).
56. ______, *Large monochromatic components in edge colorings of graphs: a survey*, pp. 77–96, Birkhäuser Boston, Boston, MA, 2011.
57. András Gyárfás, Michael Ruszinkó, Gábor N. Sárközy, and Endre Szemerédi, *An improved bound for the monochromatic cycle partition number*, J. Combin. Theory Ser. B **96** (2006), no. 6, 855–873.
58. Roland Häggkvist, *A solution to the Evans conjecture for latin squares of large size*, Colloq. Math. Soc. Janos Bolyai **18** (1978), 405–513.
59. Roland Häggkvist and Jeannette Janssen, *New bounds on the list-chromatic index of the complete graph and other simple graphs*, Combin. Probab. Comput. **6** (1997), no. 3, 295–313.
60. Denis Hanson and Bjarne Toft, *Edge-colored saturated graphs*, J. Graph Theory **11** (1987), no. 2, 191–196.
61. Frédéric Havet and Bernard Lidický, *Splitting a tournament into two subtournaments with given minimum outdegree*, Preprint (2014).
62. Penny Haxell and Yoshiharu Kohayakawa, *Partitioning by monochromatic trees*, J. Combin. Theory Ser. B **68** (1996), no. 2, 218–222.
63. Penny Haxell, Ben Seamone, and Jacques Verstraëte, *Independent dominating sets and Hamiltonian cycles*, J. Graph Theory **54** (2007), no. 3, 233–244.

64. J. Henderson, Ph.D. thesis.

65. Bill Jackson and Robin W. Whitty, *A note concerning graphs with unique f-factors*, J. Graph Theory **13** (1989), no. 5, 577–580.

66. Hee S. Jung, *Verallgemeinerung des n-fachen zusammenhangs für graphen*, Math. Ann. **187** (1970), 95–103.

67. Martin Juvan, Bojan Mohar, and Riste Škrekovski, *On list edge-colorings of subcubic graphs*, Discrete Math. **187** (1998), no. 1-3, 137–149.

68. ______, *Graphs of degree 4 are 5-edge-choosable*, J. Graph Theory **32** (1999), no. 3, 250–264.

69. Jeff Kahn, *Asymptotics of the chromatic index for multigraphs*, J. Combin. Theory Ser. B **68** (1996), no. 2, 233–254.

70. ______, *Asymptotics of the list-chromatic index for multigraphs*, Random Structures Algorithms **17** (2000), no. 2, 117–156.

71. László Kászonyi and Zsolt Tuza, *Saturated graphs with minimal number of edges*, J. Graph Theory **10** (1986), no. 2, 203–210.

72. Fiachra Knox and Robert Šámal, *Linear Bound for Majority Colourings of Digraphs*, arXiv: 1701.05715 (2017).

73. Dániel Korándi, *Rainbow saturation and graph capacities*, SIAM J. Discrete Math. **30** (2018), 1261–1264.

74. Stephan Kreutzer, Sang-il. Oum, Paul Seymour, Dominic van der Zypen, and David R. Wood, *Majority Colourings of Digraphs*, Electron. J. Combin. **24** (2017), no. 2, #P2.25.

75. Daniela Kühn, John Lapinskas, Deryk Osthus, and Viresh Patel, *Proof of a conjecture of Thomassen on hamilton cycles in highly connected tournaments*, Proc. Lond. Math. Soc. **109** (2014), no. 3, 733–762.

76. David G. Larman and Peter Mani, *On the existence of certain configurations within graphs and the* 1*-skeletons of polytopes*, Proc. Lond. Math. Soc. **20** (1974), 144–160.

77. Shoham Letzter, *Monochromatic cycle partitions of 2-coloured graphs with minimum degree* $3n/4$, preprint, arXiv:1502.07736.

78. Nicolas Lichiardopol, *Vertex-disjoint subtournaments of prescribed minimum outdegree or minimum semidegree: proof for tournaments of a conjecture of Stiebitz*, Int. J. Comb. **2012** (2012), 9 pp.

79. Henry Liu, Noah Prince, and Robert Morris, *Highly connected monochromatic subgraphs of multicoloured graphs*, J. of Graph Theory **61** (2009), no. 1, 22–44.

80. Wolfgang Mader, *Homomorphieeigenschaften und mittlere kantendichte von graphen*, Math. Ann. **174** (1967), 265–268.

81. ______, *Degree and local connectivity in digraphs*, Combinatorica **5** (1985), no. 2, 161–165.

82. Karl Menger, *Zur allgemeinen kurventhoerie*, Fundam. Math. **10** (1927), 96–115.

83. Michael Molloy and Bruce Reed, *Graph colouring and the probabilistic method*, Algorithms and Combinatorics, vol. 23, Springer-Verlag, Berlin, 2002.

84. Oystein Ore, *Note on Hamilton circuits*, Amer. Math. Monthly **67** (1960), 55.

85. Alexey Pokrovskiy, *Partitioning edge-coloured complete graphs into monochromatic cycles and paths*, J. Combin. Theory Ser. B **106** (2014), 70–97.

