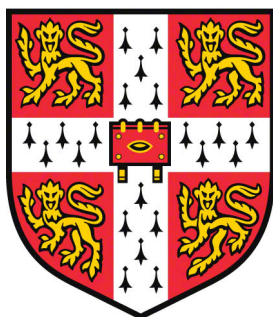


Extremal Theory of Graph Minors and Related Topics



Matthew Daniel Wales

DEPARTMENT OF PURE MATHEMATICS AND MATHEMATICAL
STATISTICS

Trinity Hall

University of Cambridge

A thesis submitted for the degree Doctor of Philosophy

25th February, 2022

Corrected 26th April, 2022

Declaration

This thesis is based on research carried out in the Department of Pure Mathematics and Mathematical Statistics.

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. I

further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Certain parts of this thesis contain work which is either submitted as a preprint or for publication; this is also declared in the text.

Chapter 2 is based on joint work with Andrew Thomason. Results from this chapter appear in a paper [66].

Chapter 7 is based on joint work with Carla Groenland, Tomas Kaiser and Oscar Treffers. Results from this chapter appear in a pre-print [21].

Chapter 8 is based on joint work with András Gyárfás, Dömötör Pálvölgyi and Balázs Patkós. Results from this chapter appear in a paper [22].

In Section 1.3, Theorem 1.10 is based on joint work with Thomason. Sections 1.6 and 1.7 are based on joint work with Groenland, Kaiser and Treffers, and Gyárfás, Pálvölgyi and Patkós respectively. Sections 1.1 and 1.2 provide a general introduction to the topic of Graph Minors, and summarise some pre-existing results. All other results are my own.

Extremal Theory of Graph Minors and Related Topics

Matthew Daniel Wales

Abstract

The core of this thesis are new asymptotically tight bounds for the maximum average degree of a graph with no H minor for several new classes of graph H . Chapter 2 derives an upper bound for almost all graphs H with t vertices and average degree d for a new sparse regime with $d = t^{o(1)}$.

Chapter 3 extends these results to use additional structural information and treat certain vertices of the graph differently to others. This derives an improved upper bound for sparse structured graphs. We also construct a lower bound for almost all such graphs.

These themes are continued in Chapter 5, which introduces the new notion of containing a multigraph as a minor. The extremal function is asymptotically derived for multigraphs rK_t , consisting of t vertices and all pairs having multiplicity r , in two regimes corresponding to $r = \omega(\log t)$ or $r = o(\log t)$. A lower bound is proven for all r which I conjecture is tight. In Chapter 6 we then consider the extremal function for fixed, small multigraphs as minors, including when the host graph is permitted to contain double edges.

In Chapter 4 we consider a related problem of finding minors in bipartite subgraphs of G , where instead of the average degree the Hadwiger number is instead fixed. We find an asymptotic result for the largest bipartite minor. We also prove bounds for the topological analogue of this question.

Moving further from minors, Chapter 7 considers the method of Entropy Compression. The main result is a construction showing that Entropy Compression, or the Rosenfeld/Wanless-Wood framework, provides a best possible bound on the block-average degree of a graph not containing an independent transversal. We also provide a new proof of this matching upper bound, which under a very slightly stronger condition gives an efficient algorithm for constructing such a transversal. Finally, Chapter 8 considers Gallai colourings, and more concretely which sequences of colour class sizes can be realized in a Gallai colouring. We prove a new restriction on such sequences, as well as two results showing certain sequences can be realized in this fashion.

Acknowledgements

I would like to thank all of those who have supported me over these challenging and unusual few years.

Firstly, Andrew Thomason who has given me so much helpful advice and helped me develop as a mathematician. Thanks also to those members of the department who have always been willing to have a chat and help me navigate all of the forms.

My experience would not have been the same without my office companions Freddie and Thomas; it is a shame the endless lockdowns cut short our conversations.

To the entire lunch group of Callum, Kasia, Liam, Robert, Sam and Vilda, for many crazy lunchtime conversations which have covered pretty much every topic.

To my parents and sister, for all of their support over the past many years. And finally to Sophie, for her support, kindness and cups of tea over the past two years.

I could not have done this without any of you.

Contents

1	Introduction	1
1.1	Graph minors: an introduction	1
1.2	The random heuristic, and the parameter $\gamma(H)$	4
1.3	Sparse minors	6
1.4	Minors in bipartite subgraphs	8
1.5	Multigraphs	9
1.6	Entropy compression	14
1.7	Gallai colourings	15
2	The extremal function for sparse minors	17
2.1	Introduction	17
2.2	Our strategy for computing extremal functions	17
2.2.1	The basic strategy for building minors	18
2.2.2	Modifications required for new results	18
2.3	Preliminaries	19
2.3.1	Minor-Minimality and the family $\mathcal{E}_{m,k}$	19
2.3.2	Breaking down our main theorems	21
2.4	Almost-compatible partitions	22
2.5	The dense case: connector and projector	26
2.5.1	Constructing the helper sets	26
2.5.2	Converting compatible partitions into minors	29
2.6	The sparse case	30
3	An extension to almost-multipartite graphs	35
3.1	The parameter $\gamma_r(H)$, and the family $\mathcal{D}(\vec{\sigma}, \mathbf{D})$	37
3.2	Breaking down our upper bounds	40
3.3	The dense case	45
3.3.1	Almost-compatible partitions	46
3.3.2	The bipartite case	48
3.3.3	The general case	50
3.4	The sparse case	53

3.5	Many non-adjacent sets in random graphs	54
3.6	Proof of lower bounds	57
4	Minors of bipartite subgraphs	63
4.1	Introduction	63
4.2	RB-bipartite graphs	64
4.3	The Hadwiger number	67
4.3.1	The lower bound	68
4.3.2	The upper bound	68
4.4	The topological clique number	71
5	Large complete multigraph minors	74
5.1	Introduction	74
5.2	Relative entropy and Chernoff bounds	76
5.3	The lower bounds	79
5.3.1	A parametric lower bound	79
5.3.2	Choosing optimal parameters for the lower bound	81
5.4	The upper bound	85
5.5	The dense case	86
5.5.1	The small r case	86
5.5.2	The large r case	89
5.5.3	Converting compatible partitions into minors	91
5.6	The sparse case	92
5.6.1	Concluding remarks	95
6	Small multigraph minors	96
6.1	Introduction and preliminaries	96
6.2	Cockades	99
6.3	An interlude: K_4 -minor-free graphs	100
6.4	Results for an edge with multiplicity	101
6.4.1	General results	102
6.4.2	The $r = 4$ case	104
6.4.3	The $r = 5$ case	105
6.4.4	The $r = 6$ case	107
6.5	Triangles	109
6.5.1	The graph ${}_2K_3^-$	109
6.5.2	The graph ${}_2K_3$	110
6.6	Further work	112
7	Entropy Compression	114
7.1	Introduction	114
7.2	A direct entropy compression proof	116

7.2.1	The graph case	116
7.2.2	A generalisation to r -graphs	121
7.3	The graph construction	124
7.4	Extension to hypergraphs	126
7.5	Maximum degree conditions	127
8	Gallai Colourings	131
8.1	Introduction	131
8.2	A necessary condition for Gallai sequences	133
8.3	Sequences with almost all entries equal are Gallai	135
8.4	Balanced sequences are Gallai	136

Chapter 1

Introduction

1.1 Graph minors: an introduction

Definition 1.1. Let G, H be graphs. We say H is a *minor* of G , $G \succ H$, if H can be obtained from G by a series of edge contractions, and vertex and edge deletions. An *edge contraction* is the operation of replacing two adjacent vertices u, v with a new vertex, joined to all vertices that were a neighbour of u or a neighbour of v . We denote the graph obtained from G by contraction of the edge uv by G/uv .

In this thesis, we will study properties of a graph G which force a fixed graph H to appear as a minor. Rather than exhibiting a series of edge contractions which exactly yield H as a minor, it is simpler to use the following characterisation (where for $u, v \in V(G)$, we say $u \sim_G v$ if there is an edge of G from u to v , and for $U, V \subset V(G)$ that $U \sim V$ if there is an edge of G with one end in U , and the other in V).

Definition 1.2. Let G, H be graphs. A collection of disjoint non-empty subsets $(V_h)_{h \in V(H)}$ of $V(G)$ is a *model* of H in G if for all edges hh' of H , $V_h \sim_G V_{h'}$, and further each $G[V_h]$ is connected. The subsets V_h are called *parts* of the model. Where G is clear from context, we will say V_h is connected, and $V_h \sim V_{h'}$.

If we do not require the subsets to be connected, we instead say $(V_h)_{h \in V(H)}$ is an *H -compatible-partition*.

If there is a model of H in G , we can contract the edges of a spanning tree of each $(G[V_h])_{h \in H}$ in any order we choose. After labelling the contraction of V_h as h , and possibly deleting some vertices and edges, we obtain an H minor in G . The converse also holds; define V_h to consist of all vertices of G which are eventually contracted to form h .

We also note that the minor relation is *transitive* — that is, if H is a minor of G , and J is a minor of H , then J is a minor of G ; this follows from our first characterisation. We can use this to combine edge contractions with models in building a minor.

The study of graph minors was initiated in the 1930s by Wagner [69] while investigating the famous 4-colour theorem (then a conjecture) which states that every graph which can be drawn in the plane has a proper vertex colouring using only 4 colours. Planar graphs can be characterised by containing no K_5 or $K_{3,3}$ minor; hence if one could show every graph without such minors can be 4-coloured, this would imply the 4-colour theorem. Wagner showed it would suffice to prove this for K_5 minor free graphs.

This line of study led to Hadwiger [25] conjecturing that every graph with no K_s minor is $(s - 1)$ colourable; a significantly stronger result. The 4-colour theorem was eventually proven by Appel and Haken [1], and Appel, Haken and Koch [2] in 1976, and while simpler proofs have since been developed, they still rely on computer verification. Robertson, Seymour and Thomas [59] proved the next case $s = 6$ by reducing it to the 4-colour theorem; all cases beyond this remain open.

Hadwiger's conjecture can be relaxed to bounding the (smallest) function $\chi(t)$ such that all graphs with no K_t minor can be $\chi(t)$ coloured. One way to bound this would be to bound the average degree of graphs not containing a K_t minor; for historical reasons we use the quantity $e(G)/|G|$ instead of the average degree (some people call this the density, though we use the convention that the density of G is $e(G)\binom{|G|}{2}^{-1}$). We note $e(G)/|G|$ is by definition half the average degree of G .

Define a function $d(t)$ to be the least d , if it exists, such that every graph G with $e(G)/|G| \geq d$ contains a K_t minor. The existence of such a function would show all K_t minor free graphs are $2d(t)$ degenerate (i.e. all subgraphs have minimum degree at most $2d(t)$) and so are $1 + 2d(t)$ colourable — this follows since we can order the vertices so each vertex follows at most $2d(t)$ neighbours, and then colour greedily. This implies $\chi(t) \leq 1 + 2d(t)$.

The first progress in this direction was made by Mader [45], who showed $d(t)$ exists and satisfies $d(t) \leq 2^{t-3}$; he improved his bound in a subsequent paper [46] to $d(t) = O(t \log t)$. Later, Bollobás, Catlin and Erdős [5] showed that $d(t) = \Omega(t\sqrt{\log t})$ holds; this was shown to be essentially tight by Kostochka [41] and independently Thomason [64] showing $d(t) = \Theta(t\sqrt{\log t})$. Finally Thomason [65] determined the exact constant by obtaining an upper bound which asymptotically matches the lower bound of Bollobás, Catlin and Erdős, showing $c(t) = (\alpha + o(1))t\sqrt{\log t}$ for the explicit constant α of Definition 1.6.

This method provided the best-known bound on $\chi(t)$ until 2019, when Norin, Postle and Song [50] improved the upper bound to $\chi(t) = O(t(\log t)^\beta)$ for any $\beta > \frac{1}{4}$. This sparked a sequence of further improvements [56, 55, 6] using varied methods such as density increments and chromatic separability, with the most recent bound $\chi(t) = O(t(\log \log t))$ by Delcourt and Postle; the bottleneck to improving their result further is the weakness of known bounds on the chromatic number of small K_t minor free graphs.

These small graphs have their chromatic number bounded by essentially a theorem of

Duchet and Meyniel [10] which shows a K_t minor free graph with n vertices contains an independent set of size at least $\frac{n+1}{2t-2}$. Iteratively removing largest independent sets, this implies there is a colouring of any n vertex K_t minor free graph using around $\frac{\log n}{\log(1-1/(2t-2))}$ colours. If Hadwiger’s conjecture is true, a version of this theorem would hold with $\frac{n}{t}$ size independent sets; however this is open. Apart from small improvements [39, 40, 54] this result remains essentially best-known.

In this thesis, we will be deriving bounds for the following generalisation of $d(t)$ to non-complete graphs.

Definition 1.3. Let H be a graph. The *minor-extremal* function $c(H)$ is the infimum of all d such that every graph G with $e(G)/|G| \geq d$ contains H as a minor.

Since $d(t)$ exists for each t , this means $c(H)$ exists for all finite simple graphs H as $H \subset K_{|H|}$ and by definition $d(t) = c(K_t)$ — recall that the minor relation is transitive, and that if H is a subgraph of G , then H is a minor of G . This setting has been considered by many authors for specific graphs H , but the case where H is a large, general graph was first studied in detail by Myers and Thomason [49]. They introduced a graph parameter $\gamma(H)$, and showed that $c(H) = (\alpha + o(1))t\gamma(H)\sqrt{\log t}$ for some explicit constant $\alpha = 0.319\dots$ to be defined later, provided that $\gamma(H)$ is bounded away from zero. We also defer details of the graph parameter until the next section, though remark that $\gamma(H) \leq \sqrt{\frac{\log d}{\log t}}$ (where $t = |H|$, and d denotes $e(H)/|H|$), and this is asymptotically tight for ‘most’ graphs H , and for all (dense) regular graphs H .

This motivates looking at bounds on $c(H)$ in terms of only the average degree d and number of vertices t . The results of [49] provide a tight bound $c(H) = (\alpha + o(1))t\sqrt{\log d}$ for most H , in the regime where d grows polynomially in t . It is natural to ask if this also holds for more sparse graphs. Reed and Wood [58, 57] showed that an upper bound of the form $c(H) = O(t\sqrt{\log d})$ holds for all pairs t, d . Norin, Reed, Thomason and Wood [51] later obtained a construction that shows $(\alpha - o(1))t\sqrt{\log d}$ remains a lower bound for almost all graphs, even when d is slow-growing, and we include a rephrasing of their result here.

Theorem 1.4 ([51, Theorem 4]). *Let $\epsilon > 0$. Then there is a $D = D_{1.4}(\epsilon)$ such that for all $t > d > D$, and all but at most 2^{-t} proportion of graphs H with t vertices and td edges, there is a graph G with $e(G)/|G| \geq (1 - \epsilon)\alpha t\sqrt{\log d}$ which does not contain H as a minor.*

They also posed the question of whether this was optimal. In Chapter 2, based on joint work with Thomason [66], we will prove this is the case (see Theorem 1.10).

The above lower bound holding for almost all graphs, and not all graphs, is a necessity — Myers and Thomason showed that the extremal function is strictly smaller for e.g. complete bipartite graphs. In fact this constraint is required more generally even for a bound in terms of $\gamma(H)$ if d is sufficiently small compared to t — recent results of Haslegrave, Kim and Liu [27] as well as Hendrey, Norin and Wood [29] showed for certain graphs, such as those

with strongly sublinear separators, the extremal function is $o(t\sqrt{\log d})$. More concretely, they showed the minor extremal function for the hypercube Q_d is $O(2^d)$. In Remark 1.8 we will show $\gamma(Q_d) = \sqrt{\frac{\log d}{\log 2^d}}$, and hence $|Q_d|\gamma(Q_d)\sqrt{\log |Q_d|} = (1 + o(1))2^d\sqrt{\log d}$ which is asymptotically larger than the upper bound for $c(Q_d)$ proven by Hendrey, Norin and Wood.

1.2 The random heuristic, and the parameter $\gamma(H)$

Suppose we have an Erdős-Renyi random graph $G = G(n, p)$. When does G contain K_t as a minor (with high probability)? It turns out to be easier to ask the (equivalent) question ‘when does $G(n, p)$ have a model of H ’, or even the weaker question ‘when does $G(n, p)$ have an H -compatible partition’. This turns out not to significantly affect the answer. Each part X in our compatible partition will have size $o(\log n)$. Initially reserving ϵ proportion of vertices (and hence replacing n with $(1 - \epsilon)n$ for the purposes of building a compatible partition), each reserved v is adjacent to every vertex in X with probability $p^{|X|}$. Thus with high probability such a v exists among our reserved vertices, and we can add it to X in order to connect it. We note that this does not affect edges between parts, since we have only made them larger.

This question was considered by Bollobás, Catlin and Erdős [5], motivated by a then-recent result of Erdős and Fajtlowicz that the Hajós conjecture (an analogue of Hadwiger’s conjecture for topological minors) was false, and moreover almost every graph was a counterexample [11]. They showed the following tight bounds on the size of such minors.

Theorem 1.5 ([5, Theorem 2]). *Let $0 < p < 1$ and $G \sim G(n, p)$. Then with high probability, G does not contain a K_m -compatible partition for $m = \frac{n}{\sqrt{\log_{1/1-p}(n)-1}}$. Further, with high probability G has a K_r minor, for $r = \frac{n}{\sqrt{\log_{1/1-p}(n)+4}}$.*

For the purposes of this thesis, we are not interested in the size of error terms: we seek to determine the asymptotic behaviour. We will thus treat n as being large, and view this result as saying ‘the order of the largest complete minor in $G(n, p)$ is with high probability $(1 + o(1))n\frac{\sqrt{\log n}}{\sqrt{\log(1/q)}}$, where $q = 1 - p$. Since also with high probability the density of $G(n, p)$ is ‘about p ’, this can be converted into a lower bound on $c(t)$ (and this connection was first observed by de la Vega [15]). More specifically, taking $t = n\sqrt{\frac{\log(1/q)}{\log n}}$, and using $\log t = (1 + o(1))\log n$, the following holds:

$$c(t) \geq (1 + o(1))\frac{np}{2} = (1 + o(1))t\sqrt{\log t} \left(\frac{p}{2\sqrt{\log(1/q)}} \right). \quad (1.1)$$

If we are interested only in $c(t)$, rather than general properties of minors in random graphs, it is natural to maximise this function over all values of p . It is not hard to check that $\frac{p}{2\sqrt{\log(1/1-p)}}$ is a concave function of $0 < p < 1$, with a unique maximum.

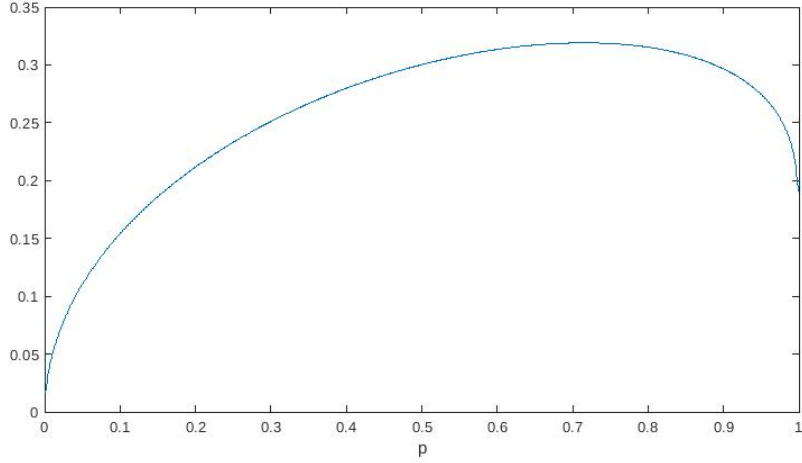


Figure 1.1: A plot of the function $p/(2\sqrt{\log(1/(1-p))})$ for $0 < p < 1$, showing the unique maximum at $p = p_0$ of α .

Definition 1.6. Let α be the quantity defined below:

$$\alpha = \max_{0 < p < 1} \frac{p}{2\sqrt{\log(1/(1-p))}}. \quad (1.2)$$

Let p_0 be the unique value of p at which this maximum is attained. We can approximate $\alpha = 0.319\dots$ and $p_0 = 0.715\dots$. A graph of the function being maximised is included as Figure 1.1.

We remark that Definition 1.6 can be rearranged into the useful form

$$\alpha\sqrt{\log(1/(1-p))} \geq p/2 \quad \forall 0 < p < 1. \quad (1.3)$$

Thomason [64] derived an upper bound which matches the (optimised) lower bound of Bollobás, Catlin and Erdős, thereby showing that $c(t) = (1 + o(1))\alpha t\sqrt{\log t}$, with random graphs of density p_0 and fixed order $n(t)$ giving examples which are close to extremal.

What can we say for non-complete graphs? A natural starting point is to investigate when a non-complete graph H occurs as a compatible partition inside $G(n, p)$. For complete graphs, the calculations are simplified since by symmetry, a partition into t parts is ‘most likely to be H compatible’ when all the parts have equal size. However, when H is not complete this argument fails — intuitively, vertices of high degree ought to be given a larger part, since more conditions need to be satisfied.

This motivated the introduction of a parameter $\gamma(H)$ by Myers and Thomason [49], which precisely picks an ‘optimal’ weighting of the vertices of H .

Definition 1.7. Let H be a graph with t vertices. Define $\gamma(H)$ to be the minimum value

of $\frac{1}{t} \sum_{v \in V(H)} w(v)$, taken over all functions $w : V(H) \rightarrow \mathbb{R}^{>0}$ satisfying the following inequality:

$$\sum_{uv \in E(H)} t^{-w(u)w(v)} \leq t. \quad (1.4)$$

Using this weighting, it is possible to compute when a graph H appears as a minor in a random graph with high probability, at least when the parameter $\gamma(H)$ is bounded away from zero. Surprisingly, it turns out that this gives an asymptotically tight lower bound with the same optimal density p_0 . Myers and Thomason matched this with an upper bound, and so proved $c(H) = (1 + o_{|H|}(1))\alpha|H|\gamma(H)\sqrt{\log |H|}$ for all graphs H where the average degree grows at least polynomially in $|H|$.

One can check that $\gamma(H) \leq \sqrt{\frac{\log d}{\log t}}$ for graphs H with t vertices and td edges (witnessed by the constant weighting) and so for all such H where d grows at least polynomially in t , it holds $c(H) \leq (\alpha + o_t(1))t\sqrt{\log d}$. This phrasing of the result was first used by Reed and Wood [58, 57] and motivates asking if this result holds more generally.

Remark 1.8. We can now prove $c(Q_d) = (1 + o(1))\sqrt{\frac{\log d}{\log 2^d}}$ holds. This is certainly an upper bound, so we need only prove every weighting $w(v)$ satisfying $\sum_{uv \in E(Q_d)} 2^{-dw(u)w(v)} \leq 1$ has $2^{-d} \sum_{u \in V(Q_d)} w(u) \geq (1 + o(1))\sqrt{\frac{\log d}{\log 2^d}}$.

To prove this, we note $E(Q_d)$ has a natural partition into d complete matchings (parallel to the co-ordinate axes). Further, every such weighting w has at most 2^k edges uv for which $2^{-dw(u)w(v)} \geq 2^{-k}$, or equivalently $w(u)w(v) \leq \frac{k}{d}$. Taking $k = \lfloor \log_2 d \rfloor - 1 = (1 - o(1))\log_2 d$, there exists some complete matching $M \subset E(Q_d)$ for which all edges have $w(u)w(v) \geq \frac{k}{d}$. By the AM-GM inequality, this implies $\frac{1}{2}(w(u) + w(v)) \geq \sqrt{\frac{k}{d}}$ for all edges of this matching. The result follows by averaging over the matching, since this also gives the average over all vertices of the cube.

Remark 1.9. This proof extends to all d -regular graphs G with t vertices; they are $d + 1$ colourable, so can be partitioned into $d + 1$ (no longer complete) matchings. But at least $d/2$ of these matchings cover at least $(1 - \frac{2}{d+1})$ proportion of the vertices, and this condition is sufficient for our result to hold; we require $d^{1-o(1)}$ matchings each covering $(1 - o(1))t$ vertices.

1.3 Sparse minors

Our first result will be deriving an extension of the upper bound $c(H) \leq (\alpha + o(1))t\sqrt{\log d}$ to a new sparse regime, where d need not grow polynomially in t . This matches a recent lower bound of Norin, Reed, Thomason and Wood (Theorem 1.4), who showed $c(H) \geq (\alpha + o_d(1))t\sqrt{\log d}$ for almost all graphs H with t vertices and td edges.

Theorem 1.10. *Let $\epsilon > 0$. Then there is a constant $C = C_{1.10}(\epsilon)$ such that for all graphs H with $e(H)/|H| = d > C$, every graph G with $e(G)/|G| \geq (\alpha + \epsilon)|H|\sqrt{\log d}$ contains H as a minor.*

The proof of this theorem is joint work with Thomason, and is the focus of Chapter 2 of this thesis. We believed it was possible to extend our new methods to incorporate some additional structure of H , in the spirit of Myers and Thomason [49]. A first non-trivial case to investigate this is sparse bipartite graphs. The results of Myers and Thomason show all subgraphs H of $K_{\beta t, (1-\beta)t}$ with average degree d (growing sufficiently quickly in t) satisfy $c(H) \leq 2\sqrt{\beta(1-\beta)}(\alpha + o(1))t\sqrt{\log d}$, provided β is bounded away from 0 and 1. This additional factor over the generic result arises from treating the two vertex classes differently.

We will extend this result to a new, sparser, regime which removes the condition on d (though we still require β to be bounded away from 0 and 1). The proof of this theorem naturally extends to a more general setting, and our main result is Theorem 1.11. This is stated in terms of a new parameter $\gamma_r(H)$, discussion of which is deferred to Section 3.1. However, we note that for subgraphs of $K_{\beta t, (1-\beta)t}$ of average degree d , it holds that $2\sqrt{\beta(1-\beta)}(\alpha + o(1))t\sqrt{\log d} \geq \gamma_r(H)(\alpha + o(1))t\sqrt{\log d}$ and so we recover the claimed generalisation.

Theorem 1.11. *Let $\epsilon > 0$, and suppose $r \geq 1$ is an integer. Then there is a constant C such that for all graphs H with $e(H)/|H| = d > C$ and $\gamma_r(H) > \epsilon$, every graph G with $e(G)/|G| \geq (\alpha + \epsilon)\gamma_r(H)|H|\sqrt{\log d}$ contains H as a minor.*

We will also prove an extension of Theorem 1.4 which shows Theorem 1.11 is tight for (some notion of) almost all graphs.

Theorem 1.11 handles one way in which $\gamma(H)$ can become small, namely that the average degree is slow-growing or constant. However, there are other ways, such as in $K_{\beta t, (1-\beta)t}$ where $\beta \rightarrow 0$. What can we say in this case?

When $\beta = o(t/\log t)$, this question has been considered by Kostochka and Prince [43, 42], and independently Kühn and Osthus [44] who showed $c(H) = (\frac{1}{2} + o(1))t$; they also have explicit bounds on the error terms. For other values of $\beta = o(1)$, the best known results are due to Kapadia, Norin and Qian [38] who show $c(K_{\beta t, t}) \leq 40(t\sqrt{\beta \log t} + (1 + \beta)t)$.

We will modify the arguments used in the proof of Theorem 1.11, and extend the methods of the lower bound, to show $c(K_{\beta t, (1-\beta)t}) = (\alpha + o(1))2t\sqrt{\beta \log t}$ for any (sequence of) $\beta < 1$ with $\beta \log t \rightarrow \infty$; we note this matches the result $c(K_{\beta t, (1-\beta)t}) = (\alpha + o(1))2\sqrt{\beta(1-\beta)}t\sqrt{\log t}$ for $\beta \neq o(1)$. In fact the argument works equally well for books $K_{s,t}^*$ and so we only state upper bounds for these graphs. We will show the following results.

Theorem 1.12. *Let $\epsilon > 0$. Then there is a constant C such that for all t, f satisfying $C < f \leq \log t$, every graph G with $e(G)/|G| \geq (2\alpha + \epsilon)t\sqrt{f}$ contains $K_{ft/\log t, t}^*$ as a minor.*

In particular, $c(K_{ft/\log t, t}^*) \leq (2\alpha + \epsilon)t\sqrt{f}$.

Theorem 1.13. *Let $\epsilon > 0$. Then there is a constant C such that for all $C < f \leq \log t$, there is a graph satisfying $e(G)/|G| \geq (2\alpha - \epsilon)t\sqrt{f}$ and no $K_{ft/\log t, t}$ minor. In particular, $c(K_{ft/\log t, t}) \geq (2\alpha - \epsilon)t\sqrt{f}$.*

1.4 Minors in bipartite subgraphs

A well-known result states that every graph with average degree at least $2d$ contains a bipartite subgraph with average degree at least d ; the same result holds for minimum degree. Do such results also hold for other graph parameters? Formally, given a (monotone) graph parameter $p(G)$ which is unbounded on bipartite graphs, does there exist some function $f(t)$ such that if $p(G) \geq f(t)$, then G has a bipartite subgraph G' with $p(G') \geq t$?

This question was posed by Hickingbotham and Wood, and answered negatively by Scott¹. Such results do hold for many natural graph parameters, however. We examine the case of minors (i.e. the Hadwiger number $h(G) = \max\{t : G \succ K_t\}$) and topological minors (the topological clique number, see Definition 1.16).

Geelen et al. [18] showed that $h(G) \geq c_1 t \sqrt{\log t}$ implies G contains a bipartite subgraph G' with $h(G') \geq t$ by combining the classical average degree result with Thomason's bound on $c(t)$ [65]. The constant c_1 can be taken to be $1.276 \dots + o(1)$, corresponding to twice the average degree required for a K_t minor. Hickingbotham and Wood re-derived this theorem, and explicitly posed the question of whether this bound is asymptotically tight. If we let $f_{\succ}(t)$ be the minimal function such that all graphs with Hadwiger number at least $f_{\succ}(t)$ contain a bipartite K_t minor, we show that $f_{\succ}(t) = (1 + o(1))ct\sqrt{\log t}$, where $c = 1.201 \dots$ is an explicit constant (strictly smaller than 4α) and so answer their question negatively.

Theorem 1.14. *For all $\epsilon > 0$, there is a t_0 such that for all $t > t_0$, if $f(t) = \lceil (1 - \epsilon)t\sqrt{\log_2 t} \rceil$, then there is a graph with a $K_{f(t)}$ minor, and in fact a $K_{f(t)}$ topological minor, which has no bipartite subgraph with Hadwiger number at least t .*

Theorem 1.15. *For any $\epsilon > 0$ there is a $t_0(\epsilon)$ such that for any $t > t_0$, any graph with Hadwiger number at least $(1 + \epsilon)t\sqrt{\log_2 t}$ contains a bipartite subgraph with Hadwiger number at least t .*

Chapter 4 uses very different methods to the earlier chapters, and provides a concrete example of an extremal problem for minors where the constant α is not involved. We will also consider an analogue of this problem for topological minors.

Definition 1.16. We say a graph is a *topological H* , or *TH*, if it consists a set of distinct branch vertices v_h labelled by the vertices of H , together with a collection $(P_e)_{e \in E(H)}$ of

¹Let $a(G)$ be the length of the shortest odd cycle in G , and 1 if none exist. Let $b(G)$ be the largest t such that G has a $K_{t,t}$ subgraph. The counterexample is given by $c(G) = \max(a(G), b(G))$ applied to odd cycles C_i .

internally vertex disjoint paths, where if $e = hh'$, then P_e is a path from v_h to $v_{h'}$. We say G contains a *topological H minor* if $G \supset TH$, and let $tcl(G)$ the *topological clique number* be the largest t such that $G \supset TK_t$.

Analogously to before, define $f_T(t)$ to be minimal such that any graph with topological clique number $f_T(t)$ contains a bipartite topological K_t minor. By first taking a bipartite subgraph with at least half the average degree, then using bounds on $tcl(G)$ in terms of average degree, Hickingbotham and Wood showed there exists a constant c (and we can take $c = \frac{20}{23} + o_t(1)$) such that $f_T(t) \leq ct^2$. We will improve the constant of this upper bound, and give a construction which shows that $f_T(t) = \Omega(t^2)$.

Theorem 1.17. *If G is a graph with topological clique number at least $(\frac{1}{2} + o_t(1))t^2$, then G has a bipartite subgraph H with $tcl(H) \geq t$.*

Theorem 1.18. *For all t , there is a graph G with $tcl(G) \geq t^2/4$ but G has no bipartite subgraph H with $tcl(H) \geq t$.*

1.5 Multigraphs

Definition 1.19. A multigraph G consists of a vertex set $V(G)$, together with a multiset $E(G)$ of unordered pairs from G . We say the *multiplicity* of an edge uv is the number of times uv appears in $E(G)$, and is denoted $m_G(uv)$ (the G can be omitted when clear from context). A multigraph is called *loopless* if $m_G(uu) = 0$ for all $u \in V(G)$. A loopless multigraph is called *simple* if all edges have multiplicity at most 1; i.e. a simple multigraph is a graph.

Definition 1.20. Let $t, r \geq 0$ be integers. The multigraph ${}_rK_t$ consists of t vertices, with all distinct pairs of vertices having r edges between them. Equivalently, it is obtained from K_t by giving all edges multiplicity r rather than 1.

Two observations led me to study multigraphs in the context of minors. The first was noting the argument of Thomason [65] that built a K_t compatible partition actually guaranteed an ${}_rK_t$ compatible partition, where $r = o(\sqrt{\log t})$.

I also observed that the argument of Bollobás, Catlin and Erdős that builds compatible partitions in a random graph (see Theorem 1.5) roughly extends to show $G(n, p)$ contains ${}_rK_t$ as a minor at about the same time it contains a $K_t = {}_1K_t$ minor for each $r = o(\log t)$. This motivates the more general question of computing the extremal function for ${}_rK_t$.

Definition 1.21. Let H be a multigraph. We say H is a *minor* of a multigraph G , $G \succ H$, if H can be obtained from G by a sequence of edge contractions, and vertex or edge deletions. An edge contraction denotes replacing two adjacent vertices u, v with a new vertex uv , and adding $m(uv) - 1$ loops at uv , together with $m(uw) + m(vw)$ edges to each other vertex w .

Definition 1.22. Let H be a multigraph. The *minor extremal function* is defined as

follows:

$$c(H) = \max\{e(G)/|G| : G \text{ is a simple graph not containing a } H \text{ minor}\}.$$

We further define $c_s(H)$ to be the maximum value of $e(G)/|G|$ over all multigraphs G with edge multiplicities at most s not containing a H minor, and $c'(H)$ to be $c_s(H)$ where s is the maximum multiplicity of an edge in H (this is a reasonable restriction, since given any multigraph containing H as a minor, deleting edges until each has multiplicity at most s preserves this minor).

We do need a restriction on the maximum multiplicity of the host graph G for the definition to make sense, since any graph with three or more vertices cannot be a minor of the multiple edge ${}_rK_2$ of arbitrarily high average degree. It certainly holds $c'(H) \geq c(H)$ for every multigraph H , and in general equality does not hold. We will later show $c({}_rK_2) = \Theta(\sqrt{r})$ holds, however $c'({}_rK_2) = \Theta(r)$ — a lower bound comes from trees where each edge is given multiplicity $r - 1$, and we will show this classifies the extremal graphs in Section 6.

We mainly focus on the case $s = 1$ where our host graphs are simple, and it is easily seen if also H is simple we recover Definition 1.3. The case $c_2(H)$ is of particular interest because of directed graphs. If H is a simple directed graph, and H' the multigraph obtained from H by ‘deleting arrows’, it holds $c(H) \geq c_2(H')$ by taking an extremal multigraph for H' and adding arrows in any way we wish. This inequality can sometimes be tight, for instance the digraph H with 3 vertices and 6 directed edges (a bidirected K_3) has $c(H) = c_2({}_2K_3)$.

To show that the function $c(H) = c_1(H)$ exists, let r be the maximum multiplicity of an edge in H . Contracting each of $|H|$ disjoint r sets in $K_{r|H|}$ gives ${}_rK_{|H|} \supset H$ as a minor; hence $c(H) \leq c({}_rK_{|H|})$. The existence of $c_s(H)$ follows from this by passing to the ‘underlying simple graph’ obtained by replacing all multiple edges in G by a single edge; this divides $e(G)/|G|$ by at most s and so $c_s(H) \leq sc(H)$.

Thomason’s proof $c(K_t) = (\alpha + o(1))t\sqrt{\log t}$ [65] can be re-analysed to show this remains the value of $c_1({}_rK_t)$ for any $r = o(\sqrt{\log t})$. We will later show this remains true up to $r = o(\log t)$, and further this condition on r is tight — for any $\epsilon > 0$, there is an $\eta > 0$ such that $c({}_rK_t) \geq (\alpha + \eta + o(1))t\sqrt{\log t}$ for any $r > \epsilon \log t$.

Classical extremal questions for multigraphs have significant additional difficulty compared to graphs, and are more difficult to ask sensibly. For instance, asking for the most edges in a graph with no ${}_rK_t$ subgraph is not well-posed, since an edge with arbitrarily large multiplicity shows this is unbounded. If we restrict each edge to have multiplicity at most r , then we should give all edges multiplicity at least $r - 1$ and add a Turán graph — there is no new theorem here.

It is therefore usual to consider a whole family of multigraphs. One family of interest is all multigraphs with at most q vertices and at least $r + 1$ edges; an n vertex graph not

containing any such subgraph is called an (n, q, r) -graph. Bondy and Tuza [67] proved a generalisation of Turán's theorem for these graphs, and this was extended by Füredi and Kündgen [16].

Graph minors, however, turn out to provide a more interesting setting because we can ask a natural extremal question for a single multigraph, with genuinely new results compared to simple graphs.

We can now state our main results, which focus on the graphs ${}_rK_t$. Understanding the extremal function for these graphs would likely be necessary to extend to more general graphs. We use a rescaling $k = r/\log t$ for lower bounds, since there is qualitatively different behaviour depending on whether k is small or large. Our first theorem gives a lower bound for all values of k , and the next two theorems give corresponding upper bounds for small and large values of k , respectively.

Theorem 1.23. *There are smooth increasing functions $p(k), q(k) : (0, \infty) \rightarrow [0, 1]$ such that*

$$c({}_{k \log t}K_t) \geq (\tfrac{1}{2} - \epsilon)pt\sqrt{(k \log t)/q}$$

holds for all sufficiently large t .

The functions p and q can be recovered from the later Theorem 5.15. We will show they have certain properties, which imply that for $k \rightarrow 0$ $c({}_{k \log t}K_t) \geq (1 - \epsilon)\alpha t\sqrt{\log t}$ holds, and for $k \rightarrow \infty$ we show $c({}_{k \log t}K_t) \geq (\tfrac{1}{2} - \epsilon)t\sqrt{k \log t}$ holds.

Theorem 1.24. *Let $\epsilon > 0$. Then there is a constant D such that for all $t > D$ and $r < (\log t)/D$, every graph G with $e(G)/|G| \geq (\alpha + \epsilon)t\sqrt{\log t}$ contains ${}_rK_t$ as a minor. In particular,*

$$c({}_rK_t) \leq (\alpha + \epsilon)t\sqrt{\log t}.$$

Theorem 1.25. *Let $\epsilon > 0$. Then there is a constant D such that for all $t > D$ and $r > D \log t$, every graph G with $e(G)/|G| \geq (\tfrac{1}{2} + \epsilon)t\sqrt{r}$ contains ${}_rK_t$ as a minor. In particular,*

$$c({}_rK_t) \leq (\tfrac{1}{2} + \epsilon)t\sqrt{r}.$$

Our lower bound is non-explicit, as it arises from optimisation of a parametric result Theorem 5.15. Graphs of the functions described are included in Figures 1.2 and 1.3.

Of particular interest is that while our lower bounds come from random graphs, their density varies. This is different to simple graphs, where almost all graphs have extremal examples with the same density — requiring a larger multiplicity does seem to require a larger density. Our proofs of the above theorems do use that t is large, which somewhat

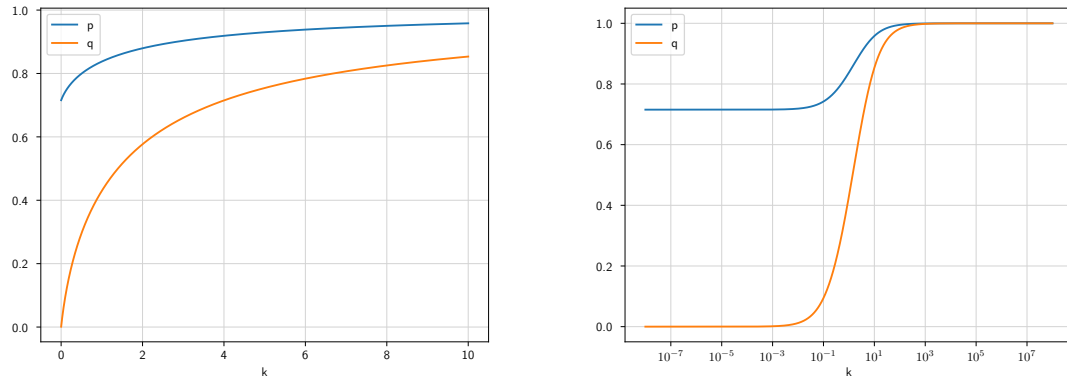


Figure 1.2: A plot of the optimised parameters p, q from Theorem 1.23 on each of linear and logarithmic axes against $k = r/\log t$.

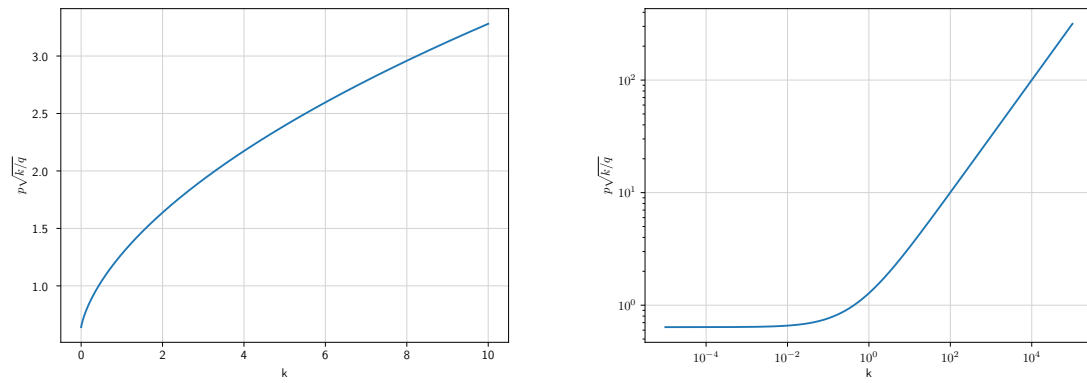


Figure 1.3: Plot of prefactor $p\sqrt{k/q}$ on each of linear/linear and logarithmic/logarithmic axes.

limits applicability. Fortunately, the general t case can be obtained from this.

Corollary 1.26. *Let $\epsilon > 0$. Then there is a constant C such that for all t and all $r > D \log t$, every graph G with $e(G)/|G| \geq (\frac{1}{2} + \epsilon)t\sqrt{r}$ contains ${}_rK_t$ as a minor.*

Proof. Let $C_1 = \lceil 2D \rceil$, where D is the constant from Theorem 1.25 applied with the current ϵ , and $C = C_1^3$. By Theorem 1.25, such a G contains ${}_{\lfloor 2r/C_1^2 \rfloor}K_{C_1 t}$ as a minor, and so also contains ${}_rK_t$ by contracting each of t disjoint collections of C_1 vertices to a single vertex. $\square$

Remark 1.27. *Any simple graph containing an ${}_rK_t$ model (see Definition 5.5; this will later be shown to be equivalent to an ${}_rK_t$ minor) exactly contains subsets $X_1, \dots, X_t$ with pairwise at least r edges between them. The minimum of $\sum |X_i|$ for which this is possible is $t\sqrt{r}$, and so $\frac{1}{2}t\sqrt{r} - 1$ is a lower bound on $c({}_rK_t)$ for all r, t .*

We will also consider the extremal function for certain small multigraphs. The case $t = 2$ is of interest to computer science — the Largest Bond Problem (see [9] for details) can be restated as finding the largest r for which ${}_rK_2$ is a minor of the graph. The question of finding graphs with large average degree not containing certain size bonds has not previously been considered. In Chapter 6 we will exactly compute the extremal function $c_s(H; n)$ for several small multigraphs, as well as classify the extremal graphs.

Definition 1.28. For H a multigraph, define the minor-extremal function $c_s(H; n)$ as the maximum number of edges in an n vertex H free multigraph where all edge multiplicities are at most s .

Computing the extremal function for small minors goes back as far as the subject of minors themselves, with early results due to Wagner [69] for K_5 , Hall for $K_{3,3}$ [26] and Mader [45] for K_5, K_6, K_7 . Dirac [8] classified all K_4 minor free graphs and extremal graphs for the graphs K_t^- obtained from K_t by deleting one edge for each of $t = 5, 6, 7$. Jakobsen [32, 33, 34] gave a classification of extremal graphs for K_8 and K_8^- minors, and for the families K_t^{--} (consisting of the two non-isomorphic graphs obtained from K_t by deleting two edges) for each of $5 \leq t \leq 8$.

This area is still of active interest; Song and Thomas [61] derived the extremal function for K_9^- , and over many recent papers extremal graphs for all simple graphs with at most 11 edges which are 3-connected have been determined; for a survey of these results see [7]. However, only a small number of graphs on more than 12 edges have their extremal function (and extremal graphs) exactly known.

The small multigraphs we consider have up to 6 edges, which is much fewer than these simple graphs. However, forbidding a multigraph as a minor corresponds to forbidding a reasonably large family of simple graphs as a minor — within each part of a model, we must consider all the ways multiple edges can interact in order to reduce to a simple graph. This simplification method also increases the number of edges greatly — even ${}_4K_2$

corresponds to a family of 3 graphs with up to 7 edges, and things grow rapidly from there. Our framework of allowing multigraphs has made these problems tractable.

We will prove a number of exact extremal results for small multigraphs. More concretely we compute $c_s(rK_2; n)$ for each of $s = 1, 2$ and each $r \leq 6$, as well as computing $c_1(2K_3; n)$ and $c_2(2K_3; n)$. We will also consider the cases $c_s(rK_2)$ where both r and s are large (and s is suitably large compared to r). In each of these cases, we are able to characterise extremal graphs.

1.6 Entropy compression

The guiding principle of entropy compression is that it is impossible to compress n random bits into some $m < n$ random bits in a lossless fashion. Indeed, for any scheme $f : \{0, 1\}^n \rightarrow \{0, 1\}^m$ and $g : \{0, 1\}^m \rightarrow \{0, 1\}^n$, the probability over uniformly random $x \in \{0, 1\}^n$ that $g(f(x)) = x$ is at most 2^{m-n} . This holds even if f and g are allowed to be randomised functions (but the randomness cannot depend on x).

The term *entropy compression* was first used by Tao [63] in a blog post describing a proof of Moser [48], which used this method to give a constructive proof of the Lovasz Local Lemma. The name is now used to describe a whole family of methods related to the local lemma, from these direct entropy compression arguments, to methods using a cluster expansion form of the local lemma, as well as the counting arguments used in the so-called Rosenfeld/Wanless-Wood framework [60, 73].

In joint work with Carla Groenland, Tomáš Kaiser and Oscar Treffers, we investigated these methods in the context of independent transversals.

Definition 1.29. Let G be a (hyper)graph with vertex partition $V(G) = \bigcup_{i=1}^n B_i$. A *transversal* of G is a choice of $(v_1, \dots, v_n)$, where $v_i \in B_i$. Such a transversal is an *independent transversal* if $\{v_1, \dots, v_n\}$ is an independent set in G .

In their paper, Wanless and Wood used entropy compression to prove that every graph G as above, with each $|B_i| = t$ and at most $t^2/4$ edges of G incident to each block, contains an independent transversal. Our main joint result was the following, which proves that this is asymptotically tight. This construction provides the first non-trivial example where entropy compression is known to be tight.

Theorem 1.30. *Let $\epsilon > 0$. Then for all t sufficiently large, there exists a graph G with vertex partition $V(G) = \bigcup B_i$ such that each B_i has size t and the average degree of vertices in B_i is at most $(\frac{1}{4} + \epsilon)t$, but G has no independent transversal.*

Our construction has several surprising properties, most notably that it is a forest. This is a very different structure to examples from the maximum degree analogue of this problem, which use complete bipartite graphs with large part sizes. We provide a hypergraph analogue of this construction which is also tight. We also consider bounding the

average degree in addition to some other parameter, either the maximum degree or the maximum number of edges between each pair of blocks. This gives rise to several interesting open problems.

We also provide a novel proof of the lower bound of Wanless and Wood (under a very slightly stronger condition) which uses a direct entropy compression argument. This gives an algorithm to generate an independent transversal with much more efficient runtime. Details are given in Theorem 7.2.

1.7 Gallai colourings

The following definition is due to Gallai [17], with the name first introduced by Gyárfás and Simonyi [24].

Definition 1.31. A (not necessarily proper) edge colouring of a graph H is called *Gallai* if it contains no ‘rainbow triangle’ — a copy of K_3 with three edges all coloured with different colours.

These colourings were originally studied in the context of Ramsey theory (see [13, 23, 47]). The Gallai-Ramsey number $gr_k(G; H)$ (introduced by Faudree et al. [13]) is the smallest integer m such that any colouring of K_m with k colours contains either a rainbow G or a monochromatic H . The case $G = K_3$ corresponds to finding a monochromatic H in a Gallai coloured graph.

Gallai colourings have interesting properties worthy of direct study. For instance, if a graph is Gallai coloured (using at least 3 colours), there must be a disconnected colour class. This result is not trivial, but not too hard to show. We will give a proof of a stronger result in Chapter 8, which gives an equivalent recursive classification of a Gallai colouring as one with a ‘Gallai Partition’ into smaller Gallai coloured graphs — this was originally proved by Gyárfás and Simonyi [24].

Definition 1.32. A partition $X_1, \dots, X_k$ of $V(G)$ is called a *Gallai Partition* if there exists a set C of at most two colours such that for each $i < j$, all edges between X_i and X_j have the same colour, and that colour is from C .

This recursive classification motivated and allowed the counting of the number of Gallai colourings of K_n , and counting those which also have some additional structure [4, 53, 52].

Motivated by these, we study the general structure of Gallai colourings, and more concretely the possible colour class sizes in a Gallai colouring. It is clear that not all sequences of class sizes are possible — for instance no colouring exists where all classes have size 1 (and there is more than 1 class). However any colouring using at most two colours is Gallai, so for any integers $e_1 \geq e_2$ with $e_1 + e_2 = \binom{n}{2}$ for some n , there is a Gallai colouring of K_n with colour classes of size e_1 and e_2 . It follows from the existence of a Gallai partition that in a Gallai colouring of K_n , some colour class must have size at least

$n - 1$. In joint work with Gyárfás, Pálvölgyi and Patkós, we derived some new results of this form. Our first such result is the following.

Theorem 1.33. *Let $n \geq l$ be integers. Then every Gallai colouring of K_n contains a set of at most l colours which together are used on at least $(n - 1) + (n - 2) + \dots + (n - l)$ edges.*

This result is tight, for instance the ‘star colouring’ of K_n where the edge ij ($i < j$) has colour j is Gallai, and attains equality for each l .

Definition 1.34. A sequence $e_1 \geq \dots \geq e_k$ of integers is called *Gallai* if there exists an $n \geq 0$ such that K_n has a Gallai colouring with k colour classes of size $e_1, \dots, e_k$ respectively.

Theorem 1.33 implies that for all Gallai sequences (e_i) with $\sum e_i = \binom{n}{2}$ and any $l \geq 1$, we have $\sum_{i=1}^l e_i \geq (n - 1) + \dots + (n - l)$. In general this condition is not sufficient; this is witnessed by the sequence $(7, 3, 2, 2, 1)$ for $n = 6$ which is not Gallai by a careful case check, see Section 8.2 for details. We will prove two positive results, showing that certain sequences taking only two distinct values are Gallai. The restrictions of the parameters in the theorem are necessary so the conditions of Theorem 1.33 are satisfied.

Theorem 1.35. *Let $p \geq n - 1$. Suppose we can express $\binom{n}{2} = kp + q$ for some integers $k, p \geq 0$. Then there exists a Gallai colouring of K_n with k colour classes of size p , and one colour class of size q . Equivalently, the sequence $(p, \dots, p, q)$ is Gallai.*

Theorem 1.36. *Let $0 \leq k \leq \lceil \frac{n}{2} \rceil$. Then K_n has a Gallai colouring, all of whose class sizes are $\lfloor \frac{n(n-1)}{2k} \rfloor$ or $\lceil \frac{n(n-1)}{2k} \rceil$. We call such a colouring a balanced Gallai k -colouring.*

We remark these two theorems are equivalent if $n = kp$. When $k \nmid \binom{n}{2}$ we will still derive Theorem 1.36 from Theorem 1.35 however additional ideas are required.

Chapter 2

The extremal function for sparse minors

2.1 Introduction

For any $\epsilon > 0$, and any graph H with t vertices and average degree $d > t^\epsilon$, Myers and Thomason showed that every graph G with $e(G)/|G| \geq (\alpha + o_t(1))t\sqrt{\log d}$ contains H as a minor, where $\alpha = 0.319\dots$ is the constant of Definition 1.6 on page 5. In this chapter, we will generalise this theorem to the regime where d is constant or slow-growing.

The aim of this section is to prove the following Theorem 1.10, showing that the theorem of Myers and Thomason holds in this more general setting, provided d is sufficiently large. The proof of this theorem is based on joint work with Thomason [66].

Theorem 1.10. *Let $\epsilon > 0$. Then there is a constant $C = C_{1.10}(\epsilon)$ such that for all graphs H with $e(H)/|H| = d > C$, every graph G with $e(G)/|G| \geq (\alpha + \epsilon)|H|\sqrt{\log d}$ contains H as a minor.*

This matches a recent lower bound of Norin et al. [51] (see Theorem 1.4) which holds for almost all graphs, and thus determines the extremal function for almost all graphs with average degree d and t vertices, up to a $(1 + o_d(1))$ multiplicative factor.

