EMBEDDING PROBLEMS FOR GRAPHS AND HYPERGRAPHS

by

OLIVER JOSEF NIKOLAUS COOLEY

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Mathematics
The University of Birmingham
October 2009

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Daniela Kühn and Deryk Osthus, for helping and guiding me through the academic side of my PhD career. Their knowledge, experience and encouragement have been invaluable.

I would also like to thank Nikolaos Fountoulakis, Daniela Kühn and Deryk Osthus for the collaborations which formed the basis for Chapters 2 and 4 of this thesis.

Undertaking a PhD would have been impossible without the generous financial support of the Engineering and Physical Sciences Research Council.

I owe a great debt of gratitude to my parents for their support through the whole of my life and academic career so far.

Finally I would like to thank my friends and colleagues the excellent and admirable members, past, present and honorary, of the Birmingham University Maths Postgraduate Society. I would have completed a PhD without them, but they have made the social side as enjoyable as the mathematical.

Abstract

This thesis deals with the problem of finding some substructure within a large graph or hypergraph. In the case of graphs, we consider the substructures consisting of fixed subgraphs or families of subgraphs, perfect graph packings and spanning subgraphs. In the case of hypergraphs we consider the substructure consisting of a hypergraph whose order is linear in the order of the large hypergraph. I will show how these problems are extensions of more basic and well-known results in graph theory. I will give full proofs of three new embedding results, two for graphs and one for hypergraphs. I will also discuss the regularity lemma for graphs and hypergraphs, an important tool which underpins these and many similar embedding results. Finally, I will also discuss graph and hypergraph Ramsey numbers, since two of the embedding results have important applications to Ramsey numbers which improve upon previously known results.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER 1

INTRODUCTION

1.1 Graph Packings

1.1.1 Notation and Preliminaries

Throughout this thesis, a *graph* refers to a simple undirected graph, that is a set V of distinct vertices and a set E of edges which is a subset of the set of unordered pairs of vertices in V. No loops or multiple edges are allowed.

Given a graph $G = (V, E)$, we write $|G| := |V|$ for the *order* of G, and $e(G) := |E|$ for the number of edges. $\chi(G)$ denotes the chromatic number of G, i.e. the least integer ℓ such that there is a partitioning of $V(G)$ into ℓ independent sets (i.e. sets containing no edges). The minimum degree of G is denoted by $\delta(G)$, and the maximum degree by $\Delta(G)$. For disjoint sets X and Y in G, we write $e(X, Y)$ for the number of edges of G with one endpoint in X and one in Y, and $d(X, Y) := \frac{e(X,Y)}{|X||Y|}$ denotes the density of edges between X and Y. We denote by $d(A) := e(A)/\binom{|A|}{2}$ the density of A. We write $x = a \pm b$ to mean $a \leq x \leq b$.

The degree of a vertex x in G is denoted by $d_G(x)$, or by $d(x)$ if this is unambiguous. The neighbourhood is denoted by $N_G(x)$ or simply by $N(x)$. For a set of vertices X the neighbourhood of X is $N(X) := \bigcup_{x \in X} N(x)$. For a subset $S \subseteq V(G)$, the number of neighbours in S of a vertex x is denoted by $d(x, S)$ or by $d_S(x)$. If we have disjoint subsets $A, B \subseteq G$, then we define $\delta(A, B) := \min_{x \in A}\{d_B(x)\}$, i.e. the minimum degree in B of a vertex in A. We denote by $G[A]$ the subgraph of G induced by the vertex set A.

By the notation $a \ll b \ll c$ we mean that we pick constants from right to left, and that there are increasing real-valued functions f and g such that our statements holds provided $b \leq f(c)$ and $a \leq g(b)$. Hierarchies with more constants are defined similarly. The necessary functions f and g could be calculated explicitly from the appropriate proofs, but for simplicity we will not do this. We simply assume that b is sufficiently small compared to c, and a sufficiently small compared to b, for all our calculations to work.

We usually use n to denote the order of a graph and think of n as very large, or indeed tending to infinity. Then given a function f, the notation $m \sim f(n)$ means that $m/f(n) \stackrel{n\to\infty}{\longrightarrow} 1$.

1.1.2 Forcing Subgraphs

The subject area of this thesis has ultimately grown out of the following question: *Which graph properties force the existence of a certain substructure in the graph?* Alternatively we can ask the contrapositive question: *How does forbidding a certain substructure influence the global properties of a graph?* Some of the early results in this field took as the substructure a particular subgraph. Perhaps the most basic result in this field is Mantel's Theorem, which gives a best possible condition on the number of edges in a graph which does not contain a triangle. Mantel's Theorem is generalised by Turán's Theorem, which concerns the number of edges in a graph which does not contain a copy of K_r, the complete graph on r vertices.

More precisely, let $T_{r-1}(n)$ denote the complete $(r-1)$-partite graph on n vertices, where the vertex classes have sizes as equal as possible. This is called the *Turán Graph*. See Figure 1.1.

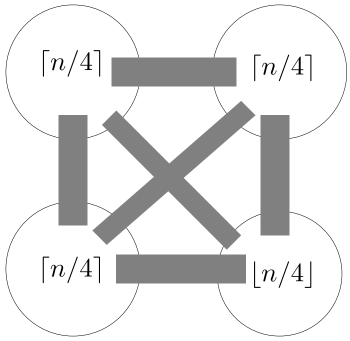


Figure 1.1: The graph $T_4(n)$ (where $n \equiv 3 \mod 4$)

Clearly $T_{r-1}(n)$ is K_r-free, i.e. it does not contain any copy of K_r as a subgraph (any copy of K_r would have to contain at least two vertices from one of the classes). It is also easy to see that $T_{r-1}(n)$ has the greatest possible number of edges of any $(r-1)$-partite graph on n vertices. What is slightly harder to prove is that in fact $T_{r-1}(n)$ contains the largest possible number of edges of any K_r-free graph on n vertices.

Let $ex(n, K_r)$ denote the greatest possible number of edges in a graph G subject to the conditions that $|G| = n$ and that G does not contain a copy of K_r.

Theorem 1.1 (Turán, 1941) *Let G be a K_r-free graph on n vertices, and suppose G satisfies $e(G) = ex(n, K_r)$. Then $G = T_{r-1}(n)$.*

Turán's theorem essentially says two distinct things. It implies firstly that the Turán Graph $T_{r-1}(n)$ achieves the maximum possible number of edges of a K_r-free graph on n vertices, and secondly that it is the unique graph which achieves this upper bound. As a corollary of this theorem we have the following result:

Corollary 1.2 $\frac{ex(n,K_r)}{\binom{n}{2}} \to 1 - \frac{1}{r-1}$ *as* $n \to \infty$.

We will generally apply Turán's Theorem in the following form:

Corollary 1.3 *For any real number $\varepsilon > 0$, there exists an integer $n_0 = n_0(\varepsilon)$ such that if G is a graph on $n \geq n_0$ vertices satisfying*

$$e(G) \geq (1 - \frac{1}{r-1} + \varepsilon)\binom{n}{2}$$

then $K_r \subseteq G$.

An extension of this result from K_r to more general graphs H was achieved by the famous Erdős-Stone-Simonovits Theorem. We define $ex(n, H)$ to be the greatest possible number of edges in an H-free graph on n vertices.

Theorem 1.4 (Erdős, Stone, 1946, Erdős, Simonovits, 1966) *Given a graph H and a real number $\varepsilon > 0$, there is an integer $n_0 = n_0(H, \varepsilon)$ such that any graph G on $n \geq n_0$ vertices satisfying $e(G) \geq (1 - \frac{1}{\chi(H)-1} + \varepsilon)\binom{n}{2}$ contains a copy of H. In particular, for any graph H, $\frac{ex(n,H)}{\binom{n}{2}} \to 1 - \frac{1}{\chi(H)-1}$ as $n \to \infty$.*

The lower bound in the limit can be deduced from the Turán graph $T_{r-1}(n)$, where $r = \chi(H)$.

1.1.3 Forcing Graph Packings

One observation which we can make from the Erdős-Stone-Simonovits Theorem is that it also provides a condition guaranteeing multiple copies of a graph H. For if we wish to find k disjoint copies of H, we let H' be the graph consisting precisely of such copies. We can apply the Erdős-Stone-Simonovits theorem to find a copy of H' in G provided $|G|$ is large enough and the density of G is at least $1 - \frac{1}{\chi(H')-1} + \varepsilon = 1 - \frac{1}{\chi(H)-1} + \varepsilon$. However, the number of copies of H we can find in this way is still only a bounded number, i.e. it does not depend on the order of G. We now wish to find a number of disjoint copies of H which cover a large proportion of the vertices of G.

We define an *H-packing in G* to be a collection of vertex-disjoint copies of H in G. A *perfect H-packing in G* is an H-packing which covers all the vertices of G. Given $\varepsilon > 0$, an *almost perfect H-packing in G* is an H-packing in G covering at least $(1-\varepsilon)|G|$ vertices.

The aim is to find natural conditions on G which guarantee a perfect H-packing. Certainly we will require $|G|$ to be divisible by $|H|$. We will also assume from now on that $\chi(H) \geq 2$ (or equivalently that $e(H) > 0$) for otherwise all packing results are trivial.

When we were looking for just one copy of H in G, Turán's Theorem and the Erdős-Stone-Simonovits Theorem gave us reasonable conditions on the number of edges in G. Such edge conditions will no longer be useful if we are looking for a perfect H-packing. For example, G may consist of a complete graph on $n-1$ vertices along with one isolated vertex. Clearly such a graph will not contain a perfect H-packing if H has no isolated vertex, yet it has a very large number of edges.

To avoid such a situation, we will now be looking at bounds on the minimum degree of G. We make the following definition:

Definition 1.5 *Given a graph H and an integer n divisible by $|H|$, let $\delta(n,H)$ denote the least integer k such that any graph G on n vertices with minimum degree $\delta(G) \geq k$ must contain a perfect H-packing.*

It is clear that $\delta(n,H)$ exists whenever n is divisible by $|H|$, since $k = n-1$ would be sufficient to guarantee a perfect H-packing, so the set of such k is non-empty. When n is not divisible by $|H|$, $\delta(n,H)$ is undefined. Whenever $\delta(n,H)$ is mentioned, we assume that n is divisible by $|H|$ without mentioning this explicitly.

To make the link between forcing subgraphs and forcing graph packings more natural, we also define a corresponding function for the subgraph case:

Definition 1.6 *Let $\delta_0(n, H)$ denote the least integer k such that any graph G on n vertices with minimum degree $\delta(G) \geq k$ contains a copy of H.*

The following can easily be deduced from the Turán and Erdős-Stone-Simonovits Theorems, and by considering the Turán graph:

Proposition 1.7 $\delta_0(n, H) \sim (1 - \frac{1}{\chi(H)-1})n$.
In particular, $\delta_0(n, K_r) \sim (1 - \frac{1}{r-1})n$.
Furthermore, if n is divisible by $r - 1$, then $\delta_0(n, K_r) = (1 - \frac{1}{r-1})n$.

When looking for perfect H-packings, one easy special case is when $H = K_2$, i.e. a single edge. In this case, an H-packing is a matching, and a perfect H-packing is a perfect matching. We can deduce an upper bound on $\delta(n, K_2)$ from Dirac's Theorem on Hamilton cycles:

Theorem 1.8 (Dirac, 1962) *Any graph G on n vertices with minimum degree $\delta(G) \geq n/2$ contains a Hamilton cycle.*

In particular, if n is even a Hamilton cycle will naturally guarantee a perfect matching.

Corollary 1.9 $\delta(n, K_2) \leq n/2$.

On the other hand, for any $n = 2k$ we can easily construct a graph on n vertices with minimum degree $n/2 - 1$ which has no perfect matching. Indeed, let G be a complete bipartite graph with vertex classes of size $k - 1$ and $k + 1$. See Figure 1.2.

It is easy to see that G satisfies all of the above conditions. This implies that $\delta(n, K_2) > n/2 - 1$, and thus:

Proposition 1.10 $\delta(n, K_2) = n/2$.

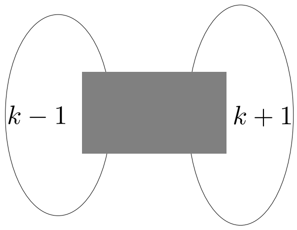


Figure 1.2: The extremal graph for K_2

Now suppose we wish to extend this observation to K_r. We might look first for an example of a graph with high minimum degree, but not containing a perfect K_r-packing. One possible example is the complete r-partite graph with $r-2$ classes of size k, one class of size $k-1$ and one class of size $k+1$, where $k = n/r$. See Figure 1.3. (Note that this example can be obtained from the extremal example for K_2 by adding on a copy of $T_{r-2}(n-2k)$ and joining all the new vertices to all the old ones.)

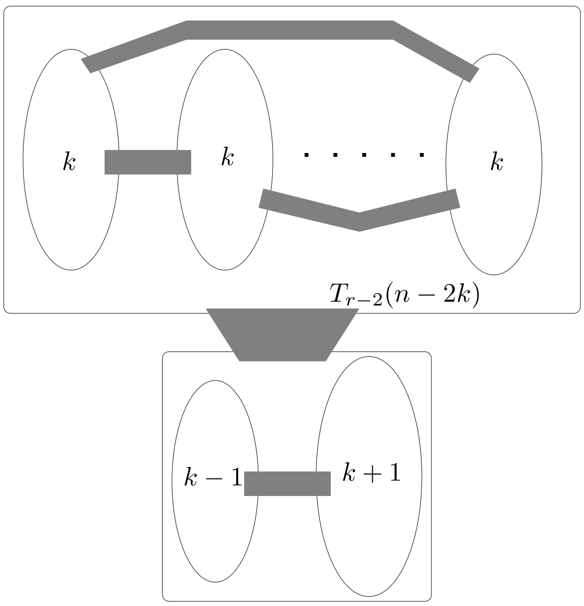


Figure 1.3: The extremal graph for K_r

This graph does not contain a perfect K_r-packing, and has minimum degree $(1-1/r)n-1$. Thus we might guess that the following result should be the correct one.

Theorem 1.11 (Hajnal, Szemerédi, [36]) *For all integers r and all integers n divisible by r, $\delta(n, K_r) = (1 - 1/r)n$.*

The example above demonstrates that $\delta(n, K_r) \geq (1 - 1/r)n$. The case $r = 3$ of Theorem 1.11 was first proved by Corrádi and Hajnal [19], and the full theorem was proved by Hajnal and Szemerédi in 1970. Comparing this with the single subgraph result which comes from Turán's Theorem, we have

$$\delta_0(n, K_r) = \left(1 - \frac{1}{r-1}\right) n$$
$$\delta(n, K_r) = \left(1 - \frac{1}{r}\right) n$$

whenever n is divisible by $r - 1$ or r respectively. Thus we might hypothesise that the extension of this result to the perfect packings analogue of the Erdős-Stone-Simonovits Theorem would be:

Possible Conjecture 1.12 *For all graphs H and all integers n divisible by $|H|$, $\delta(n, H) = (1 - \frac{1}{\chi(H)})n$.*

However, this conjecture can easily be seen to be false.

Proposition 1.13 *Let $H = K_{3,3}$, the complete bipartite graph with two classes of size 3. Then for each $k \in \mathbb{N}$, there is a graph G on $n = 6k$ vertices with minimum degree $n/2 + 1$ which does not contain a perfect $K_{3,3}$-packing.*

Proof. Construct G as follows: The vertex set of G consists of a set A of size $3k + 1$ and a set B of size $3k - 1$. The edges of G will be all the edges between A and B, along with a cycle on the vertices of A (which is possible as $|A| \geq 4$). See Figure 1.4.

Note that every vertex of G has degree $3k + 1 = n/2 + 1$. Furthermore, if G contained a perfect $K_{3,3}$-packing, one copy of $K_{3,3}$ would have to meet A in at least

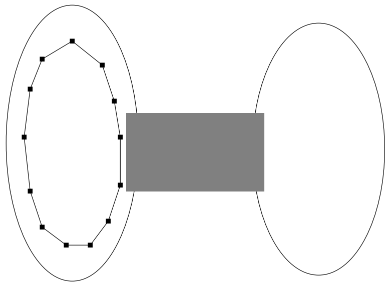


Figure 1.4: The extremal graph for $K_{3,3}$

4 vertices, and so A would contain vertices from both of the classes of $K_{3,3}$. Let X and Y be the classes of $K_{3,3}$. Furthermore, since B is independent, B could contain only vertices from one class. Without loss of generality, B contains only vertices from Y. And so X would lie completely in A, along with at least one vertex from Y. This vertex would then have degree 3 in A, which is impossible as A only contains a cycle. $\square$

However, we might attempt to modify the conjecture slightly.

Possible Conjecture 1.14 *For all graphs H and all integers n divisible by $|H|$, $\delta(n, H) \sim (1 - \frac{1}{\chi(H)})n$.*

One of the first partial results towards this conjecture is the following:

Theorem 1.15 (Alon, Yuster [4]) *Given any graph H and any real number $\varepsilon > 0$, there exists $n_0 = n_0(H, \varepsilon)$ such that for any $n \geq n_0$ divisible by $|H|$, if G is a graph on n vertices with minimum degree $\delta(G) \geq (1 - \frac{1}{\chi(H)} + \varepsilon)n$, then G contains a perfect H-packing. In other words, $\delta(n, H) \leq (1 - \frac{1}{\chi(H)} + \varepsilon)n$.*

Alon and Yuster conjectured an improvement to this theorem, which was proved by Komlós, Sárközy and Szemerédi:

Theorem 1.16 (Komlós, Sárközy, Szemerédi [48]) *Given any graph H there is a constant $C = C(H)$ dependent only on H such that for any n divisible by $|H|$, if G is a graph on n vertices with minimum degree $\delta(G) \geq (1 - \frac{1}{\chi(H)})n + C(H)$, then G contains a perfect H-packing. In other words $\delta(n, H) \leq (1 - \frac{1}{\chi(H)})n + C(H)$.*

This might suggest that the conjecture is indeed true. However, this is misleading, as in some cases we can improve substantially on this upper bound. For example, the El Zahar conjecture, proved by Abbasi, states the following:

Theorem 1.17 (Abbasi [1]) *Let $n = n_1 + n_2 + \ldots + n_k$, and let G be a graph on n vertices satisfying $\delta(G) \geq \sum \lceil n_i/2 \rceil$. Then G contains k vertex disjoint cycles of orders $n_1, n_2, \ldots, n_k$.*

In particular, if n is divisible by k and if we let $n_1 = n_2 = \ldots n_k = n/k =: h$, then the vertex disjoint cycles are precisely a perfect C_h-packing. In the case when h is odd, this means that the minimum degree required to guarantee such a packing is certainly no more than $k(h+1)/2 = n(h+1)/2h$, which is considerably less (for $h \geq 5$) than the asymptotic value of $2n/3$ given by the conjecture. Thus some more refined theorem is needed.

In order to introduce the desired theorem, we need to make some definitions. Given a graph H of chromatic number $\chi(H)$, let $\sigma(H)$ denote the smallest possible size of a colour class in a $\chi(H)$-colouring of H.

Definition 1.18 *The* critical chromatic number of H *is denoted by $\chi_{cr}(H)$, and is defined by*

$$\chi_{cr}(H) := \frac{\chi(H) - 1}{|H| - \sigma(H)}|H|.$$

Note that $\chi(H) - 1 < \chi_{cr}(H) \leq \chi(H)$, and that $\chi_{cr}(H)$ is closer to $\chi(H) - 1$ if $\sigma(H)$ is comparatively small. Thus the critical chromatic number in some sense measures how uneven the colour class sizes are, as well as how many are required.

Komlós proved that if we only require an almost perfect packing, then the critical chromatic number replaces the chromatic number as the relevant parameter in all cases.

Theorem 1.19 (Komlós, [46]) *For any graph H and any $\varepsilon > 0$ there is an integer $n_0 = n_0(H, \varepsilon)$ such that if G is a graph on $n \geq n_0$ vertices and if $\delta(G) \geq (1 - \frac{1}{\chi_{cr}(H)})n$ then G contains an H-packing covering at least $(1 - \varepsilon)n$ vertices.*

[1] The result we are aiming towards will state that for some graphs, the appropriate minimum degree to guarantee a perfect H-packing is also approximately $(1 - \frac{1}{\chi_{cr}(H)})n$. Before we can state the result formally, though, we need some more definitions.

Let $\ell := \chi(H)$. Given an ℓ-colouring c of H, let $x_1 \geq x_2 \geq \ldots \geq x_\ell$ be the sizes of the colour classes. Define $\mathcal{D}(c) = \{x_i - x_{i+1} \mid i = 1, \ldots, \ell - 1\}$. Let $\mathcal{D}(H)$ be the union of all the sets $\mathcal{D}(c)$ over all optimal colourings c of H. We define $hcf_\chi(H)$ to be the highest common factor of the elements of $\mathcal{D}(H)$ (or $hcf_\chi(H) := \infty$ if $\mathcal{D}(H) = \{0\}$). Define $hcf_c(H)$ to be the highest common factor of the orders of all the components of H.

Definition 1.20 *For any graph H, if $\chi(H) \neq 2$, we say $hcf(H) = 1$ if $hcf_\chi(H) = 1$. If $\chi(H) = 2$, we say $hcf(H) = 1$ if both $hcf_c(H) = 1$ and $hcf_\chi(H) \leq 2$.*

This may appear at first sight to be an artificial and unnatural definition, but I will briefly give a few examples to give some idea why this is an appropriate definition when attempting to characterise those graphs H for which the critical chromatic number is the parameter governing perfect H-packings. In particular, for each of the conditions required on H for $hcf(H) = 1$, if the condition does not hold

[1] In fact, Komlós' result was much more general than this, and provided asymptotically the minimum degree condition necessary to guarantee a packing covering xn vertices for any $x \in (0, 1)$.

then I will give an example of a graph G of minimum degree $\delta(G) \geq (1-1/\chi(H))n-2$ which does not contain a perfect H-packing.

If H is not a bipartite graph, then $hcf(H) = 1$ if and only if $hcf_\chi(H) = 1$. So suppose this does not hold. Let $\ell := \chi(H)$ and let G be the complete ℓ-partite graph on $n = k\ell$ vertices, where $|H|$ divides k, with $\ell - 2$ classes of size k, one class of size $k+1$ and one of size $k-1$. (Note that this is the same graph used in Section 1.1 to show that the bound in the Hajnal-Szemerédi Theorem is best possible.) It is fairly easy to see that this graph does not contain an H-packing. Roughly, when taking out copies of H we cannot even out the sizes of a class originally of size k and a class originally of size $k+1$. More precisely, we have a class A of size k and a class B of size $k+1$. Set $d = hcf_\chi(H)$. Then $|B| - |A| \equiv 1 \mod d$. Furthermore, this holds even if we have modified A and B by taking some copies of H from G, since the difference in the number of vertices taken from each of these sets is always a multiple of d. Now since $d > 1$, for any sets A' and B' obtained in this way we have $|A'| \neq |B'|$. But if a perfect H-packing existed, its removal would leave $A' = B' = \phi$, which is impossible. So no perfect H-packing exists. Yet this graph satisfies $\delta(G) = (1 - 1/\ell)n - 1$.

Now if H is a bipartite graph, to guarantee $hcf(H) = 1$ we require the weaker condition $hcf_\chi(H) \leq 2$, but we also need $hcf_c(H) = 1$. To see that the first condition is necessary we suppose it does not hold and we look at the complete bipartite graph on $n = 2k$ vertices with one set of size $k-1$ and one of size $k+1$. Now similarly as in the non-bipartite case no perfect H-packing exists (this time we cannot even out the the classes of size $k-1$ and $k+1$ because $hcf_\chi(H) > 2$), and yet $\delta(G) = n/2 - 1$.