86. ______, *Partitioning edge-coloured complete graphs into monochromatic cycles and paths*, J. Combin. Theory Ser. B **106** (2014), 70–97.

87. ______, *Highly linked tournaments*, J. Comb. Theory Ser. B **115** (2015), 339–347.

88. Frank Ramsey, *On a Problem of Formal Logic*, Proc. London Math. Soc. **30** (1930), no. 4, 264–286.

89. Daniel P. Sanders and Yue Zhao, *Planar graphs of maximum degree seven are class I*, J. Combin. Theory Ser. B **83** (2001), no. 2, 201–212.

90. Richard Schelp, *Some Ramsey-Turán type problems and related questions*, Discrete Math. **312** (2012), no. 14, 2158–2161.

91. S. V. Sevast'yanov, *On approximate solutions of scheduling problems*, Metody Diskret. Anal. (in Russian) **32** (1978), 66–75.
92. Paul D. Seymour, *On multicolourings of cubic graphs, and conjectures of Fulkerson and Tutte*, Proc. London Math. Soc. (3) **38** (1979), no. 3, 423–460.
93. Claude E. Shannon, *A theorem on coloring the lines of a network*, J. Math. Physics **28** (1949), 148–151.
94. John Sheehan, *The multiplicity of Hamiltonian circuits in a graph*, Recent advances in graph theory (Proc. Second Czechoslovak Sympos., Prague, 1974), Academia, Prague, 1975, pp. 477–480.
95. Thoralf Skolem, *Some remarks on the triple systems of Steiner*, Math. Scand. **6** (1958), 273–280.
96. Bohdan Smetanuik, *A new construction on latin squares I. Proof of the Evans conjecture*, Ars Combin. **11** (1981), 155–172.
97. Joel Spencer, *Maximal consistent families of triples*, J. Combinatorial Theory **5** (1968), 1–8.
98. Michael Stiebitz, Diego Scheide, Bjarne Toft, and Lene M. Favrholdt, *Graph edge coloring. Vizing's theorem and Goldberg's conjecture*, Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons Inc., Hoboken, NJ, 2012.
99. Robin Thomas and Paul Wollan, *An improved linear edge bound for graph linkages*, European J. Combin. **26** (2005), no. 3-4, 309–324.
100. Andrew G. Thomason, *Hamiltonian cycles and uniquely edge colourable graphs*, Ann. Discrete Math. **3** (1978), 259–268.
101. Stéphan Thomassé and Stéphan Bessy, *Partitioning a graph into a cycle and an anticycle, a proof of Lehel's conjecture*, J. Combin. Theory Ser. B **100** (2010), no. 2, 17–180.
102. Carsten Thomassen, *Even cycles in directed graphs*, European J. Combin. **6** (1851), no. 1, 85–89.

103. ———, *Connectivity in tournaments*, Graph Theory and Combinatorics, a Volume in Honour of Paul Erdős, Academic Press, London, 1984, pp. 305–313.

104. ———, *Highly connected non-2-linked digraphs*, Combinatorica **11** (1991), no. 4, 393–395.

105. ———, *Every planar graph is 5-choosable*, J. Combin. Theory Ser. B **62** (1994), no. 1, 180–181.

106. ———, *On the number of Hamiltonian cycles in bipartite graphs*, Combin. Probab. Comput. **5** (1996), no. 4, 437–442.

107. ———, *Chords of longest cycles in cubic graphs*, J. Combin. Theory Ser. B **71** (1997), no. 2, 211–214.

108. ———, *Independent dominating sets and a second Hamiltonian cycle in regular graphs*, J. Combin. Theory Ser. B **72** (1998), no. 1, 104–109.

109. Paul Turán, *On an extremal problem in graph theory (in Hungarian)*, Mat. Fiz. Lapok **48** (1941), 436–452.

110. William T. Tutte, *On Hamiltonian circuits*, J. London Math. Soc. **21** (1946), 98–101.

111. Zsolt Tuza, *Some special cases of Ryser's conjecture*, (1978), manuscript.

112. ———, *Graph colorings with local constraints—a survey*, Discuss. Math. Graph Theory **17** (1997), no. 2, 161–228.

113. Jacques Verstraëte, *Extremal problems for cycles in graphs*, Recent trends in combinatorics, IMA Vol. Math. Appl., vol. 159, Springer International, 2016, pp. 83–116.

114. Vadim G. Vizing, *On an estimate of the chromatic class of a p-graph (in Russian)*, Diskret. Analiz **3** (1964), 25–30.

115. ———, *Critical graphs with given chromatic class (in Russian)*, Diskret. Analiz **5** (1965), 9–17.

116. Walter Wessel, *Über eine Klasse paarer Graphen. I. Beweis einer Vermutung von Erdős, Hajnal und Moon*, Wiss. Z. Techn. Hochsch. Ilmenau **12** (1966), 253–256.

117. Alexandr A. Zykov, *On some properties of linear complexes (in Russian)*, Mat. Sbornik N.S. **24(66)** (1949), 163–188.