2.2 Our strategy for computing extremal functions

Thomason's asymptotic determination of $c(K_t)$ [65] showed that random graphs of an optimal density $p = 0.715\dots$ are 'close to' extremal. Myers and Thomason [49] showed the same remains true for $c(H)$ when H has average degree at least polynomial in $|H|$. This result breaks down for smaller average degree (random graphs contain too many small structures), however Norin et al. [51] showed that blowups of random graphs provide a tight lower bound for almost all H . This provides the matching lower bound for this chapter.

2.2.1 The basic strategy for building minors

In this subsection, we will outline how to build a K_t minor in graphs with $e(G) \geq m|G|$, following the method of Thomason [65], and in the next subsection we will highlight how this method can be extended and modified to apply in the new settings of this thesis.

By taking a minor if necessary, we can assume no proper minor of G has $e(G) \geq m|G|$ (we say G is minor-minimal). In fact we will take G minor minimal in a slightly different class $\mathcal{E}_{m,k}$ which ensures that neighbourhoods have reasonable density, and further that G has reasonable connectivity (at least ϵm) without losing much average degree.

We split into two cases, depending on whether $|G|$ is small or large. In the first case, the so-called ‘dense’ case, G has positive density and also a substantial amount of connectivity compared to its order. This regime is crucial, since it is where the argument is tight and $c(K_t)$ is determined. The aim is to build a complete minor at least as large as in a random graph of the same order and density. How can we do this?

The first step is to build a K_t compatible partition inside G . It turns out a (carefully chosen) random partition works — it may have a few problems, but by deleting a few vertices and accepting a slightly smaller minor these are easily resolved. Turning the compatible partition into a minor is not too difficult after some carefully chosen helper sets are reserved.

In the case where $|G|$ is large, it is less critical to be exact as results remain true up to any constant factor. In this case, we will use the local density of G to either directly obtain a K_t minor, or to build up many smaller complete minors (using the dense case result), and then use the global connectivity to combine these into a K_t minor. To join up these minors, it turns out we need a slight modification of our dense result to build ‘rooted’ minors, where we can pre-specify certain vertices.

2.2.2 Modifications required for new results

We will try to be as general as possible in this chapter to avoid having to repeat arguments in future chapters. However, to avoid making this chapter overly complex, we will sometimes prove only a simplest case here, and prove the generalisations needed separately.

For this chapter, new ideas are needed at several points. The main difficulty arises in the dense case, since the problems with our compatible partition cannot be resolved as earlier — we cannot settle for finding ‘most of H ’ as a minor! It turns out that we can use the helper sets from earlier to also fix these additional problems without throwing away the vertices; this requires introducing the new notion of *almost- H* -compatible partitions.

Our roadmap for finding minors will remain the same in future chapters, though some proofs require reworking to handle different parameters. There are also further difficulties that arise.

In Chapter 3, finding our almost-compatible partition requires a more refined approach to use the additional structure of H (which is what reduces the extremal function). Additional care is also required to break up H into smaller minors which are easier to find.

The lower bound of Chapter 3 is analogous to that of Norin et al. [51], however a significant new argument is required to use the additional structure. We will give details of these new arguments in the later chapter.

Finally, in Chapter 5 again the difficulties arise in building a compatible partition. The weakness of Thomason's proof is somewhat masked by only needing a K_t minor, rather than an rK_t minor. We use new probabilistic techniques to strengthen the argument substantially, though this only gives a tight result for suitably small or large values of r (and separate arguments are needed for these).

The lower bound returns to using random graphs, though interestingly the optimal density now varies depending on the relationship between r and t , rather than remaining a constant $0.715\dots$ as for graph minors. A large amount of new analysis is required to derive a whole-range lower bound.

2.3 Preliminaries

In this section, we aim to reduce Theorem 1.10 to two sub-theorems, Theorems 2.5 and 2.6, which will handle the 'dense' and 'sparse' cases respectively.

We first need some additional definitions and preliminary results, which we will use throughout this chapter. Rather than finding H as a minor in G under our conditions, we will instead build a model of H in (some minor of) G ; these structures are equivalent as we showed in Section 1.1. Recall from the introduction we say $u \sim_G v$ if u, v are adjacent vertices of G , and $U \sim V$ if there is an edge between the subsets U and V .

Definition 2.1. Let G, H be graphs. A collection of disjoint non-empty subsets $(V_h)_{h \in V(H)}$ of $V(G)$ is a *model* of H in G if for all edges hh' of H , $V_h \sim_G V_{h'}$, and further each $G[V_h]$ is connected. The subsets V_h are called *parts* of the model. Where G is clear from context, we will say V_h is connected, and $V_h \sim V_{h'}$.

If we do not require the subsets to be connected, we instead say $(V_h)_{h \in V(H)}$ is an *H -compatible-partition*.

2.3.1 Minor-Minimality and the family $\mathcal{E}_{m,k}$

We start by introducing the family $\mathcal{E}_{m,k}$ and the notion of minor-minimality. Our aforementioned modified graph G' will be a minor-minimal element of this family. This modification is key to our proofs, since in exchange for a small loss in average degree (which is absorbed in error terms), we gain a great number of useful properties; perhaps most importantly large connectivity.

Definition 2.2. Let $\mathcal{C}$ be a family of graphs. A graph G is *minor-minimal* in $\mathcal{C}$ if $G \in \mathcal{C}$, but no proper minor of G is an element of $\mathcal{C}$.

Definition 2.3. The family $\mathcal{E}_{m,k}$, for real numbers $1 < k < m/2$, is the set of graphs G such that $|G| \geq m$, and also $e(G) \geq m|G| - km$.

We note that every graph with $e(G) \geq m|G|$ is an element of $\mathcal{E}_{m,k}$. The usefulness of minor-minimality (although for a different class) was shown by Mader [45]. The specific choice $\mathcal{E}_{m,k}$ was used by Thomason [65], although there the parameters were taken as integers. We therefore re-derive the following properties of minor-minimal elements of $\mathcal{E}_{m,k}$. In this chapter we will later choose $m = (\alpha + \epsilon)t\sqrt{\log d}$ and $k = \epsilon m/2$; having done so elements of $\mathcal{E}_{m,k}$ will still have $e(G)/|G| \geq (\alpha + \epsilon/2)t\sqrt{\log d}$.

Lemma 2.4. *Let G be a minor-minimal element of $\mathcal{E}_{m,k}$. Then the following hold:*

- (1) $|G| \geq m + 1$,
- (2) $e(G) < m|G| - km + 1$,
- (3) $m < \delta(G) \leq 2m$,
- (4) G is k -connected,
- (5) Every edge of G is in at least $m - 1$ triangles.

Proof. Since $G \in \mathcal{E}_{m,k}$, $e(G) \geq m|G| - km \geq m(|G| - k) \geq m(|G| - m/2) \geq m|G|/2$. If also $|G| < m + 1$, we would have $e(G) \leq \binom{|G|}{2} < m|G|/2$ which is a contradiction and so (1) must hold. Property (2) holds since subgraphs are minors and therefore we cannot delete an edge while remaining in $\mathcal{E}_{m,k}$.

For (3), let v be a vertex of minimum degree. Then $G - v$ is a subgraph of G (and hence a minor), so cannot be in $\mathcal{E}_{m,k}$. Since $G - v$ has at least m vertices, we must have $e(G - v) < m(|G| - 1) - km$. But $e(G) \geq m|G| - km$ — thus v has degree more than m . By (2), G has average degree (hence minimum degree) at most $2m$ which completes the proof of property (3).

For (4), suppose that $X \subset V(G)$ is a cut-set. Let (C_1, C_2) be a partition of $V(G) - X$ into (non-empty) unions of connected components. Every vertex in C_1 has all of its neighbours (in G) within $C_1 \cup X$, and likewise for C_2 in $C_2 \cup X$. Since C_1 is non-empty, it contains a vertex — and hence $|C_1 \cup X| > m$ by the minimum degree condition. Since $G[C_1 \cup X] \notin \mathcal{E}_{m,k}$, it must hold $e(G[C_1 \cup X]) < m(|C_1| + |X|) - km$ and likewise for C_2 . The following shows this implies $|X| \geq k$, and so G is k connected:

$$\begin{aligned} m|G| - km &\leq e(G) + e(G[X]) = e(G(C_1 \cup X)) + e(G(C_2 \cup X)) \\ &\leq m|G| + m|X| - 2km. \end{aligned}$$

Finally, suppose e is an edge of G in l triangles. Then the contraction G/e has $|G| - 1$

vertices, and $e(G) - l - 1$ edges. Since this minor cannot be an element of $\mathcal{E}_{m,k}$, it follows from condition (1) that $l \geq m - 1$ and so (5) holds, completing our proof. $\square$

This notion of minor minimality will later also be applied in the multipartite case of Chapter 3 and the multigraph case of Chapter 5; the family used will be the same and although the exact parameter choice will vary we will always have $k \simeq \epsilon m$.

2.3.2 Breaking down our main theorems

We are now ready to state our sub-theorems, Theorems 2.5 and 2.6. While the exact conditions of these may seem somewhat restrictive, they have been chosen exactly so that minor-minimal elements of $\mathcal{E}_{m,k}$ will have the required properties for suitable choices of parameter. We will then use these two theorems to prove Theorem 1.10 and so complete the aim of this section. The remainder of this chapter will then focus on proving the two sub-theorems.

Theorem 2.5. *Given $\epsilon > 0$ there is a constant $D = D_{2.5}(\epsilon)$ such that for all $d > D$ and all graphs H with average degree d , the following holds.*

Let $p > \epsilon$, and suppose G is a graph with density at least $p + \epsilon$, connectivity at least $\epsilon|G|$ and $|G| \geq |H|\sqrt{\log_{1/q} d}$ where $q = 1 - p$. Then G contains H as a minor.

Theorem 2.6. *Given $\epsilon > 0$ there is a constant $D = D_{2.6}(\epsilon)$ such that for all $d > D$ and all graphs H with average degree d the following holds.*

Let $m \geq \epsilon|H|\sqrt{\log d}$, and suppose that G is a graph with $|G| \geq Dm$, $\kappa(G) \geq D|H|$, $e(G) \leq m|G|$ and every edge of G contained in more than $m - 1$ triangles. Then G contains H as a minor.

To prove Theorem 1.10 using these, we will also require the following technical lemma.

Lemma 2.7. *Let $\epsilon \leq \frac{1}{8}$ and $0 < x < 1 - \epsilon$. Then the following holds:*

$$\frac{\log(x + \epsilon)}{\log x} \leq (1 - \epsilon)^2. \quad (2.1)$$

Proof. Myers and Thomason [49] provided a proof of this under the (weaker) condition $\epsilon < \frac{1}{2}$, we provide a simpler proof of an improved bound for $\epsilon < 1/8$ instead. We recall that $x \log x$ is a negative function of $0 < x < 1$ with a global minimum value of $-1/e$. The following sequence of inequalities holds, with the first due to concavity of $\log x$, and the second to last as $(1 + \epsilon\epsilon)(1 - 2\epsilon) > 1$ for any $0 < \epsilon \leq \frac{1}{8} < \frac{1}{2} - \frac{1}{e}$:

$$\begin{aligned} \frac{\log(x + \epsilon)}{\log x} &< \frac{\log(x + \epsilon)}{\log(x + \epsilon) - \frac{\epsilon}{x + \epsilon}} \leq 1 + \frac{\epsilon}{(x + \epsilon)\log(x + \epsilon) - \epsilon} \\ &< 1 - \frac{\epsilon}{1/e + \epsilon} = \frac{1}{1 + \epsilon\epsilon} < (1 - 2\epsilon) < (1 - \epsilon)^2. \end{aligned}$$

□

Proof of Theorem 1.10, assuming Theorems 2.5 and 2.6. Replace ϵ by $\min(\epsilon, \frac{1}{8})$ so now $0 < \epsilon \leq \frac{1}{8}$, and let $\eta = \frac{\epsilon}{2D_{2.6}(0.3)}$ and $C = \max(D_{2.6}(0.3), D_{2.5}(\eta/2), 2/\epsilon)$. Let H be a graph with average degree $d > C$, $m = (\alpha + \epsilon)|H|\sqrt{\log d}$ and $k = \epsilon m/2$; note that $k > 1$.

Suppose that G is a graph with $e(G)/|G| \geq m$. Recalling Definition 2.3, such a G lies in $\mathcal{E}_{m,k}$. Replacing G with a minor, we may assume instead that G is a minor-minimal element of $\mathcal{E}_{m,k}$; recall this implies G has the properties of Lemma 2.4.

We now break into our two cases. Suppose first that $|G| > D_{2.6}(0.3)m$. Then G satisfies the hypotheses of Theorem 2.6, applied with $\epsilon = 0.3$ and the present d, H, m — recall that we chose $d > C$ sufficiently large that this application is valid. Thus G contains H as a minor in this case, and so we can assume $|G| \leq D_{2.6}(0.3)m$. But now $\kappa(G) \geq k \geq \eta|G|$, and in particular G has density at least η . We seek to apply Theorem 2.5 in this case.

For this application, let $p > \eta/2$ be such that G has density $p + \eta/2$. Recalling $d > D_{2.5}(\eta/2)$, if we can show that $|G| \geq |H|\sqrt{\log_{1/q} d}$ (where $q = 1 - p$) then by applying Theorem 2.5 with ϵ replaced with $\eta/2$, G will contain H as a minor. Recall that the following holds since $G \in \mathcal{E}_{m,k}$; the last inequality holding since $(1 - \frac{\epsilon}{2})(1 + \frac{\epsilon}{\alpha}) \geq 1 + \epsilon/2$ for any $0 < \epsilon \leq 1$:

$$(p + \eta/2)(|G| - 1) = e(G)/|G| \geq m - km/|G| \geq m - \epsilon m/2 \geq (1 + \epsilon/2)\alpha|H|\sqrt{\log d}.$$

In particular, by Inequality 1.3 and Lemma 2.7 we have $|G| \geq (1 + \epsilon/2)|H|\sqrt{\log_{1/(q-\eta/2)} d} \geq |H|\sqrt{\log_{1/q} d}$ since $(1 + \epsilon/2)(1 - \eta/2) > 1$. This completes our proof. □

2.4 Almost-compatible partitions

The aim of the next two sections is to prove our dense sub-theorem, Theorem 2.5. In this section, under slightly weaker conditions than those of Theorem 2.5 we will find some subsets in G which ‘almost’ form an H model (we will not require connectivity, and also allow a linear number of missing edges). In the following section, we will develop some helper sets which allow us to extend this relaxed notion to a model of H in G under the stronger conditions of the dense sub-theorem.

Definition 2.8. A partition $(X_h)_{h \in V(H)}$ of $V(G)$ into t non-empty parts is called *M-almost-H-compatible* if for all but at most M edges hh' of H there is an edge of G between X_h and $X_{h'}$. An *H-compatible partition* of G is a 0-almost- H -compatible partition.

We will begin this section with the following parametric lemma, which builds such an almost-compatible partition. The main result of this section, Lemma 2.11, will follow by applying this more general result with suitable parameters. Our partition will be a (carefully chosen) random one; this is discussed further in Remark 2.10.

Lemma 2.9. *Let $l, s \geq 2$ be integers, $\omega > 1$, G a graph with $n \geq sl$ vertices and density p , and $\eta < p$.*

Then $V(G)$ contains a partition $\mathcal{P} = (X_i)$ into at least $s(1 - \frac{2}{\omega\eta})$ parts (each non-empty), such that the number of pairs of parts which are non-adjacent in G is at most $2s^2(6\omega)^l(\frac{1-p}{1-\eta})^{(1-\eta)l(l-1)}$.

Proof of Lemma 2.9. Replacing G by a subgraph on sl vertices of maximal density, we can assume that $n = sl$. Since our end result is only strengthened by increasing p , we can assume that G has density p by applying the theorem with the density of G instead.

We order the vertices of G in non-increasing degree order as $(v_i)_{i=1}^n$, so that $d(v_1) \geq d(v_2) \geq \dots \geq d(v_{sl})$, and define $q_i = 1 - p_i$ by $d(v_i) = p_i(n-1)$. For simplicity of notation, if $v \in V(G)$, say $v = v_i$, let $q_v = q_i$. We partition the vertices of G into l blocks $B_1, \dots, B_l$ each of size s , where $B_i = \{v_{(i-1)s+1}, \dots, v_{is}\}$ for each $1 \leq i \leq l$.

We start by constructing s initial parts, each consisting of one vertex from each of $B_1, \dots, B_l$. At the end, we will modify some of these initial parts to obtain the properties claimed in the lemma. We will build these parts randomly and one at a time as follows. Suppose that we have already picked r parts, and let Y denote their union. By our construction method, we will have $|Y \cap B_i| = r$, and so in particular $B_i \setminus Y$ is a non-empty set for each i provided $r \leq s-1$. Our next part will consist of a uniformly randomly chosen element from each of $B_1 \setminus Y, \dots, B_l \setminus Y$. Repeat in this fashion until all s initial parts have been constructed.

From the above description, one may be concerned the distribution of a part X depends on when it was chosen. However, this is not the case. An equivalent method yielding the same distribution is to, for each i , pick a uniformly random bijection $f_i : [s] \rightarrow B_i$. We can then build parts $X_j = \{f_1(j), \dots, f_l(j)\}$. Viewing X_1 as the first part chosen and so on, we see this yields the same distribution. But now the order does not matter since the labels are arbitrary. In particular, if we fix (or condition on) some part W , each other part consists of a uniformly random element of each of $B_1 \setminus W, \dots, B_l \setminus W$.

We say that a vertex v is *bad* for a part W if $v \notin W$, and there are no edges of G from v to W — equivalently each vertex of W is one of the $q_v(n-1)$ non-neighbours of v . Let $S = V(G) \setminus N(v) \setminus \{v\}$, so $|S| = q_v(n-1)$, and let $S_i = S \cap B_i$. The probability that v is bad for W is $\prod_i \frac{|S_i|}{s}$. We can bound this using the AM-GM inequality to deduce the following:

$$\mathbb{P}(v \text{ is bad for } W) \leq \left(\frac{n-1}{n}q_v\right)^l \leq q_v^l.$$

Thus the expected number of vertices in block B_j which are bad for W is at most $sq_{j,s}^l$, since the (q_k) form an increasing sequence. We say that a part W *rejects* block B_j if more than $\omega sq_{j,s}^l$ vertices inside B_j are bad for W ; otherwise we say W *accepts* B_j . The

probability that B_j is rejected by a part W is at most $\frac{1}{\omega}$ by Markov's inequality. Let $M(W) = \{j : 1 \leq j \leq l-1, W \text{ accepts } B_j\}$. Note that block B_l is excluded here for technical reasons; inequality 2.2 requires there to be vertices after the end of each good block.

A part W is called *good* if $|M(W)| \geq (1-\eta)(l-1)$, and otherwise W is called *bad*. Again applying Markov's inequality, the probability that a part W is bad is at most $\frac{1}{\omega\eta}$. Our final partition will consist entirely of good parts.

Suppose that W is a fixed good part. Conditional on this choice of W , what is the probability some other part X is not adjacent to W ? For this to happen X must consist entirely of vertices which are bad for W ; in particular the vertex in $X \cap S_i$ is bad for W for each $i \in M(W)$. We thus bound

$$\mathbb{P}(X \not\sim W) \leq \prod_{j \in M(W)} \left(\frac{s}{s-1} \omega q_{js}^l \right) \leq (2\omega)^l \left(\prod_{j \in M(W)} q_{js} \right)^l.$$

Since the q_i are increasing, taking $m = |M(W)|$ it holds

$$\frac{1}{m} \sum_{j \in M(W)} q_{js} \leq \frac{1}{ms} \sum_k q_k \leq \frac{l}{m} q \quad (2.2)$$

, where $q = 1-p$. We can therefore apply the AM-GM inequality, and use that $(1-\eta)(l-1) \leq m \leq l$ and $\frac{q}{1-\eta} < 1$ to deduce the following holds also:

$$\mathbb{P}(X \not\sim W) \leq (2\omega)^l \left(\frac{q}{1-\eta} \frac{l}{l-1} \right)^{ml} \leq (6\omega)^l \left(\frac{q}{1-\eta} \right)^{(1-\eta)l(l-1)}.$$

Thus the expected number of pairs of parts, at least one of which is good, but there is no edge between them is at most $s^2(6\omega)^l \left(\frac{q}{1-\eta} \right)^{(1-\eta)l(l-1)}$. We now modify this initial partition to obtain our final partition. By Markov's inequality, the probability there are more than $\frac{2s}{\omega\eta}$ bad parts is at most $\frac{1}{2}$. Further, the probability there are at least $2s^2(6\omega)^l \left(\frac{q}{1-\eta} \right)^{(1-\eta)l(l-1)}$ pairs of parts X_i, X_j which are non-adjacent but at least one of X_i and X_j is good is less than $\frac{1}{2}$.

Thus by a union bound, there exists some partition $\mathcal{P}' = (X_i)$ for which neither of these occur. We form a new partition $\mathcal{P}$ by deleting all bad parts from $\mathcal{P}'$, and placing unused vertices arbitrarily among the remaining parts. We note that adding vertices to a part cannot create new problems; if two subsets are adjacent, adding vertices cannot make them non-adjacent. This new partition still has at least $(1-2s/\omega\eta)s$ parts, and further each non-adjacent pair in this new partition corresponds to a pair X_i, X_j of parts from $\mathcal{P}'$ where both X_i and X_j are good (note we deleted all the bad parts). Thus $\mathcal{P}$ has at most $2s^2(6\omega)^l \left(\frac{q}{1-\eta} \right)^{(1-\eta)l(l-1)}$ pairs of parts which are non-adjacent, and hence satisfies all the

required properties. $\square$

Remark 2.10. The key step in this proof is the two-stage removal, taking some parts to be bad. The blocking is used to give a well-defined notion of bad; we do not really use the blocked structure beyond to say our parts are ‘suitably pseudorandom’.

The two-stage removal is required for the following reason. Let G be a complete bipartite graph $K_{n/2, n/2}$. For any one-stage process, the probability a random l set is entirely contained in, say, the left-hand class is about 2^{-l} for $l \ll n$. In particular, some pair of parts are both contained in the same side, and so non-adjacent, with probability around 2^{-2l} . Our two-stage removal process gives a much smaller probability, around 2^{-l^2} , although this comes at the cost of deleting a positive proportion of parts.

Having proven the parametric form of our compatible partition lemma, we are now ready to choose suitable parameters and complete the aim of this section. The structure constructed in this lemma is our desired ‘almost H model’.

Lemma 2.11. *Given $\frac{1}{8} > \epsilon > 0$ there exists a $D = D_{2.11}(\epsilon)$ such that for all graphs H with average degree $d > D$, and all graphs G with density $p + \epsilon$ satisfying $\epsilon < p < 1 - \epsilon$ and $|G| \geq |H| \sqrt{\log_{1/q} d}$ for $q = 1 - p$, G contains a $|H|d^{-\epsilon/3}$ -almost- H -compatible partition.*

Proof. We apply Lemma 2.9 to G , except with density $p + \epsilon$ rather than p , and taking other parameters $l = \lceil (1 - \epsilon/4) \sqrt{\log_{1/q} d} \rceil$, $s = \lfloor |G|/l \rfloor$, $\omega = 20\epsilon^{-2}\eta^{-1}$, and $\eta < \epsilon$ a function of ϵ alone which is chosen later so all inequalities hold — the dependence only on ϵ ensures this is well-defined. We take D (and hence l) large enough so that all claimed inequalities hold.

We note that by our lower bound on $|G|$ we have $s(1 - 2/\omega\eta) > |H|$. Pick uniformly at random some injection from $V(H)$ to the parts (X_i) of the partition $\mathcal{P}$ from the application of Lemma 2.9. If $M = 2s^2(6\omega)^l \left(\frac{q-\epsilon}{1-\eta}\right)^{(1-\eta)l(l-1)}$, the probability (over all random injections) that a fixed pair hh' corresponds to a non-adjacent pair of parts from $\mathcal{P}$ is at most $\rho = M/|H|^2 \leq 8M/s^2$. If we can show $\rho \leq d^{-1-\epsilon/3}$ then picking an injection with fewer than $\rho|H|d$ such pairs implies the result; such an injection must exist since with positive probability a random variable is at most its expectation. We will prove this holds by bounding $\log \rho$. We have already shown the following:

$$\log \rho \leq \log(16) + l \log(6\omega) + (1 - \eta)l(l - 1)(\log(q - \epsilon) - \log(1 - \eta)).$$

Taking D sufficiently large (in terms of ω, η, ϵ) we can ensure $l - 1 \geq (1 - \eta)l$ and $\log 16 + l \log(6\omega) < \epsilon/12 \log d$ since $\log(1/q)$ is bounded — recall $q > \epsilon$. It remains only to consider the final summand. This term is at most $(1 - \eta)^4(1 - \epsilon/4)^2 \log d \left(-\frac{\log(q - \epsilon)}{\log(q)} + \frac{\log(1 - \eta)}{\log(q)}\right)$ provided $l > 2/\eta$.

By Lemma 2.7, this is at most $(1 - \eta)^4(1 - \epsilon/4)^2 \log d \left(-1 - \epsilon + \frac{\log(1 - \eta)}{\log(q)}\right)$. We are now able to pick some η depending only on ϵ such that $(1 - \eta)^4(1 - \epsilon/4)^2 \left(-1 - \epsilon + \frac{\log(1 - \eta)}{\log(1 - \epsilon)}\right) < (-1 - \epsilon/4)$;

this completes the proof. $\square$

2.5 The dense case: connector and projector

There are two main aims of this section. The first is to complete the proof of the dense sub-theorem, Theorem 2.5, by extending the almost-compatible partition built in the previous section to a minor through the use of two small helper sets. The second is to prove a slight strengthening of this theorem, Theorem 2.19 which allows us to force certain vertices to lie in the minor; we will call such an object a *rooted minor*. This will be used as an ingredient for the proof of the sparse sub-theorem (Theorem 2.6) in the following section.

Definition 2.12. Let G, H be graphs, and $R = \{r_h\}_{h \in V(H)}$ a collection of distinct vertices of G labelled by the vertices of H which we call *roots*. G contains an H minor *rooted at* R if there exists an H model $(X_h)_{h \in V(H)}$ in G with $r_h \in X_h$.

Definition 2.13. Let H be a graph. A graph G is called *H -minor-prevalent* if $|G| \geq |H|$ and for each set $R = \{r_h\}_{h \in V(H)}$ of distinct vertices of G , G contains a H minor rooted at R .

The condition $|G| \geq |H|$ is required so H -minor-prevalent graphs contain H as a minor; this is witnessed by picking an arbitrary choice of R . Our helper sets will allow us to prove suitable graphs are H -minor-prevalent just as easily as proving that they contain an H minor.

2.5.1 Constructing the helper sets

For later use in Chapter 5, it will be useful to phrase parts of our result in terms of multisets and prove a slightly stronger theorem than we require here. These slight strengthenings do not add any difficulty to the proof. We will require a version of the Chernoff bound; this form follows from [36, Remark 2.5].

Lemma 2.14. *Let X be distributed according to a Poisson binomial distribution (i.e. a sum of n independent random variables, each taking values in $\{0, 1\}$), with n trials and mean xn . Then $\mathbb{P}(X \leq qn) \leq \exp(-2n(q - x)^2)$ for any $q < x$.*

We will also require the following lemma of Myers and Thomason. We omit the proof, though note that a greedy argument works.

Lemma 2.15 ([65, Lemma 4.1]). *Let $\gamma > 0$, and suppose G is a bipartite graph with bipartition (A, B) into non-empty parts, with the property every vertex in A has at least $\gamma|B|$ neighbours in B . Then there is a set $M \subset B$ with $|M| \leq 1 + \lfloor \log_{1/(1-\gamma)} |A| \rfloor$ such that every vertex in A has a neighbour in M , i.e. $A \subset N(M)$.*

Finally, we need the following corollary of Menger's Theorem which shows that given 'positive fraction connectivity', we can find linearly many paths of constant length between any given pair of vertices.

Lemma 2.16 (Short Path Lemma, [66, Lemma 6.2]). *Let G be a graph with connectivity at least $\eta|G|$. Then for all $a \neq b \in V(G)$, G contains at least $\eta|G|/2$ internally vertex disjoint paths of length less than $2/\eta + 1$ between a and b .*

Proof. By Menger's theorem, G has at least $\eta|G|$ internally vertex disjoint paths between a and b . If at least $\eta|G|/2$ of these had length at least $2/\eta + 1$, and hence at least $2/\eta$ internal vertices, this requires at least $|G|$ distinct internal vertices among the paths. But none of these can be a or b , so only $|G| - 2$ vertices can appear as internal vertices. This is a contradiction, and so at least $\eta|G|/2$ of the paths must have length less than $(2/\eta) + 1$. $\square$

Having completed these preliminaries, we are now ready to describe our helper sets. We will construct two sets C and P , which we will call the *connector* and *projector* respectively — these names represent the properties that we will later use. The connector allows us to connect a (small) subset of G using only small subsets within C , and the projector allows us to ‘project’ a large subset in G to a much smaller one inside P . It will then suffice to connect this smaller subset. It turns out randomly chosen subsets can serve each of these roles. The set R represents a choice of ‘roots’ for a minor, which for technical reasons we need to exclude from C and P .

Theorem 2.17. *Let $\eta > 0$. Then there is a constant $D = D_{2.17}(\eta)$ such that the following holds.*

Let G be a graph with $|G| > D$ and $\kappa(G) > 8\eta|G|$, and $R \subset V(G)$ such that $|R| \leq \eta|G|$. Then there exist subsets C, P of $V(G)$ with C, P, R pairwise disjoint which have the following properties.

1. *C and P each have size at most $2\eta|G|$.*
2. *For each pair u, v of distinct vertices from $V(G) \setminus C$, there are at least $2\eta^{2+1/(2\eta)}|G|$ internally vertex-disjoint paths of length at most $1 + \frac{1}{2\eta}$ from u to v , whose internal vertices all lie in C .*
3. *Each vertex $v \in V(G) \setminus P$ has at least $2\eta^2|G|$ neighbours in P .*
4. *For any $Y \subset P$ with $|Y| \leq \eta^2|G|$, and for any $X \subset V(G) \setminus (C \cup P)$, there exists $M \subset P \setminus Y$ such that every vertex in X has a neighbour in M (i.e. $X \subset \Gamma(M)$) and $|M| \leq \lfloor \log_{1/(1-\eta/2)} |X| \rfloor + 1$.*

Proof. We start by constructing C . Place each vertex of $G \setminus R$ into C independently at random with probability η ; with probability at least $\frac{1}{2}$ it holds $|C| \leq 2\eta|G|$. Suppose that $\kappa(G) = \delta|G|$, and note $\delta \geq 8\eta$ by assumption — this means we can assume $\eta < \frac{1}{8}$. Let u, v be fixed vertices of G . By Lemma 2.16, there is a collection of $(\delta - \eta)|G|/2 \geq (\delta - \eta)|G|/4$ internally vertex disjoint paths from u to v of length at most $2/(\delta - \eta) + 1 \leq 1 + 1/2\eta$. Each such path lies internally inside C independently with probability at least $\eta^{1/2\eta}$. By the Chernoff bound, the probability fewer than $\frac{1}{2}\eta^{1+1/2\eta} \leq \eta^{2+1/2\eta}$ such paths between u

and v lie in C is at most $\exp(-\eta^{1+1/\eta}|G|)$. Thus the probability that for some u, v there are fewer than $\frac{1}{2}\eta^{1+1/2\eta}$ such paths contained in C is less than $|G|^2 \exp(-\eta^{1+1/\eta}|G|) < \frac{1}{2}$ for D sufficiently large. In particular, there exists some choice of C with $|C| < 2\eta|G|$ which also satisfies property (2). Fix some such C .

We now construct P in a similar manner. Place each vertex of $G \setminus (R \cup C)$ into P independently with probability η ; as before $\mathbb{P}(|P| \leq 2\eta|G|) \geq \frac{1}{2}$. We now move onto property (3). Let $v \in V(G)$ be fixed; since G is $8\eta|G|$ connected, v has at least $5\eta|G|$ neighbours outside of $C \cup R$. Each such neighbour lies in P independently with probability η . Using the Chernoff bound, the probability fewer than $2\eta^2|G|$ of these neighbours lie in P is at most $\exp(-\eta^3|G|)$; taking D large the probability there is a $v \in V(G)$ with fewer than $2\eta^2|G|$ neighbours in P is less than $\frac{1}{2}$, and so we can pick some P with $|P| < 2\eta|G|$ satisfying property (3). Fix some such choice.

To show that property (4) also holds for the above sets, we apply Lemma 2.15 with $A = X$, $B = P \setminus Y$ and $\gamma = \eta/2$ — each vertex has at least $2\eta^2|G|$ neighbours in P , hence at least $\eta^2|G| \geq \eta|P|/2$ neighbours in $P \setminus Y$. Thus there is a set $M \subset P$ with $|M| \leq \lfloor \log_{1/(1-\eta/2)} |X| \rfloor + 1$ and every vertex of X having a neighbour in M , concluding our proof. $\square$

The next theorem will use the helper sets in order to convert an almost- H -compatible partition and a disjoint collection of roots into a rooted H minor. For later use, this theorem has been written in terms of multisets as this adds no extra difficulty. For a multiset F , $|F|$ will denote the number of elements (counted with multiplicity), and $m_F(e)$ will denote the number of copies of e that appear in F .

Theorem 2.18. *Given $\eta > 0$ there is a constant $D = D_{2.18}(\eta)$ with the following property. Let G be a graph, and $R \subset V(G)$ such that $|G| \geq D|R|$ and $\kappa(G) > 8\eta|G|$. Let C, P be the sets given by Theorem 2.17 applied with these parameters. Suppose that $(V_r)_{r \in R}$ is a partition of $G - C - P - R$ into $|R|$ parts, and F is a multiset of pairs from R satisfying $|F| \leq |R|/\eta$. Then there are disjoint subsets $(U_r)_{r \in R}$ with the following properties.*

1. $V_r \cup \{r\} \subset U_r$ for all $r \in R$.
2. $G[U_r]$ is connected for all $r \in R$.
3. There are at least $m_F(rs)$ edges of G from $U_r \setminus V_r$ to $U_s \setminus V_s$ for all r, s .

Proof. Throughout, we will assume D is sufficiently large that all claimed inequalities hold, and also $D > 1/\eta$, but do not give an explicit choice. Our first step will be to build up sets $(M_r)_{r \in R}$ such that $|\cup M_r| \leq \eta^2|G|$, and every vertex of $V_r \cup \{r\}$ has a neighbour in M_r . We will do this one at a time, and ensuring that $|M_r| \leq \log_{1-\eta/2} |V_r \cup \{r\}| + 2$. Suppose that we have already constructed such sets for some subsets $(V_r)_{r \in R_1}$; note that $(\sum_{r \in R_1} \log_{1-\eta/2} |V_r \cup \{r\}| + 2) \leq \eta^2|G|$ by the convexity of $\log$ and taking $D > \eta^{-4}$ since $\sum_{r \in R} \log |V_r \cup \{r\}| \leq |R| \log(|G|/|R|) \leq (\log D)|G|/D$. Let $Y = \cup_{r \in R_1} M_r \cup R$; note that

$|Y| \leq \eta^2|G|$. Pick some $s \in R \setminus R_1$. By property (4) of Theorem 2.17, applied with $X = V_s \cup \{s\}$ and this Y we can obtain a subset M_s as claimed. We can thus continue the process to build all sets $(M_r)_{r \in R}$.

Our next aim will be to construct disjoint sets $N_r \subset C$ such that $G[M_r \cup N_r]$ is connected for all r ; since every vertex of $V_r \cup \{r\}$ has a neighbour in M_r this will also mean $G[\{r\} \cup V_r \cup M_r \cup N_r]$ is connected. We will construct such sets one at a time, in such a way each N_r uses at most $|M_r|/(2\eta)$ vertices. Suppose we have already constructed such sets for $r \in R_1$; let $X = \cup_{r \in R_1} N_r$ and note $|X| \leq \eta^{2+1/2\eta}|G|$ taking D large. To construct our next N_s for some $s \notin R_1$, order the vertices of M_s as $v_1, \dots, v_{|M_s|}$.

Pick some path inside C between v_1 and v_2 of length at most $1/(2\eta)$ which avoids X ; such a path exists by property (2). Add its internal vertices to both X and N_s , then repeat for v_2 and v_3 and so on — note it remains the case $|X| \leq \eta^{2+1/2\eta}|G|$. After this procedure, we have $|N_s| \leq |M_s|/(2\eta)$ as claimed, and further $G[N_s \cup M_s]$ is connected since its vertex set can be constructed as a union of paths in G .

Our final task is to add the additional edges, which we do by modifying the N_r . Let $X_1 = \cup_{r \in R} N_r$, and note $|X_1| \leq \eta^{2+1/2\eta}|G|$ as before. We will add each pair $rs \in F$ one at a time (using at most $2/\eta$ additional vertices for each) so that the union of the internal vertices of these new paths will have total size at most $|F|/(2\eta) < 16|R| < \eta^{2+1/2\eta}|G|$.

Let X_2 consist of the union of the internal vertices already used, and $X = X_1 \cup X_2$. By condition (2) again, picking arbitrarily some $v \in M_r$ and $w \in M_s$ there is a path from v to w inside C of length at most $1/(2\eta) + 1$ whose internal vertices are disjoint from X (and choose one such path which is not the edge vw). Add these internal vertices to N_r (and X_2); note that still $G[M_r \cup N_r]$ is connected. Repeat for each element of F . Taking $U_r = V_r \cup \{r\} \cup M_r \cup N_r$ completes our proof. $\square$

2.5.2 Converting compatible partitions into minors

We are now ready to use these helper sets to convert the compatible partitions constructed earlier into (rooted) minors, and so complete the aims of this section. It is easy to see that the following theorem implies our dense sub-theorem, Theorem 2.5, since being H -minor-prevalent (cf. Definition 2.13) implies containing an H minor.

Theorem 2.19. *Given $0 < \epsilon < \frac{1}{8}$ there exists $D = D_{2.19}(\epsilon)$ such that if H is a graph with average degree $d > D$ and G is a graph of density at least $p + \epsilon$ for some $\epsilon < p < 1 - \epsilon$ such that also $|G| \geq |H| \sqrt{\log_{1/q} d}$ (where $q = 1 - p$) and $\kappa(G) \geq \epsilon|G|$, then G is H -minor-prevalent.*

Proof. Throughout, we may assume that D is arbitrarily large, and hence that $|H|$ and $|G|/|H|$ may both be taken to be larger than any required function of ϵ ; in particular $|G| \geq |H|$. Let $R = (r_h)_{h \in V(H)}$ be a set of roots in G . Taking $D > D_{2.17}(\epsilon/20)$ we may apply Theorem 2.17 to G with $\eta = \epsilon/20$ and this choice of R to obtain sets C, P as

described therein.

Consider the graph $G' = G[V(G) \setminus (C \cup P \cup R)]$. We have removed at most $\epsilon/4$ proportion of vertices from G , and so $|G'| \geq (1 - \epsilon/4)|H|\sqrt{\log_{1/q} d}$. Further, $e(G') \geq e(G) - (\epsilon|G|/4)(|G| - 1)$ and so the density of G' is $2e(G')/(|G'|(|G'| - 1)) \geq p + \epsilon/2$.

We would like to apply Lemma 2.11 to G' ; but unfortunately G' has too few vertices for a naive application thereof. We avoid this problem by replacing p with $p' = p + \epsilon/4$, and ϵ with $\epsilon' = \epsilon/4$. The density of G' is at least $p' + \epsilon'$, and by Lemma 2.7 $\sqrt{\frac{\log(1/q')}{\log(1/q)}} \leq \frac{1}{(1-\epsilon/4)}$, where $q' = 1 - p'$. Thus $|G'| \geq |H|\sqrt{\log_{1/q'} d}$, and so we can apply Lemma 2.11 to obtain a $|H|d^{-\epsilon/3}$ -almost- H -compatible partition $(V_h)_{h \in V(H)}$ of $V(G')$. It remains to extend this compatible partition to a minor.

Let F be the set of pairs hh' such that hh' is an edge of H , but there is no edge of G from V_h to $V_{h'}$; by assumption $|F| \leq |H|d^{-\epsilon/3} \leq |R|/\eta$. Taking also $D > D_{2.18}$, we can apply Theorem 2.18 to extend (V_h) to a new collection $(U_h)_{h \in V(H)}$ of disjoint subsets such that each $G[U_h]$ is connected, $r_h \in U_h$ and also for each pair $hh' \in F$ there is an edge from U_h to $U_{h'}$. Thus the U_h form a H minor in G rooted at R , as required. $\square$

This theorem shows us the utility of the connector and projector. The connector is of natural use: it lets us sequentially add many disjoint paths, and thus either add missing edges or connect sets. The projector allows us to use the connector — we cannot hope to find $|G|$ disjoint paths to directly connect the sets (V_r) , but after projecting them to much smaller sets we can find sufficient paths.

2.6 The sparse case

To complete the proof of our main Theorem 1.10 it now remains only to prove the sparse sub-theorem, Theorem 2.6, and that is the aim of this section. We will be using Theorem 2.19 of the previous section as an input to our proof.

Our proof will be broken down into a sequence of lemmas. While our first proof of this theorem was not written in this way, in preparation of results from Chapter 5 it became apparent that breaking the proof into a sequence of lemmas greatly simplified proofs in more general settings, as well as serving as helpful signposts for the present case.

We will begin this section by stating our lemmas, and showing how together they imply Theorem 2.6. After this motivational interlude, we will then prove each of our three lemmas.

Lemma 2.20. *Given $0 < \epsilon < \frac{1}{8}$ there is a $D = D_{2.20}(\epsilon)$ such that for any $d > D$, the following holds.*

Let H be a graph with average degree d , and let $n \geq \epsilon|H|\sqrt{\log d}$. Let G be a bipartite graph with vertex classes A, B satisfying $|A| \geq Dn, |B| \leq n/\epsilon$ and every vertex in A has at

least ϵn neighbours in B . Then G contains a H minor.

Lemma 2.21. *Let $k > 0$. Then there is a constant $D = D_{2.21}(k)$ such for all $m > D$, the following holds. Let G be a graph with properties (1) – (4).*

1. $e(G) \leq m|G|$
2. Every edge of G is in at least $m - 1$ triangles
3. $|G| \geq Dm$
4. G has no isolated vertices

Then G contains one of the following two substructures:

- (A) A bipartite subgraph $G' = (A, B)$ with $|A| \geq |G|/3$, $|B| \leq 6km$ and every vertex in A has at least $m/7$ neighbours in B , or
- (B) Disjoint subsets $T_0, \dots, T_k$ each of size at most $6m$, with $G[T_i]$ having connectivity at least $m/40$ and minimum degree at least $3m/4$.

Lemma 2.22. *Let $k > 0$. Then there is a constant $D = D_{2.22}(k)$ such that the following holds for all integers $t, m > Dt$. Let G be a graph with connectivity at least Dt , and $(T_i)_{i=0}^k$ disjoint subsets of $V(G)$ satisfying properties (i) – (ii).*

- (i) $|T_i| \leq 6m$
- (ii) $\kappa(G[T_i]) \geq m/40$

Then G contains a collection $\{P_{i,j} : 1 \leq i \leq k, 1 \leq j \leq t\}$ of vertex-disjoint paths, where $P_{i,j}$ is from $a_{i,j} \in T_0$ to $b_{i,j} \in T_i$, such that $|\bigcup V(P_{i,j}) \cap T_i| \leq 81kt$. In addition, there exist disjoint connected subsets $C_j \subset T_0 \cup_i P_{i,j}$ such that each $P_{i,j} \subset C_j$.

Having stated these key lemmas, we can now show how they imply our sparse case Theorem 2.6; it will then only remain to prove each of our three lemmas to complete the proof of our main Theorem 1.10.

Proof of Theorem 2.6. Reduce ϵ if necessary so $\epsilon < \frac{1}{8}$, and let $l = \lceil 8/\epsilon \rceil$, $k = \binom{l}{2}$ and $D = \max(D_{2.21}(k, \epsilon), D_{2.22}(k), D_{2.20}(\epsilon/6k), 5D_{2.19}(1/300)/\epsilon)$. Since G has properties (1) – (4) of Lemma 2.21 applied with the current k and ϵ , G has one of the substructures (A) or (B) described therein.

If G has structure (A), applying Lemma 2.20 with the current ϵ replaced by $\epsilon/6k$ we deduce $G \succ H$; we can thus assume G has substructure (B). In this case, we can apply Lemma 2.22 (with $t = |H|$) to obtain sets (T_i) and paths $(P_{i,j})$ as described in that lemma. We relabel our subsets as T_0 and $T_{i,j} : 1 \leq i \leq j \leq l$ arbitrarily, this extends to relabel the paths from T_0 to $T_{i,j}$ as $P_{i,j,k} : 1 \leq k \leq t$ and also their endpoints as $b_{i,j,k}$; we also relabel the vertices of H as $\{1, \dots, |H|\}$.

Let $V_1, \dots, V_l$ be some equipartition of $V(H)$ (i.e. a partition of $V(H)$ into l parts each of size $\lfloor n/l \rfloor$ or $\lceil n/l \rceil$), and let $H_{i,j} = H[V_i \cup V_j]$ for each $1 \leq i < j \leq l$. We have

$|H_{i,j}| \leq 3n/l$ since n/l is large, and further $H_{i,j}$ has average degree at most ld . Add edges to $H_{i,j}$ if necessary so it has average degree between $\epsilon d/5$ and ld — note it remains the case $H \subset \cup_{i < j} H_{i,j}$.

Fix now some indices $i < j$, and let $X = \bigcup_{r,s,k} (T_{i,j} \cap V(P_{r,s,k}))$ — we have $|X| \leq 81k|H|$. Let $G' = G[T_{i,j} \setminus X]$. Since we have only removed a small proportion of vertices, G' is $m/80$ connected and has average degree at least $2m/3$. We would like to apply Theorem 2.19 to deduce that G' contains $H_{i,j}$ as a minor rooted at $(b_{i,j,k})_{k \in V(H_{i,j})}$. Let $p + \frac{1}{300}$ be the density of G' , and note $p \geq \frac{3m}{5|G'|}$ since G has minimum degree at least m . Thus applying Inequality 1.3 we have

$$|G'| \geq \frac{3m}{5p} \geq \frac{2}{p} |H_{i,j}| \sqrt{\log(ld)} \geq |H'| \sqrt{\log_{(1/1-p)}(ld)}.$$

Further, $H_{i,j}$ has average degree at least $\epsilon d/5 > D_{2.19}(1/300)$ so as desired there is an $H_{i,j}$ minor in G' rooted at $b_{i,j,k}$. Call the parts of such a model $X_{i,j,k}$. Recalling again our adjacency properties and the subsets C_k from Lemma 2.22, the sets $C_k \cup \bigcup_{i,j:k \in V(H_{i,j})} X_{i,j,k}$ form a H model in G . $\square$

Proof of Lemma 2.20. Let $0 < q(\epsilon) < \epsilon$ be such that $\epsilon^2 \sqrt{\log(1/2q)} \geq 1$ and so $\epsilon n \geq |H| \sqrt{\log_{1/(2q)} d}$, and take some D satisfying $Dq\epsilon^3 > 1$, and also $D > D_{2.19}(q)$. Let $G_0 = G$, and $M_0 = G[B]$ the empty graph on B . To obtain G_{i+1} from G_i , we pick some vertex $a \in A \cap V(G_i)$, and contract it to a vertex b of minimum degree in $G_i[N(a)]$. Suppose that this minimum degree is $(1 - q_i)d(a)$. Setting $M_{i+1} = G_i[B]$, M_{i+1} has at least $q_i d(a) \geq q_i \epsilon n$ more edges than M_i . Continue in this fashion to build $G_1, \dots, G_{|A|}$. Further, each G_i and hence M_i is a minor of G .

Suppose first that each $q_i \geq q$. This means $e(M_{|A|}) \geq |A|q\epsilon n > Dq\epsilon n^2 > \binom{|B|}{2}$. This is a contradiction, since each M_i is a simple graph. So it must be the case that at some stage $q_i < q$; say we contract a vertex a at step i . Then $M = M_i[N(a)]$ is a minor of G with at least ϵn vertices, and minimum degree at least $(1 - q)|M|$; in particular since $q < \frac{1}{3}$ this graph is $\frac{1}{3}|M|$ connected, and has density at least $(1 - 2q) + q$. Thus by applying Theorem 2.19 with $\epsilon = q$, we deduce M (and hence G) contains H as a minor. $\square$

Proof of Lemma 2.21. We will first attempt to build disjoint non-empty $S_0, \dots, S_k$ such that $|S_i| \leq 6m$ and $\delta(G[S_i]) \geq 5m/6$. We will do so one at a time, and suppose that we have already constructed $S_0, \dots, S_r$ with $B = \bigcup_{i \leq r} S_i$ — note $|B| \leq 6km \leq |G|/3$ by taking $D > 6k$.

Let A consist of all vertices of degree at most $6m$ which do not lie in B . There are at most $|G|/3$ vertices of G with degree at least $6m$ since $2e(G) \leq 6m \cdot |G|/3$, and hence $|A| \geq |G|/3$. If there is a vertex $a \in A$ with fewer than $m/6 - 1$ neighbours in B , we can take $S_{r+1} = N(a) \setminus B$ — note that every edge being in at least $m - 1$ triangles implies $\delta(G[N(a)]) \geq m - 1$, and since a is not isolated this graph is not empty. So either we

can build S_{r+1} , or every $a \in A$ has at least $m/6 - 1 \geq m/7$ neighbours in B . Taking $G' = G[A, B]$ this gives a subgraph as in (A). So we can assume this does not happen, and hence we can build $S_0, \dots, S_k$ in this fashion. We will now replace the S_i with subsets in order to obtain the properties of (B).

If S_i is $m/40$ connected, we can take $T_i = S_i$, and so we assume this does not occur. Then there is a cutset X of size at most $m/40$. Let S'_i be a smallest component of $S_i \setminus X$. We must have $|S'_i| \leq 3m$, and further as all neighbours in $G[S_i]$ of S'_i lie in $S'_i \cup X$, we have $\delta(G[S'_i]) \geq 5m/6 - m/40$. If S'_i is $m/40$ connected we are done, taking $T_i = S'_i$, otherwise we can obtain S''_i in the same fashion with minimum degree at least $5m/6 - 2m/40$ and at most $3m/2$ vertices. If S''_i were not $m/40$ connected, we could build an S'''_i with minimum degree at least $5m/6 - 3m/40 > 3m/4$ but at most $3m/4$ vertices — this is a contradiction. So we can take $S''_i = T_i$ in this case.

Performing this procedure for each $0 \leq i \leq k$ yields the desired disjoint subsets (T_i) with the properties of (B), completing our proof. $\square$

Proof of Lemma 2.22. We take $D = 10^6 k$. For each $1 \leq i \leq t$ pick some arbitrary subset $R_i \subset T_i$ of size t (note that $|T_i| \geq m/40 \geq t$), and let $R = \bigcup_{i \geq 1} R_i$. By Menger's theorem there exist $k|H|$ vertex disjoint paths from T_0 to R . Label the paths as $\{Q_{i,j} : 1 \leq i \leq k, 1 \leq j \leq t\}$, where $Q_{i,j}$ is a path from T_0 to some vertex $b_{i,j} \in T_i$. Taking subpaths if necessary, we may assume each path $Q_{i,j}$ meets T_0 in exactly one place, and let $a_{i,j}$ be this vertex.

It remains to modify these paths to also gain the intersection conditions. We will do so for one set T_i at a time. Certainly $|T_0 \cap \bigcup_{i,j} V(Q_{i,j})| = k|H| < 81kt$. If this also holds for all other indices we are done, so assume that l is such that $|T_l \cap \bigcup_{i,j} V(Q_{i,j})| \geq 81kt$. Let A consist of all pairs i, j for which $Q_{i,j}$ intersects T_l .

For each $a \in A$, suppose that Q_a first meets T_l at vertex x_a , and last meets it at y_a when traversing the path from a_a to b_a . Let $Q_{a,1}$ denote the subpath from a_a to x_a , and $Q_{a,3}$ the subpath between y_a and b_a . We will replace Q_a with a new path consisting of $Q_{a,1}$, $Q_{a,3}$ and a bridging path $Q_{a,2}$ between them which is entirely contained inside T_l , and further will have length at most 80. We will also require all $Q_{a,2}$ are disjoint from each other for different a , so that the paths $Q_{i,j}$ are always internally vertex disjoint. The new path intersects all other T_s in at most as many places as before, since outside T_l it is a sub-path. Therefore, this process cannot violate the condition at any new indices and so after a finite number of applications of this procedure we terminate with a collection of paths with the desired properties, which we will call $P_{i,j}$.

It remains to show we can build these paths $Q_{a,2}$. For each $a \in A$, by Lemma 2.16 there are at least $|T_k|/80 \geq m/3200$ internally vertex disjoint paths between x_a and y_a of length at most 80. Let X consist of the internal vertices of already chosen $Q_{a',2}$ together with all $x_{a'}, y_{a'}$. Together this is at most $81kt$ vertices. At least one of the $m/3200$ short

paths must avoid X (apart from at x_a and y_a) since $81kt < m/3200$, and we make such a choice as $Q_{a,2}$. Doing so for all $a \in A$ implies the first part of our result, since the $Q_{a,2}$ form a disjoint collection.