On the other hand, if $hcf_c(H) \neq 1$, we consider the graph G on $n = 2k$ vertices consisting of the disjoint union of two cliques, one of order $k-1$ and one of order $k+1$. We also choose k to be divisible by $|H|$. Once again we can show that no perfect

H-packing exists. For suppose that $c_1, c_2, \ldots, c_m$ are the sizes of the components of H. Then if it perfect H-packing exists, there are integers $a_1, a_2, \ldots, a_m$ such that $\sum_{i=1}^m a_i c_i = k+1$ (let a_i be the number of times the component of size c_i appears in the clique of size $k+1$). On the other hand, $|H| = c_1 + c_2 + \ldots + c_m$, and so $k = \sum_{i=1}^m (k/|H|)c_i$. Therefore

$$1 = \sum_{i=1}^{m} (a_i - k/|H|)c_i.$$

But since $k/|H|$ is also an integer, this shows that $hcf\{c_1, \ldots, c_m\} = hcf_c(H) = 1$, which is a contradiction. Thus no perfect H-packing exists, but still $\delta(G) = n/2 - 2$.

These examples show that if $hcf(H) \neq 1$, then $\delta(n, H) \geq (1 - 1/\chi(H))n - 1$. Together with Theorem 1.16 this shows that for such graphs, $\delta(n, H) = (1 - 1/\chi(H))n + O(1)$. The question of what happens for those graphs H with $hcf(H) = 1$ is answered by the following theorem.

Theorem 1.21 (Kühn, Osthus [53]) *For any graph H*

$$\delta(n, H) = \begin{cases} \left(1 - \frac{1}{\chi_{cr}(H)}\right) n + O(1) & \text{if } hcf(H) = 1, \\ \left(1 - \frac{1}{\chi(H)}\right) n + O(1) & \text{if } hcf(H) \neq 1. \end{cases}$$

Here the $O(1)$ error term is bounded by a constant depending only on H. Note that $\chi(H) = \chi_{cr}(H)$ would mean that $hcf_\chi(H) = \infty$, and in particular $hcf(H) \neq 1$. So when $hcf(H) = 1$, the value for $\delta(n, H)$ given by Theorem 1.21 is indeed an improvement on the upper bound given by Theorem 1.16.

Thus we now have equality in all cases, and so the result is best possible up to the $O(1)$ error term. The natural next step is to ask when this error term can be removed entirely. The first of the three main results of this thesis states that the error term can be removed completely in the case when $H = K_r^-$, the graph obtained from K_r by removing one edge. The proof of this result will form the main part of Chapter 2. Observe that $\chi(K_r^-) = r - 1$ and that $\sigma(K_r^-) = 1$. Thus $\chi_{cr}(K_r^-) = \frac{r(r-2)}{r-1}$. Note also that $hcf(K_r^-) = 1$ for $r \geq 4$, and so the result is:

Theorem 2.1 *For every integer $r \geq 4$ there exists an integer $n_0 = n_0(r)$ such that every graph G whose order $n \geq n_0$ is divisible by r and whose minimum degree is at least*

$$\left(1 - \frac{1}{\chi_{cr}(K_r^-)}\right) n$$

contains a perfect K_r^--packing.

This theorem confirms a conjecture of Kawarabayashi [42] for large n. The case $r = 4$ of the conjecture (and thus of Theorem 2.1) was proved by Kawarabayashi [42]. By a result of Enomoto, Kaneko and Tuza [25], the conjecture also holds for the case $r = 3$ under the additional assumption that G is connected. (Note that K_3^- is just a path on 3 vertices and that in this case the required minimum degree equals $n/3$.)

The proof of this theorem which appears in Chapter 2 was also given in [16]. In Section 2.6 I will give a brief sketch of how the proof can be extended to a larger class of graphs H satisfying certain conditions to remove the error term completely in these cases. I will also give some examples to show that some of the conditions are necessary, i.e. that if they do not hold then the $O(1)$ error term in Theorem 1.21 cannot be removed without making the theorem false. Although this is not yet a complete classification, it goes some way towards a classification of which graphs H require some error term and which do not.

1.2 Graph Embeddings

A natural extension of the packings question is the embedding problem. In this case, we again aim to embed a graph H into a graph G, but now the order of H might be linear in $n = |G|$ rather than fixed. The extreme of this problem is of course the case when $|H| = |G|$, i.e. when H is a spanning subgraph of G.

We could view the problem of finding a perfect H-packing as an embedding problem: If we let H' be the graph consisting of k disjoint copies of H, where $k = n/|H|$, then $|H'| = |G|$, and finding a perfect H-packing in G is equivalent to finding a copy of H' in G. In general, though, we allow H' to have a much less regular structure. We do, however, require some constraints on what H can look like. Typically we seek to bound parameters such as the maximum degree, the chromatic number or the bandwidth.

Definition 1.22 (bandwidth) *Given a graph G and an ordering, $L = v_1, v_2, \ldots, v_n$ of the vertices of G, we define $b(G, L)$ to be the largest integer k such that for some i, $v_i v_{i+k} \in E(G)$. In other words, $b(G, L)$ is the longest distance in L between two vertices which are adjacent in G. The* bandwidth *of G, denoted $b(G)$, is defined as $\min_L b(G, L)$, where the minimum is taken over all possible orderings L.*

Perhaps the simplest embedding result is Dirac's theorem, mentioned earlier, which allows us to embed a Hamilton cycle into G. (Note that a Hamilton cycle has bandwidth 2.) A generalisation of this is the Pósa-Seymour conjecture:

Conjecture 1.23 (Seymour, [68]) *For any k, if G is a graph on n vertices with minimum degree satisfying $\delta(G) \geq \frac{k}{k+1}n$, then G contains the k-th power of a Hamilton cycle.*

Here the k-th power of a Hamilton cycle is obtained from a cyclic ordering of the vertices by joining all those vertices at distance at most k in the ordering. Thus the

k-th power of a Hamilton cycle has bandwidth $2k$. The case $k = 2$ was originally conjectured by Pósa in 1962. Note that this conjecture, if true, would automatically imply the Hajnal-Szemerédi theorem, and therefore the same example as before shows that it is best possible. Conjecture 1.23 was proved by Komlós, Sárközy and Szemerédi [49] in an approximate form, in which the minimum degree condition of G had an extra factor of εn, and then the same authors proved the conjecture for large graphs [50].

These results only allowed for a constant sized bandwidth. Böttcher, Schacht and Taraz proved a result which generalises an approximate form of the Pósa-Seymour conjecture. Here the bandwidth is allowed to grow linearly, but the maximum degree and the chromatic number must still be bounded by a constant.

Theorem 1.24 (Böttcher, Schacht, Taraz [9]) *For any real number $\varepsilon > 0$, and any integers $r \geq 2$ and Δ, there is a real number $\beta > 0$ and an integer n_0 such that the following holds. If G is a graph on $n \geq n_0$ vertices with minimum degree $\delta(G) \geq (1-\frac{1}{r}+\varepsilon)n$, and if H is a graph also on n vertices with $\chi(H) = r$, $\Delta(H) \leq \Delta$ and bandwidth at most βn, then G contains a copy of H as a subgraph.*

This result was originally conjectured by Bollobás and Komlós. The case $r = 2$ of, i.e. when H is a bipartite graph, was proved by Abbasi [2]. Böttcher, Schacht and Taraz [8] proved the case $r = 3$ before proving the full conjecture for general r.

Bollobás and Eldridge [7] also conjectured the following result:

Conjecture 1.25 (Bollobás, Eldridge) *Let G and H be two graphs each on n vertices, and suppose $\delta(G) \geq \frac{\Delta(H)n-1}{\Delta(H)+1}$. Then H can be embedded into G.*

Note that this conjecture, if true, would also automatically imply the Hajnal-Szemerédi Theorem if we set H to be the disjoint union of r-cliques. Until recently only a few special cases of Conjecture 1.25 have been proved (see e.g. [20] for more details). Recently Kun has announced a proof of an asymptotic version of the conjecture, in which there is a small linear error term in the minimum degree required on G, and also a lower bound on both $\Delta(H)$ and $n - \delta(G)$.

Further areas of interest arise when we consider the problem of embedding non-spanning subgraphs H. In particular in this thesis I will be concerned with the case when H is a tree, i.e. a connected graph with no cycles. We denote by $\mathcal{T}_k$ the set of trees on $k+1$ vertices; it is a basic graph theory result that such a tree contains k edges. We write $\mathcal{T}_k \subseteq G$ if $T \subseteq G$ for all $T \in \mathcal{T}_k$, i.e. if G contains as a subgraph every tree on $k+1$ vertices. The following result is a trivial application of a greedy embedding algorithm.

Fact 1.26 *$\delta(G) > k - 1 \Rightarrow \mathcal{T}_k \subseteq G$.*

However, this can be substantially improved upon. Perhaps the most attractive potential strengthening of this result is the famous Erdős-Sós conjecture, which replaces the minimum degree by the average degree. Let $d(G) := \frac{1}{|G|} \sum_{x \in V(G)} d(x)$ denote the average degree of a vertex in G.

Note that even in the case when $|G| \gg k$, the Erdős-Stone-Simonovits Theorem (Theorem 1.4) does not provide any useful information about the number of edges (and therefore also about the average degree) required to guarantee a copy of a tree. This is because a tree is bipartite and the theorem becomes degenerate, with the $\varepsilon\binom{n}{2}$ error term becoming dominant. However, the following conjecture would provide an asymptotic condition.

Conjecture 1.27 (Erdős, Sós, 1963) *Suppose G is a graph satisfying $d(G) > k - 1$. Then $\mathcal{T}_k \subseteq G$.*

The conjecture is trivial for stars, since in order to embed a star with $k+1$ vertices we need only find a vertex of degree at least k onto which to embed the central point and then the k remaining vertices can be embedded among its neighbours; such a vertex certainly exists if $d(G) > k - 1$. On the other hand, stars also show that the bound cannot be improved in general, since we can construct a $(k-1)$-regular graph on n vertices provided that at least one of n and $k-1$ is even. Then the average degree is exactly $k-1$ and there is no vertex of degree at least k onto which to embed the central point of a star.

Some further special cases of this conjecture have been resolved. For example, McLennan [56] proved the conjecture when we restrict our attention only to trees of diameter at most 4 (the class includes stars, the only trees of diameter 2). On the other hand, Sacle and Woźniak [67] proved the conjecture with the additional assumption that G does not contain a copy of C_4, the cycle on 4 vertices. Ajtai, Komlós, Simonovits and Szemerédi have announced a proof of the conjecture in the case when k is sufficiently large.

The focus of Chapter 3 of this thesis is the Loebl-Komlós-Sós conjecture, which replaces the average degree in the Erdős-Sós conjecture with the median degree.

Conjecture 1.28 (Loebl, Komlós, Sós [26]) *Given any integers k and n, if G is a graph on n vertices in which at least $n/2$ vertices have degree at least k, then G contains as subgraphs all trees with k edges.*

Loebl's initial conjecture covered only the special case $k = n/2$, and is sometimes known as the $n/2 - n/2 - n/2$ conjecture. Komlós and Sós then extended the conjecture to all k.

Again, the conjecture is trivially true for stars, and stars show that the degree condition cannot be relaxed to $k-1$. In Chapter 3 we will also see that the number of vertices of degree k cannot be significantly reduced for large k.

Various partial results towards Conjecture 1.28 have been proved. Dobson [23] proved the conjecture with the additional assumption that the complement of G

does not contain a copy of $K_{2,3}$, while Soffer [69] proved that the conjecture is true for graphs G of girth at least 7. There have also been several partial results which make some additional assumptions about the trees to be embedded into G. Zhao [72] proved the special case when $k = n/2$, provided n is sufficiently large. Bazgan, Li and Woźniak [6] proved the conjecture for paths, i.e. that if a graph G satisfies the conditions of the conjecture, then it contains the path on $k+1$ vertices as a subgraph. In the same paper, they also proved the conjecture in the case when $k \geq n-3$. Barr and Johansson [5] and independently Sun [70] proved the conjecture when restricting attention to trees of diameter 4. Improving on this, Piguet and Stein [62] proved the Loebl-Komlós-Sós conjecture for trees of diameter at most 5, and in [61] they proved an approximate version of the full conjecture (with linear error terms in both the number of vertices with high degree and in the degree of those vertices) for n sufficiently large and for k linear in n, i.e. for large, dense graphs. The main theorem of Chapter 3 is a proof of the exact conjecture for large, dense graphs.

Theorem 3.1 *Given a positive $C' \in \mathcal{R}$ there exists $k_0 \in \mathbb{N}$ such that for any integers $k, n \in \mathbb{N}$ satisfying $k_0 \leq k \leq n \leq C'k$ the following holds: Suppose G is a graph on n vertices in which at least $n/2$ vertices have degree at least k. Then G contains as a subgraph every tree with k edges.*

The proof of this theorem presented in Chapter 3 also appeared in [15], although this thesis contains some details which were omitted from that paper. Theorem 3.1 is a partial result in the sense that it only holds for large k and n, and we demand that k is linear in n. These restrictions come about because the proof makes use of Szemerédi's regularity lemma. The same result was also proved independently by Hladký and Piguet [40].

The Loebl-Komlós-Sós conjecture has a beautiful application to Ramsey numbers of trees. For a graph H we define the Ramsey number $R(H)$ to be the least integer

n such that if the edges of K_n are 2-coloured then there is a monochromatic copy of H. Thus the usual Ramsey number $R(k)$ is just $R(K_k)$. More generally, for graphs F, H, we define $R(F, H)$ to be the least integer n such that if the edges of K_n are coloured red and blue then there is a red copy of F or a blue copy of H.

More generally still, for families of graphs $\mathcal{F}$ and $\mathcal{H}$, let $R(\mathcal{F}, \mathcal{H})$ denote the smallest integer n such that if the edges of K_n are coloured red and blue then there is a red copy of F for every $F \in \mathcal{F}$ or else a blue copy of H for every $H \in \mathcal{H}$.

For most graphs H the best known upper bounds on the Ramsey number are exponential in $|H|$. For complete graphs $H = K_k$, the lower bound is also exponential in k, a fact which was first proved by Erdős. However, Conjecture 1.28 would give the following corollary.

Conjecture 1.29 *For any positive integers p and q, $R(\mathcal{T}_p, \mathcal{T}_q) \leq p + q$.*

Since Theorem 3.1 provides a partial version of Conjecture 1.28, it also gives a partial version of Conjecture 1.29 as a corollary.

Theorem 3.2 *For any real number $C'' \geq 1$ there exists an integer p_0 such that for any integers p and q satisfying $p_0 \leq p \leq q \leq C''p$ we have $R(\mathcal{T}_p, \mathcal{T}_q) \leq p + q$. In particular, for $T_p \in \mathcal{T}_p$ and $T_q \in \mathcal{T}_q$, $R(T_p, T_q) \leq p + q$.*

In general, Ramsey numbers are famously difficult to calculate. However, as we will see in Chapter 3, Theorem 3.2 is in fact best possible up to an error term of 1 in some cases.

The proof of Theorem 3.2 given Theorem 3.1 is relatively short and simple, and will be presented in Section 3.1.

In a similar vein, Chvátal, Rödl, Szemerédi and Trotter proved an embedding result for graphs which, as a corollary, shows that graphs of bounded degree have linear Ramsey numbers.

Theorem 1.30 (Chvátal, Rödl, Szemerédi, Trotter, [12]) *For any integer* Δ *there is a real number* $c > 0$ *such that if* G *is any graph on* n *vertices, and* H *is any graph satisfying* $|V(H)| \leq cn$ *and* $\Delta(H) \leq \Delta$*, then either* $H \subseteq G$ *or* $H \subseteq \overline{G}$*.*

Corollary 1.31 *For any integer* Δ *there is a constant* $C = C(\Delta)$ *such that if* H *is a graph with maximum degree satisfying* $\Delta(H) \leq \Delta$*, then* $R(H) \leq C|H|$*.*

The aim of Chapter 4 is to prove an embedding result along the lines of Theorem 1.30, and thus also to generalise Corollary 1.31, for hypergraphs.

1.3 Hypergraph Embeddings

A *hypergraph* is a generalisation of a graph. While a graph consisted of vertices and edges, a hypergraph consists of vertices and hyperedges. The *hyperedges* of a k-uniform hypergraph are unordered k-tuples of distinct vertices in the vertex set. Thus a graph is simply a 2-uniform hypergraph. Given a k-uniform hypergraph $\mathcal{G}$, its set of vertices is usually denoted V or $V(\mathcal{G})$, and its set of hyperedges by $E(\mathcal{G})$ or $E_k(\mathcal{G})$. We denote by $|\mathcal{G}|$ the number of its vertices and write $e(\mathcal{G}) := |E(\mathcal{G})|$. We say that vertices $x, y \in \mathcal{G}$ are *neighbours* if x and y lie in a common hyperedge of $\mathcal{G}$. Just as in the graph case, the degree of a vertex $x \in V(\mathcal{H})$ is the number of neighbours of x in $\mathcal{H}$. The minimum degree and maximum degree of a hypergraph $\mathcal{H}$ are then defined in the obvious way.

Broadly speaking, we consider the problem of embedding a hypergraph $\mathcal{H}$ into a larger hypergraph $\mathcal{G}$, i.e. of finding a subhypergraph of $\mathcal{G}$ which is isomorphic to $\mathcal{H}$.

The problem is similar to the discussion in the previous section in that the order of $\mathcal{H}$ will be linear in the order of $\mathcal{G}$. In order to have any chance of embedding $\mathcal{H}$ into $\mathcal{G}$, we must have some sort of restrictions on what $\mathcal{H}$ can look like. In the last of the three main results of this thesis, we demand that the maximum degree of $\mathcal{H}$ is bounded. However, the hypergraph case is considerably more complicated than the graph case, and for this reason embedding results for hypergraphs $\mathcal{H}$ whose order is exactly $|\mathcal{G}|$, or even close to $|\mathcal{G}|$, have been out of reach until very recently. In Chapter 4 I will present the proof of an embedding result for the case when $\mathcal{H}$ has bounded degree and order $c|\mathcal{G}|$, where c is a very small positive constant and where $|\mathcal{G}|$ is large. The bulk of the proof appeared in [18], but I have added some details which were omitted in that paper.

The complete k-uniform hypergraph on n vertices (i.e. the hypergraph in which all possible k-tuples form a hyperedge) is denoted $K_n^{(k)}$, and the Ramsey number of a hypergraph $\mathcal{H}$ is the least integer n such that whenever the hyperedges of $K_n^{(k)}$ are two-coloured then there exists a monochromatic copy of $\mathcal{H}$.

For general $\mathcal{H}$, the best upper bound on $R(\mathcal{H})$ is due to Erdős and Rado [27]. Writing $|\mathcal{H}|$ for the number of vertices of $\mathcal{H}$, it implies that for any $k \geq 2$

$$R(\mathcal{H}) \leq 2^{2^{\cdot^{\cdot^{2^{c_k|\mathcal{H}|}}}}},$$

where the number of 2's is $k-1$. In the other direction, Erdős and Hajnal (see [33]) showed that if $k \geq 3$ and $\mathcal{H}$ is a complete k-uniform hypergraph then $R(\mathcal{H})$ is bounded below by a tower in which the number of 2's is $k-2$ and the top exponent is $c'_k|\mathcal{H}|^2$.

However, an application of the embedding result in [18] shows that hypergraphs of bounded degree have linear Ramsey numbers, i.e. a hypergraph analogue of Corollary 1.31.

Theorem 4.1 *For all $\Delta, k \in \mathbb{N}$ there exists a constant $C = C(\Delta, k)$ such that all k-uniform hypergraphs $\mathcal{H}$ of maximum degree at most Δ satisfy $R(\mathcal{H}) \leq C|\mathcal{H}|$.*

This is an improvement on a result of Kostochka and Rödl [52], who showed that

Ramsey numbers of k-uniform hypergraphs of bounded maximum degree are 'almost linear' in their orders. More precisely, they showed that for all $\varepsilon, \Delta, k > 0$ there is a constant C such that $R(\mathcal{H}) \leq C|\mathcal{H}|^{1+\varepsilon}$ if $\mathcal{H}$ has maximum degree at most Δ.

The case $k = 3$ of Theorem 4.1 was earlier proved in [17] and also independently in [57]. Also, Haxell, Łuczak, Peng, Rödl, Ruciński, Simonovits and Skokan [37, 38] asymptotically determined the Ramsey numbers of 3-uniform tight and loose cycles. Ramsey numbers of Berge-cycles were considered in [35] and [24].

After the submission of [18], Conlon, Fox and Sudakov [13] obtained a version of Theorem 4.1 whose proof does not use the hypergraph regularity lemma, and therefore gives a much better upper bound on the value of $C(\Delta, k)$. The same authors [14] also improved the upper and lower bounds of Erdős, Hajnal and Rado for complete hypergraphs.

The overall strategy of our proof of Theorem 4.1 is related to that of Chvátal, Rödl, Szemerédi and Trotter [12], which is based on the regularity lemma for graphs. We apply a version (due to Rödl and Schacht [64]) of the regularity lemma for k-uniform hypergraphs. Roughly speaking, it guarantees a partition of an arbitrary dense k-uniform hypergraph into 'quasi-random' subhypergraphs. Our main contribution is an embedding result (Theorem 4.2) which guarantees the existence of a copy of a hypergraph $\mathcal{H}$ of bounded maximum degree inside a suitable 'quasi-random' hypergraph $\mathcal{G}$ even if the order of $\mathcal{H}$ is linear in that of $\mathcal{G}$. In fact, we prove a stronger embedding result of independent interest (Theorem 4.3). It even counts the number of copies of such $\mathcal{H}$ in $\mathcal{G}$ and thus generalises the well-known hypergraph counting lemma (which only allows for bounded size $\mathcal{H}$).

After the submission of [18], Keevash [43] extended Theorem 4.2 to a hypergraph blow-up lemma for embeddings of spanning subhypergraphs $\mathcal{H}$. The case of 3-uniform hypergraphs in Theorem 4.1 was proved recently in [17] and independently by Nagle, Olsen, Rödl and Schacht [57]. Also, Kostochka and Rödl [52] earlier proved an approximate version of Theorem 4.1: for all $\varepsilon, \Delta, k > 0$ there is a constant C such that $R(\mathcal{H}) \leq C|\mathcal{H}|^{1+\varepsilon}$ if $\mathcal{H}$ has maximum degree at most Δ. After [18] was submitted, Conlon, Fox and Sudakov [13] obtained a proof of Theorem 4.1 which does not rely on hypergraph regularity and gives a better bound on C. Also, Ishigami [41] independently announced a proof of Theorem 4.1 using a similar approach to ours. Apart from these, the only previous results on the Ramsey numbers of sparse hypergraphs are on hypergraph cycles (see e.g. [24, 35, 37, 38]).

It would be desirable to extend Theorem 4.1 to a larger class of hypergraphs. For instance the graph analogue of Theorem 4.1 is known for so-called p-arrangeable graphs [11], which include the class of all planar graphs. However, Rödl and Kostochka [52] showed that a natural hypergraph analogue of the famous Burr-Erdős conjecture on Ramsey numbers of d-degenerate graphs fails for k-uniform hypergraphs if $k \geq 3$. (A graph is d-degenerate if the maximum average degree over all its subgraphs is at most d. If a graph is p-arrangeable, then it is also d-degenerate for some d.) But it may still be possible to generalise the Burr-Erdős conjecture to hypergraphs in a different way.

1.4 The Regularity Method

The common theme in graph packing, graph embedding and hypergraph embedding problems is the regularity lemma. Roughly speaking, the regularity lemma states that any sufficiently large and dense graph or hypergraph can have its vertex set partitioned into a small number of classes in such a way that the sub(hyper)graphs induced between the classes look very much like random (hyper)graphs. By applying the regularity lemma to a graph G or a hypergraph $\mathcal{G}$, we hope to use this pseudo-randomness to embed H' or $\mathcal{H}$ into G or $\mathcal{G}$ respectively.

The regularity lemma for hypergraphs is rather more complicated than that for graphs, and so I will introduce the two separately. The hypergraph version will be left until Chapter 4, when it is first needed. In Section 1.4.1 I will introduce the regularity lemma for graphs, originally proved by Szemerédi [71]. The subsections following will give a brief idea of the method of proof of most graph embedding results. One important step in this method uses the blow-up lemma, due to Komlós, Sárkőzy and Szemerédi, which is introduced in Section 1.4.3.

1.4.1 The Regularity Lemma for Graphs

Definition 1.32 (ε-regular pair) *Given $\varepsilon > 0$ and a bipartite graph G with vertex classes A, B, we say the pair (A, B) is* ε-regular *if for any subsets $X \subseteq A$, $Y \subseteq B$ satisfying $|X| \geq \varepsilon|A|, |Y| \geq \varepsilon|B|$ we have*

$$|d(X, Y) - d(A, B)| \leq \varepsilon.$$

In other words, for all sufficiently large subsets of A and B, the density of edges between them is roughly the same as the density between the whole of A and B. More generally, given a graph G and disjoint vertex sets A and B within G (not necessarily covering all of $V(G)$), we say the pair (A, B) is ε-regular if they form an ε-regular pair in the bipartite graph induced between them.

Occasionally we need a slightly stronger definition of regularity.

Definition 1.33 ((ε, δ)-super-regular pair) *We say the pair (A, B) is (ε, δ)-super-regular if it is ε-regular, and furthermore each vertex in A has at least $\delta|B|$ neighbours in B, and similarly each vertex in B has at least $\delta|A|$ neighbours in A.*

Super-regularity ensures that there are no 'very bad' vertices, which have almost no neighbours. It is a basic fact (see e.g. [51]) that in any ε-regular pair, at most $\varepsilon|A|$

vertices in A have degree at most $(d(A,B) - \varepsilon)|B|$ and vice versa. So by removing these vertices we can ensure that the pair is $(2\varepsilon, d - 2\varepsilon)$-super-regular.

We also sometimes use the notion of (d, ε)-regularity:

Definition 1.34 ((d, ε)-regular pair) *We say a pair (A, B) is (d, ε)-regular if for any sets $X \subseteq A$, $Y \subseteq B$ satisfying $|X| \geq \varepsilon|A|, |Y| \geq \varepsilon|B|$ we have*

$$|d(X, Y) - d| \leq \varepsilon.$$

It can easily be seen that this definition is roughly equivalent to the definition of an ε-regular pair given an appropriate choice of d. Indeed, it is exactly equivalent for $d = d(A, B)$, while for general d any (d, ε)-regular pair is also 2ε-regular. Thus subject to the deletion of a few vertices of low degree, this definition is also very similar to the definition of an (ε, δ)-super-regular pair.

Definition 1.35 (ε-regular partition) *Given a graph G and a partition P of $V(G)$ into sets $V_1, V_2, \ldots, V_k$, we say that P is an* ε-regular partition of G *if all but at most $\varepsilon\binom{k}{2}$ of the pairs (V_i, V_j) are ε-regular.*

In other words, almost all pairs are ε-regular. Roughly speaking, the regularity lemma states that any sufficiently large and dense graph has an ε-regular partition into a bounded number of sets of almost equal size.

Theorem 1.36 (Szemerédi's Regularity Lemma, 1978) *Given any integer k_0 and a real number $\varepsilon > 0$, there are integers $n_0 = n_0(k_0, \varepsilon)$ and $K_0 = K_0(k_0, \varepsilon)$ such that if G is a graph on $n \geq n_0$ vertices, then there is a partition P of $V(G)$ into sets $V_1, V_2, \ldots, V_k$ such that*

- $k_0 \leq k \leq K_0$
- $|V_i| - |V_j| \leq 1$ *for any $i, j \in [k]$*

- *P is an ε-regular partition.*

We call the classes V_i of the partition P *clusters*. Note that if we have such a partition, then the number of edges within clusters is not much more than $k(n/k)^2 = n^2/k$, and the number of edges between pairs of clusters that are not ε-regular is not much more than $\varepsilon\binom{k}{2}(n/k)^2 \leq \varepsilon n^2$. [1] Thus if we know that a graph G on n vertices has at least cn^2 edges, for some constant $c > 0$ and sufficiently large n, then we can apply the regularity lemma with sufficiently small ε and large k_0 (dependant on c) to ensure that a very large proportion of the edges of G run between ε-regular pairs. We tend to ignore all the remaining edges.

We note also that while the regularity lemma is stated for all graphs, for sparse graphs it becomes trivial. More precisely, if (G_n) is a sequence of graphs, where G_n has n vertices and $o(n^2)$ edges, then asymptotically any partition of G_n into sets of the appropriate size will satisfy the conditions of the lemma; we could have all edges being either within a cluster or between non-regular pairs. Some work has been done towards generalising the regularity lemma to an appropriate sparse version, but so far only partial results have been proved. See e.g. [30] for for a survey of the known results in this area.

For these reasons, we usually only apply the regularity to graphs with at least cn^2 edges.

1.4.2 The Reduced Graph

One very important concept in many applications of the regularity lemma is that of the *reduced graph*, which reflects the rough structure of the original graph.

Definition 1.37 (reduced graph) *Given parameters $d, \varepsilon \in (0, 1)$, a graph G and*

[1] We say 'not much more' rather than 'at most' since we need to take account of the fact that the clusters do not have size exactly n/k.

a partition P of $V(G)$, we define the reduced graph R as follows. The vertices of R are the clusters of the partition P. Two such clusters V_i and V_j are joined by an edge in R whenever the pair (V_i, V_j) is ε-regular in G, with density at least d.

We normally define the reduced graph with a parameter d which is small, but still much larger than ε. Then the reduced graph inherits many useful properties of the original graph G. The aim is to find some relatively simple structure in the reduced graph, and use this to prove the existence of a more complicated structure in the original graph.

1.4.3 The Blow-up Lemma

The blow-up lemma is one tool which enables us to transfer structure from the reduced graph back to G. Each edge of the reduced graph corresponds to an ε-regular pair of clusters in G with reasonably high density. As mentioned before, this pair can easily be made $(\sqrt{\varepsilon}, \delta)$-super-regular by deleting a few vertices. Roughly speaking, the blow-up lemma states that such an $(\sqrt{\varepsilon}, \delta)$-super-regular pair V_i, V_j will contain a copy of any bipartite graph H of bounded degree, provided H is contained in the complete bipartite graph between V_i and V_j (i.e. provided the clusters V_i and V_j are large enough to contain the vertex classes of H). In particular, this allows for spanning subgraphs H. More generally, given a subgraph in the reduced graph, the corresponding clusters in G will contain any blow-up of this subgraph which has bounded degree.

Theorem 1.38 (Blow-up Lemma [47]) *For any real number $\delta \in (0, 1]$, and integers $\Delta, k \geq 1$ there is a real number $\varepsilon > 0$ such that the following holds. Suppose G is a k-partite graph with vertex classes $V_1, V_2, \ldots, V_k$ such that all pairs V_i, V_j are (ε, δ)-super-regular. Let K denote the complete k-partite graph with vertex classes*

$V_1, V_2, \ldots, V_k$. *Suppose* H *is any graph with* $\Delta(H) \leq \Delta$. *If* K *contains a copy of* H, *then* G *also contains a copy of* H.

Thus if we have a k-clique in the reduced graph, we may discard a small number of vertices from the corresponding k clusters in G to ensure that all the $\binom{k}{2}$ pairs are not just regular but also super-regular. Then if we want to find a copy of H in G, where H is a k-chromatic graph with bounded maximum degree, we need only consider whether H is a subgraph of K, the graph obtained from G by turning all the $\binom{k}{2}$ pairs of clusters into complete bipartite graphs. Equivalently, we need only consider whether the sizes of the classes of H are small enough to fit into these k clusters of G.

So roughly the strategy for a proof of the existence of a perfect H-packing is to find copies of $K_{\chi(H)}$ in the reduced graph and expand these using the blow-up lemma to a large number of disjoint copies of H in G. There is then some tidying up to do with those few vertices that are not covered by these copies.

1.4.4 The Regularity Lemma for Hypergraphs

The regularity lemma for k-uniform hypergraphs generalises the Szemerédi's regularity lemma for graphs, and similarly aims to describe appropriate "pseudo-randomness" properties. However, the version of the hypergraph regularity lemma which we need is rather complicated. For this reason, we leave both the statement and the explanation of the lemma until Chapter 4, when it will first be required.

CHAPTER 2

GRAPH PACKINGS

The aim of this chapter is to present the proof of the following theorem.

Theorem 2.1 *For every integer $r \geq 4$ there exists an integer $n_0 = n_0(r)$ such that every graph G whose order $n \geq n_0$ is divisible by r and whose minimum degree is at least*

$$\left(1 - \frac{1}{\chi_{cr}(K_r^-)}\right) n$$

contains a perfect K_r^--packing.

Also in Section 2.6 I will indicate how this proof can be extended to a larger class of graphs H for which the error term in Theorem 1.21 can be removed entirely.

2.1 Further Notation and Preliminaries

Throughout this chapter we omit floors and ceilings whenever this does not affect the argument.

For a graph H of chromatic number ℓ, define the *bottle graph $B^*(H)$ of H*, to be the complete ℓ-partite graph which has $\ell - 1$ classes of size $|H| - \sigma(H)$ and one class of size $(\ell - 1)\sigma(H)$. (Recall that $\sigma(H)$ is the smallest possible size of a colour class in an ℓ-colouring of H.) Note that given an optimal colouring of H, then $|H| - \sigma(H)$

is the sum of all colour class sizes except the smallest one. Thus by rotating these $\ell - 1$ classes and keeping the smallest one fixed, we can see that $B^*(H)$ contains a perfect H-packing consisting of $\ell - 1$ copies of H. We will use B^* to denote $B^*(K_r^-)$ whenever this is unambiguous.

2.2 Extremal Examples

For completeness, we include the construction which shows that the bound on the minimum degree in Theorem 2.1 is best possible.

Proposition 2.2 *Let $r \geq 4$. Then for all $k \in \mathbb{N}$ there is a graph G on $n = kr$ vertices whose minimum degree is $\lceil (1 - 1/\chi_{cr}(K_r^-))n \rceil - 1$ but which does not contain a perfect K_r^--packing.*

Proof. We construct G as follows. G is a complete $(r-1)$-partite graph with vertex classes $U_0, \dots, U_{r-2}$, where $|U_0| = k - 1$ and the sizes of all other classes are as equal as possible. It is easy to check that G has the required minimum degree. Indeed,

$$\begin{aligned}
\delta(G) = n - \left\lceil \frac{n - |U_0|}{r-2} \right\rceil &= kr - \left\lceil \frac{kr - (k-1)}{r-2} \right\rceil \\
&= k(r-1) - \left\lceil \frac{2k - (k-1)}{r-2} \right\rceil \\
&\geq k(r-1) - \left(\frac{k}{r-2} + 1 \right) \\
&= \frac{r^2 - 3r + 1}{r(r-2)} n - 1 \\
&= (1 - \frac{r-1}{r(r-2)})n - 1.
\end{aligned}$$

Moreover, every copy of K_r^- in G contains at least one vertex in U_0. Thus we can find at most $|U_0|$ pairwise disjoint copies of K_r^- which therefore cover at most $(k-1)(r-1) < n - |U_0|$ vertices of $G - U_0$. Thus G does not contain a perfect

K_r^--packing. $\square$

Note that Proposition 2.2 extends to every graph H which is obtained from a K_{r-1} by adding a new vertex and joining it to at most $r-2$ vertices of the K_{r-1}. Since each such H is a subgraph of K_r^- and since $\chi_{cr}(H) = \chi_{cr}(K_r^-)$, it follows from this observation and from Theorem 2.1 that $\delta(n, H) = \lceil (1 - 1/\chi_{cr}(H))n \rceil$ if n is sufficiently large (where $\delta(n, H)$ is as defined in Chapter 1).

The following example shows that for a large class of graphs, the $O(1)$-error term in Theorem 1.21 cannot be omitted completely. The example is an extension of a similar construction in [48].

Proposition 2.3 *Suppose that H is a complete ℓ-partite graph with $\ell \geq 3$ such that every vertex class of H, except possibly its smallest class, has at least 3 vertices. Then there are infinitely many graphs G whose order n is divisible by $|H|$ and whose minimum degree satisfies $\delta(G) = (1 - \frac{1}{\chi_{cr}(H)})n$ but which do not contain a perfect H-packing.*

Proof. Let σ denote the size of the smallest vertex class of H. Given $k \in \mathbb{N}$, consider the complete ℓ-partite graph on $n := k(\ell - 1)|H|$ vertices whose vertex classes $A_1, \dots, A_\ell$ satisfy $|A_1| := (|H| - \sigma)k + 1$, $|A_\ell| := k(\ell - 1)\sigma - 1$ and $|A_i| := (|H| - \sigma)k$ for all $1 < i < \ell$. Let G be the graph obtained by adding a perfect matching into A_1 or, if $|A_1|$ is odd, a matching covering all but 3 vertices and a path of length 2 on these remaining vertices. Observe that the minimum degree of G is $(1 - \frac{1}{\chi_{cr}(H)})n$.

Consider any copy H' of H in G. Suppose that H' meets A_ℓ in at most $\sigma - 1$ vertices. Then there is a colour class X of H' which meets A_ℓ but does not lie entirely in A_ℓ. So some vertex class of G must meet at least two colour classes of H'. Since H' is complete ℓ-partite, this vertex class must have some edges in it, and so must be A_1. However, A_1 cannot meet three colour classes of H', since it is triangle free. Thus every colour class of H' except X lies completely within one A_i.

Furthermore, A_1 cannot contain two complete colour classes of H', since then $G[A_1]$ would have a vertex of degree 3, a contradiction. So A_1 meets X as well as another colour class Y of H'. Furthermore $X \setminus A_\ell \subseteq A_1$ and $Y \subseteq A_1$. Let $x \in X \cap A_1$. Then $Y \subseteq N_G(x)$ since $Y \subseteq N_{H'}(x)$. This implies that $|Y| \leq 2$ and so $\sigma = |Y| \leq 2$. Thus $|X| \geq 3$. Since at most $\sigma - 1 \leq 1$ vertices of X lie in A_ℓ this in turn implies that $|X \cap A_1| \geq 2$. As $X \cap A_1$ lies in the neighbourhood of any vertex from Y, we must have that $|X \cap A_1| = 2$. Thus $X \cap A_1$ can only lie in the neighbourhood of one vertex from Y. Hence $\sigma = |Y| = 1$. But then X avoids A_ℓ, a contradiction.

So any copy of H in G has at least σ vertices in A_ℓ. Thus any H-packing in G consists of less than $k(\ell-1)$ copies of H and therefore covers less than $k(\ell-1)(|H|-\sigma) < |G|-|A_\ell|$ vertices of $G-A_\ell$. So G does not contain a perfect H-packing. $\square$

Note that the proof of Proposition 2.3 shows that if $|H|-\sigma$ is odd then we only need that every vertex class of H (except possibly its smallest class) has at least two vertices. Moreover, it is not hard to see that the conclusion of Proposition 2.3 holds for all graphs H which do not have an optimal colouring with a vertex class of size $\sigma + 1$. See Section 2.6 for details and for further examples of graphs for which the error term is necessary.

In the proof of Theorem 2.1 we will use the following observation about packings in almost complete $(q+1)$-partite graphs. It follows easily from the blow-up lemma (see e.g. [45]), but we also sketch how it can be deduced directly from Hall's theorem.

Proposition 2.4 *For all $q, r \in \mathbb{N}$ there exists a positive constant $\tau_0 = \tau_0(q,r) = 1/(2(r+1)^{q-1})$ such that the following holds for every $\tau \leq \tau_0$ and all $k \in \mathbb{N}$. Let $H_{q,r}$ be the complete $(q+1)$-partite graph with q vertex classes of size r and one vertex class of size 1. Let G^* be a $(q+1)$-partite graph with vertex classes $V_1, \dots, V_{q+1}$ such that $|V_i| = kr$ for all $i \leq q$ and such that $|V_{q+1}| = k$. Suppose that for all distinct $i, j \leq q+1$ every vertex $x \in V_i$ of G^* is adjacent to all but at most $\tau|V_j|$ vertices in*

V_j. Then G^* has a perfect $H_{q,r}$-packing.

Proof. We proceed by induction on q. If $q = 1$ then we are looking for a perfect $K_{1,r}$-packing. The result can easily be deduced from Hall's theorem. For we have $|V_1| = kr$ and $|V_2| = k$. Let us now replace each vertex of V_2 with r new vertices, each joined to the same neighbours in V_1 as the original vertex. Now a perfect matching in the new graph corresponds to a perfect $K_{1,q}$-packing in the original graph G^*, and so we only need to check Hall's condition in the new graph. Suppose that Hall's condition fails for $A \subseteq V_1$, i.e. $|N(A)| < |A|$. Then since $\tau_0 = \frac{1}{2}$ we have $|N(A)| \geq kr/2$, and so $|A| > kr/2$. Now $V_2 \backslash N(A) \neq \emptyset$ and so $|N(V_2 \backslash N(A))| \geq kr/2$. But also $N(V_2 \backslash N(A)) \subseteq V_1 \backslash A$ and so has size at most $kr - |A| < kr/2$, which is a contradiction as required.

Now suppose that $q > 1$ and note that $\tau_0(q, r) = \tau_0(q-1, r)/(r+1) < \tau_0(q-1, r)$. As before, we can find a perfect $K_{1,r}$-packing in $G^*[V_q \cup V_{q+1}]$. Let G' be the graph obtained from G^* by replacing each copy K of such a $K_{1,r}$ with one vertex x_K and joining x_K to $y \in V_1 \cup \cdots \cup V_{q-1}$ whenever y is adjacent to every vertex of K. Let $V'_1, \ldots, V'_q$ be the classes in G'. Then in G' for all distinct $i, j \leq q$, every vertex in V'_i is adjacent to all but at most $\tau_0(q, r)/(r + 1)|V'_j|$ vertices in V'_j, and so G' contains a perfect $H_{q-1,r}$-packing by the induction hypothesis. This corresponds to a perfect $H_{q,r}$-packing in G^*. □

2.3 Overview of the Proof

Our main tool is the following result from [53]. It states that in the 'non-extremal case', where the graph G given in Theorem 1.21 satisfies certain conditions, we can find a perfect packing even if the minimum degree is slightly smaller than required in Theorem 1.21. The conditions ensure that the graph G does not look too much like one of the extremal examples of graphs whose minimum degree is just a little smaller than required in Theorem 1.21 but which do not contain a perfect H-packing.

Theorem 2.5 *Let H be a graph of chromatic number $\ell \geq 2$ with $hcf(H) = 1$. Let z_1 denote the size of the small class of the bottle graph $B^*(H)$, let z denote the size of one of the large classes, and let $\xi = z_1/z$. Let $\theta \ll \tau_0 \ll \xi, 1-\xi, 1/|B^*(H)|$ be positive constants. There exists an integer n_0 such that the following holds. Suppose G is a graph whose order $n \geq n_0$ is divisible by $|B^*(H)|$ and whose minimum degree satisfies $\delta(G) \geq (1 - \frac{1}{\chi_{cr}(H)} - \theta)n$. Suppose that G also satisfies the following conditions:*

(i) *G does not contain a vertex set A of size $zn/|B^*(H)|$ such that $d(A) \leq \tau_0$.*

(ii) *If $\ell = 2$, then G does not contain a vertex set A with $d(A, V(G) \setminus A) \leq \tau_0$.*

Then G has a perfect H-packing.

The proof of this result in [53] used the regularity lemma for graphs. In fact, during the proof of Theorem 2.1 there will be no explicit reference to the regularity lemma, since it is only needed implicitly whenever we need to apply Theorem 2.5 (which we will need to do at two separate points in the proof).

Roughly speaking, Theorem 2.5 deals with the case when there is no obvious structure in G. The regularity lemma helps in this case because it provides some sort of structure where there didn't appear to be any. But if the conditions of Theorem 2.5 do not hold, then we know that we have some structure in G, and so we will not need the regularity lemma to provide it.

More precisely, by applying this theorem with $H := K_r^-$ (where $r \geq 4$), we only need to consider the extremal case, when there are large almost independent sets. (Note that if the order of the graph G given by Theorem 2.1 is not divisible by $|B^*(K_r^-)|$, we must first greedily remove some copies of K_r^- before applying Theorem 2.5. The existence of these copies follows from the Erdős-Stone-Simonovits theorem, and since we only need to remove a bounded number of copies, this will not affect any of the properties required in Theorem 2.5 significantly.)

Suppose that we have q such large almost independent sets. Theorem 2.5 will deal with the case $q = 0$, and so we may assume that $q \geq 1$. Then we will think of the remainder of the vertices of G as the $(q + 1)$th set. We will show in Section 2.4 that by taking out a few copies of K_r^- and rearranging these $q + 1$ sets slightly, we can achieve that these sets will induce an almost complete $(q+1)$-partite graph. Furthermore, the proportion of the size of each of the first q of these modified sets to the order of the entire graph will be the same as for the large classes of the bottle graph $B^*(K_r^-)$ defined in Section 2.1.

Let B_1^* be the subgraph of $B^*(K_r^-)$ obtained by deleting q of the large vertex classes. Ideally, we would like to apply Theorem 2.5 to find a B_1^*-packing in the (remaining) subgraph of G induced by the $(q + 1)$th vertex set. In a second step we would then like to extend this B_1^*-packing to a $B^*(K_r^-)$-packing in G, using the fact that the $(q + 1)$-partite subgraph of G between the classes defined above is almost complete. This would clearly yield a K_r^--packing of G.

However, there are some difficulties. For example, Theorem 2.5 only applies to graphs H with $hcf(H) = 1$, and this may not be the case for B_1^* if it is bipartite. So instead of working with B_1^*, we consider a suitable subgraph B_1 of B_1^* which does satisfy $hcf(B_1) = 1$. Moreover, if B_1 is bipartite we may have to take out a few further carefully chosen copies of K_r^- from G to ensure that condition (ii) is also satisfied before we can apply Theorem 2.5 to the subgraph induced by the $(q+1)$th vertex set.

2.4 Tidying Up the Classes

Let n and q be integers such that n is divisible by $r(r-2) = |B^*(K_r^-)|$ and such that $1 \le q \le r-2$. Note that in the case when $H := K_r^-$ the set A in condition (i) of Theorem 2.5 has size $\frac{r-1}{r(r-2)}n$. We say that disjoint vertex sets $A_1, \dots, A_{q+1}$ are (q,n)-*canonical* if $|A_i| = \frac{r-1}{r(r-2)}n$ for all $i \le q$ and $|A_{q+1}| = \frac{n}{r} + (r-q-2)\frac{r-1}{r(r-2)}n = n - \sum_{i=1}^{q} |A_i|$. Note that in this case the graph $K(q,n)$ obtained from the complete graph on $\bigcup_{i=1}^{q+1} A_i$ by making each A_i with $i \le q$ into an independent set has a perfect $B^*(K_r^-)$-packing and thus also a perfect K_r^--packing.

Our aim in the following lemma is to remove a few disjoint copies of K_r^- from our given graph G in order to obtain a graph on n^* vertices which looks almost like $K(q,n^*)$. In the next section we will then use this property to show that this subgraph of G has a perfect K_r^--packing.

Lemma 2.6 *Let $r \ge 4$ and $0 < \tau \ll 1/r$. Then there exists an integer $n_0 = n_0(r,\tau)$ such that the following is true. Let G be a graph whose order $n \ge n_0$ is divisible by r and whose minimum degree satisfies $\delta(G) \ge (1 - \frac{1}{\chi_{cr}(K_r^-)})n$. Suppose that for some $1 \le q \le r-2$ there are q disjoint vertex sets $A_1, \dots, A_q$ in G such that $|A_i| = \lceil \frac{r-1}{r(r-2)}n \rceil$ and $d(A_i) \le \tau$ for $1 \le i \le q$. Set $A_{q+1} := V(G) \setminus (A_1 \cup \ldots \cup A_q)$. Then there exist disjoint vertex sets $A_1^*, \dots, A_{q+1}^*$ such that the following hold:*

(i) *If $G^* := G[\bigcup_{i=1}^{q+1} A_i^*]$ and $n^* := |G^*|$ then $r(r-2)$ divides n^*, and $G - G^*$ contains a perfect K_r^--packing. Furthermore, $n - n^* \le \tau^{1/3} n$.*

(ii) $|A_1^*| = |A_2^*| = \ldots = |A_q^*| = \frac{r-1}{r(r-2)}n^*$.