To build the set C_1 , we start with $C_1 = \bigcup_{i=1}^k P_{i,1}$, and let $B = \bigcup_{1 \leq i \leq k, 1 \leq j \leq t} \{a_{i,j}\}$. Since T_0 has connectivity at least $m/40$, there is some path of length at most 80 between $a_{1,1}$ and $a_{2,1}$ within T_0 which internally avoids B — applying Lemma 2.16 there are at least $m/3200$ paths of length at most 80 inside T_0 between $a_{1,1}$ and $a_{1,2}$, and each vertex in B is an internal vertex of at most one of them. Add its internal vertices to both B and C_1 , and still $|B| \leq kt + 80 < 81kt$. Repeat to build a path between $a_{2,1}$ and $a_{3,1}$ and so on up to $a_{k-1,1}$ and $a_{k,1}$. After this procedure, $|B| \leq kt + 80k$. We can therefore continue to build sets $C_2, \dots, C_t$ in this fashion, since at all steps $|B| \leq 81kt < m/3200$. $\square$

Chapter 3

An extension to almost-multipartite graphs

In this chapter, we will improve the bound on the minor extremal function given by Chapter 2 when our desired minor has some additional structure. In previous work, Myers and Thomason [49] showed that $c(H) = (\alpha\gamma(H) + o_t(1))t\sqrt{\log t}$ holds for all graphs H with t vertices, where $\gamma(H)$ is the following structural parameter; we note this result asymptotically computes the extremal function when $\gamma(H)$ is bounded away from zero for some family.

Definition 3.1. Let H be a graph with t vertices. Define $\gamma(H)$ to be the minimum value of $\frac{1}{t} \sum_{v \in V(H)} w(v)$, taken over all functions $w : V(H) \rightarrow \mathbb{R}^{>0}$ satisfying the following inequality:

$$\sum_{uv \in E(H)} t^{-w(u)w(v)} \leq t. \quad (3.1)$$

This parameter is hard to compute in general, however a constant weighting witnesses that for all graphs with t vertices and td edges, $\gamma(H) \leq \sqrt{\frac{\log d}{\log t}}$. In effect, this parameter ‘picks out’ the more important vertices for building a minor. For K_t , all vertices are given the same weight. In a complete bipartite graph $K_{\beta t, (1-\beta)t}$ (where $\beta < \frac{1}{2}$), the vertices in the smaller class are optimally given a weight $\sqrt{\frac{1-\beta}{\beta}}$, and those in the larger class $\sqrt{\frac{\beta}{1-\beta}}$ to derive $\gamma(K_{\beta t, (1-\beta)t}) = (1 + o(1))2\sqrt{\beta(1-\beta)}$.

The result of Myers and Thomason only asymptotically computes the extremal function for graphs whose average degree grows at least polynomially in their order; in the sparse regime we are interested in, their result only shows $c(H) = o(t\sqrt{\log t})$.

Our first main result of this chapter, Theorem 1.11, can be viewed as an extension of the result of Myers and Thomason to this sparser regime, where our graphs are allowed to

have constant average degree. For technical reasons we will have to introduce a modified parameter $\gamma_r(H)$, although we will later show (cf. Theorem 3.10) this parameter is equivalent to $\gamma(H)$ for almost all graphs after a suitable rescaling. We will defer discussion of this parameter until the next section.

Theorem 1.11. *Let $\epsilon > 0$, and suppose $r \geq 1$ is an integer. Then there is a constant C such that for all graphs H with $e(H)/|H| = d > C$ and $\gamma_r(H) > \epsilon$, every graph G with $e(G)/|G| \geq (\alpha + \epsilon)\gamma_r(H)|H|\sqrt{\log d}$ contains H as a minor.*

Another family for which $\gamma(H)$ converges to zero is complete bipartite graphs $K_{s,t}$ where $s \ll t$. Using a weighting which is different on the size s class to the size t class, it can be shown $\gamma(K_{s,t}) \leq \frac{2\sqrt{st}}{s+t}$ and this is tight. Thus if $s < t$ and s/t is bounded away from zero, it holds $c(K_{s,t}) = (1 + o(1))2\alpha\sqrt{st \log(s+t)} = (1 + o(1))2\alpha\sqrt{st \log t}$.

We will show this remains true under the weaker condition $s/t = \omega(1/\log t)$. This new result cannot be extended much; once $s \log t/t$ becomes small, this violates the lower bound $c(K_{s,t}) > t - 1/2$ given by complete graphs on t vertices. While our new theorem could be proven using a slight generalisation of Theorem 1.11, we provide a separate statement and proof (Theorem 1.12). Our argument in this special case is more simple, and motivates the statements and proofs of the general case. In fact our upper bound also holds for books $K_{s,t}^*$ (the graph obtained from $K_{s,t}$ by placing a clique inside the size s class) and so we state it in this more general form.

Theorem 1.12. *Let $\epsilon > 0$. Then there is a constant C such that for all t, f satisfying $C < f \leq \log t$, every graph G with $e(G)/|G| \geq (2\alpha + \epsilon)t\sqrt{f}$ contains $K_{ft/\log t, t}^*$ as a minor. In particular, $c(K_{ft/\log t, t}^*) \leq (2\alpha + \epsilon)t\sqrt{f}$.*

Our upper bounds use two key new ideas. The first occurs in the dense case, where we build up an almost-compatible partition for a graph in several ‘classes’, in order to treat certain vertices differently (and hence make a gain through our structural parameter compared to using a constant weight). The second new idea is in the sparse case, where additional care is needed to break our graph into smaller parts, each of which is easier to find as a minor.

The bounds matching is not a coincidence. In previous work, where $\gamma(H)$ is bounded away from zero, the lower bound comes from choosing a random partition of $G(n, p)$ where the part size of a vertex is proportional to the weight given to it by γ . In our new setting, this does not work directly since random graphs no longer provide good lower bounds for sparse graphs. However, if we modify this to apply γ to blowups of random graphs, we can extend to the sparse regime.

After completing the proofs of these upper bounds, we will move onto lower bounds. We will match each of our two upper bounds with a lower bound, and the proofs of these are the focus of Sections 3.5 and 3.6. These lower bounds use significant extensions of the argument of Norin et al. [51] to use some additional structure of the graphs; a key

modification is building two families of sets simultaneously in Theorem 3.2. The first of these bounds is stated in terms of a new family $\mathcal{D}(\vec{\sigma}, \mathbf{D})$ to be defined in the next section, but for now it can be thought of as holding for (a suitable notion of) almost all graphs with $\gamma_r(H) = \gamma = \gamma(\vec{\sigma}, \mathbf{D})$.

Theorem 3.2. *Let $\epsilon > 0$. Then there is a $D = D_{3.2}(\epsilon)$ such that the following holds for all $t > d > D$.*

Let $\vec{\sigma}$ be a weight vector with all $\sigma_i > \epsilon$, and $\mathbf{D}$ a matrix with entries in $[-\infty, 1]$ such that $\gamma(\vec{\sigma}, \mathbf{D}) \geq \epsilon$. Then all but at most 2^{-t} proportion of elements $H \in \mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$ have $c(H) \geq (\alpha - \epsilon)\gamma(\vec{\sigma}, \mathbf{D})t\sqrt{\log d}$, where $\alpha = 0.319\dots$ is the constant from Definition 1.6 on page 5.

Theorem 1.13. *Let $\epsilon > 0$. Then there is a constant C such that for all $C < f \leq \log t$, there is a graph satisfying $e(G)/|G| \geq (2\alpha - \epsilon)t\sqrt{f}$ and no $K_{ft/\log t, t}$ minor. In particular, $c(K_{ft/\log t, t}) \geq (2\alpha - \epsilon)t\sqrt{f}$.*

These results together form the basis of my pre-print [72]. We will begin with the bipartite case, since this contains the key new idea of giving different vertices different size ‘parts’ in the compatible partition. The size of these parts is dependent on the weight given by a (variant of) the parameter γ . For instance, in a bipartite graph with unequal class sizes, the vertices in the smaller class have higher degree, and so are more crucial for the minor. We therefore give them bigger parts relative to the vertices of the larger class. The general case then extends these ideas.

3.1 The parameter $\gamma_r(H)$, and the family $\mathcal{D}(\vec{\sigma}, \mathbf{D})$

In this section, we will provide a formal introduction to the new parameter $\gamma_r(H)$ used in our upper bounds, and the family $\mathcal{D}(\vec{\sigma}, \mathbf{D})$ used in our lower bounds. We begin with $\gamma_r(H)$.

Definition 3.3. A *weighted partition* of a graph H into r parts is a collection $(P_i, w_i)_{i=1}^r$, where $(P_i)_{i=1}^r$ is a vertex partition of H , and the w_i are non-negative real numbers (called *weights*).

Definition 3.4. Let H be a graph with t vertices and td edges. A weighted partition $(P_i, w_i)_{i=1}^r$ of $V(H)$ satisfies the *gamma inequality* if the following holds:

$$\sum_{i \leq j} e(P_i, P_j) d^{-w_i w_j} \leq t. \quad (3.2)$$

We define $\gamma_r(H)$ to be the minimum value of $\frac{1}{t} \sum_i w_i |P_i|$ over all weighted partitions (P_i, w_i) of H into r parts which satisfy the gamma inequality. We will give an example of computing this parameter at the end of this section.

This is analagous to $\gamma(H)$, except we force the weight function to take at most r

distinct values. The base of the exponent is also changed from t to d ; we do this with the benefit of hindsight and since our results then apply to graphs with $\gamma_r(H) > \epsilon > 0$.

Remark 3.5. *It is sometimes taken that H has average degree d , instead of having td edges (and hence average degree $2d$). This does not make a significant difference to any of our results, since d only appears inside logarithms in bounds, and $\log 2d \leq 1 + \log d$ — i.e. this only affects bounds by a multiplicative $(1 + o(1))$ factor which is easily absorbed in error terms.*

We now define structural parameters $\vec{\sigma}$ and $\mathbf{D}$ which will be used in our lower bounds; defining these along with the upper bound motivates them.

Definition 3.6. Given a graph H with t vertices and td edges for some $d > 1$, and a (weighted) partition $(P_i)_{i=1}^r$ of $V(H)$, we define quantities $\mathbf{D} = D_{i,j}$ and $\vec{\sigma}$ as follows. The $D_{i,j}$ take values either from $(-\infty, 1]$, or the formal symbol $-\infty$ with the property $d^{-\infty} = 0$ for any $d > 1$, and are chosen such that $e_H(P_i, P_j) = td^{D_{i,j}}$ for each $1 \leq i, j \leq r$. We define a length r vector $\vec{\sigma}$ by $\sigma_i = |P_i|/t$.

The formal symbol $-\infty$ ‘smaller than all real numbers’ is introduced to permit no edges to be present between some pairs. We can rephrase the gamma inequality for such a graph and partition in terms of the structural parameter $\mathbf{D}$ as below:

$$\sum_{i \leq j} d^{D_{i,j} - w_i w_j} \leq 1. \quad (3.3)$$

Remark 3.7. *The constant on the right-hand side of inequality 3.3 is effectively arbitrary. Suppose that a weighting w satisfies the inequality, but with 1 replaced with a constant $C > 1$. Then increasing each w_i to $(1 + \sqrt{\log_d C})w_i$, this new weighting satisfies inequality (3.3) as stated above, and so as in Remark 3.5 γ_r only changes by a $(1 + o_d(1))$ factor. The choice 1 was made for simplicity and to naturally relate to $\gamma(H)$.*

For a weighted partition to satisfy the gamma inequality (3.3), it is necessary that $w_i w_j \geq D_{i,j}$ for all i, j . Further, if we have a weighting satisfying all of these inequalities, it satisfies a version of the Gamma inequality with the quantity 1 on the right hand side replaced with $\binom{r+1}{2}$. For sufficiently slow-growing $r = d^{o(1)}$, this is equivalent by Remark 3.7. Given this, one may be tempted to replace $D_{i,j}$ by $\max(D_{i,j}, 0)$ and consider graphs with ‘up to $td^{D_{i,j}}$ edges’ to avoid introducing $-\infty$ without affecting γ_r by much. However, this would change the meaning of our lower bounds, and make them less useful — we have results for ‘almost all’ graphs with certain structure, so we need to restrict to an exact edge count since there are many more graphs with t vertices and t edges than t vertices and no edges.

We now turn our attention to lower bounds. While $\gamma_r(H)$ works well for proving upper bounds, the varied structures of graphs with fixed γ_r prove problematic for proving lower bounds. We therefore introduce a new parameter $\gamma(\vec{\sigma}, \mathbf{D})$ and families $\mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$ which

have the restricted structure mentioned earlier, allowing us to prove lower bounds with a sensible notion of ‘almost all graphs’.

Definition 3.8. Suppose that $\vec{\sigma} = (\sigma_1, \dots, \sigma_r) \in \mathbb{R}_{\geq 0}^r$ is a weight vector (i.e. $\sum \sigma_i = 1$), and $\mathbf{D} = (D_{ij})_{i=1}^r$ is a symmetric matrix, where $D_{ij} \in [-\infty, 1]$, where $[-\infty, 1]$ denotes $\{-\infty\} \cup (-\infty, 1]$. Define

$$\gamma(\vec{\sigma}, \mathbf{D}) = \min\{\vec{\sigma} \cdot \vec{\beta} : \beta_i \geq 0, \beta_i \beta_j \geq D_{ij} \text{ for } 1 \leq i, j \leq r\}$$

, where $\cdot$ denotes the usual inner product.

The definition of $\gamma(\vec{\sigma}, \mathbf{D})$ is broadly similar to that of γ_r , except that we restrict each quantity $td^{D_{ij}-w_i w_j}$ to be at most 1, rather than their sum. Bounding the contribution from each pair is often easier than bounding their sum, and for sufficiently small r this is equivalent as noted after Remark 3.7. This equivalence only holds for sufficiently slow-growing r ; we will consider only the regime where r is constant and d is large in this thesis.

Definition 3.9. Let $\vec{\sigma} = (\sigma_1, \dots, \sigma_r)$ and $\mathbf{D}$ be as in Definition 3.8. Define the class of graphs $\mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$ to be all graphs G with $V(G) = \{1, \dots, t\}$ and the following property.

Let $s_i = \lfloor \sigma_i t \rfloor$ for $i \leq i_0$, and $\lceil \sigma_i t \rceil$ for $i > i_0$, where i_0 is chosen so that $\sum s_i = t$. Let $S_1 = \{1, \dots, s_1\}$, $S_2 = \{s_1 + 1, \dots, s_1 + s_2\}$ and so on. Then G has $\lfloor td^{D_{ij}} \rfloor$ edges between S_i and S_j .

Let $\mathcal{D}(\vec{\sigma}, \mathbf{D})$ be the union of all classes $\mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$.

From the above comments, if H is a member of $\mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$, and further $\max D_{ij} = (1 + o(1))$ (this is required so that d is the base in both definitions), then we see $\gamma_r(H) \leq \gamma(\vec{\sigma}, \mathbf{D})(1 + o(1))$ by considering the partition which witnesses $H \in \mathcal{D}(\vec{\sigma}, \mathbf{D})$, and a slightly increased weighting. We have thus defined all quantities in the statement of Theorem 3.2; we remark that r is implicitly bounded by $1/\epsilon$ since all $\sigma_i \geq \epsilon$.

We finish this section with a theorem which shows Theorem 1.11 and lower bound Theorem 3.2 asymptotically match, in the sense $\gamma(\vec{\sigma}, \mathbf{D})(t, d)$ consists almost entirely of graphs for which the bounds are asymptotically equivalent. This also proves our earlier comment that γ_r and γ are equivalent for almost all graphs after a suitable rescaling. The proof of this will appear in Section 3.5.

Theorem 3.10. Let $\gamma, \epsilon, r > 0$. Then there is a $D = D_{3.10}(\gamma, \epsilon, r)$ such that for all $t > d > D$, all $\vec{\sigma}$ weight vectors of length r and $\mathbf{D}$ matrices satisfying the conditions of Definition 3.8 such that $\gamma(\vec{\sigma}, \mathbf{D}) \geq \gamma$; for all but at most ϵ proportion of graphs H from $\mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$ the following holds:

$$(1 - \epsilon)\gamma(\vec{\sigma}, \mathbf{D}) \leq \gamma(H)\sqrt{\log t / \log d} \leq \gamma_r(H) \leq \gamma(\vec{\sigma}, \mathbf{D})(1 + \epsilon).$$

Example

We now give an example where our framework can be applied. Let H be a graph with vertex partition $A \cup B \cup C$, where $|A| = |B| = |C| = t/3$. We can bound $\gamma(H)$ by consideration of $\gamma_3(H)$, and in turn bound $\gamma_3(H)$ by considering only this partition.

The case where all pairs of parts have $\Theta(t^2)$ edges between them was handled by Myers and Thomason [49], and there the only non-trivial case is complete bipartite graphs. However, suppose instead that there are only $t^{3/2}$ edges between B and C , with $\Theta(t^2)$ between A and each of B and C . In this regime, we must have $d = \Theta(t)$, and since all of our bounds only use $\log d$ it does not affect our asymptotic results to replace d with t throughout this example; we will do so for simplicity.

When considering this partition, we can view such an H as an element of $\gamma(\vec{\sigma}, \mathbf{D})(t, d)$, with $\vec{\sigma} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, and $D_{ii} = -\infty$, $D_{12} = D_{13} = 1$ and $D_{23} = \frac{1}{2}$ (strictly our parameters have an additive $o(1)$ term, though this can be neglected in the limit). We remark it does not affect our bounds to allow up to t edges within each class — this corresponds to increasing D_{ii} , but they would remain non-positive.

Suppose we assign weights w_1, w_2, w_3 to A, B, C respectively. In order for these to provide a bound on $\gamma(\vec{\sigma}, \mathbf{D})$, we require $w_1 w_2 \geq 1$, $w_1 w_3 \geq 1$, and $w_2 w_3 \geq \frac{1}{2}$ — we would like to minimise $\sum w_i$. By symmetry, it is optimal to consider $w_2 = w_3$. It is then optimal to take $w_2 = 1/w_1$, since decreasing w_1 to this value reduces the total weight while still satisfying all conditions. Writing x for the value $w_2 = 1/w_1$, it remains to minimise $2x + 1/x$ over all $x \geq \frac{1}{\sqrt{2}}$. It turns out that taking $x = \frac{1}{\sqrt{2}}$ is optimal, and so we can compute $\gamma(\vec{\sigma}, \mathbf{D}) = 2\sqrt{2}/3 < 1$. By Theorem 3.10, for almost all graphs constructed in this fashion $\gamma_3(H) = (1 + o(1))2\sqrt{2}/3$ holds.

Since $2x + 1/x$ has a global minimum at $x = 1/\sqrt{2}$, if we instead took $D_{23} = y < \frac{1}{2}$ it remains optimal to take $w_2 = 1/\sqrt{y}$. If we take $d_{23} = y \geq \frac{1}{2}$, it is optimal to take $w_2 = 1/\sqrt{y}$. We remark that we could have attained this bound using γ_2 by merging the classes B and C into a single class of size $2t/3$ with $O(t^{\frac{3}{2}})$ edges inside. In fact after constructing this optimal weighting, we can see more than t edges can be added inside each class without affecting the extremal function — up to $O(t^2)$ in A and $O(t^{3/2})$ in each of B and C .

3.2 Breaking down our upper bounds

We are now ready to start the proofs of the upper bounds, Theorems 1.11 and 1.12. Our proof will follow the broad strategy outlined in Section 2.2. The aim of this section is to reduce these two theorems to dense sub-theorems (distinct for the two cases) and a common sparse sub-theorem. The proof of the sub-theorems will then be the aim of Sections 3.2 to 3.4.

Theorem 3.11. *Let $0 < \eta, \epsilon < 1$. Then there is a constant $C_{3.11}(\epsilon, \eta) = C$ such that if $0 < p < 1$, $q = 1 - p$ and G is a graph with $n \geq 2t\sqrt{f/\log(1/q)}$ vertices, density at least $p + \epsilon$, and connectivity at least $8\eta|G|$, then G is $K_{ft/\log t, t}^*$ -minor-prevalent, provided that $C < f < \log t$.*

Theorem 3.12. *Let $0 < \eta, \epsilon < 1$ and r an integer. There is a constant $T = T_{3.12}(\epsilon, r, \eta)$ such that the following holds for any $t > T$. Let H be a graph with t vertices, and average degree d . Let $(P_i, w_i)_{i=1}^r$ be a weighted partition of H satisfying the gamma inequality, restated as follows:*

$$\sum_{i \leq j} d^{-w_i w_j} e(P_i, P_j) / t \leq 1. \quad (3.4)$$

Let $w = \sum_i |P_i|w_i$, and suppose $0 < p < 1$ and $q = 1 - p$ are such that $w \geq Tt/\sqrt{\log_{1/q} d}$. Let G have $n \geq w\sqrt{\log_{1/q} d}$ vertices, density at least $p + \epsilon$, and connectivity at least $\eta|G|$. Then G is H -minor-prevalent.

Theorem 3.13. *Let $k > 0$ be an integer, $0 < \epsilon < 1/500$. Then there is a constant $D = D_{3.13}(\epsilon, k)$ such that the following holds. Let H be a graph, and $H_1, \dots, H_k$ graphs such that $\bigcup_i H_i = H$.*

Suppose that $m > D|H|$ is such that for each $1 \leq i \leq k$, every graph $\tilde{G}$ with minimum degree at least ϵm and connectivity at least $\epsilon|\tilde{G}|$ is H_i -minor-prevalent. Suppose further that every graph $\bar{G}$ with at least $m/7$ vertices, and minimum degree at least $(1 - \epsilon)|\bar{G}|$ contains H as a minor. Then every graph G with properties (1)-(4) below contains H as a minor.

- (1) $e(G) \leq m|G|$
- (2) Every edge of G is in at least $m - 1$ triangles
- (3) $\kappa(G) \geq D|H|$
- (4) $|G| \geq Dm$.

Proof of Theorem 1.12. Reduce ϵ if necessary so that $\sqrt{\log(1/2\epsilon)} > 28/\alpha$; this only strengthens the result. Let $H = K_{ft/\log t, t}^*$, $m = (2\alpha + \epsilon)t\sqrt{f}$, and $k = \epsilon m$. Let G be a graph as in the statement, and since $e(G)/|G| \geq m$, $G \in \mathcal{E}_{m, k}$. Replacing G by a minor if necessary, we can assume G is minor-minimal in $\mathcal{E}_{m, k}$ — recall this implies G has the properties of Lemma 2.4.

Let $N = \lceil \frac{8}{\epsilon} \rceil$. Partition each vertex class of H into N almost equal size parts, and add additional vertices of H to these parts (no longer requiring that parts be disjoint) so that the parts of the left hand class have size exactly $\lceil ft/(N \log t) \rceil$, and $\lceil t/N \rceil$ for parts of the right hand class. By taking the induced subgraph on pairs of parts, we get an edge-cover of H by N^2 subgraphs (H_i) , each isomorphic to $K_{\lceil ft/(N \log t) \rceil, \lceil t/N \rceil}^*$.

We would like to apply Theorem 3.13 to H and the partition (H_i) to N^2 parts. For

this, we need to show graphs $\tilde{G}$ with minimum degree at least $2\epsilon\alpha t\sqrt{f}$ and connectivity at least $\epsilon|\tilde{G}|$ are H_i -minor-prevalent.

By the connectivity condition, $\tilde{G}$ has minimum degree at least $\epsilon|\tilde{G}|$ and hence density $p \geq \epsilon$, with $q = 1 - p$ as usual. Recalling Inequality (1.3) of page 5, the following holds:

$$\frac{p}{|\tilde{G}|} \binom{|\tilde{G}|}{2} = e(\tilde{G})/|\tilde{G}| \geq \epsilon\alpha t\sqrt{f} \geq \epsilon t\sqrt{f} \frac{p}{2\sqrt{\log(1/q)}}.$$

Therefore, $|\tilde{G}| \geq \epsilon t\sqrt{f/\log(1/q)} \geq \lceil \frac{t}{N} \rceil \sqrt{\frac{[ft/(N \log t)]}{[t/N] \log(1/q)}}$ for any $f > 4N$ and thus we can apply Theorem 3.11 to deduce $\tilde{G}$ is H_i -minor-prevalent provided that $|H_i| \geq C/N^2$ is sufficiently large.

We also need to find H directly as a minor in a very dense graph. Let $\bar{G}$ be a graph with $|\bar{G}| \geq \alpha t\sqrt{f}/7 \geq 2t\sqrt{f/\log(1/2\epsilon)}$ and minimum degree at least $(1 - \epsilon)|\bar{G}|$, and suppose $\bar{G}$ has density p . In particular, $\bar{G}$ is $|\bar{G}|/3$ connected since $\epsilon < 1/6$. Therefore, $\bar{G}$ satisfies the hypotheses of Theorem 3.11 applied with H , the current ϵ , $\eta = 1/48$ and p taken as $p - \epsilon$. To apply this theorem, assume $C > C_{3.11}(\epsilon, 1/48)$.

In particular, we satisfy the conditions of Theorem 3.13 (recall that G satisfies properties (1)-(3) by Lemma 2.4), applied with $k = N^2$ and the above choice of (H_i) provided C is sufficiently large. There is therefore a constant $D = D_{3.13}(\epsilon)(> 1)$ such that if $|G| \geq Dm$, and $m > D|H|$ then G contains H as a minor. We note that $m > 0.5t\sqrt{f} > 0.5t\sqrt{C}$ and therefore provided $C > 4D^2$ the required condition is satisfied.

The above argument shows that H is a minor of G if $|G| \geq Dm$, and so from here we may assume $|G| \leq Dm$. In this ‘dense’ case, G has connectivity at least $(\epsilon/D)|G|$ hence also minimum degree $(\epsilon/D)|G|$. Suppose that G has density p , and let $p' = p - \epsilon/4D > \epsilon/4D$. By Inequality (1.3) again, the following holds:

$$e(G)/|G| \geq 2(1 + \epsilon)p\alpha t \frac{\sqrt{f}}{p} \geq (1 + \epsilon)pt \frac{\sqrt{f}}{\sqrt{\log(1/q)}}.$$

Hence $|G| \geq 2(1 + \epsilon)t\sqrt{f/\log(1/q)} \geq 2t\sqrt{f/\log(1/(1 - p'))}$ as above, and so provided that C is sufficiently large (depending on all previous constants) G is H -minor-prevalent, and in particular contains H as a minor. $\square$

Proof of Theorem 1.11. We remark that G is an element of the family $\mathcal{E}_{m,k}$ for $m = (\alpha + \epsilon)w\sqrt{\log d}$ and $k = \epsilon m$. Replacing G by a minor if necessary, we assume G is minor minimal in $\mathcal{E}_{m,k}$.

We first handle the case where $|G|$ is very large. We fix now some N , the choice will depend only on r and ϵ and will be defined later, followed by some C depending on N . We construct a partition of the vertex set $V(H)$ as follows. Let $w : V(H) \rightarrow \mathbb{R}_{\geq 0}$ be the weight function mapping each $v \in P_i$ to w_i . This can be extended additively to a function

on $\mathcal{P}(V(H))$, with $w(H) = w$.

Let B consist of all vertices of weight at least $2w/N$, and clearly $|B| \leq N/2$. Reordering if necessary, let P_1 be a part of maximal size, so in particular $|P_1| \geq |H|/r$. We will take $C > Nr$ so that B and P_1 are disjoint (recall the function w is constant on parts). We start by partitioning $H' = H[V(H) \setminus (P_1 \cup B)]$. Let $V_1, \dots, V_N$ be N initially empty sets, which we call bags. We place the vertices of H' into these bags one at a time in non-increasing degree order. When we come to add v , we place it arbitrarily into any V_i such that $w(V_i) \leq w/N$. Such an index must exist because the total weight is at most w . Continue in this fashion until $\cup V_i = V(H')$.

At the end of this procedure, any vertex of weight at least w/N must lie in a bag on its own (consider the largest weight vertex v in such a bag; no other vertices can be added after v since the weight is too large, and v was the first vertex added since we use non-increasing degree order). In particular, the total weight $w(V_i)$ must be at most $2w/N$. We now add P_1 to these bags. We place either $\lfloor |P_1|/N \rfloor$ or $\lceil |P_1|/N \rceil$ vertices from P_1 into each V_i , so that they remain disjoint, and $\cup V_i = V(H) \setminus B$. Since $|P_1|/N \geq 4$ (for $C > 4rN$), we have that $w(P_1)/2N \leq w(V_i \cap P_1) \leq 2w(P_1)/N$. Combining these bounds, each bag V_i must have weight at most $4w/N$.

Let $(W_i)_{i=1}^N$ be some arbitrary partition of $V(H) \setminus B$ into N sets, each of size either $\lfloor \frac{|V(H) \setminus B|}{N} \rfloor$ or $\lceil \frac{|V(H) \setminus B|}{N} \rceil$. We can now construct our edge cover of H , consisting of the following $k = \binom{N}{2} + N$ graphs. We aim to apply Theorem 3.13 with these graphs, and this value of k .

- Let $H_{0,i}$ be the graph on vertex set $V(B) \cup W_i$, and all edges with at least one endpoint in B for $1 \leq i \leq n$.
- Let $H_{a,b} = H[V_a \cup V_b]$ for $1 \leq a < b \leq N$.

We would like to proceed using Theorem 3.12 and a restriction of the existing weighting for the graphs $H_{a,b}$, and directly use Theorem 3.11 for $H_{0,j}$. However, since the gamma inequality involves the order and average degree of the graph, some additional modifications must be made to the weighting. We also have to consider the case where the average degree of $H_{a,b}$ is too small to apply the desired theorem.

Let $d_{a,b} = e(H_{a,b})/|H_{a,b}|$. By either a result of [66], or equivalently applying Theorem 3.12 using a constant weighting, there is a constant D_0 such that any graph $\overline{G}$ with $e(\overline{G})/|\overline{G}| \geq |H_{a,b}| \sqrt{\log d_{a,b}}$ and $\epsilon|\overline{G}|$ connectivity is $H_{a,b}$ -minor-prevalent provided $d_{a,b} \geq D_0$ for some constant D_0 depending only on ϵ . Further, for any constant $D_1 > D_0$ (which can depend on previous constants, but cannot depend on C), by adding edges to $H_{a,b}$ if necessary and taking C sufficiently large depending on D_1 , if $d_{a,b} \leq D_1$ then any graph with $e(\overline{G})/|\overline{G}| \geq \epsilon m \geq t \sqrt{\log(2D_1)}$ and connectivity at least $\epsilon|\overline{G}|$ is $H_{a,b}$ -minor-prevalent. We will fix D_1 later, although from now we assume $d_{a,b} > 1$ — we will only use the remainder of the proof in the case $d_{a,b} > D_1$.

We define a new weighting $w_{a,b}(v) = (w(v) + \delta)\sqrt{\frac{\log d}{\log d_{a,b}}}$, where $\delta = \sqrt{\log(Nr)/\log d}$ is chosen so that the gamma inequality holds on $H_{a,b}$ for each $1 \leq a < b \leq N$ with $d_{a,b} > 1$ with the weighting $w_{a,b}$. We include the calculation below, and for simplicity let w'_i denote $w_{a,b}(v)$ for some $v \in P_i$ (recall that w is constant on parts, so our choice does not matter).

$$\begin{aligned} \sum_{i \leq j} e(P_i \cap H_{a,b}, P_j \cap H_{a,b}) d_{a,b}^{-w'_i w'_j} / |H_{a,b}| &\leq \frac{1}{|H_{a,b}| d^{\delta^2}} \sum_{i,j} e(P_i, P_j) d^{-w_i w_j} \\ &\leq \frac{t}{d^{\delta^2} |H_{a,b}|} \leq 1 \end{aligned}$$

In particular, we have $\sqrt{\log d_{a,b}} w_{a,b}(H_{a,b}) \leq (w(H_{a,b}) + \delta)\sqrt{\log d}$. We choose now N minimal such that $16\sqrt{\log Nr}/N < \epsilon$, and so any graph G' with average degree at least ϵn and connectivity $\epsilon|G'|$ is $H_{a,b}$ -minor-prevalent by Theorem 3.12 applied with η and ϵ taking the value ϵ , and the weighting $w_{a,b}$ described above — provided that $d(H_{a,b}) > T_{3.12}$. Taking $D_1 = 2T_{3.12}$ (and note that D_1 does not depend on C , so we are justified in doing so), from one of the above arguments the result whether or not $d_{a,b} > D_1$.

We next consider the graph $H_{0,j}$. Each of these is a subgraph of a book $K_{s,t'}^*$, where $s \leq N$ and $t' \leq 2|H|/N$. Applying Theorem 3.11, there is a constant C' such that for all $\log t' > f(t') > C'$, any graph with connectivity at least $\epsilon t' \sqrt{f}$ and $e(G)/|G|$ at least $(2\alpha + \epsilon)t' \sqrt{f}$ is $K_{f t' t' / \log t', t'}^*$ -minor-prevalent.

In particular, taking t sufficiently large (so that t' is sufficiently large), and N large, $H_{0,j}$ is a minor of any graph G' with average degree at least ϵn and connectivity $\epsilon|G'|$ (note that increasing to $s = 2C' \log(t')/t'$ strengthens the result). We can now fix our choice of N — note that we have not yet had to choose C .

It remains only to verify that H is a minor of suitable dense graphs. Let $\bar{G}$ be a graph with minimum degree at least $(1 - \epsilon)|\bar{G}|$ and at least $m/7$ vertices. In particular, $\bar{G}$ must be $|\bar{G}|/3$ connected since $\epsilon < 1/6$. Suppose that the density of $\bar{G}$ is p . We have $e(\bar{G})/|\bar{G}| \geq (1 - 2\epsilon)\alpha wt/7 \geq wt/\sqrt{\log(1/(2\epsilon))}$, and so applying Theorem 3.12 with p replaced with $p - \epsilon$, our value of ϵ , and $\eta = 1/3$ implies that provided d is sufficiently large, H is a minor of $\bar{G}$.

Therefore, with our value of m the conditions of Theorem 3.13 hold (recall that G is minor-minimal and hence has the properties (1)-(3) by Lemma 2.4), and so there is a constant $D_{3.13}$ such that if also $|G| \geq D_{3.13}m$ and $m \geq Dt$ then G contains H as a minor. We note $m \geq 0.3Ct$, and therefore taking $C > 4D$ will suffice for this to hold.

We can therefore assume $|G| \leq D_{3.13}m$. In this case, G is $(\epsilon/D_{3.13})|G|$ connected. Further, if G has density p (which must be at least $\epsilon/D_{3.13}$ due to connectivity), then $e(G)/|G| \geq (\alpha + \epsilon)wt \geq pwt/2\sqrt{\log(1/(1 - p + \epsilon/4))}$ by inequality (1.3). Taking $\eta = \epsilon/D_{3.13}$, and using a value of $\epsilon/4$ in place of ϵ , Theorem 3.12 directly shows that if we take

d to be sufficiently large (depending on all previous constants), then H is a minor of G . Taking D large enough all of the above conditions are satisfied implies the result. $\square$

We finish this section with a brief discussion of our arguments. We are broadly mirroring the techniques of Chapter 2, although the proofs here are significantly more involved. The general technique is to either find H directly as a minor, if the host graph G is sufficiently dense (hence sufficiently connected that we can convert compatible partitions into minors), or to break into smaller parts when G is not dense.

Additional complexity arises due to the unpredictable form of $c(H')$ over subgraphs H' of H . For instance, if H consists of a clique H' on a small number of vertices, and a remaining number of vertices which are all of degree zero, then finding a H' minor is almost as hard as finding H . This problem did not arise in Chapter 2 since $t\sqrt{\log d}$ still decreases when we pass to this subgraph.

The required resolution is then, morally, to pass to covering subgraphs each of which is easier to find as a minor, which is done by restricting attention to the original weighting, and using this to bound the total weight on subgraphs. A final step splits those vertices of small weight among the subgraphs in a uniform fashion, so they are all sufficiently large we can apply our theorems. This restriction to a single weight means a slight modification to the dense sub-theorems is required.

Rather than requiring $\gamma_r(H)$ be large, we instead require that some particular weighting is large. This allows us to apply it to these subgraphs, even though they may have $\gamma_r(H)$ small (indeed, it is possible we are applying it to an independent set) since we have a larger average degree.

In the special case of $K_{s,t}^*$, this balancing approach corresponds to all our covering subgraphs being picked in such a way they contain the same proportion of the size t class and the size s class — a naive approach may take the entire size s class, and then finding our minors becomes much harder.

3.3 The dense case

The aim of this section is to prove each of our two dense case sub-theorems, Theorems 3.11 and 3.12. Our proof will have the same broad structure as in Chapter 2; the additional work comes within these steps.

We will start by building an almost-compatible partition. We will first prove a common theorem in subsection 3.3.1, and then deduce a compatible partition form of our dense sub-theorems for the bipartite and general cases in subsections 3.3.2 and 3.3.3 respectively. We will convert these objects into our desired minors using the same helper sets (the connector and projector) from Chapter 2.

3.3.1 Almost-compatible partitions

The aim of this subsection is to prove a parametric theorem, Theorem 3.14, which builds an almost-compatible partition with certain vertices of the graph H treated differently to others. This will be used in each of our two settings with different choices of parameters. For intuition, one can consider a simple case of this theorem where there are only two classes of parts, and we are only interested in the edges between them. This corresponds to proving a result for unbalanced bipartite graphs $G[A, B]$. In the following theorem, this corresponds to taking $r = 2$, and obtaining a partition $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2)$ where $\mathcal{P}_1$ contains about $|A|$ parts, and $\mathcal{P}_2$ about $|B|$.

Theorem 3.14. *Let $a_1, b_1, \dots, a_r, b_r, l, n \geq l(a_1b_1 + \dots + a_rb_r)$ be non-negative integers with each $a_i \geq 2$. Let $\omega \geq 2$, $\eta > 0$, $\eta < p < 1$ be constants such that $\omega\eta > 2r$. Let G be a graph of density at least p with at least n vertices.*

Then $V(G)$ has a partition $\mathcal{P}$, which can itself be decomposed into subpartitions $(\mathcal{P}_i)_{i=1}^r$, where $\mathcal{P}_i$ contains at least $a_i(1 - 2r/\omega\eta)$ parts (which are non-empty if $b_i > 0$), and for each $1 \leq i \leq j \leq l$, the proportion of pairs of parts from $\mathcal{P}_i, \mathcal{P}_j$ respectively which have no edge between them is at most $4r^2 6^{lb_ab_b} \omega^{lb_a} (\frac{q}{1-\eta})^{(1-\eta)l(l-1)b_ab_b}$, where $q = 1 - p$.

Proof. We may assume $n = l(a_1b_1 + \dots + a_rb_r)$ by replacing G with a maximal density subgraph on that number of vertices — at the end, we can redistribute any extra vertices arbitrarily among the parts. We also assume that G has density exactly p by applying the theorem with the actual density, then weakening our final result.

Order the vertices of G in non-increasing degree order, so that $d(v_1) \geq d(v_2) \geq \dots \geq d(v_n)$ and let (q_i) be such that $d(v_i) = (1 - q_i)(n - 1)$. To simplify notation, if $v = v_i \in V(G)$, let $q_v = q_i$. Partition the vertices of G into l blocks $(B_i)_{i=1}^l$ of size $x = (a_1b_1 + \dots + a_rb_r)$, where $B_i = \{v_{(i-1)x+1}, \dots, v_{ix}\}$.

We will start by constructing a_1 parts for $\mathcal{P}_1, \dots$, and a_r for $\mathcal{P}_r$ randomly and one at a time as follows, with each part of $\mathcal{P}_i$ having b_i vertices in each of the l blocks for a total of lb_i vertices. Let X be the union of the vertex sets of parts already chosen, and suppose we are now picking a part W for $\mathcal{P}_a$. Then W will consist of a uniformly randomly chosen b_a vertices from each of $B_1 \setminus X, \dots, B_l \setminus X$. Each of these sets has size at least b_a by the definition of x , and so we can choose W in such a fashion. We could equally have performed this construction by picking a (random) partition of each block B_i into a_1 subsets of size b_1 , a_2 of size b_2 and so forth. We then form a part for $\mathcal{P}_a$ as a random size b_a subset from each B_i . This also shows that having fixed (or conditioned on) one part W , each other part consists of uniformly randomly chosen elements of each $B_i \setminus W$.

Fix now some indices a, b (where $a = b$ is permitted). We say that a vertex v is *bad* for a set W if v has no neighbour in W , and $v \notin W$. For two disjoint sets $W, \widetilde{W}$ to be non-adjacent, $\widetilde{W}$ must consist entirely of vertices which are bad for W . For fixed v , what is the probability a part W from $\mathcal{P}_a$ has v bad for W ? This means we have chosen W as a

subset of the $(n - 1 - d(v)) = q_v(n - 1)$ non-neighbours of v . If W was instead a uniformly random lb_a set, the probability of this would be at most $(q_v(n - 1)/n)^{lb_a}$. We show this upper bound still holds for our blocked setup.

Let S consist of the $q_v(n - 1)$ non-neighbours of v , and let $S_i = S \cap B_i$. The probability that within B_i we choose only non-neighbours of v is at most $(|S_i|/x)^{lb_a}$, and since our choices are independent we have an overall upper bound via the AM-GM inequality of

$$\mathbb{P}(v \text{ is bad for } W) \leq \prod_i \left(\frac{|S_i|}{x}\right)^{lb_a} \leq \left(\frac{n-1}{n}q_v\right)^{lb_a} \leq q_v^{lb_a}. \quad (3.5)$$

In particular, the expected number of vertices in block B_i which are bad for W is at most $x(q_{ix})^{lb_a}$, recalling that the q_i are non-decreasing. We say that a part W *rejects* block B_i (where $i < l$) if there are more than $\omega x q_{ix}^{lb_a}$ bad vertices for W in B_i — the probability of this event is at most $1/\omega$. We say that a part W is *good* if it rejects fewer than $\eta(l - 1)$ of the blocks $B_1, \dots, B_{l-1}$. The probability a part is good is at least $(1 - 1/\omega\eta)$.

Suppose we fix some choice of good part W from $\mathcal{P}_a$. Conditional on this choice of W , what is the probability a (distinct) part $\widetilde{W}$ from $\mathcal{P}_b$ is not adjacent to W ? Within each block which is not rejected by W , we have at most $\omega x q_{ix}^{lb_a}$ remaining choices of vertex which are not adjacent to W , and $x - b_a$ remaining vertices to pick from. If we let $M(W)$ be the collection of indices in $[l - 1]$ for which B_j is not rejected by W , with $m = |M(W)|$, we have:

$$\mathbb{P}(\widetilde{W} \not\sim W) \leq \prod_{i \in M(W)} (\omega x q_{ix}^{lb_a} / (x - b_a))^{b_b} \leq (2\omega)^{mb_b} \left(\prod_{i \in M(W)} q_{ix} \right)^{lb_a b_b}. \quad (3.6)$$

We used here that $x - b_a \geq x/2$ which holds as $a_a \geq 2$. Further, since the (q_i) are non-decreasing, by applying the AM-GM inequality we deduce the following (since $i \in M(W)$ implies $i < l$):

$$\prod_{i \in M(W)} q_{ix}^{1/m} \leq \frac{1}{m} \sum_{i \in M(W)} q_{ix} \leq \frac{1}{xm} \sum_{i \in M(W)} \sum_{j=1}^x q_{it+j} \leq \frac{1}{xm} \sum_{j=1}^{st} q_j \leq \frac{l}{m} q.$$

The above result combined with Inequality (3.6) implies

$$\begin{aligned} \mathbb{P}(\widetilde{W} \not\sim W) &\leq (2\omega)^{mb_b} \left(\frac{ql}{m}\right)^{mlb_a b_b} \leq (2\omega)^{lb_b} \left(\frac{l}{l-1}\right)^{l^2 b_a b_b} \left(\frac{q}{1-\eta}\right)^{(1-\eta)l(l-1)b_a b_b} \\ &\leq 6^{lb_a b_b} \omega^{lb_b} \left(\frac{q}{1-\eta}\right)^{(1-\eta)l(l-1)b_a b_b}. \end{aligned}$$

The second inequality above uses the bounds $(1 - \eta)(l - 1) \leq m \leq l - 1$ as $1 - \eta > q$.

The probability that, for some i , there are more than $2r/\omega\eta$ proportion of bad parts in $\mathcal{P}_i$ is less than $\frac{1}{2r}$ by Markov's inequality. The probability that there are more than

$2r^2(6^{lb_i b_j} \omega^{lb_i} (\frac{q}{1-\eta})^{(1-\eta)l(l-1)b_i b_j})$ pairs of parts from $\mathcal{P}_i, \mathcal{P}_j$ respectively which are non-adjacent, but at least one of $\mathcal{P}_i$ and $\mathcal{P}_j$ is good is less than $\frac{1}{2r^2}$ again by Markov's inequality. Thus by a union bound, there is at least one choice of our partitions $(\mathcal{P}_i)$ for which none of these events occur. Take such a partition.

The result follows by deleting all bad parts from each $\mathcal{P}_i$, and redistributing unused vertices arbitrarily among the good parts so that we have a genuine partition of $V(G)$. $\square$

3.3.2 The bipartite case

Recall the aim of this subsection is to prove our dense sub-theorem for the bipartite case, Theorem 3.11. We will do this in stages, starting by building an almost-compatible partition in Lemma 3.15 under slightly weaker conditions (using Theorem 3.14 with suitable parameters), and then using the connector and projector to extend such a partition to a rooted minor.

Lemma 3.15. *Let $0 < \epsilon < \frac{1}{8}$. Then there is a constant $C = C_{3.15}$ such that if $t > 0$, $C < f < \log t$, and $H = K_{ft/\log t, t}^*$ the following holds.*

Suppose that G is a graph with density at least $p + \epsilon$, where $\epsilon < p < 1 - \epsilon$, and with $n \geq 2\sqrt{ft}/\sqrt{\log(1/q)}$ vertices for $q = 1 - p$. Then G has a $t^{1-\epsilon/4}$ -almost- H -compatible partition.

Proof. The core of this proof is an application of Theorem 3.14, although work is required to choose suitable parameters. We will be taking $r = 2$, using parts from $\mathcal{P}_1$ to correspond to vertices in the class of size $ft/\log t$, and parts from $\mathcal{P}_2$ for the vertices of the other class. Our density for the application of Theorem 3.14 will be $p + \epsilon$ rather than p . Let M, ω be large constants yet to be determined, and η, δ small constants which will not depend on C .

Let $a_1 = \lfloor (1 + \delta)ft/\log t \rfloor$, $a_2 = \lfloor (1 + \delta)t \rfloor$, $b_1 = \lfloor (\log t)/M \rfloor$, $b_2 = \lfloor f/M \rfloor$ and $l = \lfloor (1 - \delta)M/\sqrt{f \log(1/q)} \rfloor$ be the remainder of the parameters. These were chosen according to some ‘optimal weighting’ to maximise $b_1 b_2$ subject to an upper bound on $\sum a_i b_i$.

It is easily seen that $l(a_1 b_1 + a_2 b_2) \leq n$ and so the conditions of Theorem 3.14 are satisfied — hence G has a partition $\mathcal{P}_1 \cup \mathcal{P}_2$ satisfying the hypotheses therein. We assume that $l, a_1, a_2, b_1, b_2 > 4/\delta$, since we can make both f and t arbitrarily large, and will choose M so that this occurs.

If ρ is the density of non-adjacent parts between $\mathcal{P}_1$ and $\mathcal{P}_2$, we have:

$$\begin{aligned} \log \rho &\leq 4 + \frac{\log 6(\sqrt{f}(\log t)/M) + (\log \omega)(\log t)/\sqrt{f}}{\sqrt{\log(1/q)}} + \\ &\quad (1 - \eta)l(l - 1)\lfloor (\log t)/M \rfloor \lfloor f/M \rfloor (\log(q - \epsilon) - \log(1 - \eta)). \end{aligned}$$

We note $l - 1 \geq (1 - \delta)(1 - \delta)M/\sqrt{f \log(1/q)}$. Recalling Lemma 2.7 and that $\epsilon <$

$p < 1 - \epsilon$, for sufficiently small η we have $\frac{\log(q-\epsilon) - \log(1-\eta)}{\log q} > \frac{1}{1-\epsilon}$. We therefore deduce the following holds:

$$\log \rho \leq 4 + \frac{(\log 6)(f \log t/M) + \log(\omega) \log t}{\sqrt{f \log(1/q)}} - \frac{(1-\eta)(1-\delta)^3}{(1-\epsilon)} \log t. \quad (3.7)$$

We now fix some choices of δ and η sufficiently small that the final summand of Equation (3.7) is at most $-(1 + \epsilon/2) \log t$.

We have $|\mathcal{P}_i| \geq a_i(1 - 2r/\omega\eta)$. Choose (and fix) some ω so that $(1 + \delta)(1 - \delta/2)(1 - 4/\omega\eta) > 1$; we then have at least $ft/\log t$ parts in $\mathcal{P}_1$, and t parts in $\mathcal{P}_2$ as given by Theorem 3.14. In order that all of our prior claims are satisfied, and in order to bound $\log \rho$ suitably we will require $M = o(\log t) = o(f) = \omega(\sqrt{f})$ — we make the choice $M = f^{3/4}$ noting that $f \leq \log t$ by assumption.

We can now choose C sufficiently large that for any $\log t \geq f \geq C$, the first two terms of Inequality (3.7) are each at most $\epsilon/12 \log t$ — hence $\rho \leq t^{-1-\epsilon/3}$. Picking randomly some $ft/\log t$ parts from $\mathcal{P}_1$, and t from $\mathcal{P}_2$, the expected number of pairs of picked parts, the first from $\mathcal{P}_1$ and the second from $\mathcal{P}_2$, is at most $ft^2\rho/\log t \leq t^2\rho$. Thus with probability more than $\frac{1}{2}$, at most $2t^{1-\epsilon/3}$ such parts are non-adjacent.

Let ρ' be the density of non-adjacent pairs in the left hand class, i.e. pairs arising from $\mathcal{P}_1$. Since $b_2 \leq b_1$ we immediately deduce (comparing with the previous calculations) that $\rho' \leq \rho \leq t^{-1-\epsilon/3}$. A union bound now implies a choice of partition satisfying the claims of our theorem exists. $\square$

Proof of Theorem 3.11. Let f, t be as in the statement, $q = 1 - p$ and $H = K_{ft/\log t, t}^*$. Reduce η if necessary so that $20\eta < \epsilon$ and also $(1 - 5\eta) \geq \sqrt{\log(1/q)/\log(1/(q - \epsilon/4))}$ — note that G still satisfies the hypotheses with this reduced value. Let R be a set of roots labelled by the vertices of H . Take C large enough that $2\sqrt{C/\log(1/\epsilon)} \geq 2/\eta$. Then since $C \leq f \leq \log t$ we have $|R| \leq |G|/D_{2.18}(\eta) \leq \eta|G|$.

Thus we can apply Theorem 2.17 with these parameters to obtain subsets C, P each of size at most $2\eta|G|$; these subsets will also have the properties of Theorem 2.18. The subgraph $G - C - P - R$ has density at least $p + \epsilon/2$ since we have only removed at most $\epsilon/4$ proportion of vertices.

However, $G - C - P - R$ may have fewer than $2t\sqrt{f/\log(1/q)}$ vertices, which prevents a naive application of Lemma 3.15. Since we only removed at most $5\eta|G|$ vertices, we can see by Lemma 2.7 $|G - C - P - R| \geq n(1 - 5\eta) \geq 2t\sqrt{f/\log(1/(q - \epsilon/4))}$ holds, and so we can apply Lemma 3.15 with p replaced by $p + \epsilon/4$, and ϵ replaced with $\epsilon/4$. This application requires $D > C_{3.15}(\epsilon/4)$.

As a consequence, $G - C - P - R$ has a $t^{1-\epsilon/16}$ -almost- H -compatible partition (V_h) . Let

F be the collection of adjacent pairs in H whose corresponding subsets are non-adjacent. Applying Theorem 2.18 with this choice of F and the partition (V_h) the result follows; note $|F| \leq t^{1-\epsilon/16} \leq t^{-\epsilon/16}|R| \leq |R|/\eta$ provided C is large enough. $\square$

3.3.3 The general case

In this subsection, we will finish our proof of the dense case by proving Theorem 3.12. Our method will mirror that used in subsection 3.3.2, although the argument is somewhat more technical due to the additional generality.

Lemma 3.16. *Let $r > 0$, $0 < \epsilon < \frac{1}{8}$, $\epsilon < p < 1 - \epsilon$. Let H be a graph with t vertices and td edges where $d > 1$. Let $(P_i, w_i)_{i=1}^r$ be a weighted partition of $V(H)$ such that*

$$\sum_{i \leq j} d^{-w_i w_j} e(P_i, P_j)/t \leq 1. \quad (3.8)$$

Let $w = \sum_i |P_i|w_i$, and suppose $w \geq 2^{10}\epsilon^{-3}t/\sqrt{\log_{1/q} d}$. Let G have density at least $p + \epsilon$, where $p = 1 - q$, and $n \geq w\sqrt{\log_{1/q} d}$ vertices. Then G contains a $64r^6t/\epsilon$ -almost- H -compatible partition.

Proof. Our proof broadly mimics that of Theorem 3.15, although the calculations are more involved. Let $\delta = \epsilon/8$, and $l = \lceil (\delta^2/8)w\sqrt{\log_{1/q} d/t} \rceil$. The lower bound on w in the statement implies $l > 2/\epsilon$, and so $l < 2w(\delta^2/8)\sqrt{\log_{1/q} d/t}$.

Replace each w_i by a new weight w'_i so that

$$\lceil (1 - \delta)w_i\sqrt{\log_{1/q} d/l} \rceil = (1 - \delta)w'_i\sqrt{\log_{1/q} d/l}. \quad (3.9)$$

We observe that $(P_i, w'_i)_{i=1}^r$ still satisfies the gamma inequality (3.8) since we have only increased the weights. We will be picking parts for each vertex of H randomly, with the size of that part proportional to the weight. Let $a_i = \lceil |P_i|(1 + \delta/4) \rceil$, and $b_i = (1 - \delta)w'_i\sqrt{\log_{1/q} d/l}$.

We now seek to apply Theorem 3.14, with $\eta = \epsilon^2/4$, and $\omega = 8r/\eta\delta$ (and hence $(1 - 2r/\omega\eta)(1 + \delta/2) > 1$) — note G has density at least $p + \epsilon$ and so it is valid to use the current value of p . We would like to assume that all the a_i and b_i are sufficiently large, so that we can ensure the earlier rounding of a_i and b_i does not make us violate the condition $|G| \geq \sum la_i b_i$. However in general this may not be the case — it is valid to take a single vertex of H as a part P_i .