(iii) *For all $i,j \le q+1$ with $i \ne j$, each vertex in A_i^* has at least $(1-\tau^{1/5})|A_j^*|$ neighbours in A_j^*.*

Proof. We first fix integers $r \geq 4$ and n divisible by r and a constant τ such that

$$0 < 1/n \ll \tau \ll 1/r.$$

Note that if n is divisible by $r(r-2)$ then the sets $A_1, \dots, A_{q+1}$ are (q,n)-canonical. If n is not divisible by $r(r-2)$ then we will change the sizes of the A_i slightly as follows. Write $n = n' + kr$ where n' is divisible by $r(r-2)$ and $0 < k < r-2$. If $k \geq q$ then we do not change the sizes of the A_i. If $k < q$ then for each i with $k < i \leq q$ we move one vertex from A_i to A_{q+1}. We still denote the sets thus obtained by $A_1, \dots, A_{q+1}$. We may choose the vertices we move in such a way that the density of each A_i with $i \leq q$ is still at most τ. Note that $\lceil \frac{r-1}{r(r-2)} n \rceil = \frac{r-1}{r(r-2)} n' + k + 1$. Thus both in the case when $k \geq q$ and in the case when $k < q$ the sets $A_1, \dots, A_{q+1}$ can be obtained from (q, n')-canonical sets by adding kr new vertices as follows. For each $i \leq \min\{k, q\}$ we add $k+1$ of the new vertices to the ith vertex set, for each i with $\min\{k, q\} < i \leq q$ we add k new vertices to the ith vertex set and all the remaining new vertices are added to A_{q+1}. Let K be the graph obtained from the complete graph on $\bigcup_{i=1}^{q+1} A_i$ by making each A_i with $i \leq q$ into an independent set. It is easy to see that $K(q, n')$ can be obtained from K by removing k vertex-disjoint copies of K_r^-.

More precisely for each $i \leq \min\{k, q\}$ we will remove a copy of K_r^- with two vertices in A_i, one vertex in A_j for each $i \neq j \leq q$ and $r - 1 - q$ vertices in A_{q+1}. Furthermore, we remove $\max\{k - q, 0\}$ copies of K_r^- with one vertex in each A_i for $i \leq q$ and $q - r$ vertices in A_{q+1}. These copies of K_r^- exist, and may be chosen to be vertex-disjoint, since K is a complete $(r-1)$-partite graph, and the graph obtained by removing them is $K(q, n')$.

In particular, since $K(q, n')$ has a perfect K_r^--packing this shows that K also has a perfect K_r^--packing. Note that if $k < q$ then this would not hold if we had not changed the sizes of the A_i. Later on we will use that in all cases we have

$$|A_i| \geq \frac{r-1}{r(r-2)}n' + k = \frac{r-1}{r(r-2)}(n-kr) + k \tag{2.1}$$

for all $i \leq q$, where we set $n' := n$ and $k := 0$ if n is divisible by $r(r-2)$. Observe that $\chi_{cr}(K_r^-) = \frac{r(r-2)}{r-1}$ and so $\delta(G) \geq (1 - \frac{r-1}{r(r-2)})n$. Thus the minimum degree condition on G implies that the neighbours of any vertex might essentially avoid one of the A_i, for $i \leq q$, but no more.

Now for each index i, call a vertex $x \in A_i$ *i-bad* if x has at least $\tau^{1/3}|A_i|$ neighbours in A_i. Note that, for $i \leq q$, the number of i-bad vertices is at most $\tau^{2/3}|A_i|$ since $d(A_i) \leq \tau$ for such i. Call a vertex $x \in A_i$ *i-useless* if, for some $j \neq i$, x has at most $(1-\tau^{1/4})|A_j|$ neighbours in A_j. In this case the minimum degree condition shows that, provided $i \neq r-1$, x must have at least a $\tau^{1/3}$-fraction of the vertices in its own class as neighbours, i.e. x is i-bad. Thus every vertex that is i-useless is also i-bad for $i \neq r-1$. In particular, for each $i \leq q$, there are at most $\tau^{2/3}|A_i|$ i-useless vertices.

For $i = q+1$ we estimate the number u_{q+1} of $(q+1)$-useless vertices by looking at the edges between A_{q+1} and $V(G)\backslash A_{q+1}$. We have

$$\begin{aligned} e(A_{q+1}, V(G)\backslash A_{q+1}) &\geq \sum_{i=1}^{q} \{|A_i|\delta(G) - 2e(A_i) - \sum_{j\neq i, j \leq q} |A_i||A_j|\} \\ &\geq q(|A_1|-1)\delta(G) - q\tau|A_1|^2 - q(q-1)|A_1|^2. \end{aligned}$$

On the other hand,

$$\begin{aligned} e(A_{q+1}, V(G)\backslash A_{q+1}) &\leq u_{q+1}\{(q-1)|A_1| + (1-\tau^{1/4})|A_1|\} + (|A_{q+1}| - u_{q+1})q|A_1| \\ &= q|A_1||A_{q+1}| - u_{q+1}\tau^{1/4}|A_1|. \end{aligned}$$

Combining these inequalities, and using the fact that $\delta(G)-(q-1)|A_1| \geq |A_{q+1}|-q$, gives

$$\begin{aligned}
\tau^{\frac{1}{4}}u_{q+1} &\leq q|A_{q+1}| + q\tau|A_1| - \frac{1}{|A_1|}[q(|A_1|-1)\delta(G) - q(q-1)|A_1|^2] \\
&\leq q|A_{q+1}| + q\tau|A_1| - q(|A_{q+1}|-q) + q\frac{\delta(G)}{|A_1|} \\
&= q^2 + q\tau|A_1| + q\frac{\delta(G)}{|A_1|} \\
&\leq q^2 + q\tau|A_1| + rq \\
&\leq 3q\tau|A_1|
\end{aligned}$$

and so $u_{q+1} \leq \tau^{2/3}|A_{q+1}|$. So in total the number of vertices which are i-useless for some i is at most $\tau^{2/3}n$.

Given $j \neq i$, call a vertex $x \in A_i$ *j-exceptional* if x has at most $\tau^{1/3}|A_j|$ neighbours in A_j. Thus every such vertex is also i-useless, and therefore i-bad if $i \leq q$. Note that if $q < r-2$ then the minimum degree condition ensures that there are no $(q+1)$-exceptional vertices. Furthermore, if $i = r-1$, then an exceptional vertex in A_i is also i-bad. So all exceptional vertices are bad.

Now if for some $i \neq j$ there exists an i-bad vertex $x \in A_i$ and an i-exceptional vertex $y \in A_j$, then let us swap x and y. (Note that a vertex is not i-exceptional for more than one i.) Having done this, since there are not too many exceptional vertices, we will still have that each non-bad vertex in A_i has at most $2\tau^{1/3}|A_i|$ neighbours in A_i, each non-useless vertex in A_i still has at least $(1-2\tau^{1/4})|A_j|$ neighbours in each A_j with $j \neq i$ and each non-i-exceptional vertex still has at least $\tau^{1/3}|A_i|/2$ neighbours in A_i. We will also have that for any i for which i-exceptional vertices exist, there are no i-bad vertices.

We now wish to remove all the exceptional vertices by taking out a few disjoint copies of K_r^- which will cover them. For simplicity, we will split the argument into two cases. In both cases we will repeatedly remove $r-2$ disjoint copies of K_r^- at a time. We say that such a collection of $r-2$ copies *respects the proportions of the* A_i if altogether these copies meet each A_i with $i \le q$ in exactly $r-1$ vertices.

Case 1. $q \le r-3$

In this case the minimum degree condition ensures that no vertex is $(q+1)$-exceptional. To deal with the j-exceptional vertices for $j \le q$ we will need the fact that we can find a reasonably large number of disjoint copies of K_{r-1-q} in $G[A_{q+1}]$. To prove this fact, observe that

$$\delta(G[A_{q+1}]) \ge \delta(G) - \sum_{i=1}^{q} |A_i| \ge |A_{q+1}| - \frac{r-1}{r(r-2)}n \tag{2.2}$$

and

$$\frac{r-1}{r(r-2)}\frac{n}{|A_{q+1}|} \overset{(2.1)}{\le} \frac{r-1}{r(r-2)}\frac{1}{\frac{1}{r}+(r-q-2)\frac{r-1}{r(r-2)}} \le \frac{1}{r-q-2} - c(r) \tag{2.3}$$

where $c(r) > 0$ is a constant depending only on r. Combining these results gives

$$\delta(G[A_{q+1}]) \ge \left(1 - \frac{1}{r-q-2} + c(r)\right)|A_{q+1}|. \tag{2.4}$$

Thus we can apply Turán's theorem repeatedly to find at least $\frac{c(r)}{r-q-1}|A_{q+1}|$ disjoint copies of K_{r-q-1} in $G[A_{q+1}]$.

Now for each $i \le q+1$ in turn, consider the exceptional vertices $x \in A_i$. Suppose that x is j-exceptional. First move x into A_j. Note that the minimum degree condition on G means that x is joined to almost all vertices in A_ℓ for every $\ell \ne j$. We greedily choose a copy of K_r^- covering x and one other vertex in A_j, $r-q-1$ vertices in A_{q+1} and one vertex in all other classes, where all vertices other than x were chosen to be non-useless. (Indeed, to find such a copy of K_r^- we first choose a copy of K_{r-q-1} in A_{q+1} which lies in the neighbourhood of x and which consists of non-useless vertices. Then we choose all the remaining vertices.) Remove this copy of K_r^-. Also greedily remove $r-3$ further disjoint copies of K_r^- such that together all these copies of K_r^- respect the proportions of the A_i. Proceed similarly for all the exceptional vertices. For each exceptional vertex we are removing $r-2$ copies of K_r^-, so in total we are removing at most $r(r-2)\tau^{2/3}n$ vertices.

Case 2. $q = r-2$

In this case, the exceptional vertices in A_{r-1} need special attention since we cannot simply move them into another class without making A_{r-1} too small. So we proceed as follows. For each $i \le r-2$, let s_i be the number of i-exceptional vertices in A_{r-1}. Whenever $s_i > 0$ we will find a matching of size s_i in $G[A_i]$. To see that such a matching exists, consider a maximal matching in A_i and let m denote the size of this matching. Note that

$$e(A_i) \le 2m\Delta(A_i) \le 2m2\tau^{1/3}|A_i|$$

since the presence of i-exceptional vertices guarantees that no vertex in A_i is i-bad. Also

$$\begin{aligned} e(A_i) &\ge \frac{1}{2}\{\delta(G)|A_i| - (n-|A_i|-s_i)|A_i| - s_i2\tau^{1/3}|A_i|\} \\ &\ge \frac{|A_i|}{2}\{|A_i| - \frac{r-1}{r(r-2)}n + s_i(1-2\tau^{1/3})\} \\ &\overset{(2.1)}{\ge} \frac{|A_i|}{2}\{s_i(1-2\tau^{1/3}) - \frac{k}{r-2}\}. \end{aligned}$$

Since $k \le r-3$ and $\tau \ll 1/r$, comparing these two bounds on $e(A_i)$ gives $m \gg s_i$ whenever $s_i > 0$. So we may pick a matching M_i with s_i edges in A_i, all of whose vertices are non-useless (since no vertices in A_i are bad). Now for each i in turn, we will remove the i-exceptional vertices in A_{r-1} using this matching. For each such vertex $x \in A_{r-1}$, pick an edge $yz \in M_i$. Swap x with y; we now no longer consider x to be exceptional. Then greedily find a copy of K_r^- which meets A_{r-1} precisely in y, which meets A_i precisely in z and which contains two vertices in some A_j with $j \neq i, r-1$ (such a j exists since $r \geq 4$), and one vertex in each other A_j. All these vertices will be chosen to be non-useless, and all (except y and z) will avoid each M_j. Remove this copy of K_r^-. Then also greedily take out $r-3$ further disjoint copies of K_r^-, avoiding the M_j and all useless vertices, in such a way that altogether they respect the proportions of the A_i. Note that we can find these copies greedily since the $(q+1)$-partite graph induced by the A_i is almost complete. We continue doing this until no exceptional vertices are left in A_{r-1}. The fact that M_i has s_i edges ensures that we will always have an edge left in the appropriate matching for each exceptional vertex in A_{r-1}.

Now for all other exceptional vertices, proceed using the argument for the case when $q \leq r-3$. In this way we will remove all the exceptional vertices.

So in both cases we will obtain sets $A'_1, \dots, A'_{q+1}$ not containing any exceptional vertices. We now want to remove any remaining useless vertices. Before dealing with the exceptional vertices, each useless but non-exceptional vertex in A_i had at least $\tau^{1/3}|A_j|/2$ neighbours in A_j for each $j \neq i$. Also, we had at most $\tau^{2/3}n$ useless vertices, and therefore also at most this many exceptional vertices. So we have taken out at most $r(r-2)\tau^{2/3}n$ vertices. Thus each remaining vertex $x \in A'_i$ still has at least $\tau^{1/3}|A'_j|/3$ neighbours in A'_j for each $j \neq i$, which is much larger than the number of j-useless vertices.

Ideally, for a useless vertex $x \in A'_i$ we would like to pick neighbours in each other class greedily so that together these vertices form a copy of K_r^- with, say, two vertices in A'_1, $r-q-1$ vertices in A'_{q+1} and one vertex in each other A'_j. The problem is that the neighbours of x may avoid a substantial proportion of A'_{q+1}, and so in particular may not include any of the copies of K_{r-q-1} which we know are contained in A_{q+1} (and therefore in A'_{q+1}).

So instead, we proceed as follows. We first deal with all the vertices which have too few neighbours in A'_{q+1}. Let U be the set of vertices in $A'_1 \cup \ldots \cup A'_q$ which originally had at most $(1-\tau^{1/4})|A_{q+1}|$ neighbours in A_{q+1}. In particular, all these vertices are useless. Note that a vertex $x \in U \cap A'_i$ (where $i \leq q$) still has at least $\tau^{1/3}|A'_i|/3$ neighbours in A'_i. For each such vertex x in turn we proceed as follows. We first move x into A'_{q+1}. Then we will greedily find a copy of K_r^- which avoids x and meets each A'_j with $j \leq q$ in precisely one vertex. Note that similarly as in (2.4) one can show that

$$\delta(G[A'_{q+1}]) \geq \left(1 - \frac{1}{r-q-2} + \frac{c(r)}{2}\right)|A'_{q+1}|. \tag{2.5}$$

So we may apply the Erdős-Stone-Simonovits theorem to find the necessary copy of K_{r-q}^- in A'_{q+1} avoiding x as well as all the $(q+1)$-useless vertices. We can extend it to the desired copy of K_r^-, also avoiding all the useless vertices. Remove this copy of K_r^-. In effect, we have removed two vertices from A'_i (one vertex in the copy of K_r^- and x), $r-q-1$ vertices from A'_{q+1} and one vertex from each other A'_j. We can also find $r-3$ further disjoint copies of K_r^- in such a way that altogether these copies respect the proportions of the A'_i. Remove these copies. Repeating this for each vertex $x \in U$, in total we move or remove at most $\tau^{1/2}n$ vertices. We denote by A''_i the sets thus obtained from the A'_i.

The effect of moving the vertices of U and taking out these copies of K_r^- is that all vertices (except those in A''_{q+1}) are joined to almost all of A''_{q+1}. The vertices in U may now be $(q+1)$-useless, but are certainly non-exceptional.

Now consider any useless vertex $x \in A''_i$ where $i \neq q+1$. Let A''_j be the vertex set in which x has the lowest number of neighbours, not including $j = i, q+1$. (Note that such a j exists since if $q = 1$, a useless vertex $x \in A_1$ would have been in U, so we would already have dealt with it.) Pick non-useless neighbours y and z of x in A''_j. (Such neighbours exist since x was not j-exceptional.) Recall that each of x, y and z is joined to almost all of A''_{q+1}. Since A''_{q+1} is almost as large as A_{q+1} it follows that many of the copies of K_{r-q-1} chosen after (2.4) lie in the common neighbourhood of x, y and z, and so form a copy of K^-_{r-q+2} together with x, y and z. Pick such a copy. Now note that the choice of j implies that x is joined to at least $|A''_\ell|/3$ vertices in A''_ℓ for each $\ell \neq i, j, q+1$. So we can greedily extend this copy of K^-_{r-q+2} to a copy of K^-_r in G by picking one non-useless vertex in every other A''_ℓ. We then greedily find $r-3$ further disjoint copies of K^-_r avoiding all the useless vertices so that together with the copy just found, these copies of K^-_r respect the proportions of the A''_i. Remove all these copies of K^-_r.

For a $(q+1)$-useless vertex x, we perform a similar process, except that x is already in A''_{q+1}, so we find non-useless neighbours y and z of x in A''_j and find a copy of K_{r-q-1} in A''_{q+1} which contains x and lies in the common neighbourhood of y and z. We can do this since (2.5) implies that

$$\delta(G[A''_{q+1}]) \geq \left(1 - \frac{1}{r-q-2} + \frac{c(r)}{3}\right) |A''_{q+1}|.$$

(Note that in particular this bound applies to the degree of x in A''_{q+1}.) So we can successively pick common non-useless neighbours of x, y and z in A''_{q+1} to construct the necessary K_{r-q-1} containing x. Together with y and z this forms a copy of K^-_{r-q+1} which we extend suitably to a copy of K^-_r. As before we then find further disjoint copies of K^-_r such that together all these copies respect the proportions of the A''_i. We can repeat this process until no useless vertices are left. The fact that there are not too many useless vertices will ensure that all our calculations remain valid.

Finally, if $k > 0$, we remove k further disjoint copies of K_r^- to ensure that the sets $A_1^*, \ldots, A_{q+1}^*$ thus obtained from the A_i'' are (q, n^*)-canonical where $n^* := |A_1^* \cup \ldots \cup A_{q+1}^*|$. This can be done because of our modification of the A_i at the beginning of the proof. Now the A_i^* contain neither exceptional nor useless vertices. Furthermore, for each exceptional vertex we removed $r(r-2)$ vertices from G, in dealing with U we moved or removed at most $\tau^{1/2}n$ vertices and then for each remaining useless vertex we removed a further $r(r-2)$ vertices. Because there were at most $2\tau^{1/3}n$ exceptional and useless vertices originally, and because $k \leq r-3$, after switching i-exceptional vertices with i-bad ones we have moved or removed at most $2r(r-2)\tau^{1/3}n + \tau^{1/2}n + r(r-3) \leq \tau^{1/4}n$ vertices. Thus each remaining vertex $x \in A_i^*$ satisfies, for each $j \neq i$,

$$\begin{aligned} d_{A_j^*}(x) &\geq (1 - 2\tau^{1/4})|A_j^*| - \tau^{1/4}n \\ &\geq (1 - 2\tau^{1/4})|A_j^*| - r\tau^{1/4}|A_j^*| \\ &\geq (1 - \tau^{1/5})|A_j^*|. \end{aligned}$$

Furthermore, the A_i^* are (q, n^*)-canonical, and $G^* = G[\bigcup_{i=1}^{q+1} A_i^*]$ was obtained from G by removing vertex-disjoint copies of K_r^-, and so $G - G^*$ contains a perfect K_r^--packing. Thus the A_i^* satisfy all the conditions of the lemma. □

2.5 Proof of Theorem 2.1

Recall that $B^* = B^*(K_r^-)$ denotes the bottle graph of K_r^-. Fix integers $r \geq 4$ and n divisible by r and constants $\tau_1, \ldots, \tau_{r-1}$ such that

$$0 < \tau_1 \ll \tau_2 \ll \ldots \ll \tau_{r-1} \ll 1/r.$$

Let G be the graph given in Theorem 2.1. Let $q \leq r-2$ be maximal such that the conditions of Lemma 2.6 are satisfied with $\tau := \tau_q$. As already observed in Section 2.3, by Theorem 2.5 we may assume that $q \geq 1$. To prove Theorem 2.1, we first apply Lemma 2.6 with this choice of q to obtain a subgraph G^* of G and a $(q, |G^*|)$-canonical partition $A_1^*, \dots, A_{q+1}^*$ of $V(G^*)$. Our definition of q will ensure that if $q \neq r-3$ then the graph induced by A_{q+1}^* does not look like one of the extremal graphs and so we can apply Theorem 2.5 to it in order to find a perfect B_1-packing, where B_1 is the spanning subgraph of B_1^* defined below. (Recall that B_1^* is the $(r-q-1)$-partite subgraph of B^* obtained by deleting q of the large vertex classes.) In the case when $q = r-3$ the graph $G^*[A_{q+1}^*]$ might violate condition (ii) of Theorem 2.5, i.e. there may exist a set $A \subseteq A_{q+1}^*$ such that $d(A, A_{q+1}^* \backslash A) \leq \tau_0$. If A is a minimal such set, we call it an *almost-component* of $G^*[A_{q+1}^*]$ (it would be a component if $d(A, A_{q+1}^* \backslash A) = 0$). If $G^*[A_{q+1}^*]$ does indeed violate condition (ii) of Theorem 2.5 then we will apply Theorem 2.5 to the almost-components of $G^*[A_{q+1}^*]$ instead.

Recall that $A_1^*, \dots, A_q^*$ all have the same size, which is a multiple of $r-1$ (the size of a large class of the bottle graph B^*). The size of A_{q+1}^* is a multiple of $|B_1^*|$. Our aim is to find a perfect B_1-packing in $G^*[A_{q+1}^*]$, where B_1 is the graph consisting of q vertex disjoint copies of K_{r-q-1} together with $r-q-2$ vertex disjoint copies of K_{r-q}^-. We think of these copies as being arranged into an $(r-q-1)$-partite graph with one vertex set of size $r-2$ and $r-q-2$ vertex sets of size $r-1$. Thus $B_1 \subseteq B_1^*$ and the vertex classes of B_1 have the same sizes as those of B_1^*. This B_1-packing in $G^*[A_{q+1}^*]$ will then be extended to a perfect K_r^--packing in G^*.

Lemma 2.7 *We can take out from G^* at most $\tau^{1/3} n^*$ vertex-disjoint copies of K_r^- to obtain subsets $A_1^\diamond, \dots, A_{q+1}^\diamond$ of $A_1^*, \dots, A_{q+1}^*$ and a subgraph $G^\diamond$ of G^* such that the sets $A_1^\diamond, \dots, A_{q+1}^\diamond$ are $(q, |G^\diamond|)$-canonical and such that $G^\diamond[A_{q+1}^\diamond]$ contains a perfect*

B_1*-packing.*

Proof. Note that in the case when $q = r - 2$ the graph B_1 just consists of $r - 2$ isolated vertices, and the existence of a perfect B_1-packing is trivial since $r - 2$ divides $|A^*_{r-1}|$. In the case when $q \le r - 3$ the proof of Lemma 2.7 will invoke the non-extremal result, Theorem 2.5, with τ_{q+1} playing the role of τ_0 there. It is for this reason that we will need the term $-\theta n$ in the minimum degree condition in Theorem 2.5. Finally, note that $\mathrm{hcf}(B_1) = 1$ (even in the case when B_1 is bipartite, i.e. when $q = r - 3$). Let $s := r - q - 1 \ge 2$. Thus B_1 is an s-partite graph. Observe that $\chi_{cr}(B_1) = \chi_{cr}(B_1^*) = \frac{s(r-1)-1}{r-1}$. Using (i) and (ii) of Lemma 2.6, similarly as in (2.2) and the first inequality in (2.3) one can show that

$$\begin{aligned} \delta(G[A^*_{q+1}]) &\ge \left(1 - \frac{1}{\chi_{cr}(B_1)} - \tau_q^{1/4}\right)|A^*_{q+1}| \\ &= \left(\frac{(s-1)(r-1)-1}{s(r-1)-1} - \tau_q^{1/4}\right)|A^*_{q+1}|. \end{aligned} \tag{2.6}$$

So the minimum degree condition of Theorem 2.5 is satisfied with $\theta := \tau_q^{1/4} \ll \tau_{q+1}$. Our choice of q implies that $G^*[A^*_{q+1}]$ satisfies condition (i) of Theorem 2.5 (with $\tau_0 := \tau_{q+1}$). Thus in the case when $s > 2$ we can apply Theorem 2.5 to find a perfect B_1-packing in $G^*[A^*_{q+1}]$.