We resolve this problem by replacing w'_i with 0 on all indices i for which $|P_i| < 4/\delta$; we call such indices terrible. Then for all non-terrible indices we have $a_i \leq (1 + \delta/4)^2 |P_i|$, and for all terrible indices we have $b_i = 0$.

Since $w'_i \sqrt{\log_{1/q} d} \leq w_i \sqrt{\log_{1/q} d} + l/(1-\delta)$ for all i , the following series of inequalities holds:

$$\begin{aligned} \sum la_i b_i &\leq \sum_i (1-\delta)(1+\delta/4)^2 |P_i| w'_i \sqrt{\log_{1/q} d} \\ &\leq \sum_i (1-\delta/3) |P_i| w_i \sqrt{\log_{1/q} d} + 2tl \\ &\leq (1-\delta/3) w \sqrt{\log_{1/q} d} + (\delta^2/2) w \sqrt{\log_{1/q} d} \leq n. \end{aligned}$$

We can therefore apply Theorem 3.14 to G as claimed, since $|G|$ is sufficient. We will later obtain our almost-compatible partition (which we will call V_v) by using parts from the $\mathcal{P}_i$ of Theorem 3.14 to represent the vertices $v \in P_i \subset H$ for each index i which is not terrible. We will return to the terrible parts at the end. We will only bound the number of non-adjacent pairs between parts from $\mathcal{P}_i$ and $\mathcal{P}_j$ when $D_{ij} > 0$, and handle the case $D_{ij} \leq 0$ at the end.

If ρ is the density of non-adjacent pairs from $\mathcal{P}_i, \mathcal{P}_j$ respectively, where $w'_i \leq w'_j$, P_i and P_j are not terrible and $D_{ij} \geq 0$, we have the following bound from the application of Theorem 3.14:

$$\begin{aligned} \log \rho &\leq \log(4r^2) + (\log 6)(lb_i b_j) + (\log \omega)(lb_i) + \\ &(\log(q - \epsilon) - \log(1 - \eta))(1 - 1/l)(-\log d / \log q) w'_i w'_j (1 - \delta)^2. \end{aligned}$$

We would now like to bound this quantity more simply; in order to obtain the desired outcome we will require $\log \rho \leq -w_i w_j \log d$ and thus $\rho \leq d^{-w_i w_j}$. We control the final two summands on the first line one at a time, and the second line as a whole. The quantity on the second line can be bounded above by $-(1 + \epsilon/3) w'_i w'_j \log d$ by an application of Lemma 2.7 as in the proof of Lemma 3.15.

For the final summand on the first line, we have $w'_i \sqrt{\log_{1/q} d} \geq l/(1-\delta)$ by Equation 3.9 since $p > \epsilon$. In particular, since also $p < 1 - \epsilon$, we have $w'_i \sqrt{\log d} > 9r/\epsilon$ by the bound on l . We therefore deduce $lb_i \leq w'_i \sqrt{\log d} \leq \epsilon/9 w_i w_j \log d$. The second term satisfies $lb_i b_j \leq (1/l)(w_i w_j \log_{1/q} d) \leq \epsilon/9 w_i w_j \sqrt{\log d}$, and finally the first term can be bounded using $\log(4r^2) \leq \log(4r^2)(\epsilon/9)^2 w'_i \sqrt{\log d} w'_j \sqrt{\log d}$. We therefore deduce $\rho \leq 4r^2 d^{\epsilon D_{ij}/9} d^{-(1+\epsilon/6)w_i w_j}$ which is in turn at most $4r^2 d^{-(1+\epsilon/9)D_{ij}}$, since $w_i w_j \leq D_{ij}$.

For each non-terrible index i , choose a uniformly random injection from the vertices $v \in P_i$ to the (non-empty) parts $V_w \in \mathcal{P}_i$. We expect $\rho \leq 4r^2 d^{-(1+\epsilon/9)D_{ij}}$ proportion of edges of H between P_i and P_j to correspond to a non-adjacent pair of parts. Take some choice of assignments for which no pairs i, j have more than $4r^4 d^{-(1+\epsilon/9)D_{ij}}$ such edges, this is possible by Markov's inequality.

By summing over the r^2 pairs, this gives at most $4r^6$ non-edges between pairs of parts,

each corresponding to a vertex in a non-terrible parts P_i, P_j respectively with $D_{ij} > 0$.

As promised, we now turn to the terrible parts and those pairs with $D_{ij} \leq 0$. We first handle vertices in terrible parts. For each of the at most $4r/\delta$ vertices in a terrible part in turn, say we are constructing a part for w , we pick some vertex from another part V_v of size at least 2, remove it from that part, and use it as a singleton part V_w for w . Of course this may introduce another up to t bad edges from removal of the vertex, and there could be up to t adjacencies at w which are not satisfied. This is, however, only at most $8rt/\delta$ total additional non edges. We are able to make such a choice since $n \geq t$.

For those pairs with $D_{ij} \leq 0$, we will not perform any calculations. In the worst case, there are only t edges between such a pair, and so certainly at most $2r^2t$ missing edges overall. Thus these two deferred cases together contribute at most $4r^2t/\delta$ bad edges.

Combining all of the above analysis by summing over pairs, we have a partition of $V(G)$ into non-empty sets labelled by vertices of H , with at most $8r^6t/\delta$ edges of H corresponding to non-adjacent pairs, as desired. $\square$

Corollary 3.17. *Let H be a graph with t vertices $td > t$ edges. Let $\epsilon < p < 1 - \epsilon$. Suppose that G has density at least $p + \epsilon$, and $n \geq \gamma_r(H)t\sqrt{\log_{1/q} d}$. Further, suppose that $\gamma_r(H)\sqrt{\log_{1/q} d} > 2^{10}\epsilon^{-3}\sqrt{\log(1/q)}$. Then G has a $64r^6t/\epsilon$ -almost- H -compatible partition.*

Proof. Let (P_i, w_i) be some weighted partition of H into r parts satisfying inequality (3.8) with $\sum_i |P_i|w_i$ minimal. Such a partition exists, because we are optimising over a finite number of partitions, and for each such partition we are solving a closed optimisation problem to pick an optimal weighting. In particular, we have $\sum_i |P_i|w_i = t\gamma_r(H)$. We now apply Lemma 3.16 to this weighted partition, with all other parameters as in the theorem. $\square$

We are now ready to prove our dense sub-theorem for the $\gamma_r(H)$ case.

Proof of Theorem 3.12. Let $q = 1 - p$, and reduce η if necessary so that $(1 - 5\eta) \geq \sqrt{\log(1/q)/\log(1/(q - \epsilon/4))}$ for all $1 \geq q \geq \epsilon$ and $\eta < \epsilon/(2^8r^6)$; note we can do so depending only on r, ϵ . Let R be a set of roots labelled by the vertices of H , and we will choose $T > D = D_{2.18}(\eta)$ so $|R| \leq |G|/D \leq \eta|G|$.

We can therefore apply Theorem 2.17 with these parameters to obtain subsets C, P each of size at most $2\eta|G|$; these subsets will also have the properties of Theorem 2.18.

The graph $G - C - P - R$ has density at least $p + \epsilon/2$ since fewer than $\epsilon|G|/4$ vertices are deleted. We also have at least $n(1 - 5\eta) \geq w\sqrt{\log_{1/(q - \epsilon/4)} d}$ vertices. In particular, we can apply Corollary 3.17, with the parameter p from that theorem replaced with $p + \epsilon/4$, and ϵ with $\epsilon/4$ provided we take $T > 2^{16}\epsilon^{-3}$. Thus $G - C - P - R$ has a $256r^6t/\epsilon$ -almost- H -compatible partition (V_h) . Let F denote the collection of adjacent pairs in H whose

corresponding subsets are non-adjacent. Applying Theorem 2.18 with this choice of F and the partition (V_h) implies the result; note $|F| \leq 256r^6t/\epsilon \leq |R|/\eta$. $\square$

We finish with some remarks about our proofs so far. As promised, the main difficulty in the dense case came in building compatible partitions; converting these into minors is easy once we have built the Connector and Projector.

Building the compatible partitions is where our argument is tight; in the sparse case, our argument would remain true even with a constant factor lower average degree, and the connector and projector could fix any number $O(t)$ of edges. The tricky part is building up classes of different sizes, while remaining sufficiently pseudorandom that the 2-stage removal can be used. The robustness of this 2-stage process is witnessed by the result being tight despite our repeated use of Markov's inequality.

3.4 The sparse case

Having completed our proofs of the dense sub-theorems, we are now ready to prove the sparse sub-theorem and hence complete our upper bounds; recall only one theorem (Theorem 3.13, which we will now restate) is required to handle both complete bipartite graphs and the more general case. Our proof will use as input the lemmas of Section 2.6 to simplify the argument, which is similar to that used in Theorem 2.6.

Theorem 3.13. *Let $k > 0$ be an integer, $0 < \epsilon < 1/500$. Then there is a constant $D = D_{3.13}(\epsilon, k)$ such that the following holds. Let H be a graph, and $H_1, \dots, H_k$ graphs such that $\bigcup_i H_i = H$.*

Suppose that $m > D|H|$ is such that for each $1 \leq i \leq k$, every graph $\tilde{G}$ with minimum degree at least ϵm and connectivity at least $\epsilon|\tilde{G}|$ is H_i -minor-prevalent. Suppose further that every graph $\bar{G}$ with at least $m/7$ vertices, and minimum degree at least $(1 - \epsilon)|\bar{G}|$ contains H as a minor. Then every graph G with properties (1)-(4) below contains H as a minor.

- (1) $e(G) \leq m|G|$
- (2) Every edge of G is in at least $m - 1$ triangles
- (3) $\kappa(G) \geq D|H|$
- (4) $|G| \geq Dm$.

Proof of Theorem 3.13. For simplicity, we will assume D is sufficiently large and neglect specific calculations. By Lemma 2.21 applied with the current m, k and $t = |H|$ we deduce that G has one of substructures (A) or (B) described therein.

Suppose first that G has, as in (A), a bipartite subgraph $G' = (A, B)$ with $|B| \leq 6km, |A| \geq |G|/3$ and every vertex in A having at least $m/7$ neighbours in B . By Lemma 2.20, applied with $M = 1$, the current value of ϵ and $d = m/7\epsilon$ we see that

G' has a minor $\overline{G}$ with minimum degree at least $(1 - \epsilon)|\overline{G}|$ and at least $m/7$ vertices, provided that D is large. In particular, H is a minor of $\overline{G}$ and hence of G .

We therefore assume that this does not occur, and therefore G contains subsets $(T_i)_{i=0}^k$ as in (B), and thus also a collection of paths $P_{i,j}$ and connected sets C_j as described in Lemma 2.22 by an application thereof with the current k, m .

Consider $G_i = G[T_i - \bigcup_{x < y} V(P_{x,y}) \cup \bigcup_{j \in V(H_i)} b_{i,j}]$. Then G_i has minimum degree at least ϵm and connectivity at least $m/40 \geq \epsilon |G_i|$. In particular, G_i is H_i -minor-prevalent and so contains a H_i minor rooted at $(b_{i,j})_{j \in V(H_i)}$. Pick such a model $(X_j^i)_{j \in V(H_i)}$.

Consider the collection $Y_j = C_j \cup \bigcup_{j \in V(H_i)} X_j^i$. These are easily seen to form an H model in G : each Y_j is connected since the C_j, X_j^i are, the Y_j are disjoint, and every edge of H appears in some H_i . $\square$

As promised, this result followed easily from the lemmas of Chapter 2. This is in a sense unsurprising, since in the sparse case our argument has a lot of slack. The freedom to choose our number of vertices to be arbitrarily large, while maintaining the same average degree, means we can lose as many constant factors as needed.

The only difficulties for this case arose in the proofs of Theorems 1.11 and 1.12. Building the subgraphs which partition our graph required some careful thought. However, having built the subgraphs, our lemmas naturally applied.

3.5 Many non-adjacent sets in random graphs

We now move to lower bounds for $c(H)$. Rather than excluding minors, we will be excluding compatible partitions which simplifies the argument. Our aim is to find a graph G such that with high probability over random H , G does not contain an H -compatible partition

The following lemma provides a graph with this property; however it does not have sufficient average degree to provide a good bound on $c(H)$. In the next section, we will use blowups of this graph to obtain our final result.

Lemma 3.18. *Let $0 < p < 1$, $0 < \epsilon$. There exists a constant $D = D_{3.18}(p, \epsilon)$ such that for all $d > D$ there is a graph $G = G_{d,p,\epsilon}$ with d vertices and density at least p , with the following property.*

Let $l = \sqrt{\log_{1/1-p} d}$, and $0 \leq x \leq l$. Let $A_1, \dots, A_s; B_1, \dots, B_s$ be disjoint subsets of G , such that $|A_i||B_j| \leq x l^2$ for all $1 \leq i, j \leq s$. Then provided $s \geq d^{\epsilon+x}$, there are at least $\frac{1}{2} d^{-x} s^2$ pairs (A_i, B_j) which are non-adjacent (i.e. have no edge between them in G).

Proof. Let $l = \sqrt{\log_{1/1-p} d}$. Since $|A_i||B_j|$ is an integer, and reducing x for fixed $|A_i||B_j|$ makes the result stronger, we only need consider each of $x = 0, 1/l^2, \dots, \lfloor l^3 \rfloor / l^2$ — the $x = 0$ case being trivial, so we also assume $x \geq 1/l^2$.

Let $G \sim G(d, p)$ be a random graph on d vertices, with edges present independently with probability p . We will show that with positive probability, G has the required properties. In [20], it is shown the probability $G(d, p)$ has density at least p is at least $1/4$, provided that $\binom{d}{2}p \geq 1$. For our result to remain self-contained, we give a sketch of a stronger result — though we require that d is large. Let $X \sim \mathcal{B}(\binom{d}{2}, p)$ be the random variable $e(G)$. Let $\mu = \mathbb{E}(X)$, and $\sigma^2 = \text{Var}(X) = \binom{d}{2}p(1-p)$. By the Central Limit Theorem, $\frac{X-\mu}{\sigma}$ converges in distribution to a standard normal $N(0, 1)$ random variable (as d increases). Let z be such that $\Phi(z) = 2/3$. Then $\mathbb{P}(X \leq \mu + z\sigma)$ converges to $2/3$, and in particular for $d > D$ sufficiently large, with probability at least $1/4$ the random variable exceeds its mean.

Fix now some choice of x , $A = (A_i)$, $B = (B_j)$. Since each edge is present independently at random, $\mathbb{P}(\text{there is no } A_i - B_j \text{ edge}) = (1-p)^{|A_i||B_j|} \geq d^{-x}$ for each i, j . In particular, the expected number of non-adjacent $A - B$ pairs is at least $s^2 d^{-x}$.

Let X_{ij} be the indicator variable for A_i and B_j being non-adjacent. Since our sets are disjoint, these depend on different edges and so are independent. Using a weak form of the Chernoff bound, namely that

$$\mathbb{P}(X \leq \frac{\mu}{2}) \leq \exp(-\frac{1}{8}\mu) \quad (3.10)$$

for any X a sum of independent Bernoulli random variables with mean μ (cf. [36, Theorem 2.1]), we can deduce that:

$$\mathbb{P}(\sum X_{ij} \leq \frac{1}{2}d^{-x}s^2) \leq \exp(-\frac{1}{8}d^{-x}s \cdot s) \leq \exp(-\frac{1}{8}d^{\epsilon}s). \quad (3.11)$$

To finish, we bound the number of choices of (A_i) and (B_j) . We remark it suffices to consider the case where all A_i and B_j are non-empty, and in particular $|A_i| \leq l^3$, $|B_j| \leq l^3$. There are at most $\binom{\binom{d}{3}}{s} \leq d^{2l^3s}$ choices of (non-empty) sets satisfying these constraints. Therefore, for each x and s , the probability that some choice of (A_i) and (B_j) have too few non-adjacent pairs is at most $\exp(-\frac{1}{16}d^{\epsilon}s)$ since $d^{2l^3s} \leq \exp(\frac{1}{16}d^{\epsilon}s)$ for large d . We can sum over the at most l^3 nonzero values of x , and the at most d choices for s (each of which is at least d^{ϵ}) to deduce that except with probability at most $\exp(-\frac{1}{32}d^{2\epsilon})$ we have the desired property for all choices of s, x . Taking D large enough, this probability is less than $\frac{1}{4}$, and hence with positive probability G satisfies all of the hypotheses of the theorem; in particular at least one graph $G_{d,p,\epsilon}$ has these properties. $\square$

We will make use of Lemma 3.18 in the following section. We finish this section by proving Theorem 3.10, again restated for convenience.

Theorem 3.10. *Let $\gamma, \epsilon, r > 0$. Then there is a $D = D_{3.10}(\gamma, \epsilon, r)$ such that for all $t > d > D$, all $\vec{\sigma}$ weight vectors of length r and $\mathbf{D}$ matrices satisfying the conditions of Definition 3.8 such that $\gamma(\vec{\sigma}, \mathbf{D}) \geq \gamma$; for all but at most ϵ proportion of graphs H from*

$\mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$ the following holds:

$$(1 - \epsilon)\gamma(\vec{\sigma}, \mathbf{D}) \leq \gamma(H)\sqrt{\log t / \log d} \leq \gamma_r(H) \leq \gamma(\vec{\sigma}, \mathbf{D})(1 + \epsilon).$$

Proof. The final two inequalities follow immediately from the definitions, and the discussion of Section 3.1 (indeed, they hold for all elements of $\mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$). We thus focus only on the first, and seek to prove this lower bound on $\gamma(H)$ holds for almost all elements of the class.

Suppose that $H \in \mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$ has $\gamma(H) \leq (1 - \epsilon)\gamma(\vec{\sigma}, \mathbf{D})\sqrt{\log d / \log t}$. Let w be a weight function with $\sum_{uv \in E(H)} t^{-w(u)w(v)} \leq t$, with average weight $\frac{1}{t} \sum_{u \in H} w(u) = \gamma(H)$. Consider the modified weighting

$$w'(v) = ((1 + \epsilon/3)w(v) + \epsilon\gamma/2r)\sqrt{\log t / \log d}.$$

If $w_i = w'(S_i)/|S_i|$ is the average weight on S_i , we have $\sum_i \sigma_i w_i \leq (1 - \epsilon/6)\gamma(\vec{\sigma}, \mathbf{D})$. By the definition of $\gamma(\vec{\sigma}, \mathbf{D})$, there must be some i, j such that $w_i w_j < D_{ij}$. Further, since $w_i \geq \epsilon\gamma/2r$ for all i , we must have $D_{ij} > \epsilon^2\gamma^2/4r^2$.

By Markov's inequality, there are sets $A_i \subset S_i$ and $A_j \subset S_j$, of size at least $\epsilon|S_i|/4$ and $\epsilon|S_j|/4$ respectively such that for all $v \in A_i$, $w'(v) \leq (1 + \epsilon/3)w(S_i)/|S_i|$; likewise for A_j . Replace A_i with a subset so that $|A_i| = \lceil \epsilon|S_i|/4 \rceil$, and in the same fashion take $|A_j| = \lceil \epsilon|S_j|/4 \rceil$.

Thus for each $v \in A_i$, since $w'(v) \leq (1 + \epsilon/3)w_i$ it follows

$$w(v) \leq w_i \sqrt{\frac{\log d}{\log t}} - \epsilon\gamma/4r \leq (w_i - \epsilon\gamma/4r) \sqrt{\frac{\log d}{\log t}}.$$

In particular, for all $v \in A_i, v' \in A_j$ we have $t^{-w(v)w(v')} \geq d^{-D_{ij} + \epsilon^2\gamma^2/16r^2}$. Since w satisfies the gamma inequality, this means there are fewer than $td^{D_{ij} - \epsilon^2\gamma^2/16r^2}$ edges between A_i and A_j . Since $D_{ij} \geq \epsilon^2\gamma^2/4r^2$, this is less than $\frac{1}{4}\lfloor td^{D_{ij}} \rfloor$. We now compute the probability that such sets A_i, A_j exist in a randomly chosen $H \in \mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$. There are r^2 choices for i and j , and at most 2^{2t} choices for A_i and A_j given such a choice of indices.

We first consider the case $i \neq j$. Let $N = |S_i||S_j|$, $M = |A_i||A_j|$ and $n = \lfloor td^{D_{ij}} \rfloor$. The number of edges between A_i and A_j is distributed according to a hypergeometric distribution $HG(N, M, n)$, with mean Mn/N . By a result of Vatutin and Mikhailov [68], the hypergeometric distribution can be written as a sum of n independent Bernoulli random variables, and hence we can apply the Chernoff bound in the form of Equation 3.10. Therefore, the probability there are fewer than $\epsilon^2 td^{D_{ij}}/64$ edges between A_i and A_j (which is less than half the mean by taking t large) is at most $\exp(-(\epsilon^2/64)^2 td^{D_{ij}}/2) \leq \exp(-\epsilon^2 td^{D_{ij}}/2^{13})$, and thus the expected number of pairs (A_i, A_j) of this kind in a randomly selected H is at most $r^2 2^{2t} \exp(-\epsilon^2 td^{D_{ij}}/2^{13}) \leq \epsilon/2r^2$ provided we take D sufficiently large.

If instead $i = j$, we follow the same argument, but only find one set $A_i \subset S_i$, and consider the edges within A_i . This time there are at most 2^t choices, each of which occurring with small probability by the Chernoff bound bound for a total probability of existence of at most ϵ/r^2 again. The result follows taking a union bound. $\square$

What have we done so far? We have built a graph $G_{d,p,\epsilon}$ with the property that given a ‘sufficient number’ of subsets $A_1, \dots, A_s$ and $B_1, \dots, B_s$, the density of non-adjacent subsets is at least $\frac{1}{2}(1-p)^{\max |A_i||B_j|}$ provided that the $|A_i||B_j|$ are not too large. Equivalently, we get at least half the expected density.

Sadly, these graphs do not have sufficient average degree to prove our lower bounds as-is — indeed in a regime of interest where d is constant, their average degree is bounded. We must modify these graphs in such a way the average degree becomes larger, without creating too many new ways for graphs to be models. This is what we do in the next section with blowups.

Our second result in this section showed us that results that are naturally phrased in terms of $\gamma(\vec{\sigma}, \mathbf{D})$ are still relevant to our upper bound setting of $\gamma_r(H)$, and also to $\gamma(H)$. Informally, we have been given a weighting and a random assignment of edges between the weight classes. On average, we will not be able to make a saving by shifting weights, since that corresponds to some large subsets having a lower density than expected — most random graphs do not have this property (and indeed, violating this means some pair would not be pseudorandom). This is unrelated to the first result, but included in this section since the techniques used in the proof were similar.

3.6 Proof of lower bounds

Our first aim in this section is to prove the following generalisation of Theorem 3.2; that theorem follows immediately by taking an appropriate choice of ϵ, γ and using only the optimal value $p = 0.715\dots$ from the definition of α (cf. Definition 1.6 on page 5).

Theorem 3.19. *Let $\epsilon, \gamma > 0$ and $0 < p < 1$. There is a constant $D = D_{3.19}(\epsilon, \gamma, p)$ such that the following holds.*

Let $\vec{\sigma}$ be a weight vector, and $\mathbf{D}$ a matrix with entries from $[-\infty, 1]$ such that $\gamma(\vec{\sigma}, \mathbf{D}) > \gamma$ and all $\sigma_i > \epsilon$. Then for all $t > d > D$, there is a graph $G = G_{t,d,p,\gamma(\vec{\sigma}, \mathbf{D}), \epsilon}$ (which does not depend on $\vec{\sigma}$ or $\mathbf{D}$ except through γ) with the following properties.

- $|G| \geq (1 - 2\epsilon)\gamma(\vec{\sigma}, \mathbf{D})t\sqrt{\log_{1/1-p} d}$
- G has density at least $p - 2\epsilon$
- All but at most 2^{-t} proportion of graphs from $\mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$ are not minors of G .

We will be building the graph for Theorem 3.19 as a blowup of a graph obtained from Lemma 3.18.

Definition 3.20. Let G_0 be a graph, and k an integer. The *balanced blowup* $G_0(k)$ is the graph constructed from $|G_0|$ disjoint independent sets I_v of size k (each corresponding to a different vertex of G_0), and an edge between $x \in I_v, y \in I_w$ exactly when v and w are adjacent in G_0 .

What does a minor look like in such a blowup? Let (V_h) be an H model in $G_0(k)$. If, for some $h \in H, v \in G_0$, V_h contains multiple vertices from I_v , we still have a H model deleting all but one such vertex. Considering only minimal models, we can assume $|V_h \cap I_v| \leq 1$ for all $h \in H, v \in G_0$.

This naturally lets us associate the parts V_h with subsets of G_0 , instead of $G_0(k)$. Having done so, each vertex of G_0 can appear in at most k of these subsets. We call a collection of subsets of $V(G_0)$ where each vertex appears at most k times a *k-blobbing* in G_0 .

Proof of Theorem 3.19. Let $G_0 = G_{d,p,\epsilon_1}$ be the graph from Lemma 3.18, applied with $\epsilon_1 = \epsilon^3 \gamma^2 / 64$ in place of ϵ , and d, p as in the theorem (we can do this provided $D > D_{3.18}(\epsilon_1, p)$).

Let $l = \sqrt{\log_{1/1-p} d}$, and $k = \lfloor (1 - \epsilon) \gamma (\vec{\sigma}, \mathbf{D}) t l / d \rfloor$. We hereon assume that D , and hence t, d, l, k are sufficiently large that all necessary inequalities hold; the exact requirements are suppressed. Let $G = G_0(k)$ be the balanced blowup of G_0 . Then since G_0 has density at least p , G has density at least $p - \epsilon$. As $\gamma t l / d > 2/\epsilon$, we have $|G| \geq (1 - 2\epsilon) \gamma t l$.

We start with some notation. We will let $H \in \mathcal{D}(\sigma, \mathbf{D})(t, d)$ be an arbitrary element, and recall H has vertex set $\{1, \dots, t\}$, and a vertex partition $S_1, \dots, S_r$ for some $r \geq 0$, where $|S_i| = \lceil \sigma_i t \rceil$ or $\lfloor \sigma_i t \rfloor$, and there are exactly $\lfloor t d^{D_{ij}} \rfloor$ edges of H between S_i and S_j . We will show that if H is a minor of G , this means H has a particular structure — and that when we later choose a random element $H \in \mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$, this structure will be unlikely to occur.

If H is a minor of G , this means there is some disjoint collection $(V_i)_{i=1}^t$ of disjoint connected non-empty subsets forming a model of H in G . We will use the properties of G_0 to deduce certain pairs of parts cannot be adjacent in the model, and therefore cannot be edges of H .

Claim 3.19.1. *For any collection $(V_i)_{i=1}^t$ of disjoint non-empty subsets of $V(G)$, there exist indices $a \leq b$ such that the density of pairs (V_i, V_j) where $i \in S_a, j \in S_b$ and V_i is (distinct from and) non-adjacent to V_j is at least $\rho_{a,b} = \epsilon^2 d^{\epsilon_1 - D_{a,b}} 2^{-13}$.*

Before proving this claim, we will show how it implies the theorem. Let H now be a randomly chosen element of $\mathcal{D}(\vec{\sigma}, \mathbf{D})(t, d)$. If H is to be a minor of G , there must be some model $(V_h)_{h \in [t]}$ of H in G . Given this model, our claim gives indices a, b as above. We take D sufficiently large that $\epsilon^2 D^{\epsilon_1/2} 2^{-13} > 1$, so we may assume $D_{a,b} > \epsilon_1/2$ as else $\rho > 1$ which cannot be the case since it is a density.

This gives us some structural information about H : all of the edges between S_a and S_b avoid the $\rho_{a,b}$ proportion of non-adjacent pairs — we would like to show this is unlikely. Since H is chosen randomly, this means the edges between S_a and S_b form a uniformly random $\lfloor td^{D_{a,b}} \rfloor$ subset. We now break into two cases for the analysis, and firstly assume $a \neq b$.

In this case, we are picking $\lfloor td^{D_{a,b}} \rfloor \geq \frac{1}{2}td^{D_{a,b}}$ edges. The probability we avoid the $\rho_{a,b}$ proportion can then be bounded as

$$\begin{aligned} \mathbb{P}(\text{we avoid these bad edges}) &\leq \binom{(1 - \rho_{a,b})|S_a||S_b|}{\lfloor td^{D_{a,b}} \rfloor} / \binom{|S_a||S_b|}{\lfloor td^{D_{a,b}} \rfloor} \leq (1 - \rho_{a,b})^{\frac{1}{2}td^{D_{a,b}}} \\ &\leq \exp(-\frac{1}{2}td^{D_{a,b}}\rho_{a,b}) \leq \exp(-\epsilon^2 td^{\epsilon_1}/2^{14}). \end{aligned}$$

If instead $a = b$, we can perform the analogous calculation for edges within S_a , and obtain the following equivalent bound:

$$\mathbb{P}(\text{we avoid these bad edges}) \leq \binom{(1 - \rho_{a,b})\binom{|S_a|}{2}}{\lfloor td^{D_{a,b}} \rfloor} / \binom{\binom{|S_a|}{2}}{\lfloor td^{D_{a,b}} \rfloor} \leq \exp(-\epsilon^2 td^{\epsilon_1}/2^{14}).$$

Therefore, for a random graph H , we have bounded the probability a fixed collection of subsets forms a model of H in G by this final quantity. If we can also show there are not too many collections of subsets which could form a model, we will be able to show our result. In fact, it suffices to consider only minimal models; if a graph does not have any minimal models, it will have no models at all. Using the discussion before our theorem, we will associate a minimal model (V_h) in G with a k -blobbing in G_0 . We can now use an encoding argument to count such blobbings.

Start by listing (in order) kd vertices of G_0 ; these are the vertices that will be allowed to be used in the blobbing (and note since every blobbing has total size at most kd , this is valid). Next, take a $\{0, 1\}$ valued sequence of length kd , where the first 1 denotes the index where we start listing V_1 , the second where we have finished listing V_1 at the previous index and started listing V_2 , and finally the $t + 1^{\text{st}}$ 1 denoting where we stopped listing V_t at the previous index. Any duplicate vertices are ignored. This gives an (injective) way to encode any given blobbing, and it is easily seen there are at most $(2d)^{kd}$ such encodings, hence at most that many blobbings.

We can now take D large enough that $(2d)^{kd} \exp(-\epsilon^2 d^{\epsilon_1} t/2^{14}) \leq 2^{-t}$ for all $t \geq d \geq D$. Having done so, except with probability at most 2^{-t} a randomly chosen H will not have any minimal model in $G_0(k)$, and so H will not be a minor of G . Thus to complete the proof of Theorem 3.19 it suffices to prove Claim 3.19.1.

Proof of Claim 3.19.1. We start by replacing each V_i with the corresponding projection to $V(G_0)$; this does not affect adjacency, and our subsets now form a k -blobbing in G_0 . Recall

that H is equipped with a vertex partition $(S_i)_{i=1}^r$, where $|S_i| = \lfloor \sigma_i t \rfloor$ or $\lceil \sigma_i t \rceil$. Suppose that X_i is the (ordered) collection of subsets $(V_{\sum_{j=1}^{i-1} s_j+1}, \dots, V_{\sum_{j=1}^i s_j})$ corresponding to S_i , and let b_i denote the average size of these $|S_i|$ subsets. Since the subsets $(V_i)_{i=1}^t$ form a k -blobbing in G_0 , $\sum_i |S_i| b_i \leq dk \leq (1 - \epsilon) \gamma t l$. In particular, $\sum_i \sigma_i t b_i < (1 - 2\epsilon/3) \gamma t l$ — since all $\sigma_i t > \gamma t > \epsilon D$, the replacing $\lfloor \sigma_i t \rfloor$ with $\sigma_i t$ cannot increase the sum by a large factor.

Let $\beta_i = b_i/l + \epsilon\gamma/3$, and so $\sum_i \sigma_i (\beta_i/(1 - \epsilon/3)) \leq \gamma(\vec{\sigma}, \mathbf{D})$. Recall $\gamma(\vec{\sigma}, \mathbf{D})$ is the minimum of $\sum \sigma_i \beta'_i$ subject to all $\beta'_i \beta'_j \geq D_{ij}$, and $\beta'_i > 0$. Therefore, there must be some indices a, b such that $\beta_a \beta_b < D_{ab}(1 - \epsilon/3)^2 \leq D_{ab}(1 - \epsilon/2)$, and in particular $b_a b_b \leq D_{ab}(1 - \epsilon/2) l^2$. Since $\beta_a, \beta_b > \epsilon\gamma/3$, we must have $D_{ab} > \epsilon^2 \gamma^2 / 9$.

Let Y_a be the set of $w \in S_a$ such that $|V_w| \leq (1 + \epsilon/8) b_a$, and analogously for Y_b . By Markov's inequality, we have $|Y_a| \geq \epsilon |S_a| / 16$. Suppose that $A = V_v$ is a part corresponding to some $v \in Y_a \subset V(H)$, and $B = V_w$ likewise for $w \in Y_b$. Let $\rho = d^{\epsilon_1 - D_{ab}}$, and let Y_a^* be the set of vertices $i \in Y_a$ such that the corresponding part $V_i \subset V(G_0)$ is disjoint from and not adjacent to at most $\epsilon |S_b| \rho / 2^8$ parts corresponding to a vertex in Y_b . In particular, elements of Y_a^* must be non-empty.

Suppose first that $|Y_a^*| \geq \epsilon |S_a| / 32$, and similarly for Y_b^* . We will build up families $(A_i)_{i=1}^{n_a}$ and $(B_j)_{j=1}^{n_b}$ of disjoint sets from Y_a^*, Y_b^* respectively, one element at a time and in a balanced fashion, such that there is at most a density $\rho/2$ of disjoint pairs between them. We would like to attain $n_a = n_b \geq d^{D_{ab} - \epsilon_1}$. Suppose we have not yet done so, and thus reordering if necessary $n_a \leq n_b \leq d^{D_{ab} - \epsilon_1}$. We would like to find another part corresponding to a vertex in Y_a^* that we can add to the family (A_i) , and therefore increase $n_a + n_b$ to get closer to our goal.

For A a part (i.e. some subset V_v) corresponding to a vertex $v \in Y_a^* \subset V(H)$, and B likewise for some $w \in Y_b^*$, we have $1 \leq |A||B| \leq l^2$, and so $|A| \leq l^2$. This means each already chosen part intersects at most kl^2 other parts, and in particular either $n_a \geq d^{D_{ab} - \epsilon_1} \geq d^{D_{ab}(1 - \epsilon/8) + \epsilon_1}$, or there are at least $\frac{\epsilon}{32} \sigma_a t - d^{D_{ab} - \epsilon_1} kl^2 \geq \frac{\epsilon}{64} |S_a|$ parts which correspond to a vertex in Y_a^* , are so far unchosen, and are disjoint from all already chosen sets. We remark that this is the only place where we use the lower bound on σ_a (except through the implicit upper bound on r).

If all of these parts are non-adjacent to at least $\rho n_b / 2$ parts corresponding to vertices from Y_b^* , by averaging some part $V_w : w \in Y_b^*$ is disjoint from and non-adjacent to at least $\epsilon \rho / 128 |S_a|$ sets from Y_a^* . This is a contradiction, since then w would not be in Y_b^* ! Hence some choice of part A is adjacent to at most $\rho n_b / 2$ such parts, and we can therefore add it to our family (A_i) . We continue in this fashion to build up such families with $n_a = n_b \geq d^{D_{ab} - \epsilon_1 / 8}$. But this now contradicts the definition of G_0 ! Hence our assumption $|Y_a^*| \geq \epsilon |S_a| / 32, |Y_b^*| \geq \epsilon |S_b| / 32$ must have been incorrect.

We therefore have (reordering if necessary) that $|Y_a \setminus Y_a^*| \geq \epsilon |S_a| / 32$. In particular,

there are at least $\epsilon^2 \rho |S_a| |S_b| / 2^{13}$ pairs $v \in Y_a \setminus Y_a^*$, $w \in Y_b$ such that V_v and V_w are disjoint and non-adjacent, completing the proof of the claim. $\square$

As discussed earlier, this completes the proof of Theorem 3.19. $\square$

We remark that our graphs (for fixed p, d, ϵ_1 but different t) are different sized blowups of the same base (pseudorandom) graph, and so will look very similar. It is interesting that the random structure required is universal.

We now move onto the lower bound for complete bipartite graphs. For $H = K_{ft/\log t, t}$, we have $\gamma_2(H) \leq 2\sqrt{f/\log t} = o(1)$ and hence we cannot apply the previous theorem — we have $\sigma_1 < \epsilon$. However, we are only looking for a complete bipartite graph, and even a single non-edge will suffice for our purposes, rather than the $\epsilon^2 \rho$ proportion required earlier. This allows us to mimic the proof of the previous theorem in this case — we leave it as a sketch, with the details able to be filled in as in Theorem 3.19.

Theorem 1.13. *Let $\epsilon > 0$. Then there is a constant C such that for all $C < f \leq \log t$, there is a graph satisfying $e(G)/|G| \geq (2\alpha - \epsilon)t\sqrt{f}$ and no $K_{ft/\log t, t}$ minor. In particular, $c(K_{ft/\log t, t}) \geq (2\alpha - \epsilon)t\sqrt{f}$.*

Proof Sketch. Let $d = ft/\log t$ (so H has td edges), and let G_0 be the graph $G_{d,p,\epsilon}$ from Lemma 3.18, with $p = 0.715\dots$ and $q = 1 - p$. Let $G = G_0(k)$, where $k = \lfloor 2(1 - \epsilon) \log t \sqrt{1/f \log(1/q)} \rfloor$. This G has at least $2(1 - 2\epsilon)t\sqrt{f/\log(1/q)}$ vertices, and density at least $p - \epsilon$ — hence average degree at least $2(\alpha - 3\epsilon)t\sqrt{f}$. We again neglect rounding throughout the proof by taking C sufficiently large.

Suppose we have an H model in G , with subsets $X_1 = \{A_1, \dots, A_{ft/\log t}\}$ representing the left hand class. and $X_2 = \{B_1, \dots, B_t\}$ the right hand. If b_i is the average size of the sets in X_i , by the definition of γ_2 , we must have $b_1 b_2 \leq (1 - \epsilon)^2 \log_{1/q} t$. Further, at least $\epsilon/8$ proportion of sets (from each of X_1 and X_2) have size at most $(1 + \epsilon/8)$ times the average size by Markov's inequality. Call these sets Y_1, Y_2 and note that if $A_i \in Y_1, B_j \in Y_2$ then $|A_i| |B_j| \leq (1 - \epsilon) \log_{1/q} d$.

We note that these sets are sufficiently small that the results of Lemma 3.18 will guarantee that for any $d^{1-\epsilon}$ sets representing the left hand class, and the same for the right hand class which are pairwise disjoint, then some non-zero density (and in particular, at least 1) of the pairs is non-adjacent.

How can we find these disjoint sets? Each set must have size at most $\log t / \sqrt{f \log(1/q)}$, since we can assume all sets under consideration are non-empty, and so intersects at most $O((\log t)^2 / f)$ other sets. It follows we can greedily find sufficiently many (of order $t^{1-\epsilon}$) disjoint sets to apply Lemma 3.18. This contradicts that $(A_i), (B_j)$ was a H model, since if a positive density of pairs are non-adjacent then in particular at least one such pair is. $\square$

We finish with some remarks about this proof. Our key new argument was building

up, in a balanced fashion, the two families (A_i) and (B_i) . This required the corresponding change in the previous section, where we built non-adjacent sets between two families, rather than within a single family. The argument is somewhat delicate: the constants involved must be chosen carefully so that we do not leave an empty set. However, the end result does have some ammount of flexibility. The exact density of non-adjacent pairs does not matter, only the asymptotic behaviour.

One limitation of our argument is the assumption $\sigma_i > \epsilon$. This is required for a subtle reason: otherwise, when we throw out the sets which intersect at least kl^2 other chosen pairs, we may be left with nothing at all. This technicality can be worked around in the argument for $K_{s,t}$, because even though $\sigma_1 = o(1)$, we are satisfied with even a single non-adjacent pair, and so can accept much worse bounds on their density (indeed, any strictly positive density will now do). The exact form of our lower bound on σ_i could likely be strengthened; however this choice gives a clean statement of the theorem while remaining quite applicable.

Chapter 4

Minors of bipartite subgraphs

4.1 Introduction

Recall that every graph G contains a bipartite subgraph G' such that the average degree of G' is at least half the average degree of G . The aim of this chapter is to prove similar results for other graph parameters, namely the Hadwiger number and topological clique number.

The motivation for these results was a recent paper of Hickingbotham and Wood, who rederived a series of related results for other graph parameters; we state below their results for our parameters of interest.

Theorem 4.1 ([30, Theorem 6]). *Let $f_{\succ}(t)$ be the smallest integer such that every graph G with $h(G) \geq f_{\succ}(t)$ contains a bipartite subgraph G' with $h(G') \geq t$. Then $f_{\succ}(t)$ exists, and is at most $(4\alpha + o_t(1))t\sqrt{\log t}$, where $\alpha = 0.319\dots$ is the constant of Definition 1.6.*

Theorem 4.2 ([30, Theorem 8]). *Let $f_T(t)$ be the smallest integer such that every graph with $tcl(G) \geq f_T(t)$ contains a bipartite subgraph G' with $tcl(G') \geq t$. Then $f_T(t)$ exists, and is at most $(1 + o_t(1))t^2$.*

For motivation, we will briefly outline their proofs. Suppose G has Hadwiger number (or topological clique number) n . We can delete edges and vertices until we are minimal with the same Hadwiger number (topological clique number), and so assume G consists of exactly n disjoint trees, with one edge between each pair of them. We can use a modified version of the classical result to find as a minor (topological minor) some graph G' with exactly n vertices and average degree at least $n/2$ within a bipartite subgraph of G . We can then use bounds on the extremal function for minors (topological minors) to deduce if n is at least $f_{\succ}(t)$ ($f_T(t)$), then G' contains in turn a K_t minor.

I believed their result for the Hadwiger number could be improved, since any graph (with suitable connectivity) which has density p and average degree at least $(p/\sqrt{\log(1/(1-p))} + o(1))t\sqrt{\log t}$ contains K_t as a minor, and this is tight — this follows

from the argument of Chapter 2, or of Thomason. The quantity α arises from maximising this over all p . However, in the argument of Hickingbotham and Wood we apply this to a graph with known density $\frac{1}{2}$ so can use a smaller constant.

If the bipartite subgraph obtained from the above argument was guaranteed to also have reasonable connectivity, this would complete our proof, however this is not the case. Fortunately, we can still use this idea to get most of the way and use the additional structure of the K_n minor to finish our proof. We will prove the following pair of matching bounds.

Theorem 1.15. *For any $\epsilon > 0$ there is a $t_0(\epsilon)$ such that for any $t > t_0$, any graph with Hadwiger number at least $(1 + \epsilon)t\sqrt{\log_2 t}$ contains a bipartite subgraph with Hadwiger number at least t .*

Theorem 1.14. *For all $\epsilon > 0$, there is a t_0 such that for all $t > t_0$, if $f(t) = \lceil (1 - \epsilon)t\sqrt{\log_2 t} \rceil$, then there is a graph with a $K_{f(t)}$ minor, and in fact a $K_{f(t)}$ topological minor, which has no bipartite subgraph with Hadwiger number at least t .*

We will also use these new techniques to improve results for topological minors. We will prove the following pair of bounds, which determine the asymptotic growth rate of $f_T(t)$.

Theorem 1.17. *If G is a graph with topological clique number at least $(\frac{1}{2} + o_t(1))t^2$, then G has a bipartite subgraph H with $tcl(H) \geq t$.*

Theorem 1.18. *For all t , there is a graph G with $tcl(G) \geq t^2/4$ but G has no bipartite subgraph H with $tcl(H) \geq t$.*

The results of this section originally appear in my pre-print [70].

4.2 RB-bipartite graphs

We will begin by introducing RB-bipartite graphs which serve as a general framework for our results. For the remainder of this section, we will be proving general properties of these graphs. The next section will focus on our results for minors, and finally in Section 4.4 we will prove results for topological minors.

Definition 4.3. Let H be a graph equipped with a 2-edge-colouring (with colours Red and Blue). Then H is *RB-bipartite* if every cycle uses an even number of Red edges (call such cycles R-even). Equivalently, H has no R-odd cycle.

This definition is not symmetric under interchanging colours: for instance, a K_3 coloured Red is not RB-bipartite, but a Blue coloured K_3 is. It is easily seen that being RB-bipartite is equivalent to being bipartite if every Blue edge is subdivided exactly once. The following lemma gives some other equivalent definitions.

Lemma 4.4. *The following statements are equivalent.*

- (1) H is RB-bipartite.

- (2) H has no circuit using an odd number of Red edges (no R -odd circuit).
- (3) There is a partition (X, Y) of $V(H)$ such that all edges between X and Y are Red, and all edges within X or within Y are Blue (we call such a partition an RB -bipartition of H).

Proof. It is easily seen that (3) $\implies$ (2) $\implies$ (1), and so it remains to prove that RB -bipartite graphs have a partition as in (3). By taking a union over components, assume H is connected. Let H_R denote the graph on the Red edges, and H_B the graph on Blue edges. It is easily seen that H_R is a bipartite graph. We will work on each component of H_R in turn, which we call Red-connected components. Let C_1 be such a component, with bipartition X_1, Y_1 . If there is a Blue edge between X_1 and Y_1 , then also taking a path in H_R from between the endpoints we get a cycle using an odd number of Red edges, contradiction. So the induced subgraph of H on C_1 is RB -bipartite.

Suppose that H has Red-connected components $C_1, \dots, C_r$; reordering if necessary assume that $H[\cup_{i \leq s} C_i]$ is connected for each $1 \leq s \leq r$. Suppose that, for given s we have an RB -bipartition (X, Y) of $H[\cup_{i \leq s} C_i]$; such a partition exists for $s = 1$ by the above. We aim to build such a partition for $H[\cup_{i \leq s+1} C_i]$ and so inductively build our partition of H . Note $H[C_{s+1}]$ has an RB -bipartition (X_{s+1}, Y_{s+1}) by the same argument as for C_1 .

By definition of Red-connected component, there are only Blue edges from $\cup_{i \leq s} C_i$ to C_{s+1} . Suppose there is a Blue edge from X_{s+1} to both X and Y . Then taking an even length Red path between the endpoints in C_{s+1} , and any path between the endpoints in X and Y (by connectedness such a path exists, and by our partition it uses an odd number of Red edges) we can obtain an R -odd cycle in H ; this is a contradiction. By the same argument, X cannot have Blue edges to both X_{s+1} and Y_{s+1} . So reordering, we can assume the only Blue edges lie between X and X_{s+1} , and Y and Y_{s+1} . But then $(X \cup X_{s+1}, Y \cup Y_{s+1})$ is our desired partition of $H[\cup_{i \leq s+1} C_i]$. This completes our proof. We remark the partition constructed is unique for connected graphs. $\square$

Theorem 4.5. *Every graph H equipped with a Red/Blue edge colouring has an RB -bipartite subgraph H' with at least $\frac{1}{2}e(H)$ edges.*

Proof. For a partition X, Y of $V(H)$, define $e_R(X)$ to be the number of Red edges with both ends inside X , $e_R(X, Y)$ the number of Red edges with one endpoint in X and one endpoint in Y , and similarly $e_B(X)$, $e_B(X, Y)$ for Blue edges.

We construct a random partition by placing each vertex independently in X or Y with probability $\frac{1}{2}$. Let $d(X, Y) = e_R(X, Y) - e_B(X, Y)$. Then since each edge is between X and Y with probability $\frac{1}{2}$, we have $\mathbb{E}(d(X, Y)) = \frac{1}{2}(e_R(H) - e_B(H))$. Pick some choice of X, Y such that $d(X, Y) \geq \mathbb{E}(d(X, Y))$. Construct the subgraph H' consisting of all Blue edges inside X or Y , and all Red edges from X to Y . By Lemma 4.4 H' is RB -bipartite,

and moreover

$$e(H') = (e_B(H) - e_B(X, Y)) + e_R(X, Y) \geq \frac{1}{2}(e_R(H) + e_B(H)) = \frac{1}{2}e(H)$$

so H' is our desired subgraph. $\square$

For our theorems, we can assume our host graph G is a minimal K_t model by passing to a minimal subgraph; in other words, G consists of some subsets $(V_i)_{i=1}^t$ with each $G[V_i]$ a tree, and exactly one edge from V_i to V_j for each $i \neq j$. We will build a 2-coloured K_t representing the lengths of paths between special vertices in this graph — for a topological minor, these special vertices will be the branch vertices.

We will be converting questions about bipartite subgraphs (of some minimal K_t model) into questions about RB-bipartite subgraphs of a Red/Blue coloured K_t . For topological minors, there is a natural way to make this conversion — if the subdivided edge between branch vertices has odd length, colour it Red, otherwise Blue. It is easily seen a cycle in the topological minor has odd length exactly when the corresponding cycle in the coloured graph is R -odd.

For minors, it is less clear this can be done. We do not have branch vertices, and further issues arise because a bipartite subgraph need not use the entirety of a part V_h . Fortunately, it turns out we can avoid the latter problem by restricting to subgraphs containing all $G[V_h]$.

Definition 4.6. Let G be a graph, with a partition into n parts V_i , such that all $G[V_i]$ are trees, and there is exactly one edge from V_i to V_j . Let $v_i \in V_i$ be an arbitrary choice of roots. The *canonical path* from v_i to v_j , where there is an edge from V_i to V_j , is the unique $v_i v_j$ path in $G[V_i \cup V_j]$.

The *auxiliary graph* $H[G]$ has vertex set $[n]$, and an edge from i to j whenever there is an edge between V_i to V_j . This edge is coloured Red if the canonical $v_i v_j$ path has odd length, and Blue if it has even length.

We remark that strictly this definition of $H[G]$ also depends on the choice of roots, however we suppress this notation as the choice does not matter.

Lemma 4.7. *Let $G, H[G]$ be as in Definition 4.6, and let H' be some (coloured) subgraph of H . Let G' be a subgraph of G containing all edges within a part V_i , together with the edge from V_i to V_j if and only if $ij \in E(H')$. Then G' is bipartite if and only if H' is RB-bipartite.*

Proof. We start by proving the ‘only if’ statement. By relabelling, we can assume we have a cycle $1, \dots, k$ in H' , where an odd number of edges $i, i+1$ are Red (and indices are taken modulo k). We aim to construct an odd cycle (or indeed circuit) in G' .

We build a circuit C in G as follows. Start by traversing the canonical path from v_1

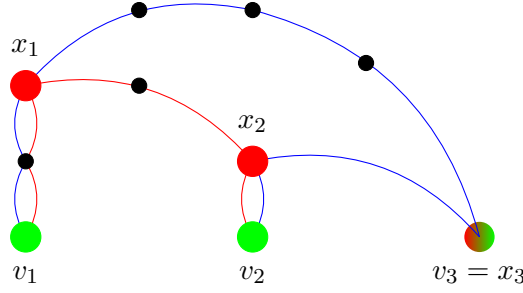


Figure 4.1: This figure illustrates how an odd cycle (or circuit) in the auxiliary graph necessarily produces an odd cycle in the main graph. The Red path is the unique odd edge, the Blue paths are even. The length of the sub-cycle of this odd circuit is also odd.

to v_2 (and call this path P_1), then continue to v_3 along canonical path P_2 , and so on. The circuit C clearly has odd length, since it has an odd number of odd length segments, and so G' cannot be bipartite (see Figure 4.1 for an illustration).

The forwards direction is similar; suppose that G' is not bipartite, and consider an odd cycle C in G' . Fix some starting point and orientation of the cycle. Relabelling, suppose that we pass through vertex classes $V_1, \dots, V_k$ (where we may have repetition) in this order — since there is only one edge between each pair of V_i , and each class is a tree, we have $k \geq 3$ and V_{i-1}, V_i, V_{i+1} are all distinct. Further, in H' we must have that i and $i+1$ are adjacent. Let Q_i be the segment of C inside V_i (between V_{i-1} and V_{i+1} , and inclusive of the edge from V_{i-1} to V_i). This is well defined since there is a unique edge from V_{i-1} to V_i and so no pair repeats.

Since V_i is connected, there is some path P'_i inside V_i from v_i to this path segment which meets the path in exactly one vertex; call its endpoint x_i . Then following P'_i from v_i to x_i , then traversing Q_i and Q_{i+1} until we reach x_{i+1} , and finally traversing P'_{i+1} to v_{i+1} must be the canonical path from v_i to v_{i+1} (it is a path contained in $V_i \cup V_{i+1}$ and so unique).

This produces an odd circuit, since each new edge we add is traversed twice. But this circuit is a union of canonical paths, an odd number of which must have odd length. In particular, the corresponding circuit $1 \dots k$ in H' has an odd number of Red edges, and therefore H' is not RB-bipartite. $\square$

4.3 The Hadwiger number

The aim of this section is to prove Theorems 1.14 and 1.15, which form a lower and upper bound for $f_{\succ}(t)$ respectively.

4.3.1 The lower bound

Proof of Theorem 1.14. Let H be a graph on vertex set $[n]$. Consider the graph $G = G(H)$ consisting of n special ‘branch vertices’ $v_1, \dots, v_n$, and for each $i < j$ a path between v_i and v_j of length either 1 (i.e. an edge between them) if ij is an edge of H , or 2 if not. Equivalently, start with a copy of K_n , and subdivide each edge $v_i v_j$ once when ij is not an edge of H . Viewing G as a K_n model in some way, with $v_i \in V_i$, the auxiliary graph $H[G]$ is a copy of K_n , with H coloured Red and H^c coloured Blue.

Clearly, the $G(H)$ has Hadwiger number exactly n for any H (since we must always contract or delete all the internal vertices of induced paths for $t \geq 3$). We will let $H \sim G(n, \frac{1}{2})$ be a random graph. We would like to say something about bipartite subgraphs of $G(H)$. Recall a subgraph of $G(H)$ consisting of a union of paths between branch vertices is bipartite exactly when we take paths corresponding to an RB-bipartite subgraph of the auxiliary graph; in particular maximal bipartite subgraphs of G are formed by taking some bipartition (X, X^c) of $[n]$, and taking all length 1 paths from X to X^c , together with all length 2 paths contained in either X or X^c . We assume our bipartite subgraph is one of the at most 2^n graphs formed in this way.

Fixing X , and sampling H from $G(n, \frac{1}{2})$, each edge of $G(H)$ within or disjoint from X has length 2 with probability $\frac{1}{2}$, and each edge from X to X^c has length 1 with probability $\frac{1}{2}$. This means that all maximal bipartite subgraphs of $G(H)$ are distributed like (subdivisions of) $G(n, \frac{1}{2})$. The subdivision does not affect the Hadwiger number. In particular, if we can show the probability $G(n, \frac{1}{2})$ has a K_t minor is below 2^{-n} , there is some choice of H for which there is no bipartite K_t minor.

Bollobás, Catlin and Erdős [5] showed that for $H \sim G(n, \frac{1}{2})$, and some fixed partition of $V(H)$ of $V(G(n, \frac{1}{2}))$ into s parts, the probability (over random H) this partition is K_s -compatible is at most $\exp(-\binom{s}{2} 2^{-n^2/s^2})$; recall that this means there is an edge between each pair of parts, and that if $(X_i)_{1 \leq i \leq s}$ is a K_s model, the (X_i) also form a K_s -compatible partition.

Taking $s = n/\sqrt{\log_2 n - 3 \log_2 \log_2 n}$, there are at most n^n partitions of $V(H)$ into s parts. Since $n^n \exp(-\binom{s}{2} 2^{-n^2/s^2}) = o(2^{-n})$, the probability $G(n, \frac{1}{2})$ has a K_s compatible partition is less than 2^{-n} for sufficiently large n . In particular, this applies for K_s minors, and so for large t there is some choice of H for which no bipartite subgraph of $G(H)$ has Hadwiger number at least $s = (1 + o(1))n/\sqrt{\log_2 n} \leq t$. $\square$

4.3.2 The upper bound

Proof of Theorem 1.15. To simplify our argument, we will prove the equivalent statement that every graph with $h(G) \geq n = t\sqrt{\log_2 t}$ contains a bipartite subgraph G' with $h(G') \geq (1 - 2\epsilon)t$. Let G be such a graph. For the rest of this proof we will neglect rounding, since this can be absorbed into the value of ϵ .

We delete as many edges as possible from G while retaining a K_n minor, and so we can assume that G consists of n disjoint subsets $V_1, \dots, V_n$, with each $G[V_i]$ a tree, as well as exactly one edge between each pair V_i and V_j ($i \neq j$).

Recall from Definition 4.6 the canonical path and auxiliary graph $H[G]$; we will use these notions here. We start by reserving a set $S \subset V(H)$ of ϵn vertices for later use, and apply our theorems with $G' = G[\cup_{i \notin S} V_i]$. Note that $H[G']$ is a 2-coloured $K_{(1-\epsilon)n}$.

By Theorem 4.5, $H[G']$ has an RB-bipartite subgraph H' (with RB-bipartition (A, B)) with at least $\frac{1}{2} \binom{(1-\epsilon)n}{2}$ edges. By Theorem 4.7, the corresponding subgraph G'' of G' (consisting of all edges inside $G'[V_i]$, as well as the edge from V_i to V_j whenever ij is an edge of H') is bipartite, and this also holds for any proper subgraph of H' . Note also that a minor in H' directly corresponds to a minor in G'' (by first contracting each V_i to a vertex), and hence to a minor in G' which is bipartite.

We aim to find a K_m minor in G , where $m = (1 - 3\epsilon/2)n/\sqrt{\log_2 n} = (1 - 3\epsilon/2)t$, which is bipartite. While we cannot necessarily find a K_m minor in H' , the following result of Thomason [65], allows us to build a structure in H' which is almost a K_m minor, provided that t is sufficiently large. The structure we obtain is called a compatible partition, c.f. Definition 2.8.

Theorem 4.8 ([65, Corollary 3.1]). *Let $\epsilon > 0$. Then there is an $n_0(\epsilon)$ such that for all $(\log \log n)^{2+\epsilon} < p < 1 - (\log n)^{-1/\epsilon}$, and all graphs F with $n > n_0$ vertices and density p , F has a K_m -compatible-partition for some $m \geq (1 - \epsilon)n/\sqrt{\log_{1/1-p} n}$.*

If in addition H' had ‘reasonably high’ connectivity, we could reserve subsets of H' (the Connector and Projector from Section 2.5) to convert this compatible partition into a minor. Since this may not be the case, we will use our reserved set S in both of these roles to extend our compatible partition in H' (hence bipartite compatible partition in G') to a bipartite minor in G .

Some care is required to ensure we remain bipartite when performing these roles. It turns out to be straightforward to perform the role of the projector, however to serve the role of the connector a nice observation is required.

Suppose our K_m compatible partition in H' consists of the pairwise adjacent disjoint subsets $U_1, \dots, U_m$ of $V(H')$, and we will start by performing the role of the projector.

Let us first restrict attention to U_1 , and suppose we have $U_A = U_1 \cap A$, $U_B = U_1 \cap B$ (recalling (A, B) is some fixed RB-bipartition of H'). Let i be some element of $S \setminus V(H')$. We want to add i to the subgraph H' in such a way we preserve RB-bipartiteness. Adding i to A , we can add all Red edges from i to B , and all Blue edges from i to A ; adding i instead to B reverses the colours. In particular, one of these choices must allow us to add edges from i to at least half of U . Make such a choice, and add i and all such edges to H' . Replace U with the set of vertices in U not adjacent in H' to i , and iterate this process.

In this way, we can construct a set $P(U_1)$ (and add its vertex set to H') of size at most $p(U_1) = (1 + \log_2 |U_1|)$, add some extra edges to H' such that every element of U_1 has a neighbour in $P(U_1)$, and extend the RB-bipartition (A, B) of H' to these new vertices and edges.

We can construct such sets in turn for each of $U_1, \dots, U_m$ using at most $\sum p(U_i) \leq t + t \log(n/t)$ elements of S overall, assuming $|S| > 2t + t \log(n/t)$. Since this holds for our value of n , we can obtain an RB-bipartite extension H'' of H' , which also contains disjoint new sets $P(U_i)$ of size at most $p(U_i)$, where each vertex of U_i has a neighbour (in H') inside $P(U_i)$. Let $S' = S \setminus \cup_i P(U_i)$, and note $|S'| \geq \epsilon n/2$.

Having shown our reserved set can serve as the projector, we need to show it also serves the role of the connector. The following lemma shows that either it can, or we can directly obtain a RB-bipartite K_t minor in H' (and hence a bipartite K_t minor in G).

Lemma 4.9. *Let H' be an RB-bipartite graph, $x, y \in V(H')$, and let $T \subset S'$ have size t , where S' is disjoint from H' . Suppose that every vertex of T is joined to every vertex of $T \cup H'$ by either a Red or a Blue edge. Then either T forms an RB-bipartite K_t , or there is some set of at most 2 vertices from T which can be added to H' to form the internal vertices of an xy path while preserving RB-bipartiteness.*

Proof. Let H' have RB-bipartition (A, B) . Suppose first that $x, y \in A$ (this corresponds to joining v_i and v_j by a path using an even number of Red edges). If some vertex is joined to both x and y by edges of the same colour, we can create a length 2 path from x to y , adding the new vertex to A or B depending on parity.

So we can assume this does not happen. Let X be the set of vertices joined to x by a Red edge, and Y those vertices joined to y by a Red edge; note this forms a partition of T . If some pair of vertices v, w in X are joined by a Red edge, then $xvwy$ is an even length path from x to y ; we can extend our bipartition by adding v to A and w to B . The same argument works if $v, w \in Y$.

If $v \in X, w \in Y$, and vw is Blue, then we likewise get a length three path. So we can assume that these cases never happen. But then (X, Y) forms an RB-bipartition of T . Since the graph on T is complete, we directly have an RB-bipartite K_t subgraph.

If instead $x \in A, y \in B$ the argument is similar. If some vertex is joined to x and y by different colours, we obtain a path of length 2 from x to y which is R-odd. Therefore, we can assume the set R of vertices joined to x and y by Red edges, and B for Blue edges form a partition of T . As above, and illustrated in Figure 4.2, if there is a Red edge within either R or B , or a Blue edge between R and B , we get an R-odd path from x to y of length at most 3. So we can assume this doesn't happen, and therefore the partition (R, B) shows that T is directly an RB-bipartite K_t . $\square$

We will now use Lemma 4.9 and the set S' to sequentially connect pairs from H'' and

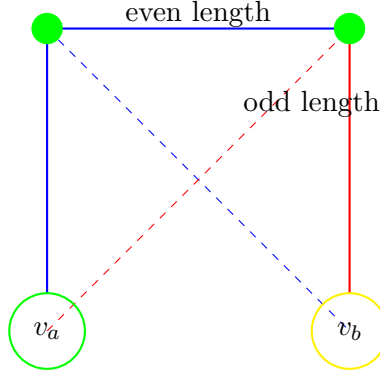


Figure 4.2: The green and yellow nodes represent A and B respectively. The thick lines show how to add the edge while remaining bipartite.

hence obtain a minor; we can assume that we never obtain an RB-bipartite K_t subgraph in the lemma as otherwise we are directly done. We will first label $P(U_1)$ as $x_1, \dots, x_{|P(U_1)|}$. Under this assumption, $t + 2l$ vertices suffice to connect each of l pairs sequentially.

Since $\sum p(U_i) + t \leq |S'|/2$, we can use two new elements of S to join x_1 to x_2 , then x_2 to x_3 and so on while remaining RB-bipartite. We can then repeat this for U_2 and so on. Suppose that the internal vertices of the paths used to connect $P(U_i)$ are $C(U_i)$. We now replace U_i with $V_i = U_i \cup P(U_i) \cup C(U_i)$, and extend the subgraph H'' [in a RB-bipartite fashion] to include $\cup C(U_i)$, with the edges along the paths also added. Then V_i forms a connected subset containing U_i , and so the V_i form an RB-bipartite K_m minor. This corresponds to a bipartite K_m minor in G . $\square$

The key observation in our proof is that graphs of density $\frac{1}{2}$ contain larger compatible partitions than guaranteed by their average degree, since their density is non-optimal. It then remained only to convert a compatible partition into a minor.

Here, the intuition is that reserving a small proportion of vertices, we have enough structure to convert a compatible partition into a minor. Of course the difficulty is then preserving bipartiteness; the proof is completed by a neat observation that if all choices fail, this is because of some significant bipartite structure in the reserved vertices, and hence we are done by a direct argument.

4.4 The topological clique number

The aim of this section is to prove Theorems 1.17 and 1.18, which form upper and lower bounds on $f_T(t)$ respectively. The proofs are shorter than those for $f_{\succ}(t)$, though the upper bound requires a careful induction statement.

Proof of Theorem 1.17. By replacing G with a minimal subgraph with the same topological

clique number, we can assume that $G = TK_t$. As before, considering the auxiliary graph we can reduce our statement to the following claim.

Claim 1.17.1. *Let $t \geq 2$, and let G be a 2-coloured copy of $K_{\binom{t+2}{2}-3} = K_{t^2/2+O(t)}$. Then H contains an RB-bipartite topological K_t minor.*

We will prove by induction on s the statement ‘For each $s \leq t$, $K_{\binom{t+2}{2}-3}$ contains a RB-bipartite topological K_s minor which uses at most $\binom{s+1}{2} - 2$ vertices’. The base case $s = 2$ is trivial since any edge works; the case $s = t$ is our desired claim. Assume the statement is true for all $s' < s \leq t$; we aim to prove it for s . Without loss of generality, assume that H contains no RB-bipartite K_t subgraph.

By induction, G contains a RB-bipartite topological K_{s-1} H using at most $\binom{s}{2} - 2$ vertices. Suppose that $V = \{v_1, \dots, v_{s-1}\}$ are the branch vertices. Let S consist all vertices of G not involved in the TK_{s-1} ; by assumption $|S| \geq \binom{t+1}{2} - 3 - ((\binom{t-1}{2} - 2)) \geq 2t - 2$. Let $w \in S$; we will aim to build a TK_s on branch vertices $V \cup \{w\}$. Let (A, B) be a RB-bipartition of $V(H)$. By the same argument used for the projector in the proof of Theorem 1.15, we can add w to either A or B along with at least half of the edges from w to H while remaining RB-bipartite. To extend to a topological K_s using at most $\binom{s+1}{2} - 2$ vertices, we aim to use at most $s - 1$ additional vertices to add the remaining (at most $\lfloor \frac{s-1}{2} \rfloor$) edges from w to V .

We will do so recursively. Let $C \subset S \setminus \{w\}$ have $|C| \geq t$. Consider a vertex v_i we do not yet have a path to. We need to either build an R-odd or R-even path, suppose first we need an R-odd path. Let $C_1 = \{v \in C : vv_i \text{ is a Red edge}\}$, and $S_2 = C \setminus S_1$. If there is a Blue edge from w to a vertex x of C_1 , the path vxw is an R-odd path. We are likewise done if there is a Red edge from w to some $y \in C_2$. We thus assume v is entirely Red-joined to C_1 and entirely Blue-joined to C_2 .

Suppose next there is a Red edge within C_1 , say xy . Then the path $wxyv_1$ is R-odd and uses two additional vertices. A similar argument works if there is a Blue edge between C_1 and C_2 , or a Red edge within C_2 . But this exactly means C with bipartition (C_1, C_2) is an RB-bipartite copy of K_t which we assumed does not exist. Hence in all cases we can build a path using at most 2 vertices. The same argument works for R-even paths by interchanging the roles of C_1 and C_2 throughout.

Since $|S \setminus \{w\}| \geq 2t - 3 \geq t + 2(\frac{s-1}{2} - 1)$, for each vertex in turn we can build such a path, starting from $C = S \setminus \{w\}$ and removing used internal vertices. This means we can build a topological K_s , using at most $\binom{s}{2} - 2 + (2\lfloor \frac{s-1}{2} \rfloor + 1) \leq \binom{s+1}{2} - 2$ vertices. This completes our inductive step, and so the proof of the theorem. $\square$

Proof of Theorem 1.18. Our lower bound relies on complete bipartite graphs being good examples of graphs with high average degree, and no topological K_t minor.

Let G be a complete bipartite graph on parts A, B , and suppose that there is a

topological K_t minor, say with s branch vertices in A and $(t - s)$ branch vertices in B . Then for each pair of vertices (in A , say), we need a distinct vertex in B to lie on the path between them. This means we need $|B| \geq (t - s) + \binom{s}{2}$, and so overall $|G| \geq t + \binom{s}{2} + \binom{t-s}{2} \geq t + 2\binom{t/2}{2} = t^2/4 + t/2$ by convexity.

If we take G to be a complete graph $K_{\lceil t^2/4 \rceil}$, bipartite subgraphs are all complete bipartite on fewer than $t^2/4 + t/2$ vertices and so cannot have a topological K_t minors. The result follows as $tcl(K_n) = n$. $\square$

It seems reasonable to conjecture that $C = \frac{1}{4}$ is tight (and indeed that cliques are extremal); our proof of $\frac{1}{2}$ for an upper bound only required this many vertices in a very structured case, which seems unlikely to occur. An existing upper bound of the shape ct^2 is twice the extremal function for topological minors. With the best known upper bound on this function (see [44]), this gives a value $c = 20/23 + o(1)$. If it were possible to improve this upper bound to the best known lower bound, we would obtain $c = 9/32 + o(1)$ (this lower bound is from bipartite random graphs, and due to Łuczak) — there remains only a small gap between this and our lower bound. This suggests the lower bound is likely to be hard to improve, since a substantial improvement would also improve bounds on the extremal function.

Chapter 5

Large complete multigraph minors

5.1 Introduction

In this chapter, we introduce the notion of containing a multigraph as a minor, and investigate the minor-extremal function for multigraphs of the form rK_t , consisting of t vertices, and all pairs of distinct vertices having r edges between them. For the purposes of this chapter, we restrict to the case where our host graph is a simple graph. This is both a natural setting, and easier to work in: there are few techniques to handle the structure of a general multigraph where the multiplicity of edges can be very large.

Our main results will be a lower bound on this extremal function, as well as a matching upper bound for two regimes: the ‘small r ’ case where $r = o(\log t)$, and the ‘large r ’ case where $r = \omega(\log t)$. We begin by recalling the definition of containing a multigraph as a minor.

Definition 5.1. A multigraph G consists of a vertex set $V(G)$, together with a multiset $E(G)$ of unordered pairs from G . We say the *multiplicity* of an edge uv is the number of times uv appears in $E(G)$, and is denoted $m_G(uv)$ (the G can be omitted when clear from context). A multigraph is called *loopless* if $m_G(uu) = 0$ for all $u \in V(G)$.

Definition 5.2. Let H be a multigraph. We say H is a *minor* of a multigraph G , $G \succ H$, if H can be obtained from G by a sequence of edge contractions, and vertex or edge deletions. An edge contraction denotes replacing two adjacent vertices u, v with a new vertex uv , and adding $m(uv) - 1$ loops at uv , together with $m(uw) + m(vw)$ edges to each other vertex w .

Definition 5.3. Let H be a multigraph. The *minor extremal function* is defined as follows:

$$c(H) = \max\{e(G)/|G| : G \text{ is a simple graph not containing an } H \text{ minor}\}.$$

The original motivation for this problem came from studying Thomason’s proof showing $c(K_t) = (\alpha + o(1))t\sqrt{\log t}$. I noticed this proof already constructed $r = o(\sqrt{\log t})$ edges

between each pair of vertices in the K_t minor, rather than just one. Thus, with a small amount of modification we can show $c({}_rK_t) = (\alpha + o(1))t\sqrt{\log t}$ for any $r = o(\sqrt{\log t})$. What happens for different values of r ?

We will begin by investigating lower bounds on $c({}_rK_t)$, which corresponds to finding graphs with large average degree and no ${}_rK_t$ minor. For simple graphs H , the best lower bounds for $c(H)$ arise from random graphs, and we will show they remain a good lower bound for multigraphs. Our random graphs will have varying density depending on the relationship between r and t , and so our lower bound is non-explicit. A graph of the optimised lower bound is included in Figure 1.3 of page 12. These graphs were obtained numerically with the SCIPY.OPTIMIZE and MATPLOTLIB python libraries.

We will prove the following properties of this lower bound; this forms our main theorem of Section 5.3. The rescaling $r = k \log t$ is used throughout this chapter, since the qualitative behaviour of the lower bound varies depending on whether k is small or large.

Theorem 1.23. *There are smooth increasing functions $p(k), q(k) : (0, \infty) \rightarrow [0, 1]$ such that*

$$c({}_{k \log t}K_t) \geq \left(\frac{1}{2} - \epsilon\right)pt\sqrt{(k \log t)/q}$$

holds for all sufficiently large t .

It is natural to conjecture this lower bound is asymptotically tight for all pairs r, t (where $rt \rightarrow \infty$). The remainder of this section will prove this lower bound is tight for the two asymptotic regimes $r = o(\log t)$ and $r = \omega(\log t)$; these results together show $c({}_rK_t) = \Theta(t\sqrt{r + \log t})$ always.

Theorem 1.24. *Let $\epsilon > 0$. Then there is a constant D such that for all $t > D$ and $r < (\log t)/D$, every graph G with $e(G)/|G| \geq (\alpha + \epsilon)t\sqrt{\log t}$ contains ${}_rK_t$ as a minor. In particular,*

$$c({}_rK_t) \leq (\alpha + \epsilon)t\sqrt{\log t}.$$

Theorem 1.25. *Let $\epsilon > 0$. Then there is a constant D such that for all $t > D$ and $r > D \log t$, every graph G with $e(G)/|G| \geq (\frac{1}{2} + \epsilon)t\sqrt{r}$ contains ${}_rK_t$ as a minor. In particular,*

$$c({}_rK_t) \leq \left(\frac{1}{2} + \epsilon\right)t\sqrt{r}.$$

The later sections of this chapter, Sections 5.4 to 5.6, will focus on proving these upper bounds. We will, as in previous chapters, break into dense and sparse case subtheorems. We will prove these sub-theorems, Theorems 1.24 and 1.25 in parallel, and where possible prove a more general result that can be specialised to both cases in order to reduce overlap.

Both of these results rely on t being large. What can we say if t is fixed, but r is large (so that we still have asymptotic results)? Even the case $t = 2$ is non-trivial, and corresponds to a natural simplest multigraph. Certainly ${}_rK_2$ is a minor of $K_{2\sqrt{r}}$, and so $c({}_rK_2) = O(\sqrt{r \log r})$. A lower bound, coming from $K_{2\sqrt{r}-1}$, is $c({}_rK_2) \geq (1 + o(1))\sqrt{r}$.

It turns out we can use the above theorems to asymptotically determine $c({}_rK_2)$ exactly, using a ‘blocking’ argument. This argument also applies in the case where t is fixed and arbitrary, see Corollary 1.26 on page 13.

Theorem 5.4. *For all $\epsilon > 0$, there is an R such that for all $r > R$, $c({}_rK_2) \leq (1 + \epsilon)\sqrt{r}$ holds.*

Proof. For simplicity we ignore rounding. Let $N = r^{1/4}$. Consider the graph $H = {}_{r/N^2}K_{2N}$. By contracting two disjoint collections of N vertices, we see H contains ${}_rK_2$ as a minor. Further, if $t = 2N$ is the number of vertices of H , it holds $r/N^2 = \omega(\log t)$. Thus Theorem 1.25 can be applied to show $c(H) \leq (\frac{1}{2} + \frac{\epsilon}{2})2N\sqrt{r/N^2} = (1 + \epsilon)\sqrt{r}$ provided that $t = 2r^{1/4}$ is sufficiently large. This completes our proof. $\square$

5.2 Relative entropy and Chernoff bounds

We will begin by introducing the relative entropy, and its application to binomial tail bounds — this quantity arises naturally when trying to find ${}_rK_t$ models in a random graph. The aim of this section is to prove some fundamental properties of the relative entropy, as well as introduce some preliminary properties of multigraphs as minors. As in the graph case, it is helpful to work with models rather than minors.

Definition 5.5. Let H, G be multigraphs. A collection $(V_h)_{h \in V(H)}$ of disjoint subsets of $V(G)$ is a *model of H in G* if all V_h are connected, and for each pair hh' of distinct vertices from H there are at least $m_H(hh')$ edges of G from V_h to $V_{h'}$, and for each h at least $m_H(hh) + |V_h| - 1$ edges within V_h .

The definition of model is a reasonably straightforward translation of the graph notion with the exception of the number of loops. This is because the notion of minor should still correspond to performing a topological contraction on the graph viewed as a simplicial complex, and means that minors and models remain equivalent as shown below.

Lemma 5.6. *Let H, G be multigraphs. Then H is a minor of G if and only if there is a model of H in G .*

Proof. Suppose that (V_h) is a model of H in G . For each $h \in V(H)$, pick a spanning tree T_h of $G[V_h]$, and in any order desired contract all edges of T_h . This results in a vertex (which we call h) obtained from T_h with by assumption at least $m_H(hh)$ loops at it. The number of edges from h to h' is likewise at least $m_H(hh')$, and so H is a subgraph of this minor of G .

For the converse direction, we reorder so first some vertices and edges of G are deleted, and then we perform a sequence of contractions $e_1, \dots, e_m$ which results in H as a minor. We partition $V(H)$ into sets $V_h = \{h\}$. Start by ‘undoing’ the contraction of e_m , and forming a new partition by replacing the vertex obtained by contracting e_m with the two endpoints of e_m . Repeating all the way back to e_1 we have disjoint connected subsets (V_h) of $V(G)$ with at least $m_H(hh')$ edges from V_h to $V_{h'}$ and the correct number of loops — i.e. a model of H in G as desired. $\square$

Our lower bounds will use the further relaxation to rK_t -compatible partitions, where as for simple graphs we drop the connectivity assumption from models. To bound the probability that some (fixed) partition of a random graph is compatible, we need to bound the probability a binomial random variable (the number of possible edges between a pair of parts which are present) does not exceed $r - 1$. An asymptotically tight bound for this is due to Chernoff and Hoeffding and requires the following definition.

Definition 5.7. Let $0 < p, q < 1$. The *Kullback-Leibler divergence* (alternatively *KL-divergence* or *relative entropy*) is the following quantity:

$$D(q||p) = q \log\left(\frac{q}{p}\right) + (1 - q) \log\left(\frac{1 - q}{1 - p}\right).$$

For clarity, throughout this section, q will not denote $(1 - p)$, but instead a different probability.

Theorem 5.8 (Chernoff-Hoeffding Theorem, [31, Theorem 1] and [3, Lemma 4.7.2]). *Let $0 < q \leq p < 1$ be real numbers, and suppose $X \sim B(n, p)$ is a binomial random variable. Then the following holds:*

$$\frac{1}{\sqrt{2n}} \exp(-nD(q||p)) \leq \mathbb{P}(X \leq qn) \leq \exp(-nD(q||p)). \quad (5.1)$$

Using a symmetry of the binomial distribution, and that $D(1 - q||1 - p) = D(q||p)$, we also have the following upper tail bounds, which hold for $0 < p \leq q < 1$:

$$1 - \frac{1}{2\sqrt{n}} \exp(-nD(q||p)) \geq \mathbb{P}(X \leq qn) \geq 1 - \exp(-nD(q||p)). \quad (5.2)$$

We will also need an application of this result to a slightly different setting, as seen in [36, Theorem 2.8]. We cannot ask for a lower tail bound, since Poisson binomial random variables can be constant.

Definition 5.9. A random variable X is *Poisson Binomial* with s trials if it can be written as a sum of s independent Bernoulli random variables $Ber(p_i)$ for some sequence of (possibly distinct) probabilities $0 \leq p_i \leq 1$.

Definition 5.10. Let $0 < q < 1$. The lower tail function $f_{l,q}$ is defined as follows.

$$f_{l,q}(x) = \begin{cases} 0 & x \leq q \\ D(q||x) & x \geq q \end{cases}$$

Lemma 5.11 ([36, Theorem 2.8]). *Let X be distributed according to a Poisson binomial distribution, with n trials and mean xn . Then $\mathbb{P}(X \leq qn) \leq \exp(-nf_{l,q}(x))$.*

We finish this section by proving some simple properties of the relative entropy, which we will need later.

Lemma 5.12 (Facts about the KL Divergence). *$D(q||p)$ satisfies the following 4 properties.*

1. $D(q||p) = D(1-q||1-p)$
2. $D(q||p) \geq 0$ with equality iff $q = p$
3. For fixed p , $D(q||p)$ is strictly increasing in $q \geq p$ and strictly decreasing in $q \leq p$.
4. For fixed q , $D(q||p)$ is strictly increasing in $p \geq q$ and strictly decreasing in $p \leq q$.

Further, the second order partial derivatives are as follows:

$$\frac{\partial^2}{\partial p^2}(D(q||p)) = \frac{p^2 - 2pq + q}{p^2(1-p)^2} \quad (5.3)$$

$$\frac{\partial^2}{\partial p \partial q}(D(q||p)) = \frac{-1}{p(1-p)} \quad (5.4)$$

$$\frac{\partial^2}{\partial q^2}(D(q||p)) = \frac{1}{q(1-q)}. \quad (5.5)$$

Proof. Property (1) holds since $D(q||p) = q \log(q/p) + (1-q) \log((1-q)/(1-p)) = D(1-q||1-p)$, and so we can move to the other properties. For these, we use that $D(q||p)$ is a smooth function of two variables $0 < p, q < 1$, and so we can calculate the partial derivatives.

$$\frac{\partial}{\partial p}(D(q||p)) = \frac{p-q}{p(1-p)} \quad (5.6)$$

$$\frac{\partial}{\partial q}(D(q||p)) = \log \frac{q(1-p)}{p(1-q)} \quad (5.7)$$

It is easily seen $q = p$ is only time either of these derivatives are zero, and so fixing either variable, that is the unique stationary point as the other varies. Further, $D(q||q) = 0$, and thus since the partial derivative in p is positive if $p > q$ and negative if $p < q$ properties (2) and (4) hold. Property (3) holds since the partial derivative in q has similar properties — note that $\frac{q}{1-q} = \frac{1}{1-q} - 1$ is increasing in q . The second derivatives are computed by differentiating Equations (5.6) and (5.7), completing our proof. $\square$

5.3 The lower bounds

We are now ready to begin the proof of our lower bounds. Recall we are trying to find a random graph which does not contain an ${}_rK_t$ minor — we will instead work with the weaker notion of compatible partitions. We will first prove a parametric lemma, finding (asymptotically tight) conditions on n, p, r, t under which with high probability $G(n, p)$ does not contain ${}_rK_t$ as a minor; this is the aim of the next subsection. We will then convert these conditions into a lower bound on $c({}_rK_t)$ in the following subsection, and prove some properties of this bound.

5.3.1 A parametric lower bound

We will be proving our condition by, for each partition of $[n]$ into t parts, finding a bound on the probability it is ${}_rK_t$ -compatible when edges are placed independently at random with probability p . Our first aim is to show that this probability is maximised when all of the parts have (almost) equal size.

Theorem 5.13. *The following inequality holds for all $0 < p < 1$, and integers $n < m$ and $r \geq 0$.*

$$\mathbb{P}(B(n+1, p) \geq r) \mathbb{P}(B(m, p) \geq r) \geq \mathbb{P}(B(n, p) \geq r) \mathbb{P}(B(m+1, p) \geq r) \quad (5.8)$$

Informally, given a fixed total number of independent equiprobable trials to be split over two events, for the greatest chance of both to succeed they should be split equally.

Proof. Let $X \sim B(n, p)$, $Y \sim B(m, p)$ and $Z \sim \text{Ber}(p)$ be independent random variables. The probability both $X + Z$ and Y are at least r is exactly:

$$\mathbb{P}(X \geq r) \mathbb{P}(Y \geq r) + p \mathbb{P}(Y \geq r) \mathbb{P}(X = r - 1).$$

We have an analogous result $Y + Z$ and X . This means to prove our theorem, it will suffice to prove

$$\mathbb{P}(Y \geq r) \mathbb{P}(X = r - 1) \geq \mathbb{P}(X \geq r) \mathbb{P}(Y = r - 1). \quad (5.9)$$

This is equivalent, by expanding out the probability mass function, to:

$$\begin{aligned} \binom{n}{r-1} p^{r-1} (1-p)^{n-r+1} \sum_{j \geq r} \binom{m}{j} p^j (1-p)^{m-j} \geq \\ \binom{m}{r-1} p^{r-1} (1-p)^{m-r+1} \sum_{j \geq r} \binom{n}{j} p^j (1-p)^{n-j}. \end{aligned}$$

We make a final transformation by clearing the common factor and letting $c = p/(1-p)$,

which can take any positive value. It thus suffices to prove, for any $c > 0$, that:

$$\sum_{j \geq r} \binom{n}{r-1} \binom{m}{j} c^j \geq \sum_{j \geq r} \binom{m}{r-1} \binom{n}{j} c^j. \quad (5.10)$$

We will prove this inequality directly. Each side is a power series in the variable c so it suffices to show the coefficient of c^j is larger on the left hand side of (5.10) than on the right-hand side for each $j \geq r$ (all coefficients for $j < r$ are zero on both sides). On the left-hand side, the c^j coefficient is $l_j = \frac{1}{j!(r-1)!} n(n-1)\dots(n-(r-2))m(m-1)\dots(m-(j-1))$, and $r_j = \frac{1}{j!(r-1)!} m(m-1)\dots m(m-(r-2))n(n-1)\dots(n-(j-1))$ on the right-hand side. Since $j \geq r$ and $m > n$, we have $r_j \leq l_j$ as desired. $\square$

Corollary 5.14. *Let $n, t, r \geq 0$ be integers, and $0 \leq p < 1$, and let $G \sim G(n, p)$ be a random graph. Let $(Y_i)_{i=1}^t$ be an equipartition of $[n] = V(G(n, p))$, and $X_1, \dots, X_t$ be some other partition. Then the probability that there are at least r edges from X_i to X_j for all $i < j$ is at most the probability there are at least r edges from Y_i to Y_j for all $i < j$.*

Proof of Corollary 5.14. Suppose that X_i is some partition of $V(G(n, p))$ which maximises the probability that ‘for all pairs $i < j$, there are at least r edges from X_i to X_j ’. Reordering if necessary, let $|X_i| = n_i$ with $n_1 \leq \dots \leq n_t$. We would like to show that the X_i are an equipartition of $V(G)$, i.e. $n_t - n_1 \leq 1$.

Suppose for contradiction $n_t - n_1 > 1$ — since these are integers, we have $n_t - 1 \geq n_1 + 1$. Replace X_1 and X_t with new parts X'_1, X'_t formed by moving a single vertex from X_t to X_1 . We will show the probability that $X'_1, X_2, \dots, X_{t-1}, X'_t$ has r edges between each pair is strictly larger than for $X_1, \dots, X_t$, which would contradict maximality. We note that $|X'_1||X'_t| > |X_1||X_t|$, and so the probability there are more than r edges between X'_1 and X'_t is strictly greater than for X_1 and X_t . We now analyse each of $X_2, \dots, X_{t-1}$ in turn.

The number of edges between X_1 and X_i is distributed as a binomial random variable with $|X_1||X_i|$ trials, and success probability p ; we have analogous distributions for the number of edges from X_i to each of X_t, X'_1, X'_t . The probability there are more than r edges from both X_1 and X_t to X_i is thus $\mathbb{P}(B(n_1 n_i, p) \geq r) \mathbb{P}(B(n_t n_i, p) \geq r)$, and the equivalent value for X'_1 and X'_t is $\mathbb{P}(B(n_1 n_i + n_i, p) \geq r) \mathbb{P}(B(n_t n_i - n_i, p) \geq r)$. Since $n_1 n_i + n_i \leq n_t n_i - n_i$, by n_i sequential applications of Theorem 5.13 this probability is at least as large. Multiplying all of our probabilities for these independent events gives the desired contradiction. $\square$

Having shown that taking a balanced partition maximises our desired probability, we can now obtain a sufficient condition for $G(n, p)$ to not contain ${}_r K_t$ as a model with high probability by taking a union bound over all possible partitions. This theorem completes the proof of our parametric lower bound.

Theorem 5.15. *Let $\epsilon > 0$, $0 < p < 1$, $0 < k$ and suppose $G \sim G(n, p)$. Then with high probability G has no ${}_rK_t$ -compatible-partition, where $q < p$ satisfies $kD(q||p)/q < 1 - \epsilon$, where the parameters r, t are defined by $t = \lceil n/\sqrt{r/q} \rceil$ and $r = k \log t$. In particular, for such a choice of parameters, $c({}_rK_t) \geq (1 + o_{r,t}(1)) \frac{p\sqrt{r}}{2\sqrt{q}}$.*

Proof of Theorem 5.15. There are at most t^n partitions of $V(G)$ into t parts. If we can show each of these is ${}_rK_t$ -compatible with probability at most $o(t^{-n})$ we will have proven our claim — if we expect $o(1)$ partitions to be compatible, with high probability no such partition exists.

By Corollary 5.14, it suffices to prove this for equipartitions. For some pair $i < j$, the probability there are at least r edges from X_i to X_j is at most $\mathbb{P}(B(\lceil \frac{n}{t} \rceil^2, p) \geq r)$. Let $M = \lceil \frac{n}{t} \rceil^2 = (1 + o(1))k \log t / q$. By Chernoff-Hoeffding (Theorem 5.8), since the events for distinct i, j are independent and $1 - x \leq \exp(-x)$ we have:

$$\mathbb{P}(\text{a fixed partition is } {}_rK_t\text{-compatible}) \leq \exp\left(-\binom{t}{2} \frac{1}{\sqrt{2M}} \exp(-MD(r/M||p))\right).$$

Using $\frac{r}{M} = (1 + o(1))q$, $M = (1 + o(1))\frac{k \log t}{q}$ and $\frac{kD(q||p)}{q} < 1 - \epsilon$, the result follows. $\square$

We will finish this subsection by showing this result is asymptotically tight. Rather than showing such graphs contain ${}_rK_t$ minors, we will instead show they contain compatible partitions — the connector and projector of Theorem 2.18 allow us to convert a compatible partition into a minor by reserving a small number of vertices.

Theorem 5.16. *Let $G \sim G(n, p)$ and $\epsilon > 0$, $k > 0$. Then with high probability G contains an ${}_rK_t$ -compatible partition, where $r = k \log t$, $q < p$ satisfies $kD(q||p)/q > 1 + \epsilon$ and $t \leq n/\sqrt{r/q}$ are such that $r, t \rightarrow \infty$.*

Proof. We neglect rounding for simplicity, and assume all quantities are integers where required. Assume ϵ is sufficiently small that all required approximations hold. Let $t' = (1 + \epsilon/4)t$, and pick some partition $V_1, \dots, V_{t'}$ of $V(G)$ into parts of size $n/t' = \frac{1}{1+\epsilon/4}\sqrt{k \log t/q}$. The probability some pair of subsets has fewer than r edges between them (and call such pairs bad) is, by Theorem 5.8, at most $\exp(-\frac{k \log t}{q(1+\epsilon/4)^2} D(q||p))$, thus the expected number of bad pairs is at most $t^2 \exp(-\frac{k \log t}{q(1+\epsilon/4)^2} D(q||p)) \leq t^{1-\epsilon/3}$. In particular all but at most $t^{-\epsilon/6}$ proportion of graphs have fewer than $t^{1-\epsilon/6}$ bad pairs by Markov's inequality. Deleting all parts incident to a bad pair finishes our proof. $\square$

5.3.2 Choosing optimal parameters for the lower bound

We are now ready to optimise our lower bound, and prove some further properties of the optimal bound. This completes the proof of our lower bound Theorem 1.23. Recall the functions described here are graphed in Figure 1.2.

Theorem 5.17. *There are unique smooth increasing functions $p(k) : (0, \infty) \rightarrow [0, 1]$, with $\lim_{k \rightarrow 0} p(k)^* = p_0 = 0.715\dots$ and $\lim_{k \rightarrow \infty} p(k) = 1$, and smooth increasing $q(k) : (0, \infty) \rightarrow [0, 1]$ with $\lim_{k \rightarrow 0} q(k) = 0$ and $\lim_{k \rightarrow \infty} q(k) = 1$, which maximise p^2/q subject to $kD(q||p) = q$, where $p_0 = 0.715\dots$ is the value of x maximising $x/\sqrt{\log(1/(1-x))}$ on $(0, 1)$.*

Proof of Theorem 1.23. Fix for now some value of k . We will use the method of Lagrange multipliers to convert our constrained optimisation problem into an unconstrained optimisation problem with the objective function. The maximum of our function lies either on the boundary, or at a stationary point of the objective function

$$f(p, q, \lambda) = p^2/q - \lambda(kD(q||p) - q). \quad (5.11)$$

Recalling the partial derivatives from Lemma 5.12, at a stationary point equations (5.12)-(5.14) hold.

$$kD(q||p) = q \quad (5.12)$$

$$\frac{2p}{q} - \lambda k \frac{p-q}{pq(1-p)} = 0 \quad (5.13)$$

$$\frac{-p^2}{q^2} + \frac{\lambda k}{q^2} \log \frac{1-q}{1-p} = 0 \quad (5.14)$$

Since we do not require the value of the Lagrange multiplier, we eliminate λ from the final two equations. This yields the following system of two equations in the two unknowns p, q .

$$kD(q||p) = q \quad (5.15)$$

$$2 \log \frac{1-q}{1-p} = \frac{p-q}{1-p} = \frac{1-q}{1-p} - 1 \quad (5.16)$$

Equation 5.16 now only depends on p, q via $s = \frac{1-q}{1-p}$, and reduces to $2 \log s = s - 1$. This is an exponential equation, which has the two roots $s = 1$ and $s = s_0 = 3.51286\dots$. The first root is impossible to attain in our system of equations, since $q < p$ implies $s > 1$. As such, all common solutions must have $q = 1 - s_0(1 - p)$.

Directly converting Equation 5.15 into a function of p and k alone yields an implicit equation which is hard to work with. It turns out to be easier to instead reduce Equation 5.15 to an explicit function $p = p_1(q, k)$, using $\frac{1-q}{1-p} = s_0$ to eliminate all but one occurrence of p . This yields the explicit function

$$p_1(q; k) = \exp(\log q + (\frac{1}{q} - 1) \log(s_0) - \frac{1}{k}).$$

We now know that stationary points occur at the intersection of this non-linear constraint

with the linear constraint $p = 1 - (1 - q)/s_0$.

Claim 5.17.1. *There is a unique intersection point between $p = 1 - (1 - q)/s_0$ and $p = p_1(q, k)$ for each $k > 0$.*

Proof. By considering the second derivative

$$\frac{d^2(\log p_1)}{dq^2} = \frac{-1}{q^2} + \frac{2}{q^3} \log s_0 = \frac{2 \log s_0 - q}{q^3} > 0,$$

we see that $\log p_1(q, k)$ is convex; and hence $p_1(q, k)$ is convex. As $q \rightarrow 0$, $p_1(q, k)$ remains positive, and so $p_1(q, k) > 1 - (1 - q)/s_0$. Conversely, as $q \rightarrow 1$, $p_1(q, k) \rightarrow \exp(-1/k) < 1 = \lim_{q \rightarrow 1} 1 - (1 - q)/s_0$. Thus by parity, since all functions are continuous there is a unique intersection point at some value $0 < q(k) < 1$; this corresponds to some function $p(k) = 1 - (1 - q(k))/s_0$. $\square$

We now allow k to vary and consider how this affects $p(k), q(k)$. The linear constraint is fixed independently of k , and $p_1(q, k)$ is strictly increasing in k for every q . Since we are intersecting with a line of positive slope, as k increases so must the value of q at the intersection point, and so $p(k)$ is strictly increasing. Since we are intersecting smooth functions (and these smooth functions always intersect exactly once), the function p must also be smooth. Since $p(k)$ is bounded between p_0 below, and 1 above there must exist a limit in both $k \rightarrow 0$ and $k \rightarrow \infty$.

We start with $k \rightarrow \infty$. For this, we return to our original Equations 5.15 and 5.16. We require $D(q||p) = q/k \rightarrow 0$. We aim to show $p(k) \rightarrow 1$, and assume for contradiction this does not hold. We have an increasing function bounded above, and let $p' < 1$ be its limit. For each ϵ , there is a δ such that if $|q - p'| > \epsilon$ then $|D(q||p')| > \delta$. By continuity again, there is a δ' such that $|D(q||p' - \delta')| > \delta/2$. But this implies that in the limit $k \rightarrow \infty$ q must exceed all values $p' - \epsilon$, and hence be asymptotic to p' . But then $(1 - q)/(1 - p') \rightarrow 1$, which does not satisfy the linear constraint, contradicting our assumption $p(k) \not\rightarrow 1$.

We now consider the case $k \rightarrow 0$. This is simpler: due to the linear constraint, p is always bounded below by $\bar{p} = 1 - 1/s_0$. We need that $D(q||p)/q \rightarrow \infty$ for some $p > \bar{p}$ and q ; since $D(q||p)$ for $q < p$ is increasing in p , it will be sufficient to show this for some fixed value, say $p(1) < 1$. In order for $D(q||p) \rightarrow \infty$ for p bounded away from 0 and 1, we need $(1/q - 1) \log s_0 \rightarrow \infty$. This can only happen if $q \rightarrow 0$ and so $p \rightarrow (1 - 1/s_0)$. It remains only to show that $\bar{p}$ is as claimed the value p maximising $x/\sqrt{\log(1/(1 - x))}$, giving a relation between s_0 and our usual constant α .

To see this, p_0 also maximises the function $x^2/\log(1/(1 - x))$ which is continuous and differentiable, and so the derivative of this function at p_0 is zero. This implies $y = 1/(1 - p_0)$ satisfies $2 \log y = y - 1$ which is the equation defining s_0 , completing our proof. $\square$

Recall Theorem 5.15 shows that $c(rK_t) \geq \frac{p\sqrt{r}}{2q}t = \frac{p\sqrt{k}}{q}t\sqrt{\log t}$ for any parameters

p, q satisfying certain inequalities. Theorem 5.17 told us how to choose these parameters appropriately. The following theorem describes how the scaling factor $p\sqrt{k/q}$ behaves for these optimal parameters, and completes the proof of Theorem 1.23.

Theorem 5.18. *The optimal prefactor $p(k)\sqrt{k/q(k)}$ is a strictly increasing function of $k > 0$, with $\lim_{k \rightarrow 0^+} p\sqrt{k/q} = 2\alpha$ and $\lim_{k \rightarrow \infty} (p\sqrt{k/q})/\sqrt{k} = 1$.*

Proof. As $k \rightarrow \infty$, both $p, q \rightarrow 1$ and so the prefactor is asymptotically equivalent to $\sqrt{k}$; we now consider the other properties. We note the prefactor is non-negative, so in order to show it is increasing it suffices to prove this for the square $p^2 k/q$; since $p(k)$ is increasing we prove this as a function of $p > p_0$ alone.

We rewrite this square of the prefactor as $f(p) = p^2/D(1 - s_0(1 - p)||p)$, recalling $kD(q||p) = q$, and compute its derivative as follows using results of Lemma 5.12. The final line follows from judicious use of $q = 1 - s_0(1 - p)$ and using $\log s_0 = (s_0 - 1)/2$.

$$\begin{aligned} f'(p) &= \frac{p}{D(q||p)^2} \left[-2D(q||p) + p \frac{d}{dp} (D(1 - s_0(1 - p)||p)) \right] \\ &= \frac{p}{D(q||p)^2} \left[2D(q||p) - p \left(\frac{p - q}{p(1 - p)} - s_0 \log \frac{q(1 - p)}{p(1 - q)} \right) \right] \\ &= \frac{p}{D(q||p)^2} \left[(2 - 2s_0 + s_0 p) \log \left(\frac{q}{p} \right) + (s_0 - 1)(-1 + s_0(1 - p/2)) \right] \end{aligned}$$

We note that $p \geq 0$ and $D(q||p) \geq 0$, so to prove the derivative is non-negative it suffices to show $(2 - 2s_0 + s_0 p) \log(q/p) > 0$ and $(s_0 - 1)(-1 + s_0(1 - p/2)) > 0$. The second of these is trivial since $p \leq 1$. For the first, we note $q \leq p$ and so $\log(q/p) < 0$; since $2 - 2s_0 + s_0(1 - 1/s_0) < 0$ this means $f'(p) > 0$ for each $1 - 1/s_0 < p < 1$ as desired. Thus the prefactor is an increasing function of p , hence of k .

For the limit in $k \rightarrow 0$, we consider the following equation satisfied by the parameters arising from expansion of $kD(q||p) = 1$:

$$k \log(q) + (k/q) \log s_0 = 1 + k \log(s_0 p). \quad (5.17)$$

The right hand side of Equation 5.17 is $1 + o_{1/k}(1)$, and so this must also hold for the left hand side. Rewriting in terms of q/k we have $k \log k + k \log(q/k) + \log(s_0)(k/q)$; the first term tending to zero in $k \rightarrow 0$. The middle term is negligible unless $q = o(k^2)$, and the final term is convergent to 1 only when $q = (1 + o(1))k \log s_0$. We only need exclude the $q = o(k)$ case, which cannot hold since then the last term is positive, and grows asymptotically faster than the middle term. Thus in the $k \rightarrow 0$ limit the prefactor must converge to $\frac{1 - 1/s_0}{\sqrt{\log s_0}} = 2\alpha$ as desired, completing our proof. $\square$

5.4 The upper bound

We now move onto the upper bounds. The aim of this section will be to state dense sub-theorems for each of the ‘small r ’ ($r = o(\log t)$) and ‘large r ’ ($r = \omega(\log t)$) cases as well as our (common) sparse sub-theorem. We will then prove both of our upper bounds, Theorems 1.24 and 1.25, assuming these sub-theorems. The arguments of this section are similar to those of other settings (and we follow the strategy outlined in Section 2.2); the new ideas appear later in building compatible partitions in the dense case, as well as in the sparse case where we need a new argument to find our target minor inside a dense bipartite graph. We begin with the two dense sub-theorems.

Theorem 5.19. *Let $\epsilon > 20\eta > 0$. Then there is a constant $T = T_{5.19}(\epsilon, \eta)$ such that for all $t > T$, every graph G with connectivity at least $8\eta|G|$ and density at least $p + \epsilon$ with $n \geq t\sqrt{\log_{(1/q)} t}$ vertices is ${}_rK_t$ -minor-prevalent for some $r \geq \delta \log t$, with $\delta = \delta_{5.27}(\epsilon/4)$ the constant of Corollary 5.27 and $q = 1 - p$.*

Theorem 5.20. *Let $\epsilon > 20\eta > 0$. Then there is a constant $T = T_{5.20}(\epsilon, \eta)$ such that for all $t > T$, and $r > T \log t$, every graph G with connectivity at least $8\eta|G|$, density at least $p + \epsilon$ and $n \geq t\sqrt{r/p}$ vertices is ${}_rK_t$ -minor-prevalent.*

Here is the sparse sub-theorem.

Theorem 5.21. *Let $\epsilon > 0$. Then there is a constant $D = D_{5.21}$ such that for any $t > D$, $r > \epsilon \log t$ and $m > \epsilon t\sqrt{r}$, every graph G with the below four properties contains ${}_rK_t$ as a minor.*

1. $e(G) \leq m|G|$
2. Every edge of G is in at least $m - 1$ triangles
3. $|G| \geq Dm$
4. $\kappa(G) \geq Dt$

We note from here on, q will always denote $1 - p$.

Proof of Theorem 1.24. Since $e(G)/|G| \geq (\alpha + \epsilon)t\sqrt{\log t}$, taking $m = (\alpha + \epsilon)t\sqrt{\log t}$ and $k = \epsilon m/2$ it holds $G \in \mathcal{E}_{m,k}$. Taking t large, it holds $k > 1$. Replacing G by a minor, we may instead assume that G is minor-minimal in $\mathcal{E}_{m,k}$. We can thus assume G has the properties of Lemma 2.4; these exactly correspond to conditions 1,2 and 4 of Theorem 5.21 holding. So there is a constant $D_1 = D_{5.21}(\frac{1}{4})$ such that if in addition $|G| \geq D_1 m$ and $t > D_1$, then G contains ${}_{\log t}K_t$ as a minor, and so taking any $C_{1.24} > 1$ this implies our result in this case. We therefore assume $|G| \leq D_1 m$.

In this case, taking $\eta = \epsilon/D_1$, G is ηm connected. Recalling also Inequality 1.3, we have $|G| \geq 2(1 + \epsilon/8)\frac{\alpha}{p}t\sqrt{\log t} \geq (1 + \epsilon/8)t\sqrt{\log_{1/(1-p)} t} \geq t\sqrt{\log_{1/(1-p+\epsilon/8)} t}$. As such we can apply Theorem 5.19 to obtain constants $\delta = \delta_{5.25}(\epsilon/8)$, $D_2 = T_{5.19}(\epsilon, \eta)$ such that provided $C > \max(1/\delta, D_1, D_2, 1)$ then G contains ${}_rK_t$ as a minor, completing our proof. $\square$

Proof of Theorem 1.25. Since $e(G)/|G| \geq (1 + \epsilon)t\sqrt{r}$, for $m = (1 + \epsilon)t\sqrt{r}$ and $k = \epsilon m/2$ we have $G \in \mathcal{E}_{m,k}$. Replacing G by a minor, we can take our graph to instead be a minor-minimal element of $\mathcal{E}_{m,k}$, and hence have the properties of Lemma 2.4.

Applying Theorem 5.21, we see there is a constant $D_1 = D_{5.21}(1/4)$ such that if in addition $|G| \geq D_1 m$ and $t > D_1$ then G contains ${}_r K_t$ as a minor. We therefore assume $|G| \leq D_1 m$, and in particular for $\eta = \epsilon/2D_1$, G is $\eta|G|$ connected. Suppose that $p > \eta$ is the density of G . We note that $|G| \geq (1 + \epsilon/3)t\sqrt{r}/p \geq t\sqrt{\frac{r}{p - \epsilon\eta/2}}$. In particular, we can apply Theorem 5.20 with ϵ replaced with $\epsilon\eta/2$, and η replaced with $\epsilon\eta/40$ to deduce that G contains ${}_r K_t$ as a minor provided that $r > T \log t$ and $t > T$ for some constant $T = T_{5.20}(\epsilon\eta/2, \epsilon\eta/40)$. Taking $C > \max(T, D_1)$ concludes the proof of our theorem. $\square$

5.5 The dense case

Our next aim is to prove our dense sub-theorems, Theorems 5.19 and 5.20. As usual we will start by building almost-compatible partitions under slightly weaker conditions, and later use the connector and projector to extend to rooted minors. We require different arguments to build compatible partitions in our two cases.

Definition 5.22. Let H be a multigraph. A collection of disjoint non-empty subsets $(V_h)_{h \in V(H)}$ of $V(G)$ is called an M -almost- H -compatible partition if

$$\sum_{hh' \in \binom{V(H)}{2}} \max(0, e_G(V_h, V_{h'}) - m_H(hh')) \leq M.$$

Equivalently, with the addition of at most M additional edges to G (possibly forming some multiple edges) the collection can be made an H -compatible partition.

5.5.1 The small r case

The aim of this subsection is to build our compatible partition in the ‘small r ’ case. We will as usual start with a parametric version, and then choose the parameters appropriately. This is stated in terms of the tail function $f_{l,q}$ of Definition 5.10.

Theorem 5.23. Let $\omega, \eta > 0$, and s, t integers. Let G be a graph with $n \geq st$ vertices, and density at least p . Then G contains an M -almost- ${}_r K_{t'}$ -compatible partition, where $t' \geq (1 - 4/\omega\eta)t$, $r = \eta^2 s^2$ and $M = 4rt^2 \exp(\eta s \log s - s^2(1 - \eta)^2 f_{l,\eta}(p - 3\eta))$.

We will require the following simple result for the proof.

Lemma 5.24. Let s, t be integers, $A \subset [s-1]$, and $1 \geq x_1 \geq \dots \geq x_{st} \geq 0$ a non-increasing sequence with $\frac{1}{st} \sum_i x_i = x$. Then $\frac{1}{|A|} \sum_{i \in A} x_{it} \geq x - \frac{s-|A|}{s}$ holds.