So we only need to consider the case when $s = 2$. In this case B_1 is the bipartite graph consisting of $r - 3$ disjoint edges and one path of length 2, and we are done if condition (ii) of Theorem 2.5 holds. So suppose not and we do have some set $C_1 \subseteq A^*_{q+1}$ with $d(C_1, A^*_{q+1} \setminus C_1) \le \tau_{q+1}$. Define $C_2 := A^*_{q+1} \setminus C_1$. Then there is a vertex $x \in C_1$ which has at most $\tau_{q+1}|C_2| \le \tau_{q+1}|A^*_{q+1}|$ neighbours in C_2. Together with (2.6) this shows that $|C_1| > \delta(G^*[A^*_{q+1}]) - \tau_{q+1}|A^*_{q+1}| \ge |A^*_{q+1}|/3$. Similarly, $|C_2| > |A^*_{q+1}|/3$. (This property shows that there cannot be more than two of these almost-components, i.e. we could not split either of these two sets into further subsets satisfying the same properties.)

We now aim to show that by moving a few vertices, we can achieve that each vertex in C_1 has few neighbours in C_2 and vice versa. (This in turn will imply that the graphs induced by both C_1 and C_2 have large minimum degree.) Call a vertex $x \in C_i$ *useless* if it has at most $|C_i|/3$ neighbours in C_i. By (2.6) every such x has at least $|C_j|/3$ neighbours in the other class C_j. This is because $s = 2$ and so (2.6) gives

$$\delta(G[A^*_{q+1}]) \geq= \left(\frac{(r-1)-1}{2(r-1)-1} - \tau_q^{1/4}\right)|A^*_{q+1}| \geq |A^*_{q+1}|/3.$$

Furthermore, the low density between C_1 and C_2 shows that there are at most $\tau_{q+1}^{3/4}|A^*_{q+1}|$ useless vertices. We move each useless vertex into the other class and still denote the classes thus obtained by C_1 and C_2. Then $d(C_1, C_2) \leq \tau_{q+1}^{2/3}$. Now call a vertex x in either class *bad* if it has at least a $\tau_{q+1}^{1/6}$-fraction of the vertices in the other class as neighbours. Clearly there are at most $\tau_{q+1}^{1/2}|A^*_{q+1}|$ bad vertices. For each bad vertex $x \in C_i$ in turn we greedily choose a copy of B_1 in C_i containing x such that these copies are disjoint for distinct bad vertices. (Use that $\delta(G^*[C_i]) \geq |C_i|/4$ for $i = 1, 2$ and the fact that B_1 consists only of edges and a path of length 2 to see that such copies can be found.) By removing these copies of B_1, we end up with two sets C'_1 and C'_2 which do not contain bad vertices. So each vertex in C'_1 has at most $2\tau_{q+1}^{1/6}|C'_2|$ neighbours in C'_2 and vice versa. Since $|C'_i| \geq |A^*_{q+1}|/4$ for $i = 1, 2$ (and thus also $|C'_i| \leq 3|A^*_{q+1}|/4$ for $i = 1, 2$) this in turn implies that

$$\delta(G^*[C'_i]) \overset{(2.6)}{\geq} \left(1 - \frac{1}{\chi_{cr}(B_1)} - \tau_{q+1}^{1/7}\right)\frac{4|C'_i|}{3} > \left(1 - \frac{1}{\chi_{cr}(B_1)}\right)|C'_i|. \qquad (2.7)$$

We now aim to take out a few further copies of K_r^- from G^* to ensure that both $|C'_1|$ and $|C'_2|$ are divisible by $|B_1|$. As observed at the beginning of this section, $|A^*_{q+1}|$ is divisible by $|B_1|$. Thus $|C'_1| + |C'_2|$ is also divisible by $|B_1|$. Assume first that $|C_1'| = m|B_1| - 1$ for some $m \in \mathbb{N}$. We aim to remove $2(r-2)$ disjoint copies of K_r^- from G^* in such a way that we remove $2(r-1)$ vertices from every A_i^* with $i \le r-3$, $(r-1)+(r-2)-1$ vertices from C_1' and $(r-1)+(r-2)+1$ vertices from C_2'. Then the sizes of the remaining subsets of C_1' and C_2' will be divisible by $|B_1|$. Moreover, since the A_i^* were $(q, |G^*|)$-canonical, and since altogether we remove $2((r-1)+(r-2))$ vertices from A_{q+1}^*, the remaining subsets will still induce a canonical partition of the remaining subgraph of G^*.

The way we remove the above copies of K_r^- is as follows: Greedily find $r-2$ disjoint copies of K_r^- with two vertices in C_1', two vertices in A_i^* and one vertex in each A_j^* with $1 \le j \le r-3$ and $j \neq i$. For each of these copies of K_r^- the index i will be different except that $i = 1$ will be chosen twice. Also find $r-4$ disjoint copies of K_r^- with two vertices in C_2', two vertices in A_i^* and one vertex in each A_j^* with $1 \le j \le r-3$ and $j \neq i$. The choices of i will be between 2 and $r-3$, and no i will be chosen twice. Finally, find two copies of K_r^- with three vertices in C_2' and one in each A_i^* for $1 \le i \le r-3$.

In the general case (i.e. when $|C_i'| \equiv t \mod |B_1|$), we simply repeat this procedure t times to even out the residues modulo $|B_1|$ between $|C_1'|$ and $|C_2'|$. We denote the remaining subsets by $A_i^\diamond$ and $C_i^\diamond$ and the remaining subgraph by $G^\diamond$. We only need to perform the above procedure at most $|B_1|-1$ times, so we are taking out a bounded number of copies of K_r^-, which will not affect any of the vertex degrees significantly. Thus each $G^\diamond[C_i^\diamond]$ satisfies the minimum degree condition in Theorem 2.5. Indeed, the first inequality in (2.7) shows that

$$\delta(G^\diamond[C_i^\diamond]) \ge \left(1 - \frac{1}{\chi_{cr}(B_1)} - \tau_{q+1}^{1/8}\right)\frac{4|C_i^\diamond|}{3} \ge \left(\frac{2}{5} - \tau_{q+1}^{1/8}\right)\frac{4|C_i^\diamond|}{3} \ge \frac{51}{100}|C_i^\diamond|. \quad (2.8)$$

This bound on the minimum degree also shows that each $C_i^\diamond$ cannot contain an almost independent set of size $|C_i^\diamond|/2$, so condition (i) of Theorem 2.5 is satisfied with room to spare. To see that condition (ii) also holds, observe that if $C_i^\diamond$ is partitioned into S_1 and S_2, where $0 < |S_1| \le |C_i^\diamond|/2 \le |S_2|$, then the neighbours of any vertex in S_1 cover a significant proportion (at least 1/50) of S_2, and so $d(S_1, S_2) \ge 1/50$. So condition (ii) is satisfied too. Thus we can apply Theorem 2.5 to each of the subgraphs of $G^\diamond$ induced by $C_1^\diamond$ and $C_2^\diamond$ to find perfect B_1-packings in $G^\diamond[C_1^\diamond]$ and $G^\diamond[C_2^\diamond]$. Adding back into $A_{q+1}^\diamond$ the vertices in the copies of B_1 which were removed when dealing with the bad vertices (and letting $G^\diamond$ denote the subgraph of G induced by the modified $A_i^\diamond$), we still have a perfect B_1-packing in $G^\diamond[A_{q+1}^\diamond]$, and $G - G^\diamond$ consists of those copies of K_r^- which we removed. Thus $G^\diamond$ and the $A_i^\diamond$ are as required in the lemma. □

Our aim now is to extend the perfect B_1-packing in $G^\diamond[A_{q+1}^\diamond]$ to a perfect K_r^--packing in $G^\diamond$. To do this, we define a $(q+1)$-partite auxiliary graph J, whose vertices are the vertices in $A_i^\diamond$ for all $1 \le i \le q$ together with all the copies of B_1 in the perfect B_1-packing of $G^\diamond[A_{q+1}^\diamond]$. There will be an edge between vertices from the $A_i^\diamond$'s whenever there was one in G, and a vertex $x \in A_i^\diamond$ for $1 \le i \le q$ will be joined to a copy of B_1 whenever x was joined to all the vertices of this copy in G.

Let $H_{q,r-1}$ denote the complete $(q+1)$-partite graph with q classes of size $r-1$ and one class of size 1. We wish to find a perfect $H_{q,r-1}$-packing in J. It is easy to see that this then yields a perfect K_r^--packing in $G^\diamond$ and thus, together with all the copies of K_r^- chosen earlier, a perfect K_r^--packing in G.

The existence of such a perfect $H_{q,r-1}$-packing follows immediately from Proposition 2.4. To see that we can apply this proposition, note that Lemma 2.6(iii) implies that in G^* each vertex is adjacent to almost all vertices in the other vertex classes and this remains true in $G^\diamond$ since we only deleted a small proportion of the vertices after applying Lemma 2.6. It follows immediately that every vertex in J is adjacent to almost all vertices in the other vertex classes of J. Note also that the vertex classes of J have the correct sizes since the sets $A_1^\diamond, \dots, A_{q+1}^\diamond$ are $(q, |G^\diamond|)$-canonical. This completes the proof of Theorem 2.1.

2.6 Generalisation of Theorem 2.1

2.6.1 Definitions and Conditions

The aim of this section is to sketch how the proof of Theorem 2.1 can be extended to a larger class of graphs H, and so to a proof that for this class, the error term in Theorem 1.21 can be removed entirely.

Recall that for a graph H we defined $\sigma(H)$ to be the smallest possible size of a colour class in a $\chi(H)$-colouring of H, and the critical chromatic number to be $\chi_{cr}(H) := \frac{\chi(H)-1}{|H|-\sigma(H)}|H|$.

We now define constants $z = z(H) := |H| - \sigma(H)$ and $z_1 = z_1(H) := (\chi(H) - 1)\sigma(H)$. Thus z_1 and z are the sizes of the small and large classes of $B^*(H)$ respectively, and $\chi_{cr}(H) = z_1|H|/z\sigma(H)$.

We call a colouring of H *optimal* if it uses exactly $\chi(H)$ colours. In any optimal colouring, we always label the colours with numbers $1, 2, \dots, \chi(H)$. We call a colouring of H *appropriate* if it is optimal and the class of colour $\chi(H)$ has size $\sigma(H)$. Given an appropriate colouring c of a graph H of chromatic number $\ell \geq 3$, let $x_1 \geq x_2 \geq \dots \geq x_\ell$ be the sizes of the colour classes. Define $\mathcal{D}^{deg}(c) = \{x_i - x_{i+1} \mid i = 1, \dots, \ell-2\}$. Let $\mathcal{D}^{deg}(H)$ be the union of the sets $\mathcal{D}^{deg}(c)$ over all appropriate colourings c of H. We define $hcf_\chi^{deg}(H)$ to be the highest common factor of the elements of $\mathcal{D}^{deg}(H)$ (or $hcf_\chi^{deg}(H) := \infty$ if $\mathcal{D}^{deg}(H) = \{0\}$). Note that if $hcf_\chi^{deg}(H) = 1$ then $hcf_\chi(H) = 1$.

In Section 2.5, we used the fact that $hcf(B_1(K_r^-)) = 1$. This may not be the case for more general H. To make a similar condition slightly easier to satisfy, we define, instead of $B^*(H)$, a larger graph $D^*(H)$. Where the vertex classes of $B^*(H)$ were formed by taking the union over cyclic permutations of the large classes of H, the vertex classes of $D^*(H)$ will be formed by taking the union over all appropriate colourings, where we distinguish colourings which differ only in the ordering of the large vertex classes.

More precisely, label all the vertices of H, and consider all appropriate colourings of H. (We also distinguish two such colourings even if their only difference is a switch between colours i and j.) Suppose there are k such colourings. Then it is clear that $(\ell-1)!$ divides k, since by permuting the first $\ell-1$ colour classes we obtain another appropriate colouring. We form $D^*(H)$ by adding up the sizes of the colour classes in each of the k colourings to form ℓ classes, and adding in all edges between classes. Thus $D^*(H)$ will be the complete ℓ-partite graph with $\ell-1$ classes of size $k(|H|-\sigma)/(\ell-1)$ and one class of size $k\sigma$. Note that this is simply a blow-up of $B^*(H)$ by a factor of $k/(\ell-1)$.

On the other hand, we form $D(H)$ by taking the disjoint union of these k copies of H. We view this graph as being arranged into ℓ classes in such a way that each one of the k appropriate colourings is induced. Thus $D(H)$ is 'built up' from these k colourings by stacking them together.

For $q<\ell-1$, define $D^*_q(H)$ to be the $(\ell-q)$-partite graph induced by the last $\ell-q$ classes of $D^*(H)$, and $D_q(H)$ is the analogous non-complete graph formed by taking the union over all k appropriate colourings of the last $\ell-q$ classes of H. $D_q(H)$ will play the role that $B_1(K_r^-)$ played in the proof of Theorem 2.1

We now introduce some conditions which will generalise the properties of K_r^- which we needed in the proof of Theorem 2.1.

Conditions:

(i) H has $hcf_\chi^{deg}(H)=1$.

(ii) H has a vertex partition into $\chi(H)$ sets $B_1, \ldots, B_\ell$, where $\ell = \chi(H)$, such that B_ℓ has size at most $\sigma - 1$, and each B_i is an independent set except for $B_{\ell-1}$. Furthermore $H[B_{\ell-1}]$ contains only vertex disjoint edges.

(iii) H has an optimal colouring c_1 with a class of size $\sigma + 1$.

(iv) $hcf(D_{\ell-3}(H)) = hcf(D_{\ell-2}(H)) = 1$

Note that condition (iv) implies $hcf(D_q(H)) = 1$ for $1 \leq q \leq \ell - 2$, since $D_i(H) \subset D_j(H)$ for $i > j$. The generalisation of Theorem 2.1 is now:

Theorem 2.8 *Let H be an ℓ-partite graph, where $\ell \geq 3$, satisfying conditions (i) to (iv). Then there exists an integer $n_0 = n_0(H)$ such that every graph G whose order $n \geq n_0$ is divisible by $|H|$ and whose minimum degree is at least*

$$\left(1 - \frac{1}{\chi_{cr}(H)}\right) n$$

contains a perfect H-packing.

Note that K_r^- itself does indeed satisfy all of the above conditions (for $r \geq 4$). To see that there are other graphs satisfying the conditions, pick any integer $\sigma \geq 2$, and any $\ell \geq 3$. (Note that our construction will demand that $\sigma(H) \geq 2$, but alternative constructions would allow $\sigma(H) = 1$, as in the case of K_r^-.) We construct an ℓ-partite graph H as follows: One class will have size σ, one class will have size $\sigma + 1$, one will have size $\sigma + 2$ and the remaining $\ell - 3$ classes will have whatever size we like (but always at least σ). The classes of size $\sigma + 2, \sigma + 1$ and σ will be called X, Y and Z respectively. Note that although we haven't yet added in any edges, these sizes will ensure that conditions (i) and (iii) are satisfied. Now to ensure that (ii) is satisfied, pick a vertex $z \in Z$ and give it exactly one neighbour $y \in Y$. (z will have neighbours in the other classes of H.) Then the vertex partition required by (ii) is obtained from the present colouring by moving z into Y. To ensure that condition (iv) is satisfied, we also demand that y has no other neighbours in Z, and that we have further distinct vertices, $y', y'' \in Y$ and $z' \in Z$ such that $N_Z(y') = N_Z(y'') = \{z'\}$ and $N_Y(z') = \{y', y''\}$. In other words, these 5 vertices induce an edge and a path of length 2 and are connected to no other vertices within Y and Z. It is fairly simple to check that these restrictions guarantee that condition (iv) is satisfied. Finally, apart from all of these restrictions, we make H a complete ℓ-partite graph with the vertex classes already given. Figure 2.1 gives a picture of such graphs in the case when $\chi(H) = 3$. It is easy to check that as well as conditions (i) to (iv), H also

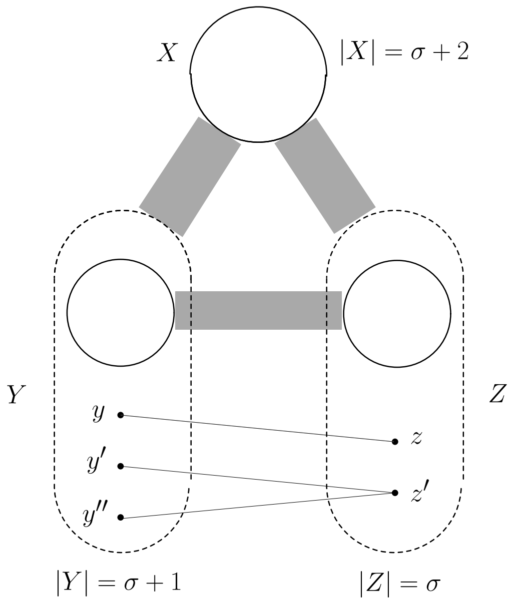


Figure 2.1: The graphs H for $\chi(H) = 3$.

satisfies $hcf(H) = 1$ and $\sigma(H) = \sigma$.

Thus there are clearly many graphs which satisfy all the necessary conditions, including K_r^-, and so Theorem 2.8 is genuinely an extension of Theorem 2.1.

2.6.2 Procedural Lemmas

Recall that by applying Theorem 2.5 we were able to assume that G had large almost independent sets, $A_1, \ldots, A_q$, and one remaining set A_{q+1}. We view G as an almost complete $(q+1)$-partite graph (with some extra edges in A_{q+1}).

In the proof of Theorem 2.1 for the graph K_r^- we repeatedly took out some copies of K_r^- in order to make these sets as nice as we would like them to be. In this subsection I aim to indicate how we can use the conditions on H required in Theorem 2.8 to do something similar in the more general case.

In particular, we will often want to 'fine-tune' the sizes of the vertex sets A_i. The following preliminary lemmas will give us the tools we need to do this. Since we are concerned only with the possible ways in which we can arrange copies of H to adjust the sizes of the A_i, I will assume for simplicity that $q = \ell - 1$, and that G is a complete ℓ-partite graph whose vertex classes are sufficiently large. (In the more general case, A_{q+1} will not be empty and will play the role of the last $\ell - q$ vertex classes, while the copies of H in G would be chosen greedily avoiding any bad vertices as in Section 2.4.)

Since the arguments for general H are rather abstract, I will also show how these lemmas correspond to the case when $H = K_r^-$, when it will be possible to give an explicit construction of the particular arrangement of copies of K_r^- which we need. In the main part of the proof of Theorem 2.1 this construction would have been stated just for the special case, without reference to the more general lemma.

In the following lemmas I will often refer to a bounded number of copies of H. This means that there is some constant $C = C(H)$, dependent only on H, which we regard as being fixed at the very beginning of the argument and a bounded number of copies of H will mean at most C copies.

Lemma 2.9 *For any distinct* $1 \leq i, j \leq \ell - 1$ *we can find a bounded number of* *copies of H which meet A_ℓ in cz_1 vertices, meet A_i in $cz+1$ vertices, meet A_j in $cz-1$ vertices and meet each other A_k in cz vertices. Here $c \le C/(\ell-1)$ is some integer dependent on H. In this case we say that we* move a residue from A_i to A_j.

The idea is that by removing these copies of H we effectively shift a vertex from A_i into A_j. In the case $H = K_r^-$ we would find two copies of K_r^- which meet A_i in exactly two vertices and each other A_k in one vertex, and then for each $1 \le s \le \ell-1$, $s \ne i, j$, a copy of K_r^- which meets A_s in two vertices and every other A_k in one vertex.

Note that we will be removing $c(\ell-1)$ copies of H meeting A_ℓ in a total of $cz_1 = c(\ell-1)\sigma$ vertices. Since G is complete ℓ-partite, it follows that each of these copies of H must meet A_ℓ in exactly σ vertices. This observation will be used in Lemma 2.10 below.

Proof. Let $d \in \mathcal{D}^{deg}(H)$. Then there is a colouring of H with colour class sizes $x_1, x_2, \ldots, x_\ell = \sigma$ such that for some $s, t \le \ell-1$ we have $x_t = x_s + d$. Without loss of generality we may assume that $s = i, t = j$.

We consider all non-trivial rotations of the $\ell-1$ large colour classes, forming copies of H with colour class sizes

$$x_2, x_3, \ldots, x_{\ell-1}, x_1, x_\ell$$

$$x_3, x_4, \ldots, x_{\ell-1}, x_1, x_2, x_\ell$$

$$\vdots$$

$$x_{\ell-1}, x_1, x_2, \ldots, x_{\ell-2}, x_\ell.$$

We also consider the colouring obtained by switching colours i and j.

$$x_1, \ldots, x_{i-1}, x_j, x_{i+1}, \ldots, x_{j-1}, x_i, x_{j+1}, \ldots, x_\ell$$

(Note that in this last line we have implicitly assumed $i < j$, which may not be the case. The proof in the case $i > j$ is identical.) By taking all these colouring arrangements (not including the initial one) together we will obtain $\ell - 1$ copies of H which meet A_ℓ in z_1 vertices, A_i in $z + d$ vertices, A_j in $z - d$ vertices and each other A_k in z vertices.

Now using the fact that $\text{hcf}(\mathcal{D}^{deg}(H)) = 1$, we can repeat this process with appropriate choices of d (note that we could have d being negative) to achieve the desired copies of H. $\square$

In the proof of Theorem 2.1 we had to deal with i-exceptional vertices in A_j. We did this by moving such a vertex into A_i. However, we then have too many vertices in A_i and too few in A_j. It is Lemma 2.9 which enables us to balance out the sizes even for the more general H in Theorem 2.8.

We also have to deal with ℓ-exceptional vertices, or vertices which are useless because they have too few neighbours in A_ℓ. Once again, in the proof of Theorem 2.8 we move such a vertex from its own class, A_j, into A_ℓ. However, we then have too many vertices in A_ℓ and not enough in A_j. The following lemma allows us to transfer the excess back, and ensure that we have the correct class sizes again.

Lemma 2.10 *For any $1 \leq j \leq \ell - 1$, we can find a bounded number of copies of H which meet A_ℓ in $cz_1 + 1$ vertices, meet A_j in $cz - 1$ vertices and meet each other A_i in cz vertices. Here $c \leq C/(\ell - 1)$ is some integer dependent on H. In this case we say that we* move a residue from A_ℓ to A_j.

Once again, the idea is to effectively move a vertex from A_ℓ into A_j. In the case

$H = K_r^-$ we would find one copy of K_r^- which has two vertices in A_ℓ and one vertex in all other classes, and for each $1 \leq s \leq \ell - 1$ where $s \neq j$ a copy of K_r^- with two vertices in A_s and one in every other class.