Proof. We note that $x_{it} \geq \frac{1}{t} \sum_{j=1}^t x_{it+j}$, and so the following holds:

$$\sum_{i \in A} x_{it} \geq \sum_{i \in A} \frac{1}{t} \sum_{j=1}^t x_{it+j} \geq \frac{1}{t} \sum_{i=0}^{s-1} \sum_{j=1}^t x_{it+j} - (s - |A|).$$

We now divide through by $|A| \leq s$ to obtain the desired result; note this is valid since the left-hand side is non-negative. $\square$

Proof of Theorem 5.23. By passing to a maximal density subgraph on st vertices we can assume $n = st$; at the end, we will distribute any extra vertices arbitrarily throughout the parts. We note our result is only strengthened by replacing p with the density of G , hence assume that G has density exactly p .

Order the vertices of G in non-increasing degree order, so that $d(v_1) \geq \dots \geq d(v_n)$, and define p_i by $d(v_i) = p_i(n-1)$. The p_i form a non-increasing sequence, and $\frac{1}{n} \sum_i p_i = p$. We partition these vertices into s blocks $(B_i)_{i=1}^s$, where $B_j = \{v_{(j-1)t+1}, \dots, v_{jt}\}$, and note that for $v \in B_j$ we have $p_{(j-1)t+1}(n-1) \geq d(v) \geq p_{jt}(n-1)$.

We begin our construction by picking t disjoint sets (called *parts*), each consisting of exactly one vertex from each block. Some of these parts will become the parts in our almost-compatible partition at the end. We pick them randomly and one at a time as follows. Suppose we have already picked k parts, and let X be their union. Form a new part by picking one vertex uniformly at random from each of $B_1 \setminus X, \dots, B_s \setminus X$. By the choice of parts having one vertex in each block, we can repeat this to build t parts as claimed. As in Lemma 2.11, the order we pick the parts does not affect the distribution.

We say that a vertex v is *bad* for a part X if $e(v, X) \leq \eta s$, and otherwise call v *good* for X . If W contains at least ηs vertices which are good for X , then $e(W, X) \geq \eta^2 s^2$. For some fixed v , what is the probability that v is bad for a (randomly chosen) part X ? This event occurs when X contains fewer than ηs of the $d(v)$ neighbours of v ; equivalently if we pick a neighbour of v from fewer than ηs blocks. Within B_i , the probability that $X \cap B_i$ is a neighbour of v is $|N(v) \cap B_i|/t$, and in particular the number of neighbours of v in X is distributed according to a Poisson Binomial distribution, with mean $d(v)/t$ and s trials. The Chernoff bound (Lemma 5.11) yields the following for $v \in B_j$, where $p'_i = p_i(1 - 1/n)$.

$$\mathbb{P}(v \text{ is bad for } X) \leq \exp(-sf_{l,\eta}(d(v)/n)) \leq \exp(-sf_{l,\eta}(p_{jt}(1 - 1/n))).$$

Thus the expected number of $v \in B_j$ which are bad for X is at most $t \exp(-sf_{l,\eta}(p'_{jt}))$. We say that X *rejects* block B_j if more than $\omega t \exp(-sf_{l,\eta}(p'_{jt}))$ elements $v \in B_j$ are bad for X , and otherwise that X *accepts* B_j ; the probability X rejects B_j at most $\frac{1}{\omega}$ by Markov's inequality. Let $M(X) = \{1 \leq j \leq s-1 : X \text{ accepts } B_j\}$. We say that X is *bad* if $|M(X)| \leq (1 - \eta)(s-1)$, else say X is *good*. The probability a part X is good is at least $(1 - \frac{1}{\omega\eta})$.

Suppose we fix some choice of good part W . Conditional on our choice of W , what is the probability some part X has fewer than $\eta^2 s^2$ edges to X ? We can bound this above by the probability that at most ηs choices of vertices in $(X \cap B_i)_{i \in M(W)}$ are good for W . Suppose that we choose fewer than ηs such good vertices. Then there is some set $A \subset M(W)$ with $|A| \geq |M(W)| - \eta s$ such that we pick a bad vertex from each $B_i : i \in A$. There are at most $(1+s)^{\eta s}$ choices of set A ; we now proceed to bound the probability of this occurring for some fixed A . By the definition of $M(W)$ the following holds.

$$\begin{aligned} \mathbb{P}(\forall i \in A, \text{ we pick a bad vertex for } X \text{ from } B_i) &\leq \prod_{i \in A} \left(\frac{\omega t}{t-1} \right) \exp(-s f_{l,\eta}(p'_{it})) \\ &\leq (2\omega)^s \exp \left(-s|A| \frac{1}{|A|} \sum_{i \in A} f_{l,\eta}(p'_{it}) \right) \end{aligned}$$

Using the convexity of the function $f_{l,\eta}$ (this follows from Equation 5.3), that $\frac{1}{n} \sum_i p'_i = p(1 - 1/n)$, Lemma 5.24 implies

$$\mathbb{P}(\forall i \in A, \text{ we pick a bad vertex for } X \text{ from } B_i) \leq (2\omega)^s \exp(-s^2(1 - 3\eta)f_{l,\eta}(p - 3\eta)).$$

Since there are at most $(2s)^s$ choices of A , except with probability at most $\bar{p}$ there are at least $r = \eta^2 s^2$ edges between W and X , where $\bar{p} = \exp(\eta s \log(2s) - s^2(1 - 3\eta)f_{l,\eta}(p - 3\eta))$. We now modify our partition to have the desired properties. By Markov's inequality, the probability there are more than $4t/\omega\eta$ bad parts is at most $\frac{1}{4}$. Further, the probability there are more than $4\bar{p}t^2$ pairs of parts, at least one of which is good, with fewer than $\eta^2 s^2$ edges between them is also at most $\frac{1}{4}$ — in particular, there is some choice of parts for which both of these bounds hold and we pick one. We now delete all bad parts, and finish by redistributing any unused vertices between the parts however we choose. In such a partition, there are $t' \geq (1 - 4/\omega\eta)t$ remaining parts, and at most $4\bar{p}t^2 r$ edges need to be added to become ${}_r K_{t'}$ -compatible, completing our proof. $\square$

Corollary 5.25. *Let $\frac{1}{8} > \epsilon > 0$. Then there are constants $T = T_{5.25}(\epsilon)$ and $\delta = \delta_{5.25}(\epsilon)$ such that for all $\epsilon \leq p \leq 1 - \epsilon$ and $t > T(\epsilon)$, every graph G with $n \geq t\sqrt{\log_{(1/1-p)} t}$ vertices has a $t^{(1-\epsilon/24)}$ -almost- ${}_r K_t$ partition, for some $r \geq \delta \log t$.*

Proof. Throughout we will assume T is sufficiently large all inequalities used hold. Let $t_1 = \lfloor (1 + \epsilon/4)t \rfloor$ and $s = \lfloor (1 - \epsilon/4)\sqrt{\log_{(1/1-p)} t} \rfloor$ — we note $|G| \geq st_1$. We apply Theorem 5.23 with t replaced with t_1 and density $p + \epsilon$, where ω and η are functions of ϵ alone yet to be determined satisfying $t_1(1 - 4/\omega\eta) \geq t$. Picking an arbitrary subset of t parts, this yields an M -almost- ${}_r K_t$ -compatible partition where $r \geq \eta^2(1 - \epsilon/3) \log t / \log(1/\epsilon)$ (and we take $\delta = \eta^2(1 - \epsilon/3) / \log(1/\epsilon)$). It remains to bound M ; we choose to do this via a bound on $\rho = (M/t_1^2)$. We have

$$\log \rho \leq \log(4r) + \eta s^2 - s^2(1 - \eta)^2 f_{l,\eta}(p + \epsilon - 3\eta).$$

From the definition, $\lim_{\eta \rightarrow 0} D(\eta||p) = \log(1/1-p)$. We choose T sufficiently large that $\log(4r) \leq \epsilon/24 \log_{1/(1-p)} t$, and choose η sufficiently small that $(1 - \epsilon/3)^2 \log_{1/(1-p)} t (1 - \eta)^2 f_{l,\eta}(p + \epsilon - 3\eta) \geq (1 + \epsilon/6) \log t$ for all $\epsilon \leq p \leq 1 - \epsilon$ — this is a continuous function on compact support so we can make such a choice. We also pick $\eta < \epsilon/24$. If we do this, we have $\log \rho \leq -(1 + \epsilon/12) \log t$. Taking T large we therefore have $M \leq t^{(1-\epsilon/24)}$ as required. $\square$

5.5.2 The large r case

The aim of this section is to build a compatible partition for the ‘large r ’ case. Unfortunately, a direct application of Theorem 5.23 from the previous subsection provides too few edges between parts to also handle the $r = \omega(\log t)$ case exactly.

We therefore need a new theorem as below, which introduces a more strict notion of being bad for a part and is thus able to force sufficient edges. This increases the value of M (and so we could not use this theorem in the small r case), however it is still small enough for our aim of proving Corollary 5.27, our result giving an almost- H -compatible partition in the large r case.

Theorem 5.26. *Let s, t be integers, $\omega > 0$, $0 < \eta < 1$ with $1/s < \eta$ and suppose G is a graph with $n \geq st$ vertices and density $0 < p < 1$. Then G has a M -almost- rK'_t -compatible partition, where $t' \geq (1 - 4/\omega\eta)t$, $r = (p - 7\eta)s^2$ and*

$$M = 4r \left(2^s (2\omega)^{\eta s} \exp(-\eta^3(1 - \eta)s(s - 1)) \right) t^2.$$

Proof. As in Theorem 5.23, assume $n = st$ and G has density p . Order the vertices of G in non-increasing degree order, so that $d(v_1) \geq \dots \geq d(v_{st})$, and define p_i by $d(v_i) = p_i(n - 1)$. If $v = v_i$ we let p_v denote p_i . In exactly the same fashion as in Theorem 5.23 we construct s blocks $B_i = \{v_{(i-1)t+1}, \dots, v_{it}\}$, and build t random parts, each consisting of one random vertex from each block.

We say that a vertex v is *problematic* for a part X if $e(v, X) \leq s(p_v - \eta)$. By Lemma 5.24, if all but at most $\eta(s - 1)$ vertices in a part W are not problematic for X , there will be at least $(1 - \eta)(p - \eta)s^2$ edges between W and X . What is the probability a vertex v is problematic for a (random) part X ? The number of neighbours of v in X has a Poisson Binomial distribution, with s trials and mean $p_v(1 - \frac{1}{n})s$. Thus using the Chernoff bound in the form Lemma 2.14, the probability v is problematic for X is at most $\exp(-2s(\eta - 1/n)^2) \leq \exp(-s\eta^2)$. We remark our above proof strictly required $p_v > \eta$, however the final result still holds if $p_v \leq \eta$ since there are at least zero neighbours in X with probability 1.

We say that X *rejects* a block B_j if more than $\omega t \exp(-s\eta^2)$ elements $v \in B_j$ are problematic for X , and otherwise *accepts* B_j . By Markov’s inequality, the probability X rejects B_j is at most $\frac{1}{\omega}$. We say a part X is *bad* if it rejects more than $\eta(s - 1)$ of blocks

$B_1, \dots, B_{s-1}$; otherwise say X is *good*. The probability X is bad is at most $\frac{1}{\omega\eta}$ as before.

Let W be a fixed good part, and let $M(W)$ be the set of indices $1 \leq i \leq s-1$ for which W accepts B_i ; by definition of a good part $|M(W)| \geq (1-\eta)(s-1)$. Conditional on this choice, we seek to bound the probability there are fewer than $(p-7\eta)s^2$ edges between W and some random part X . We will consider the indices for which the element of $X \cap B_i$ is problematic for W . By construction, these events are independent over i , and so the total number of problematic vertices for W in X can be modelled by a Poisson Binomial random variable.

If $i \in M(W)$, the probability the element of $X \cap B_i$ is problematic for W is at most $\frac{\omega t}{t-1} \exp(-\eta^2 s) \leq 2\omega \exp(-\eta^2 s)$ since W accepts B_i . We can easily bound the probability of picking more than $\eta|M(W)|$ such vertices by $2^{|M(W)|} (2\omega)^{\eta|M(W)|} \exp(-\eta^3 s |M(W)|)$ since this bounds the expected number of $\eta|M(W)|$ sets chosen. As such, except with this probability we have at least $(1-\eta)^2 s$ vertices which are not problematic. However, we need a bound on the number of edges from these to X .

Let $A = \{i \leq s-1 : i \in M(W) \text{ and } X \cap B_i \text{ is problematic for } W\}$. We note $|A| \geq (1-2\eta)(s-1)$. This corresponds to there being at least $\sum_{i \in A} p_{it}s - \eta s |A|$ edges between W and X . Applying Lemma 5.24 to the sequence p_i (where $\frac{1}{st} \sum_i p_i = p$) we deduce $\frac{1}{|A|} \sum_{i \in A} p_{it}s \geq (p-2\eta-1/s)s$. Since also $|A| \geq (1-2\eta)s$, combining these results we obtain $e(X, W) \geq (p-7\eta)s^2$ for any X consisting of at least $(1-2\eta)s$ non-problematic vertices.

We now modify this partition to have the desired property. By Markov's inequality, the probability there are more than $4t/\omega\eta$ bad sets is at most $\frac{1}{4}$. Further, the probability that there are more than $4(2^s(2\omega)^{\eta s} \exp(-\eta^3(1-\eta)s(s-1))t^2)$ pairs of sets, at least one of which is good, with fewer than $(p-7\eta)s^2$ edges between them is at most $\frac{1}{4}$. Deleting bad parts and redistributing unused vertices, we are left with $t' \geq (1-4/\omega\eta)t$ parts, and all but at most $4(2^s(2\omega)^{\eta s} \exp(-\eta^3(1-\eta)s(s-1))t^2)$ pairs of these parts have at least r edges between them as desired. $\square$

Corollary 5.27. *Let $\epsilon > 0$. Then there is a constant $T = T_{5.27}(\epsilon)$ such that for all $t > T$, $\epsilon \leq p \leq 1-\epsilon$ and $r > T \log t$ every graph G with $n \geq t\sqrt{r/p}$ vertices and density at least $p+\epsilon$ contains a t -almost- r - K_t -compatible partition.*

Proof. We will apply Theorem 5.26, where p is replaced with $p+\epsilon$, $t = \lfloor (1+\epsilon/4)t \rfloor$, and $s = \lfloor (1-\epsilon/4)\sqrt{r/p} \rfloor$. We will let $\eta = \epsilon/32$ and $\omega = 257/\epsilon$, so $\frac{4}{\omega\eta} < \epsilon/8$ and thus $t' \geq t$ by taking T large. Having done so, we obtain an M -almost- r' - $K_{t'}$ -compatible partition, and by deleting parts thus an M -almost- r' - K_t -compatible partition, where $r' = (p+\epsilon-7\eta)s^2 > r$. It remains to verify the bound on M . Let $\rho = M/t'^2$. Then it holds

$$\log \rho \leq \log r + \log 8 + s \log 2 + \eta s \log(2\omega) - \eta^3(1-\eta)s(s-1).$$

We can ensure s is as large as we wish by taking T large, and so we can neglect all terms except $\log r$ and $-\eta^2(1-\eta)s(s-1)$. We can bound $-\eta^2(1-\eta)s(s-1) \leq \eta^2(1-\eta)(1-\epsilon/3)^2 \frac{r}{1-\epsilon} \leq -(T\eta^2/2) \log t$. In particular, taking T sufficiently large, we can bound $\log \rho \leq -2 \log t$ and so we have the claimed t -almost- rK_t -compatible partition. $\square$

5.5.3 Converting compatible partitions into minors

Having built our compatible partitions, it remains only to extend these to (rooted) minors. Recall that we will use the connector and projector to do this; our lemmas of Section 2.5 were sufficiently general we can use them as stated in the multigraph setting. This subsection will therefore complete the proof of our dense sub-theorems, Theorems 5.19 and 5.20.

Theorem 5.28. *Let $\epsilon, \eta > 0$. Then for the constant $D = D_{2.18}$ the following also holds.*

Let H be a multigraph with $|H| \leq |G|/D$, and suppose that n and p are such that every graph with at least $n(1 - 5\eta)$ vertices and density at least $p + \epsilon - 10\eta$ contains a $|H|/\eta$ -almost- H -compatible partition.

Then every graph G with at least n vertices and density at least $p + \epsilon$ which is $8\eta|G|$ connected is H -minor-prevalent.

Proof. Let G be a graph as in the theorem; since $D_{2.18}(\eta) > 1$ we have $|G| \geq |H|$. Let $R = (r_h)_{h \in V(H)}$ be a set of roots. Applying Theorem 2.18, we deduce that there exist sets C, P each of size at most $2\eta|G|$ with certain properties. In particular, $|G - C - P - R| \geq |G|(1 - 5\eta) > n(1 - 5\eta)$. Further, since we deleted at most 5η proportion of vertices it must also be the case $G - C - P - R$ has density at least $p + \epsilon - 10\eta$. So by assumption, $G - C - P - R$ has a $|R|/\eta$ -almost- H -compatible partition, and call one such partition $(V_{r_h})_{h \in V(H)}$. Let F be the multiset of edges required to be added to form an H -compatible partition. Then applying Theorem 2.18 with our current η , R , (V_{r_h}) and F , we obtain new sets $(U_h)_{h \in H}$ such that $r_h \in U_h$ and with additional properties. These properties imply that the (U_h) are exactly an H minor rooted at R , implying the theorem. $\square$

Proof of Theorem 5.19. By Corollary 5.25 there is a constant $T_1 = T_{5.25}(\epsilon/4)$ such that for all $t > T_1$, every graph G with $n \geq t\sqrt{\log_{1/(q-\epsilon/4)} t}$ vertices and density at least $p + \epsilon/2$ contains a t -almost- rK_t -compatible partition, where $r \geq \delta_{5.25}(\epsilon/4) \log t$.

Recalling Lemma 2.7, we have $t\sqrt{\log_{1/(q-5\eta)} t} \leq (1 - 5\eta)t\sqrt{\log_{1/q} t}$. Applying Theorem 5.28, there is a constant $D_{2.18}$ such that if also $|G| > D|H|$ then G is H -minor-prevalent. Taking $T > \max(T_1, (3\eta)^{-D_{2.18}(\eta m)})$ implies our theorem. $\square$

Proof of Theorem 5.20. By Corollary 5.27 there is a constant $T_1 = T_{5.27}(\epsilon/4)$ such that if $t > T_1$ and $R > T_1 \log t$, every graph G with at least $t\sqrt{r/(p + 5\eta)}$ vertices and density at least $p + \epsilon/2$ contains a t -almost- rK_t -compatible partition. Again using Inequality 2.1 we

have $t\sqrt{r/(p+5\eta)} \leq (1-5\eta)t\sqrt{r}$, and also $p+\epsilon-10\eta > p+\epsilon/2$. Thus if also $|G| \geq D_{2.18}|H|$, which can be assured by taking $T > \max(T_1, D_{2.18}^2)$, we have that G is ${}_rK_t$ -minor-prevalent by direct application of Theorem 5.28. $\square$

5.6 The sparse case

We now move onto the sparse case, which is the final step required for our upper bounds and this chapter. To prove Theorem 3.13, we will require the following strengthening of Lemma 2.20. We begin this section with a proof of this strengthening, followed by applications to the small r and large r cases. These will then be combined to prove Theorem 5.21.

Theorem 5.29. *Let $\epsilon, C, M > 0$. Then there is a constant $D = D_{5.29}(\epsilon, C, M)$ such that for all $d > D$, every bipartite graph $G = (A, B)$ with $|B| < Cd$, $|A| > Dd$, with the property that every vertex in A has at least ϵd neighbours in B , contains ${}_MH$ as a minor for some simple graph H on at least ϵd vertices with minimum degree at least $(1-\epsilon)|H|$.*

Proof. We will be building up a sequence (G_i) of minors of G , together with auxiliary graphs (G'_i) . Set $G_0 = G$, and G'_0 the empty graph on vertex set B . The graph G'_i is the simple graph with vertex set $V(G_i)$, and two vertices adjacent in G'_i exactly when they are adjacent in G_i with multiplicity M . The graph G_i will be formed from G_{i-1} by contracting a single edge between A and B , and deleting edges if necessary so all edges have multiplicity at most M . We view this as contracting a vertex of A onto a vertex of B and label the contracted vertex accordingly; thus after each contraction $|A|$ decreases by 1, and $|B|$ has not changed.

Given G_i , to form G_{i+1} pick some vertex $a \in A$ (which we have not yet contracted), and pick a vertex $b \in B$ of minimum degree in $G'_i[N_G(a)]$. Form G_{i+1} from G_i by contracting the edge ab , and then reducing the multiplicity of all edges to be at most M . Repeat this process until A is empty.

Suppose first that $d_{G'_i}(b) \leq (1-\epsilon)d_G(a)$ for all i . Then $G_{i+1}[B]$ has at least $\epsilon d_G(a) - 1 \geq \epsilon^2 d/2$ more edges than $G_i[B]$. However, each $G_i[B]$ has edge multiplicities at most M , and thus in particular cannot have more than $MC^2 d^2/2$ edges! So if we perform more than Dd steps, for $D = MC^2/\epsilon^2$, we cannot always have added at least $\epsilon^2 d/2$ additional edges. So there exists some i for which $\delta(G'_i[N_G(a)]) \geq (1-\epsilon)d_G(a)$. Since ${}_MG_i$ is a subgraph of G'_i , G contains the desired minor. $\square$

Corollary 5.30. *Let $\frac{1}{3} > \epsilon > 0$. Then there is a constant $C = C_{5.30}(\epsilon)$ such that for $t > C$ and $d = t\sqrt{\log t}$, every bipartite graph $G = (A, B)$ with $|B| \leq d/\epsilon$, $|A| \geq Cd$ and every vertex in A having at least ϵd neighbours in A contains ${}_rK_t$ as a minor, where $r = (1/\epsilon) \log t$.*

Proof. Let $\delta = \delta_{5.25}(\epsilon/4)$. Applying Theorem 5.29 with the current value of ϵ , $d = t\sqrt{\log t}$, $M = 1/(\delta\epsilon)$ and $C_1 = 1/\epsilon$ there is a constant $D = D_{5.29}(\epsilon, C_1, M)$ such that if $C > D$ then G contains as a minor ${}_MH$ for some graph H with at least $\epsilon t\sqrt{\log t}$ vertices and density at least $1 - \epsilon$.

By Theorem 5.19, applied with $\eta = \epsilon/20$ and the current value of ϵ , there is a constant $T = T_{5.19}(\epsilon, \epsilon/20)$ such that if also $t > C > T$ then every graph with connectivity at least $\epsilon/2|G|$, density at least $1 - \epsilon$ and at least $t\sqrt{\log_{1/2\epsilon} t}$ vertices is ${}_{\delta \log t}K_t$ -minor-prevalent.

Since H has minimum degree at least $(1 - \epsilon)|H|$ and $\epsilon < 1/3$, in particular H is $|H|/3$ connected, hence in particular $\epsilon|H|/2$ connected. Further, if H has density $p + \epsilon$, then $p > 1 - 2\epsilon$. If we can show $|H| \geq t\sqrt{\log_{1/2\epsilon} t}$ then H contains ${}_{\delta \log t}K_t$ as a minor, and thus G will contain ${}_{(1/\epsilon) \log t}K_t$ as a minor. But this holds since $\epsilon^2/\log(1/2\epsilon) < 1$ for $\epsilon < \frac{1}{3}$, completing our proof. $\square$

Corollary 5.31. *Let $\frac{1}{3} > \epsilon > 0$. Then there is a constant $C = C_{5.31}(\epsilon)$ such that for $t > C$, $r \geq C \log t$ and $d = t\sqrt{r}$, every bipartite graph $G = (A, B)$ with $|B| \leq d/\epsilon$, $|A| \geq Cd$ and every vertex in A having at least ϵd neighbours in B contains ${}_rK_t$ as a minor.*

Proof. By Theorem 5.29, applied with $M = 1/2\epsilon^2$, $C' = 1/\epsilon$ and our current ϵ , there is a constant $D = D_{5.29}(\epsilon, M, C')$ such that if $C > D$, G contains ${}_MH$ as a minor for some graph H with $n \geq \epsilon t\sqrt{r}$ vertices and minimum degree at least $(1 - \epsilon)n$, hence connectivity at least $(1 - 2\epsilon)n \geq 8\epsilon n/40$, as well as density at least $1 - \epsilon$.

By Theorem 5.20, there is a constant $T = T_{5.20}(\epsilon, \epsilon/40)$ such that for all $t > T$ and $r > MT \log t$, every graph G' with at least $8\epsilon/20$ connectivity and density at least $1 - \epsilon$ with more than $t\sqrt{r/M} \leq \epsilon t\sqrt{r}$ vertices contains ${}_{r/M}K_t$ as a minor. Since H has these properties, H has ${}_{r/M}K_t$ as a minor and thus $G \succ {}_rK_t$ as desired. $\square$

We now re-state and prove our sparse sub-theorem. Although the statement applies in both the small and large r cases, our proof will split into these cases internally.

Theorem 5.21. *Let $\epsilon > 0$. Then there is a constant $D = D_{5.21}$ such that for any $t > D$, $r > \epsilon \log t$ and $m > \epsilon t\sqrt{r}$, every graph G with the below four properties contains ${}_rK_t$ as a minor.*

1. $e(G) \leq m|G|$
2. Every edge of G is in at least $m - 1$ triangles
3. $|G| \geq Dm$
4. $\kappa(G) \geq Dt$

Proof. Reducing ϵ if necessary, assume $\epsilon < 1/500$. Let $l = \lceil 8/\epsilon \rceil$, $\delta = \delta_{5.25}(1/300)$ and $k = \max(\lceil 4/\epsilon^2 \rceil, \lceil C_2/\delta \rceil \binom{l}{2})$ where $C_2 = (4/\epsilon^2)C_{5.20}(1/500, 1/41)$. By Lemma 2.21 applied

with this k and the current value of m , there is a constant D_1 such that if $t > D_1$ G contains one of the substructures (A) or (B) described therein. We handle each case separately.

Case (A)

Suppose that G contains a bipartite subgraph G' on vertex set $A \cup B$, where $|B| \leq 6km$, $|A| \geq |G|/3$ and with every vertex in A having at least $m/7$ neighbours in B . Applying Corollary 5.31 we see that provided $t > D_2$, and $r > C_1 \log t$ then such a G' contains ${}_r K_t$ as a minor.

We can therefore assume $r \leq C_1 \log t$, and so it suffices to find a ${}_{C_1 \log t} K_t$ minor in such a bipartite graph, where $m \geq \epsilon^{3/2} t \sqrt{\log t}$. Applying Corollary 5.30 with ϵ replaced by $1/6kC_1$ we see that for $t > D_3$ that G has a ${}_{C_1 \log t} K_t$ minor as desired. Thus in Case (A), provided $D > \max(D_1, D_2, D_3)$, G has an ${}_r K_t$ minor as desired.

Case (B)

We next suppose that, as in (B), G contains disjoint sets $(T_i)_{i=0}^k$. By Lemma 2.22 G further contains paths $(P_{i,j})$ and subsets C_j as therein described. Let $G_i = G[T_i \setminus (\bigcup_{x,y} V(P_{x,y})) \cup \bigcup_j \{b_{i,j}\}]$. The subgraph G_i has connectivity at least $m/41$ and minimum degree at least $3m/5$ for $D > D_4$, and so if $p > \frac{1}{250}$ is the density of G_i , then $|G| \geq 3m/5p$.

Overall, $|G_i| \geq 3\epsilon t/10 \sqrt{\frac{r}{p-1/500}} \geq t \sqrt{\frac{r'}{p-1/500}}$ for $r' = \epsilon^2 r/4$, and thus G_i is ${}_{r'} K_t$ -minor-prevalent provided that $r > C_2 \log t$, where $C_2 = (4/\epsilon^2)C_{5.20}(1/500, 1/41)$ is the constant from earlier. In particular, G_i contains an ${}_{r'} K_t$ model (X_j^i) rooted at the $(b_{i,j})_{j \in V(K_t)}$ for $D > D_5$. Let $Y_j = C_j \cup \bigcup_{i=1}^k X_j^i$. Then Y_j is connected, and further since there are at least r' edges between X_x^i and X_y^i for all $i, x \neq y$, there are at least r edges from X_x to X_y . In particular, the Y_j form a ${}_r K_t$ model in G as desired.

To finish our theorem it remains to handle Case (B) where $r \leq C_2 \log t$, and as in Case A we find a ${}_{C_2 \log t} K_t$ minor under the condition $m > \epsilon^{3/2} \log t$. We can assume that we have subgraphs G_i with $|G_i| \geq 3m/5p \geq 3\epsilon^{3/2} t \sqrt{\log_{1/1-p} t}/5$.

Pick some partition of $V(K_t)$ into l subsets $X_1, \dots, X_l$ each of size either $\lfloor t/l \rfloor$ or $\lceil t/l \rceil$. For each $1 \leq a < b \leq l$, let $H_{a,b}$ be the induced subgraph of ${}_{\delta C_2 \log t} K_t$ on vertex set $X_a \cup X_b$. Then the union of the $H_{a,b}$ for $1 \leq a < b \leq l$ contains ${}_{\delta \log t} K_t$ as a subgraph, and so a union of $\lceil C_2/\delta \rceil$ copies of each contains ${}_{C_2 \log t} K_t$ which is our desired minor.

Pick some function $f : [k] \rightarrow \{a, b : 1 \leq a < b \leq l\}$ such that for each $a < b$, $|f^{-1}(a, b)| \geq \lceil \frac{C_2}{\delta} \rceil$. We would like to find a $H_{f(i)}$ minor inside G_i , rooted at $(b_{i,j})_{j \in V(H_{f(i)})}$. We will do this using Theorem 5.19, noting that for p the density of G_i , $|G_i| \geq 3m/5p \geq |H_{a,b}| \sqrt{\log_{1/(1-p-1/300)} |H_{a,b}|}$ using Inequalities (2.1) and (1.3). Therefore we can apply Theorem 5.19 with $\epsilon = \frac{1}{300}$ and $\eta = \epsilon/20$ to deduce G_i is $H_{f(i)}$ -minor-prevalent by our choice of $\delta = \delta_{5.25}(1/300)$, and since $|H_{a,b}| \geq t/2l$ is large. In particular, there is some model $(X_j^i)_{j \in V(H_{f(i)})}$ of $H_{f(i)}$ in G_i rooted at $(b_{i,j})_{j \in V(H_{f(i)})}$.

Consider the collection $Y_j = C_j \cup \bigcup_{i \in V(H_i)} X_j^i$. These are connected disjoint subsets

(recall that each X_j^i and C_j is connected, and that C_j intersects X_j^i at $b_{i,j}$). Further, each pair of vertices appears in at least $\frac{C_2}{\delta}$ graphs $H_{f(x)}$, and therefore there are at least $\frac{C_2}{\delta} \delta \log t = C_2 \log t$ edges between X_i and X_j , as desired. $\square$

5.6.1 Concluding remarks

It is natural to conjecture that the lower bound, from random graphs, is asymptotically tight.

Conjecture 5.32. Let $k > 0$. Then $c_{\lfloor k \log t \rfloor} K_t \leq (1 + o_t(1))p(k)\sqrt{k/q(k)t\sqrt{\log t}}$, where $p(k), q(k)$ are the functions described in Theorem 5.17.

We have used two broadly similar arguments to prove two special cases of this conjecture, where $k \rightarrow 0$ and $k \rightarrow \infty$. What prevents us from combining these into a single argument and so proving our conjecture in full?

The main obstacle is in the dense case, and more specifically in building almost-compatible partitions (see Theorems 5.23 and 5.26). The two arguments for the small r and large r cases are each quite weak — we apply Markov’s inequality several times, and are able to make this argument sufficiently strong by placing strong conditions on r .

In the small r case, we are able to prove this since we only require a small number of edges between each pair; even if our argument only gives us a millionth of the true number of edges between each pair, this is acceptable since we can take $r/\log t$ arbitrarily small.

Conversely, in the large r case our partition involves nearly the true number of edges, but our estimate for the probabilities is too small. This issue can be worked around by taking $r/\log t$ large, but prevents an extension to the general r case.

If our argument can be extended to prove Conjecture 5.32, it would be necessary to consider when the number of edges deviates above the mean, rather than just below. It also seems necessary to consider several notions of ‘badness’ simultaneously — recall in our proof for $k = o(1)$ we considered bad vertices to have $e(v, X) \leq \eta s$, and for $k = \omega(1)$ we considered a vertex ‘problematic’ if $e(v, X) \leq s(p_v - \eta)$, where sp_v was the expected number of edges over uniformly random X . It seems likely an interpolation between these would be needed.

Chapter 6

Small multigraph minors

6.1 Introduction and preliminaries

In this chapter, we will be studying the structure of extremal graphs for containing fixed small multigraphs as minors. The corresponding problem for graphs has been well-studied, and there small graphs behave very differently to large graphs. For instance, Mader [46] showed $c(K_t) = t - 2$ for $t \leq 7$, which contrasts to the results of Kostochka and independently Thomason that $c(K_t) = \Theta(t\sqrt{\log t})$.

Having introduced the notion of a multigraph minor it is thus natural to investigate small cases, in particular graphs of the form ${}_rK_t$ where both r and t are small. We will work in a slightly more general setting than the previous chapter, allowing multiple edges in our host graph (although still bounding the maximum multiplicity of an edge).

Definition 6.1. Let H be a multigraph. The *minor extremal function* is defined as follows:

$$c(H) = \max\{e(G)/|G| : G \text{ is a simple graph not containing a } H \text{ minor}\}.$$

We further define $c_s(H)$ to be the maximum value of $e(G)/|G|$ over all multigraphs G with edge multiplicities at most s not containing a H minor, and define $c'(H)$ to be $c_s(H)$, where s is the maximum multiplicity of an edge in H .

Recall from the previous chapter that $c({}_rK_2) = \Theta(\sqrt{r})$, with our extremal examples being graphs of high density. Some of our results in this chapter complement this, for instance we show $c({}_3K_2)$ arises from (cockades of) complete graphs on 3 vertices, and $c({}_5K_2)$ and $c({}_6K_2)$ from complete graphs on 4 vertices.

Why do cockades appear here? Consider a 2-connected graph G . If v is a vertex of degree at least r , then $\{v\}, G \setminus \{v\}$ forms a ${}_rK_2$ minor in G . Thus every ${}_rK_2$ minor free graph either has maximum degree at most $r - 1$, or contains a cutvertex. Applying this recursively, such a graph is (formally) a $(\mathcal{H}; 1)$ cockade, where $\mathcal{H}$ is the family of graphs

with maximum degree at most $r - 1$.

When building a cockade, the average degree is determined by $e(H)/(|H| - 1)$ over constituent graphs H . Since $E(H) \leq (r - 1)|H|$, if we can get close to this bound the extremal graphs are forced to be small. These ideas will allow us to exactly compute $c({}_rK_2)$ in several cases.

We can also compute $c'({}_rK_2)$ using this idea. The same inequalities as above hold, and we can attain equality on the graph ${}_{(r-1)}K_2$. Thus $c'({}_{(r-1)}K_2) = r - 1$. This shows c' and c can asymptotically have very different behaviour.

The aim of this chapter is to compute exactly the value of $c_s(H)$ for certain small values of s in the special case where H is a small multigraph, in particular where $H = {}_rK_2$ an edge with multiplicity, or where H is a copy of K_3 with some number of double edges.

An independent motivation for the ${}_rK_2$ case comes from computer science. The Largest Bond problem, or equivalently connected max-cut problem, is finding the largest r for which G contains 2 disjoint connected subsets with at least r edges between them. This is equivalent to the largest r for which ${}_rK_2$ is a minor. Our results for ${}_rK_2$ can therefore be viewed as describing extremal graphs for containing a certain size bond.

The following theorems provide an outline of our results where H and s are fixed graphs. In most of these cases, we will actually prove a slightly stronger result which also classifies extremal graphs.

Theorem 6.2. *The following results hold for $n \geq 1$.*

- (A) $c_1({}_2K_2; n) = c_2({}_2K_2; n) = n - 1$.
- (B) $c_1({}_3K_2; n) = \lfloor 3(n - 1)/2 \rfloor$.
- (C) $c_2({}_3K_2; n) = 2n - 2$.
- (D) $c_1({}_4K_2; n) = \lfloor \frac{5}{3}(n - 1) \rfloor$.
- (E) $c_2({}_4K_2; n) = 2n - 2$.
- (F) $c_1({}_5K_2; n) = 2n - 2$ if $n \equiv 1(3)$, and $2n - 3$ otherwise.
- (G) $c_2({}_5K_2; n) = 3n - 3$ if $n \equiv 1(2)$, and $3n - 4$ otherwise.
- (H) $c_1({}_6K_2; n) = 2n - 2$ unless $n \in \{2, 3, 6\}$ in which case $c_1({}_6K_2; n) = 2n - 3$.
- (I) $c_2({}_6K_2; n) = 3n - 3$ unless $n = 2$.

Theorem 6.3. *Let K_3^+ denote K_3 with one doubled edge, and ${}_2K_3^-$ denote ${}_2K_3$ with one edge removed. Then the following hold for $n \geq 1$.*

- (A) $c_1(K_3^+; n) = \lfloor 3(n - 1)/2 \rfloor$.
- (B) $c_2(K_3^+; n) = 2n - 2$.
- (C) $c_1({}_2K_3^-; n) = 2n - 3$.

$$(D) \quad c_2(2K_3^-; n) = 2n - 2.$$

$$(E) \quad c_1(2K_3; n) = 2n - 2 \text{ for } n \equiv 1(3), \text{ and } 2n - 3 \text{ otherwise.}$$

$$(F) \quad c_2(2K_3; n) = 3n - 4.$$

Our bound for $c_2(2K_3; n)$ matches a result of Janaqi et al. [35] for $c(*K_3)$, where $*K_3$ denotes a copy of $2K_3$ with all edges directed as opposing pairs. This shows restricting orientation does not yield any additional obstacles to containing $*K_3$ relative to its underlying multigraph $2K_3$.

Theorem 6.4. *For each $n \geq 1$, the following hold.*

$$(A) \quad c_3(4K_2; n) = 3n - 3.$$

$$(B) \quad c_3(5K_2; n) = 3n - 3.$$

$$(C) \quad c_3(6K_2; n) = \lfloor \frac{7(n-1)}{2} \rfloor.$$

Before stating our results for general r and s , we first introduce the notion of a cockade. These will be used to describe our extremal graphs, both in the general case and in specific cases. We will prove some basic properties of cockades in the next section.

Definition 6.5. Let $\mathcal{F}$ be a family of graphs, and $k \geq 0$ an integer. The family of $(\mathcal{F}; k)$ -cockades is defined as follows.

- Every element of $\mathcal{F}$ is an $(\mathcal{F}; k)$ -cockade.
- If H_1 and H_2 are $(\mathcal{F}; k)$ -cockades, and K^1, K^2 (labelled) copies of K_l in H_1, H_2 respectively for some $l \leq k$, every graph obtained from a disjoint union of H_1 and H_2 by identifying K^1 and K^2 then deleting a (possibly empty) set of edges from this clique is a $(\mathcal{F}; k)$ -cockade.

Where it does not cause confusion, we will drop set notation and refer to, for instance, an $(H_1, H_2; 1)$ -cockade.

Theorem 6.6. *Let $r - 1 \geq s > \frac{3}{4}(r - 1)$ be integers. Then $c_s(rK_2; n) = s(n - 1)$, and the extremal graphs are exactly $(sK_2; 1)$ -cockades.*

As a special case, this theorem proves $c'(rK_2; n) = (r - 1)(n - 1)$ with extremal graphs being ‘trees with multiplicity r ’, as asserted in Chapter 5 — note that no edge with multiplicity r can appear in a rK_2 minor free graph, hence $c'(rK_2; n) = c_r(rK_2; n) = c_{r-1}(rK_2; n)$.

Theorem 6.7. *Let r be an odd integer, and s such that $\frac{3}{4}(r - 1) \geq s > \frac{r-1}{2}$. Then*

$$c_s(rK_2; n) = \begin{cases} \frac{3}{4}(r - 1)(n - 1), & n \text{ odd} \\ \frac{3}{4}(r - 1)(n - 2) + s, & n \text{ even} \end{cases}$$

holds. The extremal graphs are exactly $(sK_2, (r-1)/2 K_3; 1)$ -cockades.

Theorem 6.8. *Let r be an even integer, and s such that $\frac{3}{4}(r-1) \geq s > \frac{r-1}{2}$. Then*

$$c_s({}_rK_2; n) = \begin{cases} \left(\frac{3}{2}r - 2\right)\frac{n-1}{2}, & n \text{ odd} \\ \left(\frac{3}{2}r - 2\right)\frac{n-2}{2} + s, & n \text{ even} \end{cases}$$

holds. The extremal graphs are exactly $({}_sK_2, (\lfloor (r-1)/2 \rfloor K_3)^+; 1)$ -cockades.

We have now stated our main results, and so can give a broad overview of this chapter. In the next Section 6.2, we will prove some basic properties of cockades; these are required for all of our proofs. Then, in Section 6.3, we will discuss K_4 minor free graphs. This is useful for several of our theorems, since the minors of interest are ‘close to’ appearing inside K_4 minors.

Section 6.4 will contain proofs of all the results in Theorem 6.2, as well as those from Theorems 6.4 and 6.6 to 6.8 — i.e. all of the results relating to $c_s({}_rK_2)$. Section 6.5 then proves Theorem 6.3. We finish the chapter with a discussion of further cases, in particular for ${}_2K_4$.

6.2 Cockades

The aim of this section is to prove some basic properties of cockades. These will be used to simplify descriptions of extremal graphs — all of our extremal graphs will be $(\mathcal{F}; 1)$ -cockades for some (reasonably simple) family $\mathcal{F}$. The implicit reduction to 2-connected graphs in the next lemma is perhaps the most important property.

Lemma 6.9. *Let $H \neq K_2$ be a connected multigraph with no cut vertex. The extremal H -minor free graphs with edge multiplicities at most s are $(\{{}_lK_2\} \cup \mathcal{F}(H); 1)$ -cockades, where $\mathcal{F}(H)$ is the family of edge-maximal 2-connected H -minor-free graphs and l is either s , if H is not an edge with multiplicity, or $\min(r-1, s)$ if $H = {}_rK_2$.*

Proof. Suppose that G is an H -minor free graph with the maximum number of edges. We will proceed by induction on $|G|$, the base case $|G| = 1$ being trivial. If G is disconnected, suppose C_1 and C_2 are two components, and $v \in C_1, w \in C_2$. Consider the graph $G + vw$. Since G is maximal, this graph must contain an H model (V_h) . Further, it cannot contain both a part entirely contained in C_1 and a part entirely contained in C_2 , since then either $H = K_2$, H is disconnected or if x is such that $v \in V_x$, then x is a cutvertex. Such an x must exist since otherwise the same subsets form a minor in G . So without loss of generality the model only has parts contained in C_1 . But then the sets $V_h \cap C_1$ form an H model in G and we have a contradiction, so G must be connected.

If G contains a cutvertex v , suppose that C_i are the components of $G \setminus \{v\}$. If each $G[C_i \cup \{v\}]$ is maximal H minor free, then by induction G is an $(\mathcal{F}(H); 1)$ -cockade. So assume that $G[C_1 \cup \{v\}]$ is not maximal. Let G' be the graph obtained from G by deleting

all edges in $C_1 \cup \{v\}$, and placing a maximal H minor free graph on the same vertices. Since G has the maximal number of edges among H minor free graphs, this new graph G' has an H model.

As above, if all of such a model (V_h) intersected $C_1 \cup \{v\}$, then the intersections $V_h \cap (C_1 \cup \{v\})$ form an H minor in $G'[C_1 \cup \{v\}]$ which is a contradiction. So at least one part V_h is entirely contained in C_2 , and likewise some $V_{h'}$ within C_1 . Then since H is connected, some part $V_{h''}$ contains v ; but this means h'' must be a cutvertex which is a contradiction.

We can therefore assume G is 2-connected, but then since G has the maximal number of edges among H minor free graphs, we deduce $G \in \mathcal{F}(H)$ as desired. $\square$

The next lemma is immediate from the definition of a cockade.

Lemma 6.10. *Let G be an $(H_1, \dots, H_k; 1)$ -cockade, where we use n_i copies of H_i . Then G has exactly $n = \sum_{i=1}^k n_i(|H_i| - 1)$ vertices, and $\sum_{i=1}^k n_i e(H_i)$ edges.*

Lemma 6.11. *Let $H = {}_r K_2$. Then every 2-connected graph G without an H minor has maximum degree at most $r - 1$.*

Proof. Suppose instead $d(v) \geq r$. Then $\{v\}, G \setminus \{v\}$ is an H minor in G . $\square$

These lemmas are useful since complete graphs do not have cutvertices, and so we can describe our extremal graphs as cockades. The second lemma is incredibly simple, but also very useful for edges with multiplicity; in fact we can use it to bound the order of a 2-connected extremal graph for small values of r . We remark this bound does not generalise to larger r , since the maximum average degree of ${}_r K_2$ minor free graphs is $(1 + o(1))\sqrt{r} \ll r - 1$ as shown in Chapter 5.

6.3 An interlude: K_4 -minor-free graphs

Many of our multigraphs are ‘close to’ appearing inside a K_4 minor — for instance, ${}_4 K_2$ is a minor of K_4 , ${}_5 K_2$ and ${}_2 K_3$ appear if we double a single edge, and ${}_6 K_2$ appears if we add any two edges. This means a K_4 minor will impose significant additional structure. We therefore introduce the following definition, used to classify K_4 minor free graphs.

Definition 6.12. A graph G is called a *series-parallel (SP) graph* with terminals x, y according to the following recursive definitions.

1. The edge xy is a series-parallel graph with terminals x, y .
2. If $G_1, \dots, G_k$ are series-parallel graphs on disjoint vertex sets with terminals x_i, y_i respectively, the graph obtained by identifying y_i and x_{i+1} is series-parallel with terminals x_1 and y_k (series composition).

3. If $G_1, \dots, G_k$ are series-parallel graphs on disjoint vertex sets with terminals x_i, y_i respectively, and at most one G_i contains the edge xy , the graph obtained by identifying all x_i as a new vertex x and all y_i as y is series-parallel with terminals x, y (parallel composition).

We say a series-parallel graph *has type S* if it is a series composition of two or more SP graphs, and *has type P* if it is a parallel composition.

A folklore result states that a 2-connected K_4 minor free graph is a series-parallel graph (of type P). This result can be proven inductively, using the stronger statement that in a 2-connected graph the terminals x, y can be chosen arbitrarily. Another characterisation is that 3-connected graphs contain K_4 as a minor. For a proof of this, pick a cycle C not using every vertex, and a vertex v not in the cycle. There are 3 paths from v to C by Menger's theorem, and this gives a K_4 minor.

We note every type S graph has a cut vertex. We can write every type P graph as a parallel composition of $k \geq 2$ graphs which are type S or an edge, and every type S graph as a series composition of $k \geq 2$ graphs, which are either type P or an edge. The edge xy is neither type P nor type S (and is the only series-parallel graph of neither type).

Lemma 6.13. *A type P series-parallel graph has at most $2n - 3$ edges. A type S series-parallel graph has at most $2n - 4$ edges.*

Proof. We proceed by induction on n , the base cases $n = 2, 3$ being trivial and also hold treating an edge as a type P graph. If G is a series composition of $G_1, \dots, G_k$ where $|G_i| = n_i$, then $|G| = \left(\sum_{i=1}^k n_i\right) - (k - 1)$ and $e(G) \leq \left(\sum_{i=1}^k 2n_i\right) - 3k \leq 2n - 4$.

Conversely, if G is a parallel composition of $G_1, \dots, G_k$, by adding additional edges if necessary assume that G_1 is the edge xy . Then if $|G_i| = n_i$ for $i \geq 2$ we have $|G| = 2 + \left(\sum_{i=2}^k n_i\right) - 2k + 2$ and $e(G) \leq 1 + \sum_{i=2}^k (2n_i - 4) \leq 2|G| - 3$ as required. $\square$

We finish this section by noting K_4 appears as a topological minor exactly when it appears as a minor since it has maximum degree at most 3. To prove this, consider any (minimal) K_4 model; this consists of four trees and exactly one edge between each pair. We can pick a branch vertex for a part by picking a vertex that separates the three edges incident to it.

6.4 Results for an edge with multiplicity

Recall the aim of this section is to prove our results for graphs ${}_r K_2$. In subsection 6.4.1 we will prove our more general results for $c_s({}_r K_2)$, namely Theorems 6.6-6.8. We will also use these results to derive some cases of Theorem 6.2 and prove Theorem 6.4. The later subsections will all be dedicated to Theorem 6.2, with the results proven in increasing order of r .

6.4.1 General results

Proof of Theorem 6.6. We proceed by induction on n , the base cases $n = 1, 2$ being trivial. Let G with $|G| = n$ be maximal rK_2 minor free, and by Lemmas 6.9 and 6.11 we can assume G is 2-connected on n vertices with maximum degree at most $r - 1$. If $(r - 1)n/2 \leq s(n - 1)$ the result holds, and so we can assume $s(1 - 1/n) \leq (r - 1)/2$; in particular we have $(1 - 1/n) \leq (r - 1)/2s$. If $n \leq 2$ the result holds as earlier, so assume $n \geq 3$. Then this result can be rewritten as $\frac{2}{3} \leq \frac{r-1}{2s}$, i.e. $s \leq \frac{3}{4}(r - 1)$; this contradicts the statement of our theorem. $\square$

This is essentially tight; for instance if $r = 5$ and $s = 3$ the graph ${}_2K_3$ provides a different extremal example; for $s = 2$ this becomes uniquely extremal up to parity constraints. The following two corollaries are immediate since the conditions of Theorem 6.6 hold.

Corollary 6.14 (Theorem 6.2(A)). $c_1({}_2K_2; n) = n - 1$, and the extremal graphs are $(K_2; 1)$ -cockades (equivalently, extremal graphs are trees).

Corollary 6.15 (Theorem 6.2(C)). $c_2({}_3K_2; n) = 2n - 2$ and further the extremal graphs are exactly $({}_2K_2; 1)$ -cockades, equivalently trees with multiplicity 2.

Corollary 6.16 (Theorem 6.4(A)). $c_3({}_4K_2; n) = 3n - 3$, and the extremal graphs are exactly $({}_3K_2; 1)$ -cockades.

Proof of Theorem 6.7. As before, let G with $|G| = n$ be rK_2 -minor free with the maximal number of edges; we assume that G is 2-connected and hence has maximum degree at most $r - 1$. If n is odd, such a graph has at most $(r - 1)n/2$ edges, compared to the $3(r - 1)(n - 1)/4$ edges in a ${}_{(r-1)/2}K_3$ cockade. For G to be extremal, we require $n/2 \geq 3(r - 1)/4$ and so $n = 3$ and G is $(r - 1)$ regular to match this. If m is the multiplicity of the largest edge, the other two edges have multiplicity $r - 1 - m$. So G has $2r - 2 - m$ edges; this is maximised uniquely when $m = \frac{(r-1)}{2}$ and so G is the claimed 3 vertex graph.

If instead n is even, we are comparing $(r - 1)n/2$ to the $\frac{3(r-1)}{4}(n - 2) + s$ in our extremal graphs. Since $s > \frac{r-1}{2}$ we cannot have as many edges as the graphs in the statement, which completes our proof. $\square$

Corollary 6.17 (Theorem 6.4(B)). $c_3({}_5K_2; n) = 3n - 3$, and the extremal graphs are exactly $({}_3K_{2,2} K_3; 1)$ -cockades.

Remark 6.18. The requirement $s > \frac{r-1}{2}$ is tight, because taking ${}_{(r-1)/2}C_4$ matches these examples when $s = \frac{r-1}{2}$.

Proof of Theorem 6.8. Suppose for contradiction G is an H minor free graph with $|G| = n$ and at least as many edges the claimed extremal examples. As before we assume G is 2-connected and hence has maximum degree at most $r - 1$. Suppose first that n is odd. G has at most $n(r - 1)/2$ edges, and this must be at least $((r - 1) + \frac{(r-2)}{2})(n - 1)/2$ edges. We can reduce this inequality to $n(1 - \frac{r}{2}) \geq 2 - \frac{3r}{2}$, and so either $r = 2$ (and hence no such

s exists) or $n \leq 3 + \frac{2}{r-2} < 4$. The $n = 2$ case is trivial, and for $n = 3$ our example has 2 vertices of degree $r - 1$ and one of degree $r - 2$; this is uniquely extremal since there is no graph on an odd number of vertices with all degrees odd.

We therefore assume n is even. This time our extremal examples have $(\frac{3}{2}(r - 1) - \frac{1}{2})\frac{n-2}{2} + s$ edges. Comparing this to our bound $e(G) \leq \frac{r-1}{2}n$ and using $s \geq \frac{r}{2}$ we deduce that $4(2 - r) \leq n(2 - r)$, which holds for $r = 2$ and any n (however in this case no valid s exists) or $r > 2$ and $n \leq 4$. The $n = 4$ case holds with equality, and so we need $s = \frac{r}{2}$ and G is $(r - 1)$ regular to have a contradiction. The only 4 vertex graphs with no cut-vertices and minimum degree at least 2 (recall $s < r - 1$) contain a C_4 subgraph. Some edge of this C_4 must have multiplicity less than $\frac{r}{2}$ as $_{r/2}C_4$ contains the forbidden minor. Such an edge has multiplicity between 1 and $\frac{r-2}{2}$ and does not form a cut-set; in particular it forms one half of an $_rK_2$ minor which is a contradiction. $\square$

Corollary 6.19 (Theorem 6.2(E)). $c_2({}_4K_2; n) = 2n - 2$ and the extremal graphs are exactly $({}_2K_2, K_3^+; 1)$ -cockades.

Corollary 6.20 (Theorem 6.4(C)). $c_3({}_6K_2; n) = \lfloor \frac{7(n-1)}{2} \rfloor$ and the extremal graphs are exactly $({}_3K_{2,2}, K_3^+; 1)$ -cockades.

General statements of this form can likely be extended further, however as r/s increases the cases we need to consider become more involved — even the cases $s = 1$ and any r will become part of such a theorem eventually.

Lemma 6.21. *Let r, s, t be integers. Then $c_{ts}({}_{t(r-1)+1}K_2; n) \geq tc_s({}_rK_2; n)$.*

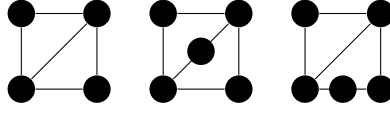
Proof. Let G be a graph with edge multiplicities at most s and no $_rK_2$ model. Then multiplying all edge multiplicities by t , we obtain a graph with edge multiplicities at most st . This graph cannot contain a $_{rt}K_2$ model. But also all models have a multiple of t edges between the parts, so in fact it contains no $_{(r-1)t+1}K_2$ model. $\square$

The corresponding upper bound from this argument is not very good, however; we can only obtain $c_{st}({}_{rt}K_2; n) \leq tsc_1(\lceil r/s \rceil K_2; n)$ (recall every edge need not have multiplicity equal to s). Our best asymptotic results are in Chapter 5 and show $c_1({}_rK_2) = (\frac{1}{2} + o(1))\sqrt{r}$; hence taking r/s and s both large, we have $c_s({}_rK_2) \geq (\frac{1}{2} + o(1))\sqrt{rs}$.

Theorem 6.22. $c_1({}_3K_2; n) = \lfloor 3(n-1)/2 \rfloor$, and the extremal graphs are exactly $(K_2, K_3, C_4; 1)$ -cockades.

Proof. As usual, by Lemmas 6.9 and 6.11 we can assume G is a 2-connected graph with maximum degree at most 2. The only such graphs are cycles C_l . We note that a $(K_3, K_2; 1)$ -cockade on $n \geq 5$ vertices using at least two copies of K_3 has strictly more than n edges; hence extremal graphs are exactly $(K_2, C_3, C_4; 1)$ -cockades. $\square$

Our arguments so far have used the maximum degree to infer an upper bound on the

Figure 6.1: In left to right order, the graphs K_4^- , L_1 and L_2

number of vertices of (2-connected) extremal graphs. This argument is not strong enough for cases beyond this, and further ideas are needed. We will still be using the maximum degree bound, however.

6.4.2 The $r = 4$ case

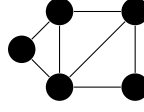
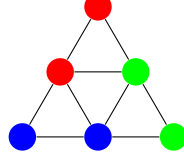
Theorem 6.23. $c_1({}_4K_2; n) = \lfloor \frac{5}{3}(n-1) \rfloor$. Further, extremal graphs are $(\mathcal{F}; 1)$ -cockades for the family $\mathcal{F} = \{K_2, K_3, K_4^-, L_1, L_2\}$, where the graph K_4^- denotes K_4 with an edge removed, and L_1 and L_2 are the graphs obtained from K_4^- by subdividing a single edge. These are depicted in Figure 6.1.

Proof. The claimed value is a lower bound since all of $K_2, K_3, K_4^-, L_1, L_2$ are ${}_4K_2$ minor free. We seek to show that all graphs with at least $\lfloor \frac{5}{3}(n-1) \rfloor$ edges and no ${}_4K_2$ minor are cockades of these graphs. We prove this by induction, and as usual assume G is 2-connected with maximum degree at most 3. The base cases $n \leq 3$ are trivial. We note that for $n \equiv 1(3)$ we have $e(G) \geq \frac{5}{3}(n-1)$, and likewise $\frac{5}{3}(n-1) - \frac{2}{3}$ for $n \equiv 2(3)$ and $\frac{5}{3}(n-1) - \frac{1}{3}$ for $n \equiv 0(3)$. The maximum degree bound does tell us $n \leq 12$ however this is still many cases to check; we therefore use a different argument.

Claim 6.23.1. Every degree 2 vertex of G is contained in a triangle.

Proof. If not, suppose v is joined to vertices a, b . Then G is a subdivision of the graph G' obtained from G by deleting v and adding the edge ab . Since G' has $n-1$ vertices and $e-1$ edges, and is a minor of G , we must have $n \equiv 2(3)$, and G' is a $(K_4^-; 1)$ -cockade. Thus G is obtained from a $(K_4^-; 1)$ -cockade by subdividing a single edge. This cannot stop a vertex from being a cutvertex, so in fact $G' = K_4^-$ and thus $G = L_1$ or $G = L_2$. We used here that $(\mathcal{F}; 1)$ -cockades cannot be 2-connected unless they are a member of $\mathcal{F}$. $\square$

Now G cannot contain a degree 1 vertex, else its neighbour is a cut-vertex as $n \geq 3$. So G contains at least one degree 3 vertex, else G is a cycle and we are done. Let v be such a vertex. If G has 3 edges in its neighbourhood, we have a K_4 subgraph and hence ${}_4K_2$ minor; so assume it is at most 2. If there are exactly two edges, G contains two adjacent degree three vertices v, w with two common neighbours x, y . If this is the entire graph, $G = K_4^-$ so assume not; to avoid a cutvertex both x and y have degree 3. There must be some path outside the K_4^- from a neighbour of x to a neighbour of y ; let P be such a path. Then $\{xwP\}, \{yv\}$ is a ${}_4K_2$ minor. We therefore assume there is at most one neighbour; in particular v has a neighbour w with no common neighbour. If w has degree 2, we are


 Figure 6.2: The graph H_3 from Theorem 6.24.

 Figure 6.3: The ‘triforce graph’, with monochrome subsets forming a ${}_2K_3$ model.

a subdivision of a $(K_4^-; 1)$ -cockade and hence a $(K_4^-, L_1, L_2; 1)$ -cockade, so assume w has degree 3.

Consider the graph $G/\{vw\}$. This has 1 fewer edge and 1 fewer vertex, so by induction is a $(K_4^-; 1)$ -cockade and G is obtained from this cockade by ‘splitting’ a vertex. To avoid having a cutvertex, we deduce G is obtained from a cockade of two copies of K_4^- with their degree 2 vertices joined at vw . To avoid v or w being a cut-vertex, each must have one edge to each copy of K_4^- . Using two paths inside each gives a ${}_4K_2$ minor and hence a contradiction. $\square$

This argument is a simplest case of our new methodology: iteratively gain more information about a 2-connected extremal graph, until we have reduced to a small set of graphs, none of which contain the forbidden minor. We will also use the reduction to minimum degree at least 3 throughout.

6.4.3 The $r = 5$ case

Theorem 6.24. $c_1({}_5K_2; n) = \begin{cases} 2n - 2, n \equiv 1(3) \\ 2n - 3, n \not\equiv 1(3) \end{cases}$, and the extremal graphs are $(\mathcal{F}; 1)$ -cockades where $\mathcal{F} = \{K_4, K_2, K_3, K_{2,3}^*, H_1, H_2, H_3\}$ and H_1 is a copy of K_4 with a single edge subdivided, H_2 is the triforce graph from Figure 6.3 and H_3 is the graph of Figure 6.2. In order to be extremal, all but at most one graph of the cockade must be a K_4 .

Proof. Let G be a 2-connected graph on n vertices with at least the claimed number of edges. Suppose first that G contains a K_4 minor (and hence a K_4 subdivision). Adding any edge to a K_4 subdivision allows us to contract to K_4^+ and hence to ${}_5K_2$ which is a contradiction. If there is any vertex not contained in the K_4 subdivision, by 2-connectivity it has 2 paths to the subdivision and so we can contract to gain an extra edge as above. Thus G is a topological K_4 and hence by the edge bound G is either K_4 or H_1 .

Therefore, using results from Section 6.3 G must be a (maximal) series-parallel graph

of type P . Let G be a parallel composition of graphs $G_1, \dots, G_k$, and by maximality we can take G_1 to be the edge xy . Consider the maximum l such that G contains an ${}_lK_2$ minor with x in one part, and y in the other. Such an l must be the sum of the corresponding l_i over $G_1, \dots, G_k$. By assumption $l < 5$ and hence $\sum_{i=2}^k l_i < 4$. We note $l_i = 1$ if and only if G_i is a series composition of edges. Further, edge-maximal series-parallel graphs only ever take a series composition of two graphs, and so $l_i = 1$ if and only if G_i is a series composition of exactly two edges (i.e. an edge subdivided once). In particular, if $k = 4$ then $G = K_{2,3}^*$.

Suppose next $k = 3$ and G_2 is an edge subdivided once — this must exist up to reordering as it can't hold $l_2, l_3 \geq 2$. We consider G_3 . This must be a series composition of two graphs, each not containing a ${}_3K_2$ minor between their respective terminals. In particular, each such graph is either an edge or a triangle. If both are edges, $|G| = 4$ and we are not extremal. If exactly one is a triangle, $G = H_3$, and if both are triangles, $G = H_2$. It remains only to consider the $k = 2$ case.

Suppose the graph G_2 is a series composition of G'_1 and G'_2 . If either of these is an edge say G'_2 , G can equivalently be expressed as a parallel composition of G'_1 and the series composition of two edges. This is either a series composition of three or more graphs, as handled before, or G is exactly K_3 . It remains to handle the case where neither G'_1 nor G'_2 is an edge.

Since each of G'_1 and G'_2 is not an edge, they each contain a ${}_2K_2$ model with their terminals in different parts of the model. This means we cannot have a ${}_3K_2$ minor in either — otherwise we can build a ${}_5K_2$ model with the vertex of $G'_1 \cap G'_2$ in one part, and x and y in the other. This means each of G'_1 and G'_2 is exactly a copy of K_3 , and hence $G = H_3$. This completes our proof since we have classified all 2-connected extremal graphs. $\square$

Theorem 6.25. $c_2({}_5K_2; n) = \begin{cases} 3n - 3, n \equiv 1(2) \\ 3n - 4, n \equiv 0(2) \end{cases}$, and the extremal examples are exactly $({}_2K_3, {}_2C_4, {}_2K_2; 1)$ -cockades.

Proof. As usual it suffices to consider 2-connected graphs (which have maximum degree at most 4), and all such graphs have $e(G) \leq 2n$. Thus extremal graphs have $n = 3$ or $n = 4$. If $n = 3$ then $G = {}_2K_3$ so assume $n = 4$ and G is 4-regular. Then G has at least 2 double edges. If two of these are incident, our maximum degree condition forces $G = {}_2C_4$. The only remaining case is G being a copy of K_4 with two non-incident double edges ab, cd . In this case, $\{ac\}, \{bd\}$ is a ${}_6K_2$ minor which is a contradiction. $\square$

The proof of Theorem 6.24 is the first case where we use the structure of K_4 minor free graphs. We will be using this structure for all arguments from now on. The realisation that K_4 minor free proofs greatly simplify our arguments came after an original, longer, proof

along the lines of Theorem 6.23. An important step is that edge maximal series-parallel graphs can only ever take a series composition of two graphs.

6.4.4 The $r = 6$ case

Theorem 6.26. $c_1({}_6K_2; n) = \begin{cases} 2n - 2, n \neq 6 \\ 2n - 3 = 9, n = 6 \end{cases}$, and the extremal graphs for $n \geq 7$

are $(K_4, K_5^{--}; 1)$ -cockades, where K_5^{--} denotes the family of two non-isomorphic graphs obtained from K_5 by deleting two edges.

Proof. The cases $n \leq 5$ are trivial, and it is not hard to verify the $n = 6$ case. By considering such graphs which maximise $\frac{e(H)}{|H|-1}$, by Lemma 6.9 it suffices to prove that there are no 2-connected graphs on $n \geq 6$ vertices with at least $2n - 2$ edges and no ${}_6K_2$ minor. We will prove this by induction on n , with the base case $n = 6$ holding so assume $n \geq 7$.

If G contains an edge in no triangle, contracting this edge contradicts our extremal function; thus we can assume every edge is in a triangle. Further, we can assume G contains no degree 2 vertex: if it did, deleting such a vertex v we obtain an $n - 1$ vertex graph with $2n - 4$ edges and no ${}_6K_2$ minor; further this graph is 2-connected (recall every edge is in a triangle, and so the neighbours of v are adjacent) and so contradicts our induction hypothesis.

Since $e(G) > 2n - 3$, G contains a K_4 topological minor. If two or more edges of the K_4 topological minor are subdivided, using that each edge is in a triangle we can contract to a minor which is a K_4 with two additional edges; this in turn contains a ${}_6K_2$ minor and gives a contradiction.

If our topological minor is a K_4 subgraph, then as $n \geq 5$ we can pick a vertex v outside this subgraph. By 2-connectivity there are 2 vertex disjoint paths from v to K_4 , which provides an alternative subdivided path between some pair of vertices. Therefore we can assume exactly one edge of the topological minor is subdivided.

Let $abcd$ be the branch vertices of such a topological minor, consisting of edges between all pairs except ab , which is replaced with a path $x_0 \dots x_l$ with $l \geq 2$, $x_0 = a$ and $x_l = b$. If at least two of $x_1, \dots, x_{l-1}$ have edges to a, b, c or d (excluding the necessary ax_1 and $x_{l-1}b$), then contracting this edge appropriately we can obtain a K_4 with either two doubled or a single tripled edge and hence a ${}_6K_2$ minor; we therefore assume this never happens.

Suppose next that some vertex x_k has no edge to a branch vertex (apart from our used path) for $0 < k < l$. This means that $x_{k-1}x_k$ is contained in a triangle with some vertex y (not among a, b, c, d). If there were some path outside this triangle and c, d between some pair from x_{k-1}, x_k, y then we can contract suitably to obtain a K_4 with a triple edge; this is a contradiction. In particular no pair have a common neighbour. But this means we can

contract the triple $x_{k-1}x_ky$ to obtain a minor with $n - 2$ vertices and at least $2n - 5$ edges, and hence containing a ${}_6K_2$ minor.

Therefore neither of the above cases can occur, and thus $l = 2$ with x_1 having an edge to one of c and d , relabelling assume it is c . Then G contains $K_1 + C_4$ as a minor, with c the degree 4 vertex. Further, any multigraph obtained from $K_1 + C_4$ by adding an edge contains a ${}_6K_2$ minor. Therefore taking any vertex outside this subgraph and two paths to it we obtain a ${}_6K_2$ minor. This eliminates all cases, and so there are no n vertex 2-connected extremal graphs with $n \geq 6$ as desired. $\square$

Theorem 6.27. $c_2({}_6K_2; n) = 3n - 3$, and the extremal graphs are exactly $({}_2K_3, H_4; 1)$ -cockades, where H_4 is the graph obtained from ${}_2C_4$ by addition of a diagonal edge.

Proof. It suffices to show there are no other 2-connected ${}_6K_2$ -minor free multigraphs with $3n - 3$ vertices, since ${}_2K_3$ and H_4 are maximal ${}_6K_2$ minor free. Since the maximum degree is at most 5, we have $e(G) \leq 5n/2$ and so $n \leq 6$. We now rule out each case in turn.

Suppose first $n = 4$. Then G has 9 edges, which must include at most 4 double edges since $c_1({}_3K_2; 4) = 4$. Our maximum degree condition means no vertex is incident to three double edges. Our graph is therefore either H_4 (the only possible graph with 4 double edges and maximum degree less than 6), or a copy of K_4 with three doubled edges; even having two doubled edges already gives a ${}_6K_2$ minor.

Suppose next $n = 5$, and G has 12 edges thus degree sequence 5, 5, 5, 5, 4. This implies G contains at least 2 single edges. Define the *underlying graph* of G to be the simple graph obtained by replacing all multiple edges with single edges. Since $c_1({}_6K_2; n) = 8$, the underlying graph has at most 8 edges, and since G has at least two single edges it must have at least 7. If the underlying graph has 8 edges, it must be a K_5^- . But doubling any edge of this gives a ${}_6K_2$ minor and a contradiction. So the underlying graph has 7 edges, and so G has 5 double edges which must form a 2-regular graph. Thus G is a copy of ${}_2C_5$ with two extra non-incident edges; such a graph has a ${}_6K_2$ minor.

We can therefore assume $n = 6$ and G is 5-regular. By considering the underlying graph (which must have at most 9 edges), and that each vertex of G has edges to at least 3 distinct vertices, both of these hold with equality (and so the double edges form a 2-regular graph). There are two cases. If the double edges form two triangles abc, def , our other edges must be ad, be, cf up to relabelling. Then $\{ad\}, \{becf\}$ forms a ${}_6K_2$ minor. Thus assume the double edges form a 6-cycle $abcdef$, and so the single edges must be ad, be, cf . Then $\{abc\}, \{def\}$ forms a ${}_6K_2$ minor. This completes our proof. $\square$

The methods of the $r = 6$ case are somewhat more subtle; we now have to consider additional paths relative to a K_4 topological minor. This idea will return later, in the proof of Theorem 6.32.

It is natural to ask if these ideas can be pushed further. Asymptotically, stronger

methods will be needed, considering larger and larger subsets, since we know from Chapter 5 that asymptotically $c(rK_2) = \Theta(\sqrt{r})$, so subsets of size $\Theta(\sqrt{r})$ need to be considered as parts of minors. Stopping at $r = 6$ is natural since beyond this the K_4 minor simplifications start to break down: K_5 is $7K_2$ minor free.

Does anything more general hold? It seems reasonable to conjecture that when $r = t^2 + 1$ is one more than a square, then K_{2t} -cockades are uniquely extremal (under divisibility constraints). This result may also hold when $r = (t - \frac{1}{2})(t + \frac{1}{2}) + 1$, giving cockades of odd complete graphs.

6.5 Triangles

Having proven all of the results for edges, it remains only to prove Theorem 6.3 which is the aim of this section. We will proceed in increasing order of edge count, starting with K_3^+ , then moving onto $2K_3^-$ and finally $2K_3$.

Theorem 6.28. $c_1(K_3^+; n) = \lfloor 3/2(n-1) \rfloor$, and the extremal examples are, as for $3K_2$, $(K_2, K_3, 1)$ -cockades. Similarly, $c_2(K_3^+; 2) = 2n - 2$ and the extremal examples are $(2K_2, 1)$ -cockades.

Proof. Suppose that a connected simple graph G (with $|G| \geq 3$) contains $3K_2$ as a minor; let (A, B) be some $3K_2$ model in G and we may assume $|A| \geq 2$ since G is simple. Take a spanning tree for A and some leaf a . Consider $\{a\}$ and $A \setminus \{a\}$. At least one of these has two edges to B , and both are connected, so G contains a K_3^+ minor. Conversely, contracting either single edge of K_3^+ yields $3K_2$. Thus extremal graphs for K_3^+ and $3K_2$ are identical. The same argument works to show $c_2(K_3^+; 2) = c_2(3K_2; 2) = 2n - 2$. $\square$

6.5.1 The graph $2K_3^-$

Theorem 6.29. $c_1(2K_3^-; n) = 2n - 3$ for each $n \geq 2$. Further, all graphs with $2n - 3$ vertices and no $2K_3^-$ model are isomorphic to books $K_{2,n-2}^*$.

Proof. We will proceed by induction on n ; the base cases $n \leq 3$ hold trivially. Let G be an n vertex $2K_3^-$ minor free graph with at least $2n - 3$ edges; we aim to show G is a book. G cannot contain a K_4 minor (contracting any edge of K_4 gives $2K_3^-$) and so G must be a maximal type P series-parallel graph by the discussion of Section 6.3. Let u, v be the terminals of this graph, and suppose we take a parallel composition of graphs $G_1, \dots, G_l$ with $l \geq 2$ and G_1 the edge uv , all other G_i are type S series parallel graphs.

Each G_i for $i \geq 2$ must be a series composition of exactly two graphs $G_{i,1}$ and $G_{i,2}$, each either a parallel composition or an edge using maximality. Since any parallel composition can be contracted to a double edge, it cannot be the case that both $G_{i,1}$ and $G_{i,2}$ are parallel compositions. If no such graph $G_{i,j}$ is a parallel composition, then G is exactly a book $K_{2,n-2}^*$. So, relabelling if necessary, assume $G_{2,1}$ is a parallel composition.

If $l \geq 3$ then contracting G_3 we can obtain a double edge from u to v ; after contracting $G_{2,1}$ also we obtain a contradiction. So we must have $l = 2$. But now we can re-express G as a series-parallel graph with terminals u and the vertex of $G_{i,1} \cap G_{i,2}$; further this expression has $l \leq 3$. Thus the earlier case shows G is a book as desired. $\square$

Remark 6.30. *In fact this argument shows that ${}_2K_3^-$ minor free graphs can be exactly characterised as $(\{K_{2,n-2}^* : n \geq 0\}; 1)$ -cockades.*

Theorem 6.31. $c_2({}_2K_3^-; n) = 2n - 2$. Extremal graphs are $({}_2K_2, K_{2,n}^{**}; 1)$ -cockades, where $K_{2,n}^{**}$ denotes a book with the base edge having multiplicity 2.

Proof. As before we proceed by induction on n , the base cases $n = 2$ and $n = 3$ are trivial. Let G be a graph with n vertices and at least $2n - 2$ edges. As in Theorem 6.29 we can assume G is connected. If G has a cutvertex v , by induction each component must have $e(G[C \cup \{v\}]) = 2|C|$ and so is a cockade; thus G is also a cockade as claimed. Since the result that 3-connected graphs contain a K_4 minor also extends to multigraphs, we deduce that G has a cut-set $\{u, v\}$ of size 2. Further, we can assume every single edge is in a triangle (else we can contract it and get a contradiction).

Suppose first that some component C_i of $G \setminus \{u, v\}$ has three or more edges to $\{u, v\}$. Contracting this component to a point we have a new vertex w with two edges to u , and at least one to v up to reordering. This implies $m_G(uv) \leq 1$; further if this multiplicity is equal to 1 contracting any other component yields the forbidden minor and so actually $m_G(uv) = 0$. This means there are exactly two components C, C' of $G \setminus \{u, v\}$ and C' has exactly one edge to each of u, v as before. But now the edge from u to C' is in no triangle which gives a contradiction.

Thus every component of $G \setminus \{u, v\}$ has exactly two edges to $\{u, v\}$. To avoid both a single edge in no triangle, and existence of a cutvertex, we must have $m_G(uv) \geq 1$ and in each component there is a common neighbour w_i of u and v in C_i ; to avoid w_i being a cutvertex we have $C_i = \{w_i\}$. Thus either $G = K_{2,n-2}^*$ or $G = K_{2,n-2}^{**}$ only the latter of which is extremal. $\square$

6.5.2 The graph ${}_2K_3$

We will again use the structure of K_4 minor free graphs. While we cannot directly argue G does not contain a K_4 minor (indeed K_4 is ${}_2K_3$ minor free), either G does not contain such a minor, or it contains a K_4 minor and in particular a topological minor. We can then forbid many paths within G and hence deduce that G must be ‘close to’ K_4 .

Theorem 6.32 (Theorem 6.3(E)). $c_1({}_2K_3; n) = 2n - 3$ for $n \not\equiv 1(3)$, and $2n - 2$ for $n \equiv 2(3)$ or $n \equiv 0(3)$. All graphs with $e(G) \geq 2n - 3$ and no ${}_2K_3$ minor are $(\mathcal{F}; 1)$ -cockades where $\mathcal{F}$ consists of:

1. K_4

2. K_4 with one edge subdivided

3. Type P series-parallel graphs (i.e. 2-connected K_4 -minor-free graphs).

To be an extremal example, all but at most one graph in the cockade must be a copy of K_4 .

Proof. We proceed by induction on $|G|$; the base cases $n = 2, 3$ being trivial. As usual we can assume that G is 2-connected with at least $2n - 3$ edges. If G is K_4 minor free, to have at least $2n - 3$ edges G must be a maximal type P series-parallel graph and hence we are done, so assume that G contains a K_4 minor (and hence topological minor). Let T be such a topological minor, with branch vertices a, b, c, d and a unique ‘canonical path’ between each pair; each of these are internally vertex disjoint. We aim to show that G is equal to this topological minor, and so to attain the edge bound G is K_4 or K_4 with one edge subdivided.

Suppose that there is some path outside of $E(T)$ between some pair of vertices in T . Contracting this path, we can view this as an extra edge between some pair of vertices.

Assume first these vertices lie in the same canonical path, say the path from a to b . Let $x_0 \dots x_l$ with $x_0 = a, x_l = b$ be this path, and $x_i x_j$ the edge, where $i < j$. Then contracting $x_0 \dots x_{j-1}$ and $x_j \dots x_l$ and all other paths arbitrarily we obtain K_4 with a doubled edge as a minor; this in turn contains ${}_2K_3$ as a minor.

If they lie in different paths, say the edge is $x_i y_j$ where $j > 0$, $x_0 \dots x_l$ and $y_0 \dots y_l$ are canonical paths and $y_l \neq x_l$. Then contracting $x_1 \dots x_l$ and $y_1 \dots y_l$ gives K_4 with a doubled edge as a minor, once again a contradiction. Thus the topological K_4 is in particular induced.

If there is any vertex v outside $V(T)$, since G is 2-connected there are two disjoint paths from v to T ; taking minimal such paths they are disjoint from $E(T)$. Concatenating these paths gives a path outside $E(T)$ between two vertices of T ; we ruled this out above. So as desired G is a topological K_4 completing our proof. $\square$

Note that not all series-parallel graphs are ${}_2K_3$ -minor free, for instance the Triforce graph from Figure 6.3 of page 105 is series-parallel but has this minor. One can naturally ask the question of which series parallel graphs do not contain such a minor, and I have resolved this question in a preprint [71]. The proof is omitted from this thesis for brevity.

Theorem 6.33. $c_2({}_2K_3; n) = 3n - 4$ for each $n \geq 2$. Extremal graphs are ${}_2K_2$, or obtained from a ${}_2K_3$ minor free maximal type P series parallel graph by doubling the edges of some tree.

Proof. Suppose first that G is 2-connected, has no ${}_2K_3$ minor and maximum edge multiplicity at most 2 with at least $3n - 4$ edges. The graph of double edges is acyclic (since ${}_2C_l$ contracts to ${}_2K_3$) and so G has at most $n - 1$ double edges. Thus the underlying graph has at least $2n - 3$ edges. By Theorem 6.32 this implies the underlying graph is one of a

maximal type P series parallel graph, or some subdivision of K_4 . For a K_4 subdivision, we still need at least one double edge, and hence a K_4 with a doubled edge is a minor of G , contradicting that there is no ${}_2K_3$ minor.

So G is a type P series parallel graph with exactly $n - 1$ double edges; these must span a tree. This completes our proof, since any ${}_2K_3$ minor free graph is a cockade of 2-connected examples, and if we are not 2-connected we must have strictly fewer than $3n - 4$ edges (or be K_2). $\square$

As in the previous theorem, not all graphs obtained from a type P series parallel graph not containing ${}_2K_3$ as a minor by doubling the edges of a spanning tree remain ${}_2K_3$ minor free. To show that this can be done sometimes, and hence our extremal function is correct, consider any graph where a single vertex v is joined to an $n - 1$ vertex tree T entirely by double edges.

6.6 Further work

There seem to be a couple of natural directions to proceed from here. The first is to ask about more small cases; for instance, what can we say about ${}_2K_4$?

One family of ${}_2K_4$ -minor-free graphs are $(K_6; 1)$ -cockades. These have $n \equiv 1(5)$ vertices, and $3n - 3$ edges.

Conjecture 6.34. If G is a graph with n vertices and at least $3n - 3$ edges, then either G is a $(K_6; 1)$ -cockade, or G contains a ${}_2K_4$ minor.

The arguments of above applied to a natural induction hypotheses were able to rule out many cases, though I have been unable to show this result holds. The best upper bound I have comes from a result of Jakobsen [32].

Theorem 6.35 ([32, Theorem 1]). *Let G be a graph with $n \geq 7$ vertices and $e \geq 4n - 9$ edges. Then either G is a $(K_6; 3)$ -cockade, or G contains a graph obtained from K_7 by removing two edges as a minor.*

We remark that doubling any edge in a K_6 yields a ${}_2K_4$ minor, and likewise that any graph with 7 vertices and 19 edges does. So we deduce that any graph with at least $4n - 9$ edges contains a ${}_2K_4$ minor, unless $G = K_6$.

The above conjecture that $(K_6; 1)$ -cockades are uniquely optimal arise from time experimenting with small graphs, and attempts at proving such a conjecture.

Conjecture 6.34 would also prove a near-tight result for $c_2({}_2K_4; n)$. The graph of double edges in a ${}_2K_4$ minor free must span a K_4 minor free graph, hence has at most $2n - 3$ edges. Likewise, viewing it as a simple graph we must be ${}_2K_4$ minor free. This would imply $c_2({}_2K_4; n) \leq 5n - 6$ for large n . For a lower bound, the graph obtained from ${}_2K_3$ by sequentially adding vertices joined by double edges to 2 of the vertices, and a single

edge to the remaining shows $c_2({}_2K_4; n) \geq 5n - 9$. Without assuming the conjecture, this proof method shows $c_2({}_2K_4; n) \leq 6n - 13$.

Another direction is to improve asymptotic results. The results of Chapter 5 imply the maximum value of $e(G)/|G|$ in ${}_rK_t$ minor free graphs is $(\frac{1}{2} + o(1))t\sqrt{r}$ for large r and t fixed. The error terms here are poor, and it would be interesting to improve them for sufficiently small t (or even $t = 2$).

Chapter 7

Entropy Compression

7.1 Introduction

Entropy compression describes a broad range of related methods, but originally (and most accurately) describes an argument of Moser [48], used to prove a constructive version of the Lovász Local Lemma. A randomised algorithm is described, and the proof the algorithm terminates uses incompressibility of random bits. We now describe this argument more carefully.

It is impossible to compress n random bits into some $m < n$ random bits in a lossless fashion. Indeed, for any scheme $f : \{0, 1\}^n \rightarrow \{0, 1\}^m$ and $g : \{0, 1\}^m \rightarrow \{0, 1\}^n$, the probability over uniformly random $x \in \{0, 1\}^n$ that $g(f(x)) = x$ is at most 2^{m-n} by considering the number of possible outputs. This remains true even if f and g are allowed to be randomised functions (but this randomness cannot depend on x).

Suppose we have a randomised algorithm which takes steps $A_0, A_1, \dots$. The algorithm might continue indefinitely, or it can terminate at some desirable structure. One way of describing this algorithm up to step M is to describe the randomness used to reach steps $A_0, \dots, A_M$ in turn; we assume our algorithm uses this randomness in an injective fashion.

Suppose we have some other way of reconstructing this algorithm up to step M which uses strictly fewer bits (for large M). Then since we can recover the random bits from the algorithm, this gives a way of compressing random noise! Since that cannot be done, we deduce the algorithm terminates with positive probability in at most M steps (and so the desirable structure, in particular, exists). This method will be explicated in our proof of Theorem 7.2.

Other methods, such as the Rosenfeld / Wanless-Wood framework and the cluster-expansion local lemma, have since been included under the entropy compression umbrella since they give equivalent bounds.

To describe our main result, we require some preliminaries. Let G be a graph equipped

with a vertex partition into n ‘blocks’ B_i , each of size t . Recall a *transversal* of G is a choice of a vertex v_i from each B_i ; such a transversal is independent if no two v_i are adjacent. In this chapter we will investigate the number of edges incident to each block in a graph with no independent transversal. Wanless and Wood showed the following lower bound.

Theorem 7.1 ([73, Theorem 9]). *Let G be a graph with vertex partition (B_i) into blocks of size t . Then if each block is incident to at most $t^2/4$ edges of G , G has an independent transversal. Equivalently, if the average degree of vertices from each block is at most $t/4$, then G has an independent transversal.*

At a workshop, I investigated this result jointly with Carla Groenland, Tomáš Kaiser and Oscar Treffers. The original question posed was whether the average degree bound of $t/4$ could be increased, perhaps as far as $t/2$ (this would then match a result of Haxell showing that if each vertex has degree at most $t/2$ then G contains an independent transversal). Surprisingly, we showed the opposite — Theorem 7.1 is asymptotically tight. This was the first non-trivial example of a problem where the Rosenfeld / Wanless-Wood framework gives a tight result, and is the focus of a joint preprint [21].

This chapter has two aims. The first is to prove a variant, Theorem 7.2, of Theorem 7.1 which, under a very slightly stronger condition, provides an efficient algorithm to produce such a transversal. Our proof uses an encoding argument, and has an interesting method of fractional bits within the proof. To work with bits, for the remainder of this chapter $\log$ will denote the logarithm to base 2. We state our theorem now in a more general, hypergraph, form.

Theorem 7.2. *Let $r \geq 2$ and t be integers. Let G be an r -uniform hypergraph, equipped with a vertex partition $(B_i)_{i=1}^n$ into blocks of size t such that fewer than $\frac{(r-1)^{r-1}}{r^r} t^r$ edges of G meet each B_i . Then G has an independent transversal, and further there exists a randomised algorithm for finding such a transversal in $O(nr^2 t^{2r})$ expected time.*

The second aim is to give a construction which shows this bound on ‘block average degree’ is essentially tight; this forms the basis of Section 7.3 and is formalised as Theorem 1.30. We will finish the chapter by posing many modifications of this problem, combining the average degree bound with other conditions, and show how our construction can be modified to give good bounds for these problems.

Theorem 1.30. *Let $\epsilon > 0$. Then for all t sufficiently large, there exists a graph G with vertex partition $V(G) = \cup B_i$ such that each B_i has size t and the average degree of vertices in B_i is at most $(\frac{1}{4} + \epsilon)t$, but G has no independent transversal.*

We also prove a hypergraph analogue of this theorem in Section 7.4. The key ideas are broadly similar, however additional complications arise and so we prove our results first for graphs to give motivation and highlight the most important parts of the proof.

7.2 A direct entropy compression proof

The aim of this section is to prove our lower bound, Theorem 7.2. We will prove this theorem by constructing an algorithm; it will take a series of randomised steps, and terminate only when it has constructed such an independent transversal. We will use entropy compression to prove the algorithm must terminate in the claimed number of steps with positive probability. We start this section with a proof for the $r = 2$ case, which gives an idea of the algorithm while deferring discussion of complexities arising in the hypergraph case. After proving the graph case, we will give an example of running our algorithm.

7.2.1 The graph case

Proof of Theorem 7.2 for $r = 2$. Since the proof is somewhat involved, we will break it up into stages. Let G be a 2-uniform hypergraph (i.e. a graph) as in the statement of the theorem. Let $e < t^2/4$ be the maximum number of edges incident to any block B_i ; if $e = 0$ the result is trivial so assume $e \geq 1$. Since edges within a block (for r -graphs, not between r distinct blocks) cannot appear in a transversal, we assume that all edges of G are between blocks, and call such edges *stretched*. Fix now some ordering of each B_i as well as of the edges incident to each B_i .

There are three stages of our proof. We will first describe our randomised algorithm. This algorithm might not terminate, but if it does terminate, it will have constructed an independent transversal. The next stage of our proof is to describe a method of ‘logging’ steps of the proof, and show how this log file can be used to recreate the randomised steps the algorithm took. The final step is to combine these into a proof by ‘entropy compression’ that after some (known) number of steps, the algorithm has terminated with positive probability.

Stage 1. The algorithm

Our algorithm will proceed in a sequence of steps. It starts, at step 0, by choosing a transversal Σ_0 which consists of a uniformly random vertex from each block; note this need not be independent. The transversal Σ_0 can be described with $n \log t$ bits of randomness. From here, we will proceed according to a structure we call the *helper tree*.

The helper tree T is a rooted, ordered, vertex labelled tree. This tree will be build up as the algorithm proceeds. The root will be labelled specially as R , and all other vertices of the tree will be labelled by an index $1, \dots, n$ corresponding to a block. This tree will be modified as the algorithm proceeds, and at each step we think of our algorithm as ‘at’ some particular vertex of T . While we think of T as a single tree which we build up in stages, we will use the notation T_i for the helper tree built up as of step i .

At step zero, our helper tree $T = T_0$ will consist of our root vertex (labelled R) together with n descendants, labelled $1, \dots, n$ in that order. At step zero, the algorithm is ‘at’ the

descendant of the root labelled 1.

We can now describe our algorithm. Suppose we are given a transversal $\Sigma_{m-1} = \{v_1, \dots, v_n\}$ after $m-1$ steps, together with a helper tree T_{m-1} , and the vertex x of the helper tree the algorithm is ‘at’ after $m-1$ steps. It will suffice to describe the transversal after m steps, our new helper tree T_m , and a vertex x of T_m . We for now discount the case $x = R$; our algorithm will ensure this never happens.

Suppose that x is labelled by i . We have two cases depending on whether or not there is an edge of G within Σ_{m-1} incident to v_i . If there is no such edge, we take $T_m = T_{m-1}$, $\Sigma_m = \Sigma_{m-1}$. It remains only to choose our new vertex. If the ancestor of x is a non-root vertex y , we say the algorithm is at y after this step. If instead the ancestor is the root, we move to the highest unvisited descendant of the root. If we have visited all descendants of the root, the algorithm terminates. We will show this only happens when Σ_m is an independent transversal at the end of this proof stage.

We thus suppose there is an edge of G within Σ_{m-1} incident to v_i . Suppose the first such edge (according to our ordering of edges incident to B_i) is $v_i v_j$. In this case, we create a new last descendant y of x in the helper tree and label it j . This forms our tree T_m . Our algorithm will be at this new vertex y after m steps. To construct our new transversal Σ_m , we take Σ_{m-1} , remove v_i and v_j , and add a uniformly random vertex from each of B_i, B_j . We call a step of this form an *edge-creation step*, and note that we have used $2 \log t$ bits of randomness here.

It remains only to show that the algorithm can only terminate when the transversal Σ_m is independent. In order to prove this, we require the following lemma.

Lemma 7.3. *Suppose that $t_1 < t_2$ are such that the algorithm is at the same vertex x of the helper tree, with label i . Then for every edge of G contained in Σ_{t_2} , either this edge is contained in Σ_{t_1} also, or this edge is incident to block i . In particular, if t_2 is the last time the algorithm is at x , then $\Sigma_{t_2} \subsetneq \Sigma_{t_1}$.*

This implies our result by the following brief argument. The last time we visit a descendant labelled i of the root, there are no edges of the transversal incident to $B_1, \dots, B_i$. If there are no edges incident to $B_1, \dots, B_n$ this exactly means the transversal is independent.

Proof. After visiting vertex x in step t_1 , we must have an edge-creation step; recall we only move ‘right’ through our tree by creating new nodes of the helper tree. If our new node is labelled j , this edge creation step can only create new edges inside the transversal incident to v_i or v_j . We construct an auxiliary walk (initially $B_i B_j$), and all new bad edges (edges that appear in the present transversal) are incident to a block within this walk. At each edge-creation step, we append the corresponding edge to the end of our walk. If we ever have a step in which we backtrack through the helper tree, the block at the end of our walk is incident to no bad edges; in particular no new bad edges. So we delete the last edge of

our walk and continue. We only return to x when our walk has completely backtracked (indeed, the length of the walk represents how many ‘descendants deeper than x ’ we are in the helper tree), and this means the only new bad edges are as claimed. $\square$

We finish this stage with a remark. Viewing the root vertex as the ‘left’ of the helper tree, we move right exactly in edge-creation steps, and left in all other steps (ignoring the root for now). In particular, whenever we move right it is to a newly created vertex of the helper tree.

Stage 2. The log file

In order to show our algorithm terminates with positive probability, we require an efficient encoding of its steps. This is the purpose of our log file. For each step of our algorithm, we will append some bits to the end of the log file. These bits will be ‘charged to’ a particular edge of the spanning tree to make counting easy. Our log file after zero steps is empty. We will be using a prefix-free encoding method; this ensures we can uniquely decode the log file.

If step m of the algorithm is not an edge creation step (i.e. $\Sigma_m = \Sigma_{m-1}$ and the algorithm is at x), we log a 1 and charge this bit to the edge from x to its ancestor. Since we only move right through the tree by creating new vertices, each edge of T is charged with at most one bit in this fashion.

If step m is an edge creation step, suppose we are at a vertex x , and move to a new vertex y . We log a 0 and charge it to this edge; since each edge is created exactly once, it is charged with one bit in this fashion. We also need to encode why we made this change. To do this, let i be the label of x , and suppose that $v_i v_j$ is the (first) edge of G within Σ_{m-1} incident to v_i . We log the index of $v_i v_j$ among the edges from v_i ; this uses at most $\log e$ bits (and so each edge has now been charged with at most $2 + \log e = \log(4e)$ bits). This completes our description of the log file.

Rather than building the helper tree T , we will build the *augmented helper tree* T' . The augmented helper tree T' is a copy of T , with each non-root incident edge labelled by an edge of G . In each edge-creation step, we label the new edge of T' by the edge $v_i v_j$ which was re-randomised. This additional information does not help in running the algorithm, and so this augmented tree is a purely technical tool. The following lemma will be very useful for us, and is essentially proven by reading the log file in a forwards direction.

Lemma 7.4. *Given the log file after M steps of the algorithm, the augmented helper tree T'_M can be reconstructed, together with which vertex of T'_M the algorithm is at after M steps.*

Proof. We will proceed by induction on M . The base case $M = 0$ is trivial, since our helper tree (and hence augmented helper tree) consists a single root vertex with n descendants,

and the algorithm is at the descendant labelled 1. We thus assume $M \geq 1$, and suppose we have constructed the augmented helper tree T'_{M-1} from the log file for the first $M - 1$ steps, with the algorithm at a vertex x_{M-1} of T'_{M-1} .

Suppose that the next entry in the log file is a 1. Then the algorithm did not make an edge creation step, and the augmented helper tree $T'_M = T'_{M-1}$. Since the algorithm was at vertex x_{M-1} after $M - 1$ steps, it must be at the ancestor x_M of x_{M-1} after M steps (unless this is the root, in which case we are at the next unvisited descendant).

So, assume the next entry in the log file starts with zero, and so the algorithm made an edge-creation step. Suppose the label of x_{M-1} is i . Our log file tells us which of the at most e edges incident to B_i is used; suppose it is $v_i v_j$. We create a new last descendant $y = x_M$, labelled j , of x_{M-1} , and label the edge between x_{M-1} and x_M by $v_i v_j$. The augmented helper tree T'_M and vertex x_M are as claimed. $\square$

Recall from the introduction that we need to reconstruct the algorithm, not just some internal structure we used. To do this, we will also use the transversal after N steps, and show that given this final transversal and the augmented helper tree, we can reconstruct all earlier transversals! This will be done by reading our log file backwards.

Lemma 7.5. *Given the augmented helper tree T'_M , the sequence of vertices $x_0, \dots, x_M$ of T'_M that the algorithm is at and the transversal Σ_M , it is possible to construct all transversals $\Sigma_0, \dots, \Sigma_M$.*

Proof. It suffices to prove that we can reconstruct Σ_{M-1}, x_{M-1} and T_{M-1} from this information by recursion.

Suppose first that x_M has (at least one) descendant. Let x_{M-1} be the last such descendant. This must be the vertex the algorithm was at the step before; we can only have arrived at x_M either from an edge-creation step or from its last descendant. In this case, also $\Sigma_{M-1} = \Sigma_M$, and $T'_{M-1} = T'_M$. The same argument works where x_M is a descendant of the root with no descendants; we cannot arrive there from an edge-creation step, so we must have $\Sigma_{M-1} = \Sigma_M$ and x_{M-1} the descendant of the root directly above x_M .

We thus assume x_M is not a descendant of the root, and has no descendants. This means that step M was an edge-creation step. Let x_{M-1} be the ancestor of x_M ; this must be the vertex the algorithm was at one step earlier. The structure of the augmented helper tree tells us which edge was re-randomised in this step; exactly the label of $x_{M-1}x_M$! Let x_{M-1} be labelled i , x_M labelled j and this edge of G f . The transversal Σ_{M-1} is obtained by deleting the vertices v_i, v_j in $B_i \cap \Sigma_M, B_j \cap \Sigma_M$ and adding the edges of f . T'_{M-1} is obtained from T'_M by deleting the leaf x_M . This completes our proof. $\square$

Stage 3. Termination by entropy compression

For simplicity, we restrict attention to edge-creation steps here; each non edge-creation step is backtracking a previously used step, so if there are M' edge-creation steps, there are at most $n + 2M'$ steps overall.

Suppose we have run the algorithm for M steps, $M' \leq M$ of which are edge-creation steps. Our algorithm has used $n \log t + 2M' \log t$ bits of randomness to run. However, we can recover all of these bits from the log file and the final transversal. How many bits does this use?

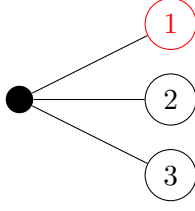
Encoding the final transversal Σ_M uses $n \log t$ bits. Each edge of the helper tree incident to the root is charged with at most 1 bit for a total of n . Each edge which is not incident to the root is, as mentioned earlier, charged with at most $\log(4e)$ bits. Since there are M' edge-creation steps, there are M' non root-incident edges, and so at most $n(1 + \log t) + M' \log(4e)$ bits used in this encoding.

Any scheme for encoding k random bits in at most $k - 1$ random bits must fail to successfully encode with probability at least $\frac{1}{2}$, by a counting argument. Suppose that we take some number k of random bits, and run our algorithm until these bits run out. We then use the log file as our encoding of these random bits. If the algorithm terminates, we log some random (valid) log file instead. If we choose M' such that $n(1 + \log t) + M' \log(4e) + 1 \leq n \log t + 2M' \log t$, we deduce the algorithm must terminate in at most M' edge-creation steps with probability at least $\frac{1}{2}$. Running it for this many steps over and over gives that the algorithm (with high probability) terminates in $O(M')$ edge-creation steps.

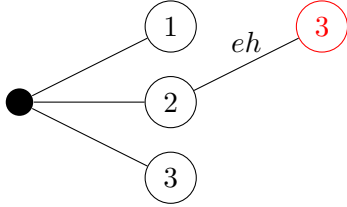
The required inequality holds provided $M' \geq 1 + n/(\log(t^2/4e))$. Since $e < t^2/4$, we have $t^2/4 - e > \frac{1}{4}$ and so this is, as claimed, $O(nt^2)$ steps. Each step of our algorithm required $O(t^2)$ time and so with high probability the algorithm constructs an independent transversal in time $O(nt^4)$. $\square$

We remark that if $e < (\frac{1}{4} - \epsilon)t^2$ this gives an $O(nt^2)$ time algorithm for finding an independent transversal. In the generic case, we have an $O(nt^4)$ time algorithm; both of these are better than the $O(c^n)$ time algorithm implied by the argument of Wanless and Wood although they only apply in a slightly more restrictive setting.

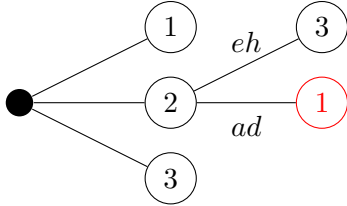
Example. Let G be the graph with vertex partition $B_1 = \{a, b, c\}, B_2 = \{d, e, f\}, B_3 = \{g, h, i\}$ and edge set $\{ad, eh, ia\}$. We will show how our algorithm works, and build up a log file simultaneously. The algorithm starts by choosing a uniformly random transversal, let us assume we chose $\Sigma_0 = \{a, e, h\}$. Our initial helper tree T_0 will be as below, with the algorithm occupying the red node.



Our algorithm started at R and immediately moved to the highest descendant, labelled 1 above. There is no edge of G contained in the current transversal incident to $a \in B_1$, we thus move the algorithm back to the root R and log a 1; again proceeding to 2. This time there is an edge contained in the transversal, namely eh , incident to $e \in B_2$. We thus log a zero, followed by an encoding of the edge eh . We create a new last descendant of the current vertex, which is labelled by 3, and have the algorithm occupy it. We then re-randomise the vertices from B_2 and B_3 to form a new transversal Σ_2 , and suppose we choose $\{a, d, g\}$.



There are no edges in the current transversal incident to g , so we log a one and proceed back to the ancestor. This time there is an edge incident to d , namely da , in the current transversal. We therefore log a 0 followed by an encoding of da , create a new last descendant labelled 1, move the algorithm there and re-randomise B_1 and B_2 to obtain a new transversal, say $\Sigma_4 = \{b, e, g\}$ is chosen. The transversal is now independent, so after further logging of ones the algorithm will terminate.



As in Stage 2, reading the log file in a forwards direction we can recover the above augmented helper tree. Also, given this labelled tree and the final transversal, we can recover all randomised steps the algorithm took.

7.2.2 A generalisation to r -graphs

We would like to use the above method for r graphs also. While we can use essentially the same algorithm (with suitable modifications to the helper structure), the logging method causes problems. Each edge now involves r parts, and we need to log both creation of an

edge of the helper structure, and then a return from each of the other $r - 1$ blocks involved. If we use the smallest possible amount of 1 bit to log each of these, we can only hope to prove a result of the form ‘block average degree at most $t^{r-1}2^{-r}$ implies an independent transversal exists’, which is too weak.

To beat this we will need to use *less than 1 bit* sometimes, in order to use an additional at most $r \log r - (r - 1) \log(r - 1)$ bits on each edge of the helper tree. This sounds initially absurd, but in fact we can make this approach rigorous.

For integers $a_1, \dots, a_l$, we can use a sequence of codewords of lengths $a_1, \dots, a_l$ respectively in a decodable code exactly when $\sum_{i=1}^l 2^{-a_i} \leq 1$; this is Kraft’s inequality, see [3, Theorems 2.3.1 and 2.4.1]. If we can in some way extend this result to non-integers, we could encode a one using $\log r$ bits and a zero using $(\log r - \log(r - 1))$ bits, and so have a sufficiently efficient logging method.

Recall that what really matters is that the log file uses fewer bits than the amount of randomness used in our algorithm, not the absolute length of the log file. With this in mind, we introduce some auxiliary randomness which we also encode into the log file; this will also be recoverable when we decode the log. This may lengthen our log file, but it also increases the amount of recovered randomness. We are able to make a gain here by combining information about the next state of the algorithm with some auxiliary randomness.

We will use this method to encode our ones and zeros for traversing through the helper tree. Our codewords will all have fixed length s , but when we encode a 1 we will also encode ‘about $n - \log r$ ’ auxiliary random bits with it, and likewise ‘about $s - (\log r - \log(r - 1))$ ’ auxiliary random bits with each zero. Assuming we can do this, we have only used a *marginal* $\log r$ bits to encode a one, and $\log r - \log(r - 1)$ to encode a zero (i.e. our log file has only increased in length by this much more than the random bits encoded). These marginal bits are really the quantity we are counting. We now show this can be done.

Let x be a positive real number and $\epsilon > 0$. Pick some integers n, m such that $s - \epsilon < x + \log m \leq s$. Such an integer exists because the values of $\log(m + 1) - \log m$ converge to zero, and $\log m$ is unbounded.

Pick some m sequences of length s arbitrarily to form a set A , and let the remaining $2^s - m \geq (2^x - 1)m$ sequences of length n form $A^c = B$. To encode a one, we take $\log |A|$ auxiliary random bits, and use these to pick a random element of A . To encode a zero, we do the same with $\log |B|$ random bits and choose an element of B . We note that given this codeword, we can recover both whether a zero or one was encoded (depending on whether or not the codeword is in A) and all the auxiliary random bits used. The number of marginal bits used is at most $x + \epsilon$ and $-\log(1 - 2^{-x}) + \epsilon$ for a one and zero respectively. Taking $x = \log r$ gives our claimed encoding. We remark this encoding is essentially tight in terms of Kraft’s inequality.

In the following proof, we will use this method to encode our bits to some incredibly small error term. Throughout we will use the term ‘encode in k marginal bits’ to describe an encoding according to the above scheme. We also note that these technicalities only arise in the log file, and so are only used in analysis and not required to run the algorithm — they thus do not affect the runtime at all.

Proof of Theorem 7.2. As before, our proof will be in three stages. Let G be an r -graph with vertex partition $(B_i)_{i=1}^n$ as in the statement, and let $e < \frac{(r-1)^{r-1}}{r^r} t^r$ be the maximum number of edges incident to a block; as before assume all edges are stretched. Let η be an (arbitrary, fixed) constant satisfying $0 < 3r\eta < \log(\frac{(r-1)^{r-1}}{r^r} t^r) - \log(e)$; we will consider the $\eta \rightarrow 0$ limit in Stage 3.

Stage 1. The algorithm

Our algorithm is similar to the $r = 2$ case, but the required *helper structure* is slightly different. We start with, at step zero, a uniformly random transversal Σ_0 . Our algorithm will now take a series of randomised steps, with form described by the helper structure.

Our helper structure T is analogous to a rooted ordered vertex-labelled tree, except that edges not incident to the root are r edges, consisting of a single ancestor, and $r - 1$ ordered descendants. Our initial structure T_0 will again be a root node labelled R , and n descendants $1, \dots, n$ each joined to the root by an edge. The algorithm will be at a non-root vertex of the helper tree after each step.

Suppose that after $M - 1$ steps the helper structure is T_{M-1} , and the algorithm is at a non-root vertex x_{M-1} labelled i . Suppose our transversal is $\Sigma_{M-1} = \{v_1, \dots, v_n\}$ with $v_i \in B_i$. If there is no edge of G contained in Σ_{M-1} and incident to v_i , we set $T_M = T_{M-1}$, $\Sigma_M = \Sigma_{M-1}$. To deduce the next vertex x_M we are at, consider the edge from x_{M-1} to its ancestor y . If this ancestor is the root, we set x_M the next unvisited descendant of R , or terminate if we have visited them all. If this ancestor is not the root, consider the (hyper)edge involving y and x_{M-1} . If x_{M-1} is the last vertex in this edge, we set $x_M = y$. Otherwise, suppose the descendants of y in this edge are, in order, $z_1, \dots, z_{r-1}$ with $x_{M-1} = z_l$. Then we set $x_M = z_{l+1}$.

So suppose instead there is an edge of G in Σ_{M-1} incident to B_i , and let the first such edge be $f = \{v_{i_1}, \dots, v_{i_r}\}$. In this case, we make an edge-creation step: suppose that $i = i_l$ for some $1 \leq l \leq r$. Then we create a new edge of T_M incident to x_{M-1} , with $r - 1$ new descendants $z_1, \dots, z_{r-1}$ labelled in order $i_1, i_2, \dots, i_{l-1}, i_{l+1}, \dots, i_{r-1}, i_r$. We move the algorithm to $x_M = z_1$. We construct Σ_M by replacing the vertices of f with new, uniformly random vertices from $B_{i_1}, \dots, B_{i_r}$. This completes our description.