Proof. By condition (iii) we can find a copy of H which meets A_ℓ in $\sigma + 1$ vertices, and so meets all the other classes in a total of $|H| - \sigma - 1$ vertices. Finding $\ell - 2$ further copies of H which intersect the vertex classes in any arbitrary way, except that they all have σ vertices in A_ℓ, we obtain $\ell - 1$ copies of H which meet A_ℓ in $z_1 + 1$ vertices, and so meet all the other vertex classes in a total of $(\ell - 1)z - 1$ vertices. By Lemma 2.9 we can then find further copies of H such that each meets A_ℓ in σ vertices, and we move residues between the other classes until the sizes are as equal as possible. In particular, these copies of H will meet $\ell - 2$ classes in $c'z$ vertices and one, A_i say, in $c'z - 1$ vertices. If $i \neq j$, apply Lemma 2.9 once more to move a residue from A_i to A_j, and we have the desired copies of H. $\square$

Lemma 2.11 *Let $h := hcf(|H| - \sigma, \ell - 1)$. We can find c of copies of H, for some integer c which depends only on H and which satisfies $cz \equiv h \mod \ell - 1$, such that these copies of H meet A_ℓ in $c\sigma$ vertices, A_i in $\lceil cz/(\ell - 1) \rceil$ vertices for each $1 \leq i \leq h$ and each other A_j in $\lfloor cz/(\ell - 1) \rfloor$ vertices. In this case we say that we* remove a residue from the first h classes.

This lemma will be used to overcome the problem that although the order of G will always be divisible by $|H|$, it may not be divisible by $|B^*(H)|$, and so we may not be able to arrange the vertices of G neatly into a bottle shape.

In the case $H = K_r^-$, we have $h = 1$, and the arrangement we are looking for consists of a single copy of K_r^- with two vertices in A_1 and one vertex in every other A_i.

Proof. There is some positive integer $c' \leq \ell - 2$ such that $c'(|H| - \sigma) \equiv h$

mod $\ell - 1$. Pick c' copies of H which meet the A_i in any way we like (except that they must all meet A_ℓ in σ vertices). Then choose further copies of H to move residues between classes until the classes are as even as possible. Note that when applying Lemma 2.9, the number c'' of additional copies of H that we take out is automatically divisible by $\ell - 1$, and so the total number we have taken out is still congruent to $h \mod \ell - 1$. Let $k := (c'' + c')(|H| - \sigma)/(\ell - 1)$. Then we will be removing $\lceil k \rceil$ vertices from h of the large classes, $\lfloor k \rfloor$ vertices from the remaining $\ell - 1 - h$ large classes and $kz_1/z = (c'' + c)\sigma$ vertices from A_ℓ.

Now by applying Lemma 2.9 at most h times more, we can ensure that we remove the extra vertices from the A_i for $1 \leq i \leq h$. These copies of H are then as desired.

$\square$

Note that in the same way we can remove extra vertices from any h of the $\ell - 1$ large classes. In the proof of Theorem 2.8 we apply this lemma repeatedly until the size of the remaining subgraph of G is divisible by $|B^*(H)|$. The fact that the copies of H guaranteed by the lemma meet each A_i in about the correct number of vertices means that once we have a graph divisible by $|B^*|$ the vertex classes will form a canonical partition - i.e. they have the correct sizes as a proportion of the whole graph.

2.6.3 Sketch of the Proof of Theorem 2.8

With the help of the three procedural lemmas, we can now sketch the proof of Theorem 2.8. In essence it is exactly the same as the proof in the case $H = K_r^-$. At various points in that argument we used the specific structure of K_r^-. At these points, the lemmas above will allow us to do something similar for more general H.

So as before, in the nonextremal case, when there are no large almost independent sets, we simply invoke Theorem 2.5 to guarantee a perfect H-packing. Therefore we may assume that G does have large almost independent sets, $A_1, \ldots, A_q$, along with the remaining set A_{q+1}. As before we tidy up the classes by moving/removing exceptional and useless vertices. It is here that the procedural lemmas will be needed.

For an i-exceptional vertex $x \in A_j$ we move x into A_i and use Lemma 2.9 to remove some copies of H and in so doing effectively transfer the extra vertex back from A_i to A_j.

Condition (ii) will be used when $q = \ell$ and we have i-exceptional vertices in A_ℓ. Recall that we were able to find a large matching in A_i whenever i-exceptional vertices exist. So we aim to move the exceptional vertex from A_ℓ to A_i, and then use condition (ii) to remove a copy of H with only $\sigma - 1$ vertices in A_ℓ, thus making sure A_ℓ is not too small.

To deal with the useless vertices we again use a procedure similar to the case $H = K_r^-$. This time we apply the Erdős-Stone-Simonovits theorem to find a large number of copies of $K_{\ell-q}(s)$ in A_{q+1}, where $s \gg |H|$. For a useless but non-exceptional vertex, we pick the correct number of neighbours in the class in which it has the fewest neighbours and extend this to a copy of H. Once again we may need to deal separately with those vertices which have few neighbours in A_{q+1}, but just as in the case when $H = K_r^-$ we move these into A_{q+1} before using Lemma 2.10 to even out the class sizes again. Then we proceed as before.

Once the classes have been tidied up, we aim to find a perfect D_q-packing in A_{q+1}. Once again the argument is essentially the same as the case when $H = K_r^-$. It is for this section of the argument that we require condition (iv), since we will want to apply Theorem 2.5 for D_q to the 'almost-components' within A^*_{q+1}. In the more general case there may be more than 2 of these 'almost-components', but the same arguments as before generalise fairly naturally.

Having found a perfect D_q-packing in A^*_{q+1}, we extend it to a perfect D-packing (and therefore also a perfect H-packing) in G^* by applying Lemma 2.4. Together with the copies of H which have already been removed, this forms a perfect H-packing in G, as required.

2.6.4 Extremal Examples

In this section, I will discuss the importance of the conditions imposed on H, including which are necessary, and which could possibly be weakened without making Theorem 2.8 false.

The following proposition shows that condition (iii) is necessary in Theorem 2.8.

Proposition 2.12 *Suppose H is an ℓ-partite graph which does not satisfy condition (iii). Then for each n divisible by $|B^*(H)|$ there is a graph G on n vertices, with minimum degree $(1 - \frac{1}{\chi_{cr}(H)})n$, which does not contain a perfect H-packing.*

Proof. Let $n = k|B^*(H)| = k(\ell - 1)|H|$ for some integer k. Construct a graph G on n vertices as follows. G will be a complete ℓ-partite graph with one class, A_ℓ, of size $kz_1 + 1$, one of size $kz - 1$, and $\ell - 2$ of size kz. Then G has minimum degree

$$\delta(G) = n - kz = \left(1 - \frac{|H| - \sigma(H)}{(\ell - 1)|H|}\right) n = (1 - \frac{1}{\chi_{cr}(H)})n$$

as required.

Suppose there is a perfect H-packing in G. Then this packing must contain $k(\ell - 1)$ copies of H, and so there must be at least one copy with more than σ vertices in A_ℓ. Since there is no colouring of H with a class of size $\sigma + 1$, this copy must contain at least $\sigma + 2$ vertices in A_ℓ. But then, removing this copy, we are left with at most $(k(\ell - 1) - 1)\sigma - 1$ vertices in A_ℓ, and since each of the other copies of H meets A_ℓ in at least σ vertices, there are at most $k(\ell - 1) - 2$ further copies of

H, and so in total there were fewer than $k(\ell-1)$ copies of H, and the packing was not perfect, a contradiction. Thus no perfect H-packing exists in G. $\square$

Condition (ii) is almost necessary. The following slightly weaker condition is certainly necessary, as the proposition below shows.

(ii′) H has a vertex partition into $\chi(H)$ sets $B_1, \ldots, B_\ell$ such that B_ℓ has size at most $\sigma - 1$, and each B_i is an independent set except for $B_{\ell-1}$. Furthermore $H[B_{\ell-1}]$ contains only vertex disjoint edges if $|H| - \sigma$ is odd, and at most one vertex of degree 2 and no vertices of degree greater than 2 if $|H| - \sigma$ is even.

Proposition 2.13 *Suppose H is an ℓ-partite graph which does not satisfy condition (ii′). Then for each $n = k|B^*(H)|$, where k is an odd integer, there is a graph G on n vertices, with minimum degree $(1 - \frac{1}{\chi_{cr}(H)})n$, which does not contain a perfect H-packing.*

Proof.

Case 1: $z = |H| - \sigma$ is odd.

We construct G as follows: Let G have $\ell - 2$ classes of size kz, one class, $A_{\ell-1}$, of size $kz+1$ and one class, A_ℓ, of size $kz_1 - 1$. In order to satisfy the minimum degree condition, we add a perfect matching into $A_{\ell-1}$. (This is possible since k and z are odd, and so $kz+1$ is even.) Apart from this matching, let G be a complete ℓ-partite graph with these vertex classes. Once again, G has minimum degree

$$\delta(G) = n - kz = \left(1 - \frac{|H| - \sigma(H)}{(\ell-1)|H|}\right) n = (1 - \frac{1}{\chi_{cr}(H)})n$$

as required.

Now suppose that G contains a perfect H-packing. Then this packing would contain $k(\ell-1)$ copies of H. Since A_ℓ has size $kz_1 - 1 = k(\ell-1)\sigma - 1$, one of these copies of H must meet A_ℓ in at most $\sigma - 1$ vertices. Let $B_1, \ldots, B_\ell$ be the vertex classes of H induced by this copy in G. Since $A_1, \dots, A_{\ell-2}$ and A_ℓ are independent, so are $B_1, \dots, B_{\ell-2}$ and B_ℓ. But then, since condition (ii′) does not hold, $H[B_{\ell-1}]$ must contain a vertex of degree at least 2. But $G[A_{\ell-1}]$ does not contain such a vertex – a contradiction. So G does not contain a perfect H-packing.

Case 2: $z = |H| - \sigma$ is even.

Once again, let G have $\ell - 2$ classes of size kz, one class, $A_{\ell-1}$, of size $kz + 1$ and one class, A_ℓ, of size $kz_1 - 1$. Now $A_{\ell-1}$ has odd size, and in order to satisfy the minimum degree condition we add into $A_{\ell-1}$ a path covering 3 vertices, and a matching covering all the remaining vertices. By the same calculation as in Case 1, G does satisfy the minimum degree condition.

Now just as in the case when z was odd, we assume that a perfect H-packing exists, and we obtain $B_1, \dots, B_\ell$ where once again B_ℓ has size at most $\sigma - 1$, and all apart from $B_{\ell-1}$ are independent. But now the fact that condition (ii′) is not satisfied guarantees either two vertices of degree 2 or a vertex of degree at least 3 in $H[B_{\ell-1}]$. This is a contradiction since $G[A_{\ell-1}]$ only contains one vertex of degree 2 and none of degree greater than 2, and so no perfect matching exists. □

Unfortunately, condition (ii′) is not quite enough to make the proof work. The problem is that we might be able to adapt the extremal example above in Case 2, when $|H| - \sigma$ is even. If we no longer demand that $n = |G|$ is divisible by $|B^*(H)|$ but only by $|H|$, then we start with a graph with one class of size $\sigma n/|H|$, and all other classes as equal as possible. Now some of the large classes have size $\lfloor n/\chi_{cr}(H) \rfloor$, and some have size $\lfloor n/\chi_{cr}(H) \rfloor + 1$. We move one vertex from the small class into a class of size $\lfloor n/\chi_{cr}(H) \rfloor$. We make this graph complete ℓ-partite, and if $\lfloor n/\chi_{cr}(H) \rfloor$ is odd, then we can add a matching into the sets of size $\lfloor n/\chi_{cr}(H) \rfloor + 1$ to ensure that the minimum degree condition is satisfied. Now similarly to case 2, we cannot guarantee a perfect H-packing unless we have some modification of condition (ii′)

which allows for a vertex partition with several sets with only disjoint edges in them.

However, this problem only arises if there is such an n with $\lfloor n/\chi_{cr}(H)\rfloor$ odd, which is not always the case. Therefore a condition which is both necessary and sufficient for the proof to work would have to consider several different possible cases. Such conditions turn out to be very complicated, and so I will not consider them here.

We have seen that condition (iii) is certainly necessary, and condition (ii) is almost necessary, so let us consider conditions (i) and (iv).

The importance of condition (i) is that it allows us to prove the three procedural lemmas, each of which is vital to the argument. However, consideration of an extremal example indicates how (i) might be weakened slightly.

We would like to consider a complete $\chi(H)$-partite graph G on $n = k|B^*(H)|$ vertices, with one class of size $\sigma n/|H|$, one class of size $k(|H| - \sigma) + 1$, one class of size $k(|H| - \sigma) - 1$ and all other classes of size $k(|H| - \sigma)$. By arguments similar to those used in Section 1.1.3, it is impossible to even out classes of size $k(|H| - \sigma)$ and $k(|H| - \sigma) + 1$ by removing copies of H unless $hcf_{\chi}^{deg}(H) = 1$.

The problem is that a vertex in the class of size $k(|H| - \sigma) + 1$ does not meet the required degree condition; we would have to add some edges into this class to ensure that all of its vertices have the correct degree. So we might look for a condition similar to condition (ii), guaranteeing a partition in which only one class has any edges, and at most one vertex in this class has degree 2. If the sizes of classes of this partition worked together nicely with the colour class sizes of the appropriate colourings, then we might be able to use this property to take out one copy of H before evening out the class sizes of G.

However such a condition would already be quite complicated to state, and the situation is complicated still further by the case $\chi(H) = 3$. Informally we ignore the small classes of H and G and look for a perfect packing of a bipartite graph H' in G'. So we might look for a condition guaranteeing $hcf(H') = 1$. Since H' is bipartite, we might guess that this condition should be something like: $hcf_c^{deg}(H) = 1$ and $hcf_\chi^{deg}(H) \leq 2$. (Note that $hcf_c^{deg}(H)$ has not been defined, but it is the extension of $hcf_c(H)$ in the same way that $hcf_\chi^{deg}(H)$ is the extension of $hcf_\chi(H)$.) However, to make the condition best possible, we would once again have to consider the possibility of having vertex partitions with just one non-independent class, which itself has few edges. I believe that by the time all of this has been taken into account, any such conditions would be too complicated to be useful.

In summary, condition (i) could indeed be weakened, and Theorem 2.8 therefore strengthened, but not without sacrificing simplicity. Let us now consider condition (iv).

Condition (iv) automatically implies that $hcf_c(D_{\ell-2}(H)) = 1$, and that $hcf(D_q) = 1$ if $q = 1, 2, \ldots, \ell - 3$.

This condition could also be weakened slightly, but again this would make it much more complicated. The problem is that the last $\ell - q$ classes of an appropriate colouring do not always induce the same subgraph of H. D_q itself is the union of all the possible such subgraphs (with multiplicity). Our aim is to find a perfect D_q-packing in A_{q+1}, but it might be possible to find a perfect H-packing in G which does not induce a perfect D_q packing in A_{q+1}, if we do not demand that each colouring of H must be used an equal number of times. However, this complicates the issue considerably, and I believe any condition designed to take this into account would again be too complicated to be useful.

In short, while Theorem 2.8 is not as strong as it could be, any substantial strengthening seems to involve very technical conditions on H, so I have not studied this problem further.

CHAPTER 3

EMBEDDINGS OF TREES

The main objective in this chapter is to prove Theorem 3.1.

Theorem 3.1 *Given a positive* $C' \in \mathcal{R}$ *there exists* $k_0 \in \mathbb{N}$ *such that for any integers* $k, n \in \mathbb{N}$ *satisfying* $k_0 \leq k \leq n \leq C'k$ *the following holds: Suppose* G *is a graph on* n *vertices in which at least* $n/2$ *vertices have degree at least* k. *Then* G *contains as a subgraph every tree with* k *edges.*

Our proof of Theorem 3.1 follows the general strategy of Zhao's proof in [72] of the special case when $k = n/2$, but substantial additional difficulties arise in the more general case. The proof will be split into two main parts. In Section 3.5 we will prove that the theorem holds provided that G does not look too similar to certain extremal graphs - graphs which are close to satisfying the conditions of Theorem 3.1, but which fail to contain some tree T with k edges. We call this the non-extremal case. In that section we will use regularity arguments to embed any $T \in \mathcal{T}_k$ into G.

In Section 3.6 we will show that in the extremal case, i.e. when G is similar to some extremal graph, we have sufficient structure in G to be able to embed any tree $T \in \mathcal{T}_k$ directly. Before all this, in Section 3.4 we will prove Theorem 3.1 when G satisfies a certain special case. This is considerably shorter than either the extremal case or the non-extremal case, and will be needed in the proof of the non-extremal case.

3.1 Ramsey numbers of trees

Before going on to prove Theorem 3.1, we show how it can be used to prove Theorem 3.2 as a corollary.

Theorem 3.2 *For any real number $C'' \geq 1$ there exists an integer p_0 such that for any integers p and q satisfying $p_0 \leq p \leq q \leq C''p$ we have $R(\mathcal{T}_p, \mathcal{T}_q) \leq p + q$. In particular, for $T_p \in \mathcal{T}_p$ and $T_q \in \mathcal{T}_q$, $R(T_p, T_q) \leq p + q$.*

Proof. Given C'', let $C' = C'' + 1$, and let k_0 be the integer given by Theorem 3.1. We set $p_0 = k_0$, and let p, q be any integers such that $p_0 \leq p \leq q \leq C''p$. Suppose we have a colouring of the edges of K_n, where $n = p + q$, with two colours, red and blue. Let G_{red} denote the monochromatic red subgraph. Suppose that at least $n/2$ vertices have degree at least p in G_{red}. Then the conditions of Theorem 3.1 hold in G_{red}, where $k = p$. Thus $\mathcal{T}_p \subseteq G_{red}$ as required.

On the other hand, suppose that fewer than $n/2$ vertices have degree at least p in G. Let G_{blue} be the monochromatic blue subgraph of K_n, i.e. the complement of G_{red}. We have at least $n/2$ vertices in G_{red} with degree at most $p - 1$, and so in G_{blue} we have at least $n/2$ vertices with degree at least $n - 1 - (p - 1) = q$. Thus in G_{blue} the conditions of Theorem 3.1 hold with $k = q$, and so $\mathcal{T}_q \subseteq G_{blue}$, as required. Thus $R(\mathcal{T}_p, \mathcal{T}_q) \leq p + q$. □

We observe that this result is close to best possible. For example, if S_p and S_q are stars with p and q edges respectively, then we have

$$R(S_p, S_q) = \begin{cases} p + q - 1 & \text{if } p, q \text{ are even,} \\ p + q & \text{otherwise.} \end{cases}$$

The lower bound is seen easily by constructing a $(p-1)$-regular graph on $p+q-1$ vertices, whose complement is $(q-1)$-regular. This is not possible when p and q are both even, since then $p-1$ and $p+q-1$ are both odd, leading to the case distinction above. Thus the upper bound in Theorem 3.2 is best possible up to an error of 1.

On the other hand, if P_m denotes a path with m edges, then $R(P_m, P_m) = \lfloor (3m+1)/2 \rfloor$, as proved in [29].[1] Thus for some specific trees the bound in Theorem 3.2 is a long way from best possible.

3.2 Notation, Definitions and Preliminaries

We first introduce some more notation and definitions. Some of these definitions will be recalled later, when they are first needed. We introduce them all together here so that they can be easily found and referred to if necessary.

For a set S we write $S = A + B$ to mean that A and B form a partition of S, i.e. that $A \cup B = S$ and $A \cap B = \emptyset$.

Recall from Chapter 1 that for a disjoint pair of subsets $X, Y \subseteq V(G)$, $e(X, Y)$ denotes the number of edges with one endpoint in X and one endpoint in Y. Then $d(X, Y) := \frac{e(X,Y)}{|X||Y|}$ denotes the *density* of the pair.

While it is certainly true that $e(X, Y) = e(Y, X)$, we will occasionally distinguish between the two in order to indicate how our value or bound has been calculated. For example, if there are constants d_X, d_Y such that for all $x \in X$ we have $d(x, Y) \geq d_X$ and for all $y \in Y$ we have $d(y, X) \leq d_Y$, then we say that $e(X, Y) \geq d_X|X|$ and that $e(Y, X) \leq d_Y|Y|$. Thus the order of X and Y indicates that we calculate the bound based on some property of vertices in the first set listed.

In this chapter we will also need to consider weighted graphs, in which each edge

[1]Note that this notation may be different to other conventions, when P_m may denote a path with m vertices rather than m edges.

is assigned a weight. Then for a vertex x we will denote by $\mathbf{d}(x)$ the *weighted degree* of x, i.e. the sum of the weights of all the edges incident to x. Then for a set of vertices S, $\mathbf{d}_S(x)$ is defined analogously to the unweighted case. Also for two vertex sets X and Y we will denote by $\mathbf{e}(X, Y)$ the sum of the weights of all the edges with one endvertex in X and the other in Y.

For graphs H and G we write $H \to G$ to mean that H can be embedded into G, i.e. that G contains a copy of H as a subgraph. We also use this notation for a subset $S \subseteq V(G)$. Then $H \to S$ means that $H \to G[S]$. In this case, the graph G is implicitly understood, and it will be obvious from the context what G should be.

Given a tree T rooted at a vertex r, we define T_{odd} to be the set of vertices of T whose distance from r is odd. Similarly we define T_{even}. We consider the root to be at the top of the tree, with all other vertices hanging below it. Then for a vertex $x \neq r$ the *parent* $P(x)$ of x is the neighbour immediately above x in the tree. In other words, $P(x)$ is the neighbour of x on the unique path in T from x to r. Similarly for a set of vertices $X \subseteq V(T) - r$ we define $P(X) := \{P(x) : x \in X\}$. The *children* of x are all the neighbours immediately below x, i.e. those vertices y such that $x = P(y)$. We define $T(x)$ to be the subtree below x, i.e. the subgraph of T induced by all those vertices y for which the (unique) path between y and the root r includes the vertex x. Note that $x \in V(T(x))$.

A *skew-partition* of a tree T is a partition of $V(T)$ into sets U_1 and U_2 such that $|U_1| \leq |U_2|$ and U_2 is independent. (Note that in particular, T_{odd} and T_{even} form a skew-partition in some order.) The *gap* of a skew-partition is $g(U_1, U_2) := |U_2| - |U_1|$. The *gap of* T is defined to be $g(T) := ||T_{odd}| - |T_{even}||$.

We define the *ratio* of a tree T to be $ratio(T) := |T_{odd}|/|T|$. Given a real number $c \in (0, 1/2)$ we say that a tree T is *c-balanced*, or simply *balanced*, if $ratio(T) \in (c, 1 - c)$. We will generally use this concept for $c \ll 1$.

We call vertices of a graph G which have degree at least k *large* vertices, and vertices of degree less than k are called *small* vertices. We denote by $L(G)$ the set of vertices in G which are large, and $S(G)$ denotes the set of small vertices. Thus for the graph G which we consider in Theorem 3.1 we have $|L(G)| \geq |G|/2$.

In many places during the proof, we will observe that if we have a vertex b in T which is adjacent to a leaf c, and if b has been embedded onto a vertex y in $L(G)$, then we can always embed the leaf c onto a neighbour of y greedily after performing any other necessary embedding. This is because y has at least k neighbours, and T has $k+1$ vertices. Therefore if all the vertices of T except for c have already been embedded, and b has been embedded onto y, at most $k-1$ neighbours of y have already been used in the embedding, and so at least one neighbour remains onto which we can embed c. From now on, if such a situation occurs, we will simply state that we can embed the appropriate leaves greedily at the end.

Throughout our proof we will omit floors and ceilings where these do not affect the argument significantly.

3.3 Outline of the Proof

We now fix various constants that we will need during our proof. First of all let C' be the constant given in the statement of Theorem 3.1. We now pick k_0 to be sufficiently large, and let k and n be the integers given in Theorem 3.1. We define $C := n/k$, and note that $1 \leq C \leq C' \ll k$.

Throughout the rest of the chapter we fix further constants satisfying the following hierarchy.

$$0 < \frac{1}{k_0} \ll \varepsilon \ll \delta \ll d \ll \theta_1 \ll \theta_2 \ll \ldots \ll \theta_{\lfloor C \rfloor + 4}$$

$$\ll \tau \ll \tau' \ll \theta^{\dagger}_1 \ll \ldots \ll \theta^{\dagger}_{\lfloor C \rfloor + 2} \ll \nu' \ll c \ll \nu \ll \frac{1}{C'}.$$

Note that if we chose k_0 to be sufficiently large compared to C' then it is possible to find these constants. Note also that we have $0 < 1/n \leq 1/k \leq 1/k_0$.