The algorithm can only terminate at an independent transversal; the proof of this is similar to in the $r = 2$ case and so we omit the details. The natural generalisation of Lemma 7.4 holds with essentially the same proof; the only variant is rather than an

auxiliary path, we create an auxiliary path of hyperedges.

Stage 2. The log file

Our logging method is similar to the $r = 2$ case, though with additional complexity due to fractional bits.. This time, we log a one using $\log r - \log(r - 1) + \eta$ marginal bits when we have a non-edge-creation step, and log a zero using $\log r + \eta$ marginal bits followed by an encoding of the edge we re-randomise.

As before, when we perform an edge creation step, the $\log r + \log e + \eta$ marginal bits used in the log file are charged to the new edge of T_M relative to T_{M-1} . When we perform a non-edge-creation step, supposing our previous vertex was x_{M-1} and the new vertex x_M , we charge the $(\log r - \log(r - 1))$ bits to the (unique) edge involving both vertices. In this fashion, each edge of T is charged with at most $r \log r - (r - 1) \log(r - 1) + \log(e) + r\eta = \log(\frac{r^r}{(r-1)^{r-1}}e) + r\eta$ bits.

Defining the augmented helper tree analogously to the $r = 2$ case, this log file suffices to reconstruct the augmented helper tree by reading the log file forwards. We can also use the augmented helper tree T'_M , the sequence $x_1, \dots, x_M$ of vertices occupied and the transversals $\Sigma_M, \dots, \Sigma_0$ at each step; we have thus fully reconstructed the algorithm.

Stage 3. Entropy compression

If there are M' edge creation steps, there are at most $rM' + n$ steps overall so as before we restrict attention to these. Running the algorithm for M' edge creation steps uses $n \log t + M'r \log t$ bits of randomness. The log file up to M' edge creation steps, which can reconstruct all of this randomness, used at most $n + n \log t + M'(\log(\frac{r^r}{(r-1)^{r-1}}e) + r\eta)$ marginal bits. If this is at least one fewer bit (which holds when $M' \geq 1 + \frac{n+1}{r \log t - (\log(\frac{r^r}{(r-1)^{r-1}}e) + r\eta)} > 0$), the algorithm terminates in M' edge creation steps with probability at least $\frac{1}{2}$. The result follows by taking a limit $\eta \rightarrow 0$, since each step can be performed in $O(t^r)$ time (this is necessary to check if each edge is in the transversal). Taking this limit is valid, since η does not affect the steps we take! It only affects the log file, but for an implementation of the algorithm we do not need to make this log. $\square$

7.3 The graph construction

The aim of this section is to prove Theorem 1.30 (restated below), giving our matching upper bound for the graph case. Our motivation for the construction came from investigating constructions of graphs with low maximum degree not containing independent transversals. These usually contained large complete bipartite graphs between pairs of blocks. Is this a necessary property of graphs with no independent transversals? Our construction answers this question negatively; it is a union of stars.

Theorem 1.30. *Let $\epsilon > 0$. Then for all t sufficiently large, there exists a graph G with vertex partition $V(G) = \cup B_i$ such that each B_i has size t and the average degree of vertices in B_i is at most $(\frac{1}{4} + \epsilon)t$, but G has no independent transversal.*

Proof. Reducing ϵ if necessary, assume $\epsilon < 1$. We will build up our graph by constructing blocks (each containing t vertices) of different grades. A block of grade i will have edges only to blocks of grades $i - 1$ and blocks of grade $i + 1$, further different grade i blocks have edges to disjoint collections of grade $i - 1$ blocks (we say the grade $i - 1$ blocks are *private* to the grade i block). Let N be a large integer yet to be determined. We start with N blocks of grade zero, and let $(\frac{1}{4} + \frac{\epsilon}{2})t < n_1 < (\frac{1}{4} + \frac{3\epsilon}{4})t$ be an integer; if $t > 4/\epsilon$ such an n_1 must exist since the bounds differ by more than 1.

Let $m_1 = N/n_1$. We now construct m_1 blocks of grade 1, each containing n_1 heavy vertices which are each joined to all vertices from a grade zero block; these blocks are distinct for different vertices. Thus each of the N grade 0 blocks is joined to exactly one of the $m_1 n_1$ heavy vertices in the grade 1 blocks.

The remaining $t - n_1$ vertices in each grade 1 block will be called light. This construction has the following nice property. Suppose we pick some heavy vertex v as part of a transversal of G . Considering the block v is completely joined to, this means the transversal is not independent! So if G has an independent transversal, we must use a light vertex in each grade 1 block. We now iterate this phenomenon.

To build a grade 2 block, let $n_2 < t$ be such that $(\frac{1}{4} + \frac{\epsilon}{2})t^2 < n_2(t - n_1) < (\frac{1}{4} + \frac{3\epsilon}{4})t^2$. We will have $m_2 = m_1/n_2$ grade 2 blocks, each containing n_2 heavy vertices, each completely joined to the $t - n_1$ light vertices of a distinct grade 1 block, and $t - n_2$ light vertices. If we have an independent transversal in G using a heavy vertex in a grade 2 block, this means we must use a heavy vertex in the grade 1 block it is adjacent to; this contradicts the above. Such an n_2 exists unless $(\frac{1}{4} + \frac{\epsilon}{2})t^2 \geq t(t - n_1)$. Continue in this fashion until at some step no such n_J exists (and we later show this must eventually occur).

Since t is large, there is an integer m satisfying $(\frac{1}{4} + \frac{\epsilon}{2})t^2 < m(t - n_{J-1}) < (\frac{1}{4} + \frac{3\epsilon}{4})t^2$, and so it must be the case $(\frac{1}{4} + \frac{\epsilon}{2})t^2 > t(t - n_{J-1})$. At this step we instead take $n_J = t$ and construct exactly one grade J block consisting entirely of heavy vertices, and hence take $N = n_1 \dots n_{J-1}$. By construction, each block is incident to $n_j(t - n_{j-1}) + (t - n_j) < (\frac{1}{4} + \epsilon)t^2$ edges and so has average degree at most $(\frac{1}{4} + \epsilon)t$.

It remains to show there is some stage where no n_J exists, so the construction algorithm terminates and is well defined. By construction, $\frac{n_j}{t}(1 - \frac{n_{j-1}}{t}) > (\frac{1}{4} + \frac{\epsilon}{2})$. Further, $x(1 - x)$ has a unique maximum at the point $x = \frac{1}{2}$ attaining $\frac{1}{4}$. Thus $\frac{n_j}{t(1+2\epsilon)} > \frac{n_{j-1}}{t}$ and so $n_j > (1 + 2\epsilon)n_{j-1}$. This implies our construction does indeed terminate, and further must do so in at most $\log_{(1+2\epsilon)} 4$ grades.

Finally, note that G contains no independent transversal. This follows by an inductive application of earlier remarks: in an independent transversal, we must use only light vertices

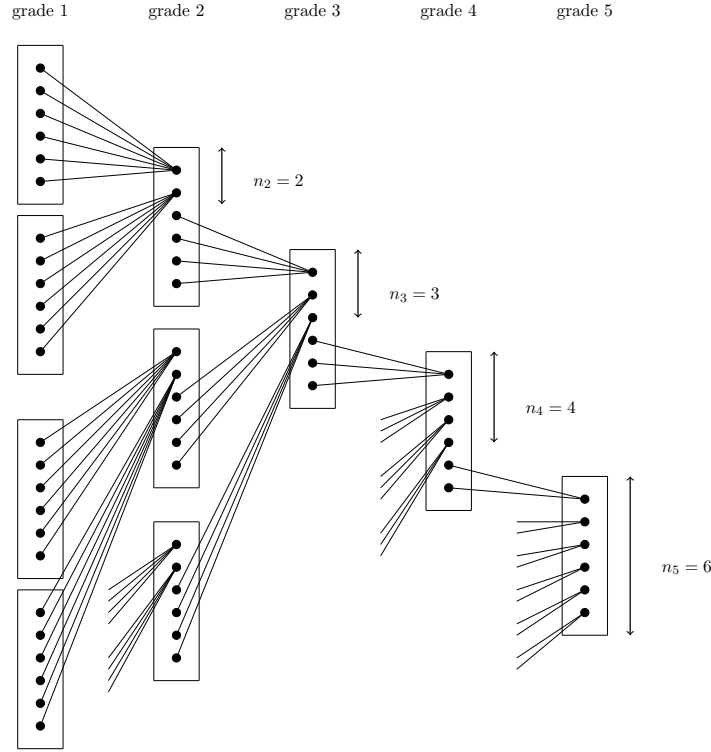


Figure 7.1: A diagram of part of the construction. The rectangles represent different blocks. Each ‘heavy’ vertex in a grade 2 block is joined to the entirety of a grade 1 block, and hence cannot appear in an independent transversal. Our construction iterates this phenomenon: the heavy grade 3 vertices are joined to a (smaller) forced set, and so cannot appear in a transversal.

in each block of grade 1. This means we must use only light vertices in each block of grade 2, and so on for each grade up to J . However, the grade J block has no light vertices, so this is a contradiction. We include a sketch of (part of) our construction in Figure 7.1.

□

7.4 Extension to hypergraphs

In this section, we will extend the construction of Section 7.3 to the hypergraph case. While the ideas are similar, the analysis becomes more technical.

Theorem 7.6. *Let $\epsilon > 0$, and $r \geq 2$ an integer. Then for t sufficiently large, there exists an r uniform hypergraph G , with vertex partition $V(G) = \bigcup_i B_i$ such that $|B_i| = t$ and each B_i meets at most $(\frac{(r-1)^{r-1}}{r^r} + \epsilon)t^r$ edges but G has no independent transversal.*

Proof sketch. Our construction will as before consist of blocks of different grades. Each hyperedge will consist of a single vertex from a grade j block, together with $r - 1$ vertices from distinct grade $j - 1$ blocks. These blocks will be private to the grade j block. Let N

be a yet undetermined integer, and begin as before with N grade 0 blocks. Let n_1 be such that $(\frac{(r-1)^{r-1}}{r^r} + \epsilon/2)t < n_1 < (\frac{(r-1)^{r-1}}{r^r} + \frac{3\epsilon}{4})t$; taking t large enough such an integer exists. We construct N/n_1 grade 1 blocks, each consisting of n_1 heavy vertices. Each heavy vertex is joined to all $(r-1)$ -tuples from $r-1$ private grade 0 blocks (and so each heavy vertex in a grade 1 block has degree t^{r-1}). The remaining $t - n_1$ vertices will be light, and joined to no vertices of grade 0.

Continue in this fashion, picking n_J such that $(\frac{(r-1)^{r-1}}{r^r} + \epsilon/2)t^r \leq n_J(t - n_{j-1})^{r-1} \leq (\frac{(r-1)^{r-1}}{r^r} + \frac{3\epsilon}{4})t^r$. By the same argument as before, using that $x(1-x)^{r-1}$ has a maximum value of $\frac{(r-1)^{r-1}}{r^r}$ over $0 < x < 1$, the n_J form an increasing sequence and so eventually we must require $n_J \geq t$. At this step construct a single grade J block with t heavy vertices, and hence define $N = (r-1)^{J-1}n_1 \dots n_{J-1}$. It is easy to see any independent transversal must consist entirely of light vertices; but the grade J block has no light vertices, so this graph has no independent transversals. Taking t large enough, the number of edges incident to each block is sufficiently small. $\square$

7.5 Maximum degree conditions

Our results above consider only the average degree of vertices in a block. However, we could combine this with other restrictions. For instance, in our constructions heavy vertices in a block of grade 1 had ‘maximum possible degree’ t^{r-1} to some $(r-1)$ -tuple of blocks. Can we also bound this *maximum local degree*?

Definition 7.7. Let G be an r -graph equipped with a partition (B_i) into blocks. The *local degree* of a vertex v to blocks $B_{i_1}, \dots, B_{i_{r-1}}$ is the number of edges of G involving v and exactly one other vertex from each of $B_{i_1}, \dots, B_{i_{r-1}}$. The *maximum local degree* of G is the maximum value attained by the local degree, taken over all vertices v , and all blocks $B_{i_1}, \dots, B_{i_{r-1}}$.

Alternatively, we can consider a combination of our result with a maximum degree condition — for simplicity we only consider this for graphs. Haxell [28] showed that maximum degree $t/2$ yields an independent transversal; however this allows the average degree in a block to be as high as $t/2$ also. Our construction has $(\frac{1}{4} + \epsilon)t$ average degree within blocks, but the maximum degree is now t .

In this section, we will describe how our graph construction can be modified to also bound these properties. We first investigate the maximum degree, and then move onto other possible combinations. We finish by stating some intriguing open problems related to these.

Lemma 7.8. *For any $\epsilon > 0$, there exists a T such that for all $t > T$, there is a graph with partition (B_i) into parts of size t , maximum degree at most $(\frac{1}{2} + \frac{1}{2\sqrt{2}} + \epsilon)t$ and each block incident to at most $(\frac{1}{4} + \epsilon)t^2$ edges with no independent transversal.*

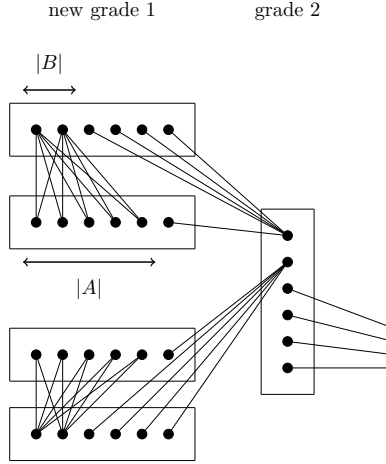


Figure 7.2: Modified graph construction to obtain a maximum degree bound.

Proof. Consider the complete bipartite graph $K_{\lceil \beta t \rceil, \lceil t/(4\beta) \rceil}$ with at most $(1 + \epsilon)t^2/4$ edges for fixed β and large t . Place the parts X, Y of this graph into blocks B_1, B_2 respectively. For an independent transversal, we cannot pick a vertex from both X and Y , and so must pick a vertex from $(X \setminus B_1) \cup (Y \setminus B_2)$. Thinking back to our earlier construction, rather than joining each vertex of a grade 1 block to the entirety of a private grade 0 block (and thus having degree 2), we can instead join it to each of $B_1 \setminus X$ and $B_2 \setminus X$ in a private copy of the above B_1, B_2 . This yields degree at most $2t - \beta t - t/4\beta$, which is strictly less than t provided $\beta \neq \frac{1}{2}$. Since the vertices in a grade 0 block have degree at most $\max(\beta t, t/4\beta)$, taking $\beta = \frac{1}{2} + \frac{1}{2\sqrt{2}} \sim 0.854\dots$ yields a maximum degree over grade 0 and 1 blocks of $0.854\dots$. Blocks of higher grade have lower maximum degree; since even in the original construction it was always below $0.75t$. A diagram of our modified construction appears as Figure 7.2 $\square$

What about local degree? Since we can obtain $0.854t$ maximum degree, this also bounds the maximum local degree. By modifying the proof of Lemma 7.8 to use balanced bipartite graphs at grade 0, we can bound the maximum local degree by $0.75t$, and using the same construction with a slightly unbalanced bipartite graph can improve this to $0.73t$. These choices are optimal for this particular construction, although it seems reasonable that other constructions exist which can beat this.

This leaves two natural open questions. Firstly for maximum degree, let C_1 denote the set of pairs (x, y) such that for some t , there is a graph with average degree at most xt and maximum degree at most yt which does not have an independent transversal. The set C_1 by construction has the property that if $(x, y) \in C, x' \geq x, y' \geq t$ then $(x', y') \in C$. This means we can define $y(x) = \inf\{y : (x, y) \in C\}$ for each $\frac{1}{4} < x$. The result of Wanless and Wood, or equivalently Theorem 7.2 shows C contains no points with $x \leq \frac{1}{4}$; the result of Haxell [28] shows that it contains no points with $y \leq \frac{1}{2}$. Further, the result of Haxell is

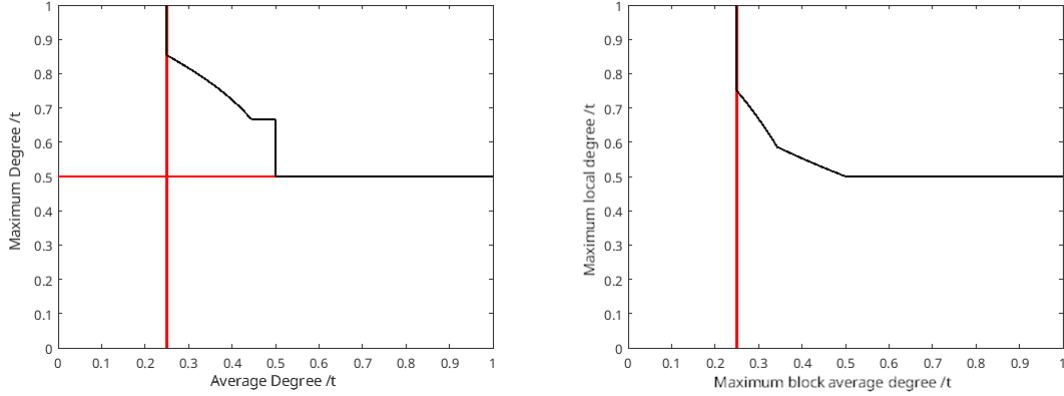


Figure 7.3: A plot of known bounds on the infeasible region for maximum degree or maximum local degree against average degree.

tight, and so using a construction of for instance Szabo and Tardos [62] we deduce $y(x) = \frac{1}{2}$ for each $x > \frac{1}{2}$. Optimising our constructions above, we can show $y(x) \leq \frac{1}{2} + \frac{1}{2}\sqrt{1-2x}$ for each $\frac{1}{4} < x < \frac{4}{9}$, and $y(x) \leq \frac{2}{3}$ for each $\frac{4}{9} \leq x$.

Figure 7.3 gives a graphical description of our region. C_1 cannot meet the red lines, and contains everything above and to the right of a black line. However, we do not know anything about the in-between region. One natural conjecture would be that for any $\epsilon > 0$, there exist graphs with maximum degree at most $(\frac{1}{2} + \epsilon)t$ and average degree at most $(\frac{1}{4} + \epsilon)t$ which do not have independent transversals, but this is open. If this fails, it remains of interest to ask for the largest quantity c_1 such that for all $\epsilon > 0$, all graphs with maximum degree at most $(c_1 - \epsilon)t$ and block average degree at most $(\frac{1}{4} + \epsilon)t$ contain independent transversals. Our results earlier show that $c_1 \leq \frac{1}{2} + \frac{1}{2\sqrt{2}}$; the conjecture above is exactly $c_1 = \frac{1}{2}$.

A key difficulty in proving such a result is that optimal examples for the maximum degree often use large complete bipartite graphs, and therefore the average degree is equal to (or very close to) the maximum degree. Conversely, our example for average degree uses stars, and so while the average degree is controlled our maximum degree is very large. The blending approach of placing complete bipartite graphs inside grade zero allows us to gain features of both of these, but it seems unlikely these are optimal.

Of course, one could also consider a version of this problem in terms of maximum local degree rather than maximum degree. As before we have a feasible region C and infimum $y(x)$ with all $y(\frac{1}{4} + \epsilon) \leq 0.73\dots$. Further, taking a complete bipartite graph with equal part sizes, we deduce all $y(\frac{1}{2} + \epsilon) \leq \frac{1}{2}$. We can use optimisation arguments to interpolate between these results and obtain the second graph of Figure 7.3. Our bounding curves are $1 - \frac{x}{2-2\sqrt{x}}$ for the smaller values of x , switching to $1 - \sqrt{\frac{x}{2}}$ for larger values. We remark that since maximum degree also bounds the maximum local degree, our feasible region

here is a superset of the feasible region from before. We give a graphical description of our known bounds in Figure 7.3. Results on local degree forcing independent transversals are very much open.

To the best of our knowledge, the best lower bound on this region is due to Kang and Kelly [37] and independently Glock and Sudakov [19], who showed that for each value x , there is a function $y(x)$ such that block average degree at most xt and maximum local degree at most yt implies existence of an independent transversal, thus our region C does not have $y(x) = 0$ for any value of x . We remark that there exist examples of graphs with any t with at most one edge between each pair of blocks that do not contain an independent transversal; this proves that our function $y(x)$ tends to zero as x approaches ∞ .

One such construction is due to Erdős, Gyárfás and Łuczak [12] who showed such a graph exists with $n = (t+1)^2$ whenever an affine plane of order $t+1$ exists; a special case of their construction takes some prime p , and considers the 2-dimensional vector space $\mathbb{F}_p^2$. It consists of vertices $v_{i,j}$ where $1 \leq i \leq p-1$ and $1 \leq j \leq p^2$. The columns $v_{-,j}$ form the p^2 blocks of G . Label the vertices of $\mathbb{F}_p^2$ arbitrarily as $y_1, \dots, y_{p^2}$ and pick some $p-1$ vectors in different directions of $\mathbb{F}_p^2$, say $x_1, \dots, x_{p-1}$. We join $v_{i,j}$ to $v_{i,l}$ if y_j and y_l differ by a multiple of x_i . As desired each pair of blocks has at most one edge between them. Each row induces p disjoint cliques of order p , corresponding to equivalence classes. Thus from each row we can only choose p vertices for a total $p(p-1) < p^2$, and so the result follows. Since the average degree in this construction is exactly $p-1$, this implies $f(p-1) \leq \frac{1}{p}$ for each prime p .

Of course, all of these results could equally well be asked in combination with hypergraph questions. However, we do not have as strong results here. A common generalisation of Theorem 7.6 and Lemma 7.8 gives a hypergraph with maximum degree at most ηt^{r-1} (where $\eta \sim 1 - \frac{1}{3er}$), block average degree at most $(\frac{(r-1)^{r-1}}{r^r} + \epsilon)t^{r-1}$ and no independent transversals. The maximum degree bound is now much higher than the $o_r(1)t^{r-1}$ block average degree. We ask if this can be improved, either to some fixed $0 < \eta < 1$ or possibly to $\eta = o(1)$; both of these are open.

Chapter 8

Gallai Colourings

8.1 Introduction

We begin this section by recalling the definition of a Gallai colouring.

Definition 8.1. An edge-colouring of a graph G is called *Gallai* if it contains no rainbow triangles.

A sequence $e_1 \geq \dots \geq e_k$ of positive integers is called *Gallai* if there exists an $n \geq 1$ and a Gallai colouring c of K_n whose colour classes have size $e_1, \dots, e_l$ respectively.

Why do we study these? While Gallai colourings have been historically studied in the context of Ramsey theory (where we seek the minimum n for which all Gallai colourings of K_n contain a monochromatic H), Gyárfás and Simonyi showed that Gallai colourings have intrinsically interesting properties. For instance, every Gallai colouring has a so-called Gallai partition (see Theorem 8.4), which shows these colourings have a recursive structure. This also shows each Gallai colouring contains a connected colour class. In particular, each Gallai colouring contains a colour used on at least $n - 1$ edges.

Is this condition alone sufficient? In joint work with Gyárfás, Pálvölgyi and Patkós [22], we investigated which sequences are Gallai. We were able to find a stronger condition that the class sizes of a Gallai colouring must satisfy, as well as show that two natural sequences form Gallai sequences. Our two positive theorems, which we now state, will be proven in Sections 8.3 and 8.4 respectively.

Theorem 1.35. *Let $p \geq n - 1$. Suppose we can express $\binom{n}{2} = kp + q$ for some integers $k, p \geq 0$. Then there exists a Gallai colouring of K_n with k colour classes of size p , and one colour class of size q . Equivalently, the sequence $(p, \dots, p, q)$ is Gallai.*

Theorem 1.36. *Let $0 \leq k \leq \lceil \frac{n}{2} \rceil$. Then K_n has a Gallai colouring, all of whose class sizes are $\lfloor \frac{n(n-1)}{2k} \rfloor$ or $\lceil \frac{n(n-1)}{2k} \rceil$. We call such a colouring a balanced Gallai k -colouring.*

We will also prove the below strengthening of the statement ‘Every Gallai colouring

contains a colour used on at least $n - 1$ edges', which gives the minimum number of edges some set of at most k colours must appear on. This result is tight, as seen by the Star colouring (cf. Definition 8.2).

Theorem 1.33. *Let $n \geq l$ be integers. Then every Gallai colouring of K_n contains a set of at most l colours which together are used on at least $(n - 1) + (n - 2) + \dots + (n - l)$ edges.*

Both of our positive theorems will in fact be slightly stronger: rather than a Gallai colouring, we will build a special Gallai colouring under the same conditions.

Definition 8.2. A colouring of $V(K_n)$ is a *special Gallai colouring* if labelling the vertices of K_n as $[0, n - 1] = \{0, \dots, n - 1\}$, there exists some partition $T_1, \dots, T_k$ of $[1, n - 1]$ such that colour class i consists $\bigcup_{j \in T_i} S(j)$, where the star $S(j)$ has edges $\{lj : l < j\}$. Equivalently, a special Gallai colouring is a colouring that can be refined to the Star colouring, where $S(i)$ is given colour i . The choice of vertex set is so $S(j)$ has j edges.

Every special Gallai colouring is a Gallai colouring, since the star colouring is Gallai (if $i < j < k$ is some triangle in K_n , then both ik and jk have colour k), and merging colour classes cannot create new rainbow triangles. This simplifies our argument since we do not need to verify that the colourings we construct are Gallai, only that we can sequentially add stars to colour classes.

We remark this is a genuine restriction: not all Gallai sequences can be realised as special Gallai colourings. A concrete example of this, as used in [22], is the sequence 8, 3, 3, 1 for $n = 6$. In a special Gallai colouring with these class sizes, the size 1 class must be $S(1)$ alone. But now consider the star $S(2)$; this cannot be placed in a class of size 3 since we do not have another size 1 class. But this means the two size 3 classes are unions of $S(3), S(4), S(5)$; only one of these has size at most 3 which is a contradiction. This sequence is however Gallai, as witnessed by any colouring starting with a Gallai partition with $n_1 = 2, n_2 = 4$ (see the next section for details).

The above results form part of a longer paper [22], which also asks about the following function $g(k)$.

Definition 8.3. Let $k \geq 0$ be an integer. The function $g(k)$ is the least n , if it exists, such that every sequence $e_1 \geq \dots \geq e_k$ with $\sum e_i = \binom{n}{2}$ is Gallai.

In our joint paper [22], it is shown that $g(k)$ exists, and further $2k - 2 \leq g(k) \leq 8k^2 + 1$. Determining the asymptotic growth rate of this problem is left as an open problem therein; this was mostly resolved by Feffer, Fu and Yan [14] who showed that $g(k)$ is $\Omega(k^{1.5} / \log k)$ and $O(k^{1.5})$.

8.2 A necessary condition for Gallai sequences

The aim of this section is to prove Theorem 1.33. We will be using a different, recursive, classification of Gallai colourings. This equivalence was proven by Gyarfás and Simonyi; for completeness we will give a sketch proof.

Theorem 8.4 (Theorem A, [24]). *For any Gallai colouring of K_n there exist at most two colors, say 1, 2, and a decomposition of K_n into $m \geq 2$ vertex disjoint complete graphs K_{n_i} ($1 \leq i \leq m$) so that all edges between $V(K_{n_i})$ and $V(K_{n_j})$ are coloured with the same color and that color is either 1 or 2. Such a decomposition is called a Gallai partition.*

Proof Sketch. If the colouring uses at most two distinct colours, the result is trivial, so assume at least three colours are used. We will first show that there is always a colour which spans a disconnected graph in such a colouring, and then use this to prove the theorem.

We prove the existence of a disconnected colour class by induction; consider a minimal counterexample. All colour classes are connected and so appear at every vertex. Pick some $v \in V(G)$. By minimality there is a disconnected colour class in $G \setminus \{v\}$, suppose this colour is Red. Let $C_1, \dots, C_m$ be the Red components of this subgraph.

Let Blue and Yellow be two other colours, and let x, y be the endpoints of Blue and Yellow edges from v respectively. If x and y lie in the same component C_i , by considering a Red path $v_1, \dots, v_k$ within C_i from x to y , we have xv_1 is Blue, and xv_k is Yellow. Thus some xv_i has a different colour from xv_{i+1} , and this forms a rainbow triangle and gives a contradiction.

So we can assume x and y lie in different components C_1, C_2 , and since there is no rainbow triangle assume xy is Blue. Let z be the endpoint of a Red edge from v to C_1 , and let $v_1 \dots v_k$ be a Red path in C_1 from x to v . To avoid a rainbow triangle, we must have yv is a Yellow edge (recall C_1 and C_2 are Red components. Now we can use the same argument as above to build a rainbow triangle and obtain our contradiction.

We can now perform this operation recursively. If colour i forms a disconnected colour class, let $C_1, \dots, C_m$ be the components. Any two incident edges vx, vy between the same pair of components must have the same colour; consider a colour i path within the component between x and y . Applying this result twice, this holds for all pairs of edges e_1, e_2 . Thus we are a blowup of a (necessarily Gallai coloured) graph on m vertices. This either uses two colours, in which case we are done, or we can recursively apply our theorem to find a new disconnected colour class. This procedure must terminate at a Gallai partition. $\square$

Remark 8.5. *We can now verify the claim from Chapter 1 that $(7, 3, 2, 2, 1)$ is not a Gallai sequence. Any Gallai partition must be into at most 2 parts to have at most 10 ($=7+3$) edges between them; thus only one colour is used and in fact the part sizes are 1 and 5 respectively, with all edges between them of colour 1. It remains to find a Gallai*

colouring of K_5 with classes of size $(3, 2, 2, 2, 1)$ which cannot exist, since no colour can span a connected subgraph.

Before proving our main theorem, we will prove a couple of useful properties of Gallai colourings.

Theorem 8.6. *Every Gallai colouring of K_n contains a monochromatic spanning tree. Every Gallai colouring of K_n uses at most $n - 1$ colours.*

Proof. Consider a Gallai partition $V_1, \dots, V_m$ of our colouring. The ‘reduced graph’ obtained by replacing each V_i with a single point is now a 2-coloured copy of K_m . Then one of the two colours used on this K_m must span a connected graph (if one colour is disconnected, the other colour is complete between components). Returning to our original graph, that colour class must also span a connected graph, and hence it contains a monochromatic spanning tree.

For the second result, we proceed by induction on n , the base case $n = 2$ being trivial. Taking the Gallai partition, at most 2 colours appear between the V_i , and at most $|V_i| - 1$ colours within V_i . Overall, this is at most $2 + n - m$ colours. So if $m \geq 3$ the result follows. If $m = 2$, then actually only one colour appears between the V_i and again the result holds. $\square$

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

Proof of Theorem 1.33. We proceed by induction on n ; the base case $n = 1$ and any l being trivial so assume $n \geq 2$. Considering a Gallai partition, either the colour 1 or colour 2 edges span a connected graph, and that class has at least $n - 1$ edges as desired; we thus assume $l \geq 2$. If G uses exactly 2 colours, the result holds as $e_1 + e_2 = \binom{n}{2} \geq 2n - 3$ so we can assume at least 3 colours are used. We replace l with $\min(l, n)$ as this only strengthens the result we need to prove.

Let $V_1, \dots, V_m$ be a Gallai partition, with $|V_i| = n_i$. Suppose first that $m \geq 3$. We aim to pick a sequence $(l_i)_{i=1}^m$ of integers with $\sum l_i = l - 2$ and each $l_i \leq n_i - 1$. If we can do so, we will pick the l_i colours which span the most edges within V_i , together with the colours that appear on the partition.

If we cannot pick such a sequence, by Theorem 8.6 then actually our colouring has at most l non-trivial colour classes, and the result holds since $\binom{n}{2} = (n - 1) + \dots + 1$. We therefore assume such a sequence exists, and fix one now.

We will form our l colour classes as the l_i classes which use the most colours in $G[V_i]$, together with the at most two colours used on the Gallai partition. We remark this may be fewer than l colours due to overlap.

By induction, the l_i largest classes of $G[V_i]$ have at least $(n_i - 1) + \dots + (n_i - l_i)$ edges, and so there are at most $1 + \dots + (n_i - l_i - 1)$ edges inside V_i of a colour we did

not use. This is overall at most $\binom{n-m+3-l}{2}$ edges, and so we must have used at least $(n-1) + \dots + (n-m+3-l)$ as required. It remains only to handle the case $m = 2$; this is similar except we only use one colour on the Gallai partition, and so can repeat the above argument looking for l_i with $\sum l_i = l - 1$. $\square$

8.3 Sequences with almost all entries equal are Gallai

Having proven our necessary condition, we now turn our attention to the positive theorems 1.35 and 1.36. We will prove Theorem 1.36 in the next section, using Theorem 1.35 as a proof ingredient. The aim of this section is to prove the ‘two-class’ theorem, Theorem 1.35.

Proof of Theorem 1.35. We will proceed by induction, proving the stronger statement *For all $p \geq n - 1$ and non-negative integers k, q such that $n = kp + q$, there is a special Gallai colouring of K_n into k classes of size p and one of size q .* The base case $n = 1$ is trivial and so assume $n \geq 2$. Our proof method will be to use a series of different ‘moves’ to either build colour classes and reduce to a smaller value of n with the same p, q , or to build some partial colour classes and reduce to a smaller n with different p, q . We will assume $q \leq p - 1$, with the general result following by merging a class of size $q \pmod p$ with $\lfloor \frac{q}{p} \rfloor$ classes of size p .

Move 1. Removing a single star

Suppose first that $q \geq n - 1$. We will remove the single star $S(n - 1)$ to obtain a copy of K_{n-1} . By induction K_{n-1} has a special Gallai colouring into k classes of size p and a single class of size $q - n + 1$. Adding $S(n - 1)$ to the size $q - n + 1$ class forms our desired special Gallai colouring. We can therefore assume $q \leq n - 2$ also.

Move 2. Removing star pairs

Suppose now that p is odd, and also $p \leq 2n - 3$. Under these conditions we can write $p = (n - 1) + p - n + 1$ and so we can form a class of size p for each of $1 \leq i \leq \frac{p+1}{2}$ the ‘star pairs’ $S(n - 1 - i) \cup S(p - n + i)$. Removing the central vertices of these stars, we are left with a copy of K_{p-n+1} in which we seek to find $k - (n - \frac{p-1}{2})$ classes of size p , and one of size q . By induction such a special Gallai colouring exists.

Move 3. Removing star pairs and a star

Assume now that $p \leq 2n - 2$ is an even integer. The above argument does not quite work, since after removing star pairs we are left with K_{p-n+1} and also the star $S(p/2)$. However we do have that p is quite small, and in particular $p/2 \geq p - (n - 1) > (p - n + 1) - 1$. Thus by induction there is a special Gallai colouring of K_{p-n+1} into $2(k - (n - p/2)) - 1$ classes of size $p/2$ and a single class of size q . Taking the star $S(p/2)$ as a last size $p/2$ class and pairing these sets arbitrarily gives our desired special Gallai colouring.

Forming partial classes

We can now assume $p \geq 2n - 1$ and so $k < n/4$. We will be using the above moves, but this time rather than forming complete classes we will instead form k equal size star pairs, and find the remainder using the induction statement. More concretely, as in Move 2 we form star pairs $S(n - i) \cup S(n - 2k - 1 + i)$ for each $1 \leq i \leq k$, and place these into different classes of size p . We now seek to find k classes of size $p - (2n - 2k - 1)$ and one class of size q in the remaining graph K_{n-2k} . If $p - (2n - 2k - 1) \geq n - 2k - 1$ then we can find a special Gallai colouring with these class sizes by induction, so assume not and so $2n - 1 \leq p \leq 3n - 4k - 3$ (and hence $4k \leq n - 2$). Since also $q \leq n - 2$ we have $\binom{n}{2} \leq k(3n - 4k - 3) + (n - 2)$; writing $n = 4k + \delta$ for some $\delta \geq 2$ we deduce $\delta = 2$, and hence $p = 8k + 3$, $q = 3k + 1$. We now use move 1 to place the star $S(n - 2k - 1)$ into the class of size q , since $3k + 1 \geq n - 2k - 1$ we can do this. It remains only to find a special Gallai colouring of K_{2k+1} with 1 class of size k and k classes of size $2k$, which exists by induction. $\square$

8.4 Balanced sequences are Gallai

We are now ready to prove our final theorem, Theorem 1.36, and thus complete the aims of this chapter. Like the proof of Theorem 1.35, we will be reducing to a smaller case by a series of moves, and associating removed edges to colour classes in a suitable fashion.

Proof of Theorem 1.36. We will again proceed by induction on n , proving the statement *For all $k \leq \lfloor \frac{n}{2} \rfloor$ there is a balanced special Gallai colouring of K_n with k colours.* The base case $n = 1$ and any k is trivial, and so assume $n \geq 2$ and the result holds for all smaller complete graphs (and any permitted k). We will be using Move 2 of the previous subsection repeatedly.

Claim 1.35.1. *If there is a balanced special Gallai colouring of K_n with $2k$ colours, then there is a balanced special Gallai colouring of K_n using k colours.*

Proof. Suppose that there is a balanced special Gallai colouring of $E(K_n)$ using $2k$ parts. Suppose there are m classes of size $Z = \lfloor \frac{n(n-1)}{4k} \rfloor$ and $2k - m$ parts of size $Z + 1$. We build our k classes from pairs of these classes, first taking $\min(m, 2k - m)$ pairs of size Z and $Z + 1$. The remaining classes all have size either Z or $Z + 1$, and can be paired in any way we desire. Since these classes have size $2Z$ or $2Z + 2$, our final special Gallai colouring is balanced. $\square$

Our reduction statement follows by picking the least integer r such that $\frac{n}{4} \leq 2^r k$ and applying Claim 1.35.1 r times, starting from $2^r k$. We therefore assume $\frac{n}{4} \leq k \leq \lfloor \frac{n}{2} \rfloor$. Let $Z = \lfloor \frac{n(n-1)}{2k} \rfloor$, and by the bounds on k we have $n - 1 \leq Z \leq 2n - 2$. We will seek to pull off star pairs (and perhaps a single star) in order to form parts of our classes. We seek to build $m \geq 1$ colour classes of size Z , and $k - m$ of size $Z + 1$.

Case A. Z is odd

Suppose first that Z is odd, and in particular $Z \leq 2n - 3$. We will now use Move 2 to build star pairs $S(n - i) \cup S(Z - n + i)$ for each of $1 \leq i \leq \max(m, n - \frac{Z+1}{2})$ to form $\max(m, n - \frac{Z+1}{2})$ colour classes of size Z . If also $m \geq \frac{2n-1-Z}{2}$, we can now apply our induction hypothesis to the remaining K_{Z-n+1} with $k' = k - \frac{2n-1-Z}{2}$ to complete our proof — note that our sought classes have size at least $n - 1 \geq Z - n$ and so k' is small enough this is valid.

We therefore assume $m < \frac{2n-1-Z}{2}$, and so after forming the above star pairs it remains to build classes of size $Z+1$ inside the union of K_{Z-n+1} and the stars $S(n-m-1), \dots, S(Z-n+m+1)$. To do so we will use move 3. We will remove star pairs $S(n-m-i) \cup S(Z-n+m+i+1)$ for each of $1 \leq i \leq n-m-\frac{Z+1}{2}-1$. This builds some classes of size $Z+1$, and leaves only the clique K_{Z-n+1} and stars $S(Z-n+m+1)$ and $S(\frac{Z+1}{2})$. We use these as remaining stars to build a star pair of total size $Z - (n-m-1-\frac{Z+1}{2}) \leq Z$. It remains now to find a special Gallai colouring of K_{Z-n+1} into a single class of size $n-m-\frac{Z+3}{2}$ and all remaining classes of size $Z+1$; such a colouring exists by our earlier Theorem 1.35, and so the result follows merging our classes appropriately.

Case B. Z is even.

It remains only to handle the case where Z is even. In this case $Z+1$ is odd and so we can once again use Move 2 to build star pairs $S(n-i) \cup S(Z+1-n+i)$ for each $1 \leq i \leq \min(n - \frac{Z+2}{2}, k-m)$ star pairs of size $Z+1$. This gives a non-trivial number of classes of size $Z+1$ unless $Z = 2n - 2$. However in this case we seek to find $n/4$ classes of size exactly $2n - 2$ and the result follows by Theorem 1.35 directly; thus assume this is not the case.

If also $k-m \geq n - \frac{Z+2}{2}$ then we can apply induction to build a special Gallai colouring of the remainder of the graph into classes of size Z or $Z+1$ as in Case A, and then use these star pairs to form the remaining classes. We therefore assume $k-m < n - \frac{Z+2}{2}$. We now proceed by skipping the star $S(n-k+m-1)$ and forming further star pairs $S(n-k+m-1-i) \cup S(Z+1-n+k-m+i)$ of size Z for each of $1 \leq i \leq n-k+m-1-\frac{Z+1}{2}-1$. This leaves only the stars $S(\frac{Z}{2})$ and $S(n-m+k-1)$ together with the complete graph on $Z-n+2$ vertices. By Theorem 1.35 there is a special Gallai colouring of this complete graph with one class of size $\frac{Z}{2} - (n-m+k-1)$ and all remaining classes having size exactly $\frac{Z}{2}$. Since $Z \leq 2n-2$, it holds $\frac{Z}{2} \geq (Z-n+1)-1$ and so the application was valid. Merging the classes of size $\frac{Z}{2}$ in pairs, and the remaining classes of size $\frac{Z}{2} - (n-m+k-1)$ and $n-m+k-1$ completes the proof. $\square$

Bibliography

- [1] K. Appel and W. Haken. “Every planar map is four colorable. Part I: Discharging”. *Illinois Journal of Mathematics* 21.3 (1977), pp. 429–490.
- [2] K. Appel, W. Haken, and J. Koch. “Every planar map is four colorable. Part II: Reducibility”. *Illinois Journal of Mathematics* 21.3 (1977), pp. 491–567.
- [3] R.B. Ash. *Information Theory*. Dover, 1965.
- [4] J. Balogh and L. Li. “The typical structure of Gallai colorings and their extremal graphs”. *SIAM Journal on Discrete Mathematics* 33 (4 2019), pp. 2416–2443.
- [5] B. Bollobás, P.A. Catlin, and P. Erdős. “Hadwiger’s Conjecture is True for Almost Every Graph”. *European Journal of Combinatorics* 1.3 (1980), pp. 195–199.
- [6] M. Delcourt and L. Postle. *Reducing Linear Hadwiger’s Conjecture to Coloring Small Graphs*. 2021. arXiv: 2108.01633 [math.CO].
- [7] G. Ding and C. Liu. “Excluding a small minor”. *Discrete Applied Mathematics* 161 (3 2013), pp. 355–368.
- [8] G.A. Dirac. “Homomorphism Theorems for Graphs”. *Mathematische Annalen* 153 (1964), pp. 69–80. URL: <http://eudml.org/doc/161119>.
- [9] G.L. Duarte, H. Eto, T. Hanaka, Y. Kobayashi, Y. Kobayashi, D. Lokshantov, L.L.C. Pedrosa, R.C.S. Schouery, and U.S. Souza. “Computing the Largest Bond and the Maximum Connected Cut of a Graph”. *Algorithmica* 83 (5 2021), pp. 1421–1458.
- [10] P. Duchet and H. Meyniel. “On Hadwiger’s number and the stability number”. *Annals of Discrete Mathematics* 13 (1982), pp. 71–74.
- [11] P. Erdős and S. Fajtlowicz. “On the conjecture of Hajós”. *Combinatorica* 1 (2 1981), pp. 141–143.
- [12] P. Erdős, A. Gyárfás, and T. Łuczak. “Independent Transversals in sparse partite hypergraphs”. *Combinatorics, Probability and Computing* 3.3 (1994), pp. 293–296.
- [13] R. Faudree, R. Gould, M. Jacobson, and C. Magnant. “Ramsey numbers in rainbow triangle free colorings”. *The Australasian Journal of Combinatorics* 46 (Feb. 2010), pp. 269–284.
- [14] J. Feffer, Y. Fu, and J. Yan. “Remarks on the distribution of colours in Gallai colourings”. *Discrete Mathematics* 343 (10 2020).
- [15] W. Fernandez De La Vega. “On the Maximum Density of Graphs Which Have No Subcontraction to K^s ”. *Discrete Math.* 46.1 (1983), pp. 109–110.

- [16] Z. Füredi and A. Kündgen. “Turán problems for integer-weighted graphs”. *J. Graph Theory* 40 (2002), pp. 195–225.
- [17] T. Gallai. “Transitiv orientierbare Graphen (in German)”. *Acta Mathematica Academiae Scientiarum Hungarica* 18 (1967), pp. 25–66.
- [18] J. Geelen, B. Gerards, B. Reed, P. Seymour, and A. Vetta. “On the odd-minor variant of Hadwiger’s conjecture”. *Journal of Combinatorial Theory, Series B* 99.1 (2009), pp. 20–29.
- [19] S. Glock and B. Sudakov. “An average degree condition for independent transversals”. *Journal of Combinatorial Theory, Series B* 154 (2022), pp. 370–391.
- [20] S. Greenberg and M. Mohri. “Tight Lower Bound on the Probability of a Binomial Exceeding its Expectation”. *Statistics & Probability Letters* 86 (June 2013).
- [21] C. Groenland, T. Kaiser, O. Treffers, and M. Wales. “Graphs of low average degree without independent transversals”. *Journal of Graph Theory (to appear)* (2022).
- [22] A. Gyárfás, D. Pálvölgyi, B. Pátkos, and M. Wáles. “Distribution of colors in Gallai Colourings”. *European Journal of Combinatorics* 86 (2020).
- [23] A. Gyárfás, G.N. Sárközy, A. Sebő, and S. Selkow. “Ramsey-type results for Gallai colorings”. *Journal of Graph Theory* 64 (2010), pp. 233–242.
- [24] A. Gyárfás and G. Simonyi. “Edge colorings of complete graphs without tricolored triangles”. *Journal of Graph Theory* 46 (2004), pp. 211–216.
- [25] H. Hadwiger. “Über eine Klassifikation der Streckenkomplexe”. *Vierteljahrsschrift der Naturf. Gesellschaft in Zürich* 88 (1943), pp. 133–142.
- [26] D.W. Hall. “A note on primitive skew curves”. *Bulletin of the American Mathematical Society* 49 (12 1943), pp. 935–936.
- [27] J. Haslegrave, J. Kim, and H. Liu. “Extremal density for sparse minors and subdivisions”. *International Mathematics Research Notices* (June 2021).
- [28] P. Haxell. “A note on vertex list colouring”. *Combinatorics, Probability and Computing* 10.4 (2001), pp. 345–347.
- [29] K. Hendrey, S. Norin, and D. Wood. *Extremal functions for sparse minors*. 2021. arXiv: 2107.08658 [math.CO].
- [30] R. Hickingbotham and D. Wood. *Structural properties of bipartite subgraphs*. 2021. arXiv: 2106.12099 [math.CO].
- [31] W. Hoeffding. “Probability Inequalities for Sums of Bounded Random Variables”. *Journal of the American Statistical Association* 58.301 (1963), pp. 13–30.
- [32] I. T. Jakobsen. “A homomorphism theorem with an application to the conjecture of Hadwiger”. *Stud. Scient. Math. Hungarica* 6 (1971), pp. 151–160.
- [33] I. T. Jakobsen. “On certain homomorphism properties of graphs I”. *Mathematica Scandinavica* 31.2 (1972), pp. 379–404.
- [34] I. T. Jakobsen. “On certain homomorphism properties of graphs II”. *Mathematica Scandinavica* 52.2 (1983), pp. 229–261.

- [35] S. Janaqi, F. Lescure, M. Maamoun, and H. Meyniel. “Digraphs contractible onto $*K_3$ ”. *Mathematica Bohemica* 123.4 (1998), pp. 365–369.
- [36] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley, 2000. ISBN: 0-471-17451-2.
- [37] R. Kang and T. Kelly. “Colorings, transversals and local sparsity”. *Random Structures and Algorithms* (2021).
- [38] R. Kapadia, S. Norin, and Y. Qian. “Asymptotic density of graphs excluding disconnected minors”. *J. Comb. Theory, Ser. B* 146 (2021), pp. 151–175.
- [39] K. Kawarabayashi, M. Plummer, and B. Toft. *Journal of Combinatorial Theory, Series B* 95 (2005), pp. 152–167.
- [40] K. Kawarabayashi and Z-X. Song. “Independence Number and Clique Minors”. *Journal of Graph Theory* 56.3 (2007), pp. 219–226.
- [41] A. V. Kostochka. “The minimum hadwiger number for graphs with a given mean degree of vertices (in russian)”. *Metody Diskretnogo Analiza* 38 (01 1982).
- [42] A. V. Kostochka and N. Prince. “On $K_{s,t}$ -minors in graphs with given average degree, II”. *Discrete Mathematics* 312.24 (2012), pp. 3517–3522.
- [43] A.V. Kostochka and N. Prince. “On $K_{s,t}$ -minors in graphs with given average degree”. *Discrete Mathematics* 309 (19 2008), pp. 4435–4445.
- [44] D. Kühn and D. Osthus. “Extremal connectivity for topological cliques in bipartite graphs”. *Journal of Combinatorial Theory, Series B* 96.1 (2006), pp. 73–99.
- [45] W. Mader. “Homomorphieeigenschaften und mittlere Kantendichte von Graphen”. *Mathematische Annalen* 174 (1967), pp. 265–268.
- [46] W. Mader. “Homomorphiesätze für Graphen”. *Mathematische Annalen* 178 (1968), pp. 154–168.
- [47] C. Magnant and I. Schiermeyer. “Gallai-Ramsey number for K_5 ”. *Journal of Graph Theory (to appear)* (2022).
- [48] R.A. Moser. “A Constructive Proof of the Lovász Local Lemma”. In: *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*. STOC ’09. Association for Computing Machinery, 2009, pp. 343–350. ISBN: 9781605585062.
- [49] J. Myers and A. Thomason. “The Extremal Function For Noncomplete Minors”. *Combinatorica* 25 (Dec. 2005), pp. 725–753.
- [50] S. Norin, L. Postle, and Z-X. Song. *Breaking the degeneracy barrier for coloring graphs with no K_t minor*. 2020. arXiv: 1910.09378 [math.CO].
- [51] S. Norin, B. Reed, A. Thomason, and D. Wood. “A Lower Bound on the Average Degree Forcing a Minor”. *Electron. J. Comb.* 27 (2020), P2.4.
- [52] J. de Oliveira Bastos, F.S. Benevides, and J. Han. “The number of Gallai k -colourings of complete graphs”. *Journal of Combinatorial Theory, Series B* 144 (1 2020), pp. 1–13.
- [53] J. de Oliveira Bastos, F.S. Benevides, G. O. Mota, and I. Sau. “Counting Gallai 3-colorings of complete graphs”. *Discrete Mathematics* 342 (9 2019), pp. 2618–2631.

- [54] A. Pedersen and B. Toft. *A Basic Elementary Extension of the Duchet-Meyniel Theorem*. 2008. arXiv: 0810.0846 [math.CO].
- [55] L. Postle. *An even better Density Increment Theorem and its application to Hadwiger's Conjecture*. 2020. arXiv: 2006.14945 [math.CO].
- [56] L. Postle. *Further progress towards Hadwiger's conjecture*. 2020. arXiv: 2006.11798 [math.CO].
- [57] B. Reed and D. Wood. "Forcing a sparse minor" — CORRIGENDUM". *Combinatorics, Probability and Computing* 25.2 (2016), pp. 323–323.
- [58] B. Reed and D. Wood. "Forcing a sparse minor". *Combinatorics, Probability and Computing* 25 (2015), pp. 300–322.
- [59] N. Robertson, P. Seymour, and R. Thomas. "Hadwiger's conjecture for K_6 -free graphs". *Combinatorica* 13.3 (1993), pp. 279–361.
- [60] M. Rosenfeld. "Another approach to non-repetitive colorings of graphs of bounded degree". *Electronic Journal of Combinatorics* (2020), P3.43. ISSN: 1077-8926.
- [61] Z-X. Song and R. Thomas. "The extremal function for K_9 minors". *J. Comb. Theory, Ser. B* 96 (2006), pp. 240–252.
- [62] T. Szabó and G. Tardos. "Extremal problems for transversals in graphs with bounded degree". *Combinatorica* 26.3 (2006), pp. 333–351.
- [63] T. Tao. *Moser's entropy compression argument*. Available at <https://terrytao.wordpress.com/2009/08/05/mosers-entropy-compression-argument/>, accessed 2021-11-09. 2009.
- [64] A. Thomason. "An extremal function for contractions of graphs". *Math. Proc. Cam. Phil. Soc.* 95.2 (1984), pp. 261–265.
- [65] A. Thomason. "The Extremal Function for Complete Minors". *Journal of Combinatorial Theory, Series B* 81.2 (2001), pp. 318–338.
- [66] Andrew Thomason and Matthew Wales. "On the extremal function for graph minors". *Journal of Graph Theory (to appear)* (2022).
- [67] Z. Tuza and J. A. Bondy. "A weighted generalization of Turán's Theorem". *Journal of Graph Theory* 25 (1997), pp. 267–275.
- [68] V. A. Vatutin and V. G. Mikhailov. "Limit Theorems for the Number of Empty Cells in an Equiprobable Scheme for Group Allocation of Particles". *Theory of Probability & Its Applications* 27.4 (1983), pp. 734–743.
- [69] K. Wagner. "Über eine Eigenschaft der ebenen Komplexe (in German)". *Mathematische Annalen* 114 (1937), pp. 570–590.
- [70] M. Wales. *Bipartite clique minors in graphs of large Hadwiger number*. 2021. arXiv: 2107.06716 [math.CO].
- [71] M. Wales. *Extremal functions for small multigraph minors*. (in preparation).
- [72] M. Wales. *The extremal function for structured sparse minors*. 2021. arXiv: 2110.08008 [math.CO].

- [73] I. Wanless and D. Wood. *A general framework for hypergraph colouring*. 2021. arXiv: 2008.00775 [math.CO].