We will also have some further constants which are not fixed, since we will need to apply the appropriate lemmas with different values of these constants. Most importantly, the statement of the theorem which covers the non-extremal case uses constants α_1 and α_2. The theorem will be applied with α_1 depending on θ_{i+1} and α_2 on θ_i for some i. We will also have another similar situation for α_i depending on $\theta^{\dagger}_{i'}$. In either case we will therefore have $d \ll \alpha_2 \ll \alpha_1 \ll \nu' \ll \nu \ll 1/C'$. We then define further constants to satisfy:

$$\alpha_2 \ll \eta \ll \beta \ll \rho \ll \alpha_1.$$

We observe that we may make a few preliminary assumptions about the structure of the graph G. Firstly, the conditions of Theorem 3.1 remain true if we delete any edges between small vertices. If in this modified graph we can find a copy of a tree T, then we can certainly find a copy of T in the original graph. We therefore assume that $S(G)$ is an independent set. More generally, we assume that G is edge-minimal subject to satisfying the conditions of Theorem 3.1. In particular, if there are at least $n/2 + 2$ vertices of degree at least k, then we could remove any edge from the graph and still leave at least $n/2$ vertices with degree at least k. So we may assume that $n/2 \leq |L(G)| < n/2 + 2$.

As well as edge-minimality, we will also assume that G is vertex-minimal subject to satisfying the conditions of Theorem 3.1. So we note that if there exists a set $S' \subseteq S(G)$ with the property that $|N(S')| \leq |S'|/2$, then we could delete S' and move some vertices of $N(S')$ into $S(G)$ if necessary (i.e. if they were large but now have degree less than k). This gives a new graph G' with $|G'| = |G| - |S'|$ and $|L(G')| \geq |L(G)| - |S'|/2 \geq |G'|/2$. In particular, $L(G')$ is non-empty and so $|G'| \geq k+1$. So G' also satisfies the conditions of Theorem 3.1, and by repeating the argument as often as possible, we may assume that there does not exist a set $S' \subseteq S(G)$ such that $|N(S')| \leq |S'|/2$.

Recall that in Section 3.2 we assumed that there is no set $S' \subseteq S(G)$ such that $|N(S')| \leq |S'|/2$. Slightly more generally than this, suppose that there is a set $S' \subseteq S(G)$, and a set $L' \subseteq L(G)$ such that $|L'| \leq (2/5)|S'|$, $|S'| \geq k/4$ and such that $e(S', L(G)\backslash L') \leq \tau k^2$, where τ is the constant defined in the hierarchy above. We may assume that L' is minimal given S', and in particular that every vertex of L' has at least one neighbour in S'. Then deleting S' and moving vertices of L' into $S(G)$ if necessary, we obtain a new graph $G^\dagger$ with the following property. Observe that $\tau \ll \tau' \ll \nu \ll 1$ and let $L^*(G^\dagger, G) := \{v \in V(G^\dagger) : d_{G^\dagger}(v) \geq (1-\tau')k \text{ and } d_G(v) \geq k\}$. Then we can see that $|L^*(G^\dagger, G)| \geq (1+\nu)|G^\dagger|/2$. For if not, then in $V(G^\dagger)$ we have at least $|S'|/10 - \nu|G^\dagger|/2 \geq k/41$ vertices which lie in $L(G)\backslash L'$ but not in $L^*(G^\dagger, G)$, and therefore have degree at most $(1-\tau')k$ in $G^\dagger$. They must therefore have degree at least $\tau' k$ in S', and so

$$e(S', L(G)\backslash L') \geq (\tau' k)(k/41) > \tau k^2$$

which is a contradiction.

Thus we have a new subset of the large vertices, and although they don't quite have degree k in $G^\dagger$, there are substantially more than $|G^\dagger|/2$ of them. This will enable us to embed T into G.

Theorem 3.3 *Let $0 < \tau \ll \tau' \ll \nu$ and suppose that we have subgraphs $G^\dagger \subseteq G^* \subseteq G$. Let $L^* = L^*(G^\dagger, G^*) := \{v \in V(G^\dagger) : d_{G^\dagger}(v) \geq (1-\tau')k$ and $d_{G^*}(v) \geq k\}$. Suppose furthermore that*

- *G^* was obtained from G by removing some edges between $V(G)\backslash V(G^\dagger)$ and $V(G^\dagger)\backslash L^*$ (and in particular, $V(G^*) = V(G)$);*

- $e_{G^*}\big(V(G^\dagger), V(G^*)\backslash V(G^\dagger)\big) \le \tau k^2$*;*

- $|L^*| \ge (1+\nu)|G^\dagger|/2$*.*

Then $\mathcal{T}_k \subseteq G$.

In order to apply Theorem 3.3 given sets S' and L' as above we define G^* to be the graph obtained from G by removing all edges between S' and L', and define $G^\dagger := G - S'$. We just need to check that, with L^* defined as in Theorem 3.3, $L' \cap L^* = \emptyset$ and therefore that G^* has the form described above. But recall that originally G was edge-minimal subject to satisfying the conditions of Theorem 3.1. Therefore since any vertex x in L' had a neighbour y in S' then x had degree exactly k, since $d_G(x) \ge k$ but if $d_G(x) \ge k+1$ then we could have deleted xy from G without violating the conditions of Theorem 3.1. Therefore once the edges between L' and S' are deleted, every vertex in L' has degree at most $k-1$ in G^* and so cannot lie in L^*, as required.

In our proof we will generally identify G with G^*, since if we can prove $\mathcal{T}_k \subseteq G^*$ then certainly $\mathcal{T}_k \subseteq G$. Note that it is not true that at least $n/2$ vertices in G^* have degree at least k, but we will not use this assumption in the proof of Theorem 3.3. However, it is also not necessarily true that G^* satisfies some of the assumptions that we made on G regarding edge or vertex minimality. In particular there may be a set $S'' \subseteq S(G^*)$ such that $|N_{G^*}(S'')| \le |S''|/2$, which is not the case in G. We will need this assumption on G in Section 3.6.2 and so in that section we will once again distinguish between G^* and G.

As mentioned in Section 4.1, the proof of Theorem 3.1 proceeds in two main steps, which will constitute Sections 3.5 (which covers the non-extremal case) and 3.6

(which covers the extremal case). In this section we introduce the main results of these two sections, as well as giving an outline of how they will be proved. In both sections we will further distinguish whether or not the conditions of Theorem 3.3 hold. Thus in both Sections 3.5 and 3.6 we will essentially have two subcases, one where the conditions of Theorem 3.3 hold, and one where they do not and thus there are no sets S' and L' as defined above. In the non-extremal case, this will lead us to two separate theorems, one of which will be required to prove Theorem 3.3, and one in which we will need to apply Theorem 3.3. Although the statements are distinct, the two proofs are, until the very end, essentially identical, and so we will prove them together. In the extremal case in Section 3.6 the two proofs will be slightly more distinct. It will be here that we need the constants $\theta_i^\dagger$, which play a similar role to the constants θ_i. However we need to introduce these different constants so that τ and τ' have the correct place in the hierarchy for the proof to work.

3.3.1 The Non-Extremal Case

Let us first define an extremal graph. As mentioned at the beginning of this chapter, this is a graph which is close to satisfying the conditions of Theorem 3.1, but which does not contain some tree T with k edges. In fact, the graph which we define will not contain any tree with k edges. The construction is an extension of one given in [72].

Definition. The *half-complete graph on k vertices* is a graph H_k on vertex set $V = V_1 + V_2$ where $|V_1| = \lceil k/2 \rceil$ and $|V_2| = \lfloor k/2 \rfloor$, and with edge set consisting of all pairs within V_1 and all pairs between V_1 and V_2.

Definition. Let $G_{ex}(n)$ be the graph consisting of $\lfloor C \rfloor$ disjoint copies of H_k together with further copy of $H_{n-\lfloor C \rfloor k}$ (recall that $C = n/k$).

Now $G_{ex}(n)$ does not contain any tree with k edges ($k+1$ vertices), since its components all have size at most k. However, when C is very close to an integer, we can also see that almost $n/2$ vertices have degree almost k (at least $n/2$ if k divides n exactly). $G_{ex}(n)$ therefore comes very close to satisfying the conditions of Theorem 3.1, but nevertheless fails to satisfy the conclusion. In this sense it is an extremal graph.

$G_{ex}(n)$ shows that we cannot weaken the conditions in Theorem 3.1 to demand at least $n/2$ vertices of degree $k-1$. A slight modification of $G_{ex}(n)$ shows that we also cannot substantially decrease the $n/2$ bound, i.e. we cannot get away with substantially fewer than $n/2$ vertices of degree k.

Definition. We define H'_k to be the graph on vertex set $V' = V'_1 + V'_2$ where $|V'_1| = \lfloor (k+1)/2 \rfloor - 1$ and $|V'_2| = \lceil (k+1)/2 \rceil + 1$, and with edge set consisting of all pairs within V_1 and all pairs between V_1 and V_2.

Definition. Let $G'_{ex}(n)$ be the graph consisting of $\lfloor n/(k+1) \rfloor$ disjoint copies of H'_k together with further copy of $H'_{n-\lfloor n/(k+1) \rfloor (k+1)}$.

Note that $G'_{ex}(n)$ has $\lfloor n/(k+1) \rfloor (\lfloor (k+1)/2 \rfloor - 1)$ vertices of degree k. In particular, if $0 < 1/k, 1/n \ll \varepsilon \ll 1$ then $G'_{ex}(n)$ has at least $(1-\varepsilon)n/2$ vertices of degree k. However, note that $G'_{ex}(n)$ does not contain a copy of P_k, the path on $k+1$ vertices, since such a copy of P_k would have to lie within a copy of H'_k, and so would have to contain at least $\lfloor (k+1)/2 \rfloor$ vertices within V'_1, which is impossible since V'_1 is not large enough.

$G'_{ex}(n)$ is very similar to $G_{ex}(n)$, and could be considered to be an approximation of $G_{ex}(n)$. In fact, it turns out that in some sense $G_{ex}(n)$ is the unique extremal graph.[1] This fact is captured by the Stability Theorem which will be introduced

[1]Of course we could modify $G_{ex}(n)$ by making each copy of H_k complete. However, the resulting graph would clearly contradict our assumption that there are no edges between small vertices and that there are at most $n/2+1$ large vertices.

in Section 3.3.2. However, I will not explicitly prove the Stability Theorem in this thesis - it is simply an implicit consequence of the proof of Theorem 3.1.

The extremal case partly describes the structure of the extremal graph. We denote the extremal case by EC or $EC(\alpha)$, where $\alpha \ll 1$ will be some appropriate parameter:

Definition 3.4 *$EC(\alpha)$: G contains a set of vertices A of size k such that $e(A, V(G)\backslash A) \leq \alpha k^2$.*

However, we will need to be slightly more careful than this, and so we define EC_j for $1 \leq j \leq \lfloor C \rfloor$. Recall that $\theta_1 \ll \theta_2 \ll \ldots \ll \theta_{\lfloor C \rfloor} \ll 1$.

Definition 3.5 *EC_j: G contains disjoint sets of vertices $V_1, \ldots, V_j$ each of size k such that $e(V_i, V(G)\backslash V_i) \leq \theta_j k^2$ for $1 \leq i \leq j$.*

For the proof of Theorem 3.3 we will need a similar condition, but with θ_j replaced by $\theta_j^\dagger$.

Definition 3.6 *$EC_j^\dagger$: $G^\dagger$ contains disjoint sets of vertices $V_1, \ldots, V_j$ each of size k such that $e_{G^\dagger}(V_i, V(G^\dagger)\backslash V_i) \leq \theta_j^\dagger k^2$ for $1 \leq i \leq j$.*

Definition 3.7 *If a graph G does not satisfy EC_j for any $1 \leq j \leq \lfloor C \rfloor$, we say that we are in the non-extremal case.*

Then it turns out that G is sufficiently different from $G_{ex}(n)$ that we can embed T into G even if we relax the degree conditions of Theorem 3.1 slightly. As mentioned before, we will need two versions of the non-extremal theorem. Theorem 3.9 will be required for the proof of Theorem 3.3 which in turn is required to guarantee the conditions of Theorem 3.8. However, since it is only towards the end of the proofs that the two differ significantly, we go through most of the proof for both results together.

Theorem 3.8 *Suppose we have constants satisfying* $0 < \alpha_2 \ll \alpha_1 \ll \tau \ll 1/C' \leq 1$, *and an integer* k_0 *satisfying* $0 < 1/k_0 \ll \alpha_2$. *Then for any integers* $k, n \in \mathbb{N}$ *satisfying* $k_0 \leq k \leq n \leq C'k$ *the following holds: Let* G *be a graph on* n *vertices, let* $G' \subseteq G$ *be an induced subgraph on* $n' \leq n$ *vertices, and let*

$$L = L(G', G) := \{v \in V(G') : d_{G'}(v) \geq (1 - \alpha_2)k \text{ and } d_G(v) \geq k\}.$$

Suppose that $|L| \geq (1 - \alpha_2)n'/2$, *that* $e_G(V(G'), V(G)\backslash V(G')) \leq \alpha_2^2 k^2$ *and that* G' *does not satisfy* $EC(\alpha_1)$. *Suppose furthermore that there do not exist sets* $S' \subseteq S := V(G')\backslash L$ *and* $L' \subseteq L$ *such that* $|S'| \geq k/4$, $|L'| \leq (2/5)|S'|$ *and* $e(S', L\backslash L') \leq \tau k^2$. *Then* $\mathcal{T}_k \subseteq G$.

Theorem 3.9 *Suppose that we have constants satisfying*

$$1/k_0 \ll \tau' \ll \alpha_2 \ll \alpha_1 \ll \nu \ll 1/C' \leq 1$$

and integers $k, n \in \mathbb{N}$ *such that* $k_0 \leq k \leq n \leq C'k$. *Let* $G^\dagger \subseteq G^* \subseteq G$ *be subgraphs as in the statement of Theorem 3.3, and let* $G' \subseteq G^\dagger$ *be a further subgraph on* n' *vertices. Let*

$$L = L(G', G^\dagger, G^*) := \{v \in L^*(G^\dagger, G^*) : d_{G'}(v) \geq (1 - \alpha_2)k\}.$$

Suppose that $|L| \geq (1 + \nu/2)n'/2$, *that* $e_{G^\dagger}(V(G'), V(G^\dagger)\backslash V(G')) \leq \alpha_2^2 k^2$ *and that* G' *does not satisfy* $EC(\alpha_1)$. *Then* $\mathcal{T}_k \subseteq G$.

The crucial difference between these two theorems is that in Theorem 3.8 we have the condition that there are no sets $S' \subseteq S(G)$ and $L' \subseteq L(G)$ which would have led to the existence of $G^\dagger$, while in Theorem 3.9 we assume that $G^\dagger$ exists, and thus we have the extra condition that L^* covers substantially more than half of the vertices of $G^\dagger$. This in turn leads to the condition that L covers more than half of the vertices of G', as in the statement of Theorem 3.9. To enable us to go through most of the proof of both together, we will not use either of these extra conditions until near the end of the proof. Although α_1 and α_2 appear in both Theorems, they will not be the same (as mentioned before, they will be chosen later to depend either on θ_i or on $\theta_i^\dagger$). However, we use the same notation because they will play similar roles in the two theorems, and by using the same notation we can go through both proofs together.

The interdependence of the main results in this chapter is shown in Figure 3.1. We distinguish two cases based on whether there is a set S' with the properties

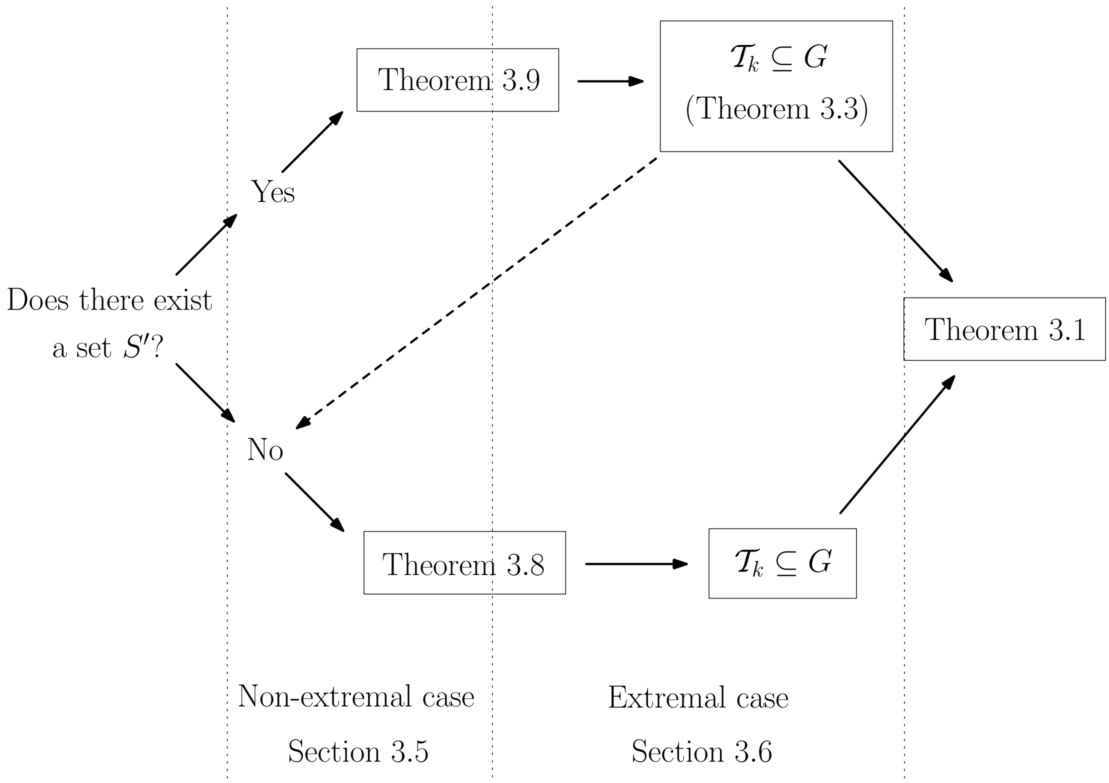


Figure 3.1: The interdependence of the main results.

described in Theorem 3.8, i.e. a set of small vertices whose neighbourhood, apart from very few edges which we ignore, lies within a set L' of large vertices where

$|L'| \leq (2/5)|S'|$. If there is such a set S', then we remove it. A few large vertices (in L') may now become small, but nevertheless we now have significantly more large vertices than small ones, and the conditions of Theorem 3.3 are satisfied. In this situation if we need to apply the non-extremal case, then the conditions of Theorem 3.9 will be satisfied and so this theorem, proved in Section 3.5, will be the non-extremal theorem corresponding to Theorem 3.3. We then go on to complete the proof of Theorem 3.3 in Section 3.6.2, and so in this case $\mathcal{T}_k \subseteq G$.

On the other hand if there is no such set S', then the non-extremal case will be covered by Theorem 3.8, which will also be proved in Section 3.5. We then go on to complete the proof in the extremal case of this situation, and therefore complete the proof of Theorem 3.1, in Section 3.6.

Although it is natural to think of the argument as composed of two parallel situations which together cover the whole proof, as shown in Figure 3.1, we can also think of it as a linear argument. Since we aim ultimately to prove Theorem 3.1 by contradiction, Theorem 3.3 can be interpreted as stating that the conditions of the theorem are impossible under the assumption that $\mathcal{T}_k \nsubseteq G$, and so ultimately there is no set S' which would have given rise to those conditions, as indicated by the dashed arrow in Figure 3.1.

In both situations, the extremal case will use the corresponding non-extremal theorem. Roughly, if $V_1, \ldots, V_j$ are sets as in the definition of the extremal case, then we look to apply the appropriate non-extremal theorem to $G[V_0]$, where $V_0 = V(G) \setminus \bigcup_{i=1}^{j} V_i$. This will ensure either that $\mathcal{T}_k \subseteq G$, in which case we are done, or that the conditions of the non-extremal theorem do not hold and so $V_1, \ldots, V_j$ satisfy certain conditions which will be useful to us in the extremal case.

More precisely, note that if G does not satisfy EC, then $G' = G$ will satisfy the conditions of Theorem 3.8 (or similarly if $G^\dagger$ does not satisfy EC, then $G' = G^\dagger$ will satisfy the conditions of Theorem 3.9). However, we need the stronger statement here because in our proof of Theorem 3.1 (or Theorem 3.3) we will consider the maximal j for which EC_j (or correspondingly $EC_j^\dagger$) holds in G (in $G^\dagger$). We will then apply Theorem 3.8 (Theorem 3.9) to $G' = G - \bigcup_{i=1}^{j} V_i$ (or $G' = G^\dagger - \bigcup_{i=1}^{j} V_i$). Since we assumed that j was maximal, $EC(\theta_{j+1})$ (or $EC(\theta_{j+1}^\dagger)$) will not hold in G', and we will show that the remaining conditions of the theorem also hold unless $j = \lfloor C \rfloor = \lfloor n/k \rfloor$ (or $j = \lfloor |G^\dagger|/k \rfloor$). Thus we may assume that G (or $G^\dagger$) splits completely into "almost components" of size k and one leftover set of size less than k. In Section 3.6 we will go on to use this structure to embed T directly into G.

The proof of Theorems 3.8 and 3.9 will make use of Szemerédi's regularity lemma. Using the standard fact that various properties of the original graph are inherited by the reduced graph, we will be able to prove a structure lemma (Lemma 3.16). This will give us two adjacent clusters A and B in the reduced graph, together with a matching $\mathcal{M}$ into which both A and B have appropriately high degree. We will then split the tree T into a (small) number of sub-trees in an appropriate way. (To recover T, we re-connect the roots of these trees to their original parent vertices in T.) The roots and the parent vertices will be embedded into A and B, while the remaining vertices will be embedded into $\mathcal{M}$.

3.3.2 The Extremal Case

In the extremal case we need to be more careful, since G may be close to a graph which does not contain some $T \in \mathcal{T}_k$, and using the regularity lemma would remove some edges.

Instead, we use the structure which we already know we have in G. Recall that we could assume that G splits completely into $\lfloor C \rfloor$ almost components of size k and one leftover set of size at most k. We can show first that this left-over set has size almost k or almost zero (relative to k). For the proof of Theorem 3.3 this will already be enough. Secondly, we prove that in fact we may assume that a stronger version of the extremal case holds, which we call EC'. We delay the precise definition of EC' until Section 3.6.1. Roughly it guarantees that in addition to the properties of EC, we may also assume that every vertex of $L \cap V_i$ has almost all its neighbours in V_i. We will then use this stronger structure to embed the tree T directly into G (with most of T generally being embedded into just one of the V_i).

Together with Theorem 3.8, this gives us the following Stability Theorem (c.f. Theorem 1.9 in [72]).

Theorem 3.10 (Stability Theorem) *For every $\mu > 0$ and $C \geq 1$ there exist $\varepsilon > 0$ and $k_0 \in \mathbb{N}$ such that for any $k_0 \leq k \leq n \leq Ck$ the following holds: Suppose that G is a graph on n vertices with $|L(G)| \geq (1-\varepsilon)n/2$. Suppose furthermore that no proper subgraph $G' \subset G$ satisfies $L(G') \geq (1-\varepsilon)|G'|/2$, and that G does not contain some $T \in \mathcal{T}_k$. Then G can be transformed into $G_{ex}(\lfloor C \rfloor k)$ or $G_{ex}(\lceil C \rceil k)$ by adding or deleting at most μk vertices and at most μk^2 edges.*

In particular either $C - \lfloor C \rfloor \leq \mu$ or $\lceil C \rceil - C \leq \mu$.

I will not prove this theorem explicitly in this thesis. However, it is an implicit consequence of the proof of Theorem 3.1.

Theorem 3.10 roughly says that $G_{ex}(n)$ is the only extremal graph when the bound on the number of large vertices is decreased by a small amount. The same cannot be said if we decrease the degree of the large vertices. For example, suppose we demand that at least $n/2$ vertices of G have degree at least $(1-\varepsilon)k$. We will assume for now that n is even. Then we partition $V(G)$ into V_1 and V_2, where $|V_1| = |V_2| = n/2$ and construct a random $\lceil (1-\varepsilon)k/2 \rceil$-regular graph within V_1 and a random $\lfloor (1-\varepsilon)k/2 \rfloor$-regular bipartite graph between V_1 and V_2. Note that the maximum degree of G is $(1-\varepsilon)k$, and so G does not contain the star S_k on

$k + 1$ vertices. So with high probability G will satisfy analogous conditions to Theorem 3.10 while the $n/2$ vertices of V_1 have degree $(1-\varepsilon)k$. However, with high probability G will not look like $G_{ex}(n)$; in fact G will be an expander. We omit the proof of these assertions. The case when n is odd is similar.

3.4 The Special Case

We first consider the following special case.

- SC: $e(L(G)) \leq \nu k^2$.

We will show that in this case we can embed the tree T into G directly, and thus we may assume that SC does not hold. We need this assumption in the non-extremal case, and therefore the following lemma has a similar form to Theorems 3.8 and 3.9, in that we have a graph G' which was obtained from G or $G^\dagger$ by removing "almost components" of size k.

Lemma 3.11 *Suppose we have constants such that* $0 < 1/k_0 \ll \nu'' \ll \nu \ll 1/C' \leq 1$ *and integers* $k, n \in \mathbb{N}$ *satisfying* $k_0 \leq k \leq n \leq C'k$. *Let* G *be a graph on* n *vertices, let* $G' \subseteq G$ *be a graph on* $n' \leq n$ *vertices, and let* $L = L(G', G) := \{v \in V(G) : d_{G'}(v) \geq (1-\nu'')k$ *and* $d_G(v) \geq k\}$. *Suppose* $|L| \geq (1-\nu'')n'/2$. *Suppose further that* $e(L) \leq \nu k^2$. *Then* $\mathcal{T}_k \subseteq G$.

We will use k_0, ν and C' as defined in our hierarchy, while ν'' may be chosen as required later. Observe that $1/k \ll \nu'' \ll \nu \ll 1/C$, where $C = n/k \leq C'$ as before. For the proof we will need the following simple fact, which appears in [72] as Fact 5.13.

Fact 3.12 *If the vertex set of a tree* T *is partitioned into two subsets* U_1 *and* U_2 *such that* U_2 *is an independent set, then* U_2 *contains at least* $|U_2| - |U_1| + 1$ *leaves of* T. □

Proof of Lemma 3.11. Let $S := V(G')\backslash L$. Note that since $e(L) \leq \nu k^2$, there are at most $2\sqrt{\nu}k$ vertices of L with more than $\sqrt{\nu}k$ neighbours in L. Removing such vertices, we obtain L' with $\delta(L', S) \geq (1 - \nu'' - \sqrt{\nu})k \geq (1 - 2\sqrt{\nu})k$, and $|L'| \geq (1 - \nu'')n'/2 - 2\sqrt{\nu}k \geq (1 - 3\sqrt{\nu})n'/2$. Let $S' \subseteq S$ be the set of vertices with at least $(1 - \nu^{1/5})k$ neighbours in L'. Now

$$e_{G'}(L', S) \geq (1 - 2\sqrt{\nu})(1 - 3\sqrt{\nu})kn'/2 \geq (1 - 5\sqrt{\nu})kn'/2.$$

Conversely, since no vertex in S has degree more than k in G',

$$\begin{aligned} e_{G'}(S, L') &\leq |S'|k + |S\backslash S'|(1 - \nu^{1/5})k \\ &\leq |S'|k + [(1 + \nu'')n'/2 - |S'|](1 - \nu^{1/5})k \\ &= \nu^{1/5}k|S'| + (1 - \nu^{1/5})(1 + \nu'')kn'/2. \end{aligned}$$

Combining these two inequalities gives

$$\begin{aligned} \nu^{1/5}|S'| &\geq ((1 - 5\sqrt{\nu}) - (1 - \nu^{1/5})(1 + \nu''))n'/2 \\ &\geq (\nu^{1/5}(1 + \nu'') - 6\sqrt{\nu})n'/2 \\ &\geq \nu^{1/5}((1 + \nu'')n'/2 - \nu^{1/5}k) \\ &\geq \nu^{1/5}(|S| - \nu^{1/5}k) \end{aligned}$$

and thus we have at most $\nu^{1/5}k$ vertices in S which have fewer than $(1 - \nu^{1/5})k$ neighbours in L'. Removing these, we obtain a bipartite subgraph $G'' \subseteq G[L' \cup S']$ with minimum degree at least $(1 - \nu^{1/6})k$.

Now T is also bipartite with classes U_1 and U_2, where without loss of generality $|U_1| \leq |U_2|$. Suppose $|U_1| \geq k/3$. Then since $|U_1| \leq |U_2| \leq 2k/3 + 1$, by the minimum degree of G'' we can embed T greedily. On the other hand if $|U_1| \leq k/3$

then by Fact 3.12, U_2 contains at least $|U_2| - |U_1| + 1 \geq k/3$ leaves. Removing these leaves gives a set U_2' of size at most $2k/3 + 1$. So we can embed U_1 and U_2' greedily into L' and S' respectively. Then since vertices of U_1 were embedded into L', whose vertices are large in G, we can embed the remaining leaves of U_2 greedily. In either case, we embed T into G as required. □

3.5 The Non-Extremal Case

We now aim to prove Theorems 3.8 and 3.9. Since the proofs are almost identical, for most of this section we will go through both together. Only towards the end of the argument will we distinguish the two proofs. The main tool that we use is Szemerédi's regularity lemma

3.5.1 The regularity lemma

In this section we will introduce another version of regularity lemma which is slightly different from the version given in the introduction, as well as defining the reduced graph again. Some of the notation here is slightly different to the notation used in the Chapters 1 and 2; this is to ensure that the notation is consistent with [15]. We will also state some standard properties of both the regularised graph and the reduced graph. Recall that given a bipartite graph with vertex classes X and Y, and given $\varepsilon > 0$, we say that the pair (X, Y) is *ε-regular* if for all subsets $X' \subseteq X$ and $Y' \subseteq Y$ which satisfy $|X'| \geq \varepsilon|X|$ and $|Y'| \geq \varepsilon|Y|$ we have

$$d(X', Y') = d(X, Y) \pm \varepsilon.$$

The version of the regularity lemma which we use is the degree form (see e.g. [51]).

Lemma 3.13 (Regularity Lemma (degree form)) *For every $\varepsilon > 0$ there is an $N_0 = N_0(\varepsilon)$ and an $n_0 = n_0(\varepsilon)$ such that for any $d \in [0,1]$ and for any graph G' on $n' \geq n_0$ vertices, there is a partition of $V(G')$ into $V_0, V_1, \ldots, V_N$ and a subgraph G'' of G' such that the following holds:*

- $N \leq N_0$
- $|V_0| \leq \varepsilon n'$
- $|V_1| = |V_2| = \ldots = |V_N| \leq \lceil \varepsilon n' \rceil$
- $e(G''[V_i]) = 0$ *for* $0 \leq i \leq N$
- *All pairs (V_i, V_j) for $1 \leq i < j \leq N$ are ε-regular in G'', with density either 0 or at least d.*
- $d_{G''}(v) \geq d_{G'}(v) - (d+\varepsilon)n$ *for every vertex $v \in V(G')$.* □

The V_i are usually called *clusters*. We apply the regularity lemma to the graph G' in Theorems 3.8 and 3.9, with constants d and ε as given in the hierarchy at the start of Section 3.3. (Note that since we had $1/k \ll \varepsilon$, and since $n' \gtrsim k$, we will have $n' \geq n_0$.) We thus obtain a regularised graph G'', which is simply G' with some edges removed. (The edges that have been removed are the edges within clusters V_i, edges between clusters (V_i, V_j) forming a non-ε-regular pair and edges between clusters (V_i, V_j) forming a regular pair of density less than d.) V_0 is the exceptional set. We generally ignore V_0, removing it from the graph G'', but still denote the "pure" graph thus obtained by G''. Note that now in G'' we still have that for each vertex v in $V(G')\backslash V_0$,

$$d_{G''}(v) \geq d_G(v) - (d+\varepsilon)n - |V_0| \geq d_G(v) - (d+2\varepsilon)n \geq d_G(v) - 2dn.$$

As in Chapter 1 we also obtain a reduced graph H on N vertices (where $N \leq N_0(\varepsilon)$). The vertices of H will be the clusters $V_1, \dots, V_N$. There will be an edge in H between two such clusters if they form an ε-regular pair of density at least d in G''. (This is equivalent to saying that there is at least one edge between these two clusters in G''.)

Note that each cluster contains approximately n'/N vertices of G'. For simplicity, we will assume that each cluster contains exactly $M := n'/N$ vertices (and in particular we assume that n'/N is an integer). This assumption does not affect any calculations significantly.

When appropriate, we will consider H to be a weighted graph. It will be clear from the context when this is intended. We define the *weight* of an edge XY in the reduced graph to be $\mathbf{d}(X, Y) := Md_{G''}(X, Y) = e_{G''}(X, Y)/M$. Thus the weight of an edge is the average number of neighbours in one cluster of a vertex in the other. If XY is not an edge then we define $\mathbf{d}(X, Y) := 0$. Recall from Section 3.2 that the weighted degree $\mathbf{d}(X)$ of a cluster X in the reduced graph is defined to be the sum of the weights of all edges incident to that cluster, i.e. $\mathbf{d}(X) = \sum_{Y \in V(H)-X} \mathbf{d}(X, Y)$.

Suppose that we have an ε-regular pair (A, B) with density d'. Then we say that a vertex $x \in A$ is *typical* with respect to B if $d_B(x)/M \in (d' - \varepsilon, d' + \varepsilon)$. By the definition of an ε-regular pair, all but at most $2\varepsilon M$ vertices of A are typical with respect to B. More generally, if we have a cluster set $\mathcal{B} = \{B_1, B_2, \dots, B_s\}$ and each pair (A, B_i) is ε-regular with density d_i, then we say that a vertex $x \in A$ is *typical with respect to* $\mathcal{B}$ if for all but $\sqrt{\varepsilon}s$ of the clusters B_i, $d_{B_i}(x)/M \in (d_i - \varepsilon, d_i + \varepsilon)$ (i.e. x is typical with respect to B_i). Then that at most $2\sqrt{\varepsilon}M$ vertices of A are not typical with respect to $\mathcal{B}$ (for each i there are at most $2\varepsilon M$ vertices atypical to B_i, giving at most $2\varepsilon Ms$ pairs (x, B_i) of vertices x and sets B_i with respect to which x is atypical). Even more generally, suppose we have subsets $B'_i \subseteq B_i$ for

$i = 1, \ldots, s$, and we define $b_i := |B'_i|/M$. Let $\mathcal{B}' = \{B'_1, B'_2, \ldots, B'_s\}$. Then we say that a vertex $x \in A$ is *typical with respect to* $\mathcal{B}'$ if for all but $\sqrt{\varepsilon}s$ of the sets B'_i, $d_{B'_i}(x)/(b_i M) \in (d_i - \varepsilon, d_i + \varepsilon)$. If $b_i \gg \varepsilon$ for each i, then it is easy to see that all but at most $\varepsilon^{1/3}M$ vertices of A are typical with respect to $\mathcal{B}'$.

3.5.2 Outline of the non-extremal case

In this section we present a short overview of the main ideas in the non-extremal case. Since the proofs of Theorem 3.8 and 3.9 are very similar, we will go through both proofs together until near the end of the argument when we need to distinguish them. The proof proceeds by contradiction and therefore we assume that there is some tree $T \in \mathcal{T}_k$ such that $T \nsubseteq G$. From this assumption we will go on to prove several properties of the tree T and the graph G, and eventually derive a contradiction.

We will apply the regularity lemma to the graph G' defined in Theorem 3.8 or 3.9 to obtain a reduced graph H. In H we define $\mathcal{L}$ to be the set of clusters which contain many vertices of L, and $\mathcal{S} := V(H)\backslash\mathcal{L}$. We think of $\mathcal{L}$ as being the "large" clusters of H, and indeed $\mathcal{L}$ inherits many of the properties of L.

We will then prove a Structure Lemma (Lemma 3.16) and apply it to H to find two adjacent clusters A and B and a cluster matching $\mathcal{M}$ such that $\mathbf{d}_{\mathcal{M}}(A)$ and $\mathbf{d}_{\mathcal{M}}(B)$ are appropriately large (recall that $\mathbf{d}_{\mathcal{M}}(A)$ denotes the total weight of edges between A and $V(\mathcal{M})$). Our aim will be to embed T primarily into $A \cup B \cup \mathcal{M}$.

In order to help us to do this, in Section 3.5.3 we split the tree T into smaller subtrees, giving us a forest in which each tree has its own root. These roots will be embedded into A or B, while the remaining vertices of a subtree will be embedded into an edge e of $\mathcal{M}$. Since all trees are bipartite, and since the subtrees are small, we will be able to use standard regularity arguments to perform this embedding.

Thus any particular subtree can be embedded easily, but we need to work to show that we can embed all of the subtrees without re-using any vertices.

The Structure Lemma in fact gives two cases, which we deal with separately.

Case 1: $\mathbf{d}_{\mathcal{M}}(A), \mathbf{d}_{\mathcal{M}}(B) \simeq k$.

In this case we can almost embed the whole tree straight away, with standard regularity arguments, but small error terms mean we fall just short of a complete embedding. Thus more work is needed.

We first show that the weighted neighbourhood of A is essentially the same as that of B, i.e. that for almost all edges $e \in \mathcal{M}$, $\mathbf{d}_e(A) \simeq \mathbf{d}_e(B)$ (Corollary 3.21).

We then show that for almost every vertex X in the the cluster matching $\mathcal{M}$, $d(A, X) \simeq 0$ or $d(A, X) \simeq 1$. (Claim 3.24).

Thirdly, we show that for almost every edge $e = (X, Y) \in \mathcal{M}$, either $\mathbf{d}_e(A)/M \simeq 0$ or $\mathbf{d}_e(A)/M \simeq 2$ (recall that M is the number of vertices in a cluster given by the regularity lemma). In other words, we do not have $d(A, X) \simeq 0$ and $d(A, Y) \simeq 1$ or vice versa (Claim 3.25).

We now consider those edges $e \in \mathcal{M}$ such that $\mathbf{d}_e(A) \simeq \mathbf{d}_e(B) \simeq 2M$. The vertices in these edges form a set of size approximately k/M, which we call $\mathcal{V}_1$. The corresponding vertices in G' form a set V_1 of size approximately k. We set $V_2 := V(G') \backslash V_1$, and $\mathcal{V}_2 := V(H) \backslash \mathcal{V}_1$. Now since G' does not satisfy the conditions of the extremal case, we know that $e_{G'}(V_1, V_2)$ is reasonably large, and so correspondingly we deduce that $e_H(\mathcal{V}_1, \mathcal{V}_2) > \rho(k/M)^2$.

On the other hand, we will split $\mathcal{V}_1$ into disjoint cluster sets $\mathcal{L}_1$ and $\mathcal{S}_1$, and show that $e(\mathcal{S}_1, \mathcal{V}_2)$ and $e(\mathcal{L}_1, \mathcal{V}_2)$ are both small (Claims 3.29 and 3.30). This will give us the required contradiction.

Case 2: $\mathbf{d}_{\mathcal{M}}(A) \simeq k$ and $\mathbf{d}_{\mathcal{L} \cup \mathcal{M}}(B) \simeq k/2$. Furthermore, every edge of $\mathcal{M}$ has at most one endvertex in the neighbourhood of A.

In this case we will first use the properties of $\mathcal{L}$ to construct a matching $\mathcal{M}_{\mathcal{L}}$ attached to $N_{\mathcal{L}\backslash\mathcal{M}}(B)$. Using this matching to augment the original matching $\mathcal{M}$, we show that we can embed T unless $\mathcal{M}_{\mathcal{L}}$ is very small. Since the former would give us the required contradiction, we can assume that $\mathcal{M}_{\mathcal{L}}$ is very small. We will use this to show that $\mathbf{d}_{\mathcal{L}\backslash\mathcal{M}}(B)$ is very small and therefore $\mathbf{d}_{\mathcal{M}}(B) \simeq k/2$.

We then observe that under the conditions of Theorem 3.9, we would actually obtain something even stronger than this. In particular, we can prove that $\mathbf{d}_{\mathcal{M}}(B)$ is significantly larger than $k/2$, and this will allow us to complete the proof of Theorem 3.9 easily, and we turn our attention to the proof of Theorem 3.8.

From here we use arguments similar to those in Case 1 to show that for almost all edges $e \in \mathcal{M}$, either $\mathbf{d}_e(B)/M \simeq 0$ or $\mathbf{d}_e(B)/M \simeq 2$ (Claims 3.32 and 3.33).

We now consider the set of clusters in those edges $e \in \mathcal{M}$ such that $\mathbf{d}_e(B)/M \simeq 2$, and we split the clusters in these edges into two sets, $\mathcal{S}_0$ and $\mathcal{L}_0$. We also consider $\mathcal{R}_0 := N_H(B)\backslash(\mathcal{L} \cup \mathcal{M})$ (see Figure 3.2). Our bounds on the degree in H of B, together with the previous results, will show that altogether $\mathcal{V}_0 := \mathcal{R}_0 \cup \mathcal{S}_0 \cup \mathcal{L}_0$ has size approximately k/M, and that $|\mathcal{L}_0| \simeq k/(4M)$.

Next we prove that there is no large matching between $\mathcal{S}_0$ and $V(H)\backslash\mathcal{V}_0$ (Claim 3.34), and no large matching between $\mathcal{R}_0$ and $V(H)\backslash\mathcal{V}_0$ (Claim 3.36). This will also imply that $\mathcal{R}_0 \cup \mathcal{S}_0$ is made up almost entirely of clusters from $\mathcal{S}$ (Corollary 3.35).

But then by considering the clusters of $\mathcal{R}_0 \cup \mathcal{S}_0$ which are also in $\mathcal{S}$ and which lie outside a maximum matching between $\mathcal{R}_0 \cup \mathcal{S}_0$ and $\mathcal{V}(H)\backslash\mathcal{V}_0$, we obtain a set $\mathcal{S}_1' \subseteq \mathcal{R}_0 \cup \mathcal{S}_0$ of size approximately $3k/(4M)$, and whose neighbourhood outside $\mathcal{V}_0$ lies only among the other endpoints of the maximum matching. Thus we can show that $N_H(\mathcal{S}_1')$ lies essentially within $\mathcal{L}_0$, and so has size less than $7|\mathcal{S}_1'|/20$.

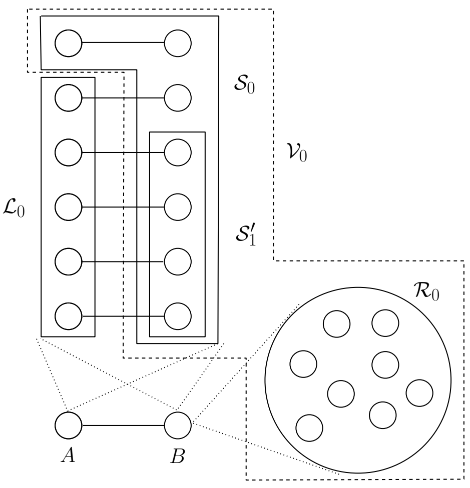


Figure 3.2: The structure of H in Case 2

However, we can show that in G' this gives rise to a set $S' \subseteq S$ of size approximately $3k/4$ and a set $L' \subseteq L$ of size at most $2|S'|/5$ such that $e_{G'}(S', L \backslash L')$ is very small. We then denote the current G' by $G^\dagger$ before deleting S', and moving some vertices of L to S if necessary, to obtain a new G' which satisfies the conditions of Theorem 3.9, which we have already proved. Thus the proof of Theorem 3.8 will also be complete.

3.5.3 Preparing the tree T

We will be attempting to embed the tree T into G using the regularity lemma. In order to help us do this, we first split T up into smaller trees.

A tree T' rooted at a vertex r' is called an *εM-tree* if it has at least εM vertices, but if every tree in the forest $T' - r'$ has fewer than εM vertices. Now if T is rooted at r and has at least εM vertices, then there must be some vertex r' such that $T(r')$ is an εM-tree (for consider the lowest r', i.e. furthest from the root r of T, such that $T(r')$ has at least εM vertices). As long as T still has more than εM vertices, we remove such an εM-tree. This process gives us a sequence of trees

$T_1, T_2, \ldots, T_t$ with roots $r_1, r_2, \ldots, r_t$, where each T_i (except possibly T_t) is an εM-tree. Let $p_i := P(r_i)$ (or if $r_i = r$, then we do not define p_i). We denote the resulting forest by $F := T_1 \cup \ldots \cup T_t$, and we can recover T from F by connecting r_i to p_i for each i. Note that the p_i are not necessarily distinct.

We now perform some extra splitting to ensure that the forest F has the sort of structure that we will need. For $i > j$ we call the two roots r_i and r_j *close roots* if $p_j \in T_i$. (This is equivalent to saying that two roots are close if the unique path between them in T contains no other roots.) It will be useful later on to have the property that any two close roots are either at even distance in T or are in fact adjacent in T (i.e. are at distance 1). Therefore if r_i and r_j are close roots, with $i > j$, and if they are at an odd distance greater than 1, we will split the forest F still further by turning p_j into a root in its own right, and deleting the edge between p_j and $P(p_j)$. Note that since p_j and r_i have even distance, this process does not create any new pairs of close roots at odd distance greater than 1. We need therefore perform the process at most once for each of the original roots. Thus in total we increase the number of roots by at most a factor of 2, and so the number of roots is still relatively small. This will be important later on.

We now no longer have that F is composed almost entirely of εM-trees. Instead, F is composed of εM-trees and trees with fewer than εM vertices. For simplicity we will generally refer to εM-trees, even though the trees may have fewer than εM vertices.

We now redefine t to be the number of trees we have after this extra splitting, and we re-enumerate the roots r_i in an appropriate way; in particular we require that if $r_j \in T(r_i)$, then $j \leq i$. Then as before we define $p_i := P(r_i)$ (unless r_i is the root r of the whole tree T). Thus we obtain a sequence of trees $T_1, T_2, \ldots, T_t$ with roots $r_1, r_2, \ldots, r_t$, where

$$t \leq 2(k+1)/(\varepsilon M) = 2(k+1)N/(\varepsilon n') \leq 2N/\varepsilon \leq f(\varepsilon) \ll k, dM. \tag{3.1}$$

Here f is some function arising from the regularity lemma.

We will always start embedding at the root r of T, and so will embed the T_i in reverse order. In this way, whenever we come to embed a vertex x of T, the only neighbour of x already embedded is $P(x)$, i.e. none of the children of x will be embedded before x. Therefore we will only need to find an image vertex for x in the neighbourhood of one vertex of G', namely the vertex chosen for $P(x)$. Sometimes we appear to embed the trees of F in some other order. However, in such cases we will actually only pick some trees, reserve some clusters of H into which we will embed them and show that they can be embedded at the appropriate time. We will always be able perform the actual embedding in reverse order of T_i, although we will not mention this explicitly from now on.

We will be attempting to embed T into clusters A and B, which are adjacent in the reduced graph, and a matching $\mathcal{M}$ in the reduced graph into which both A and B have appropriately high degree. The roots r_i will be embedded into A and B, while the remaining vertices will be embedded into the clusters of $\mathcal{M}$. It is important therefore to observe that as stated in (3.1) the number of roots and parents t is considerably smaller than M, the size of the clusters A and B.

During the non-extremal case, to ease notation we will sometimes abuse notation by writing, for example, $A \cup \mathcal{M}$, where A is a cluster and $\mathcal{M}$ is a cluster matching. In this case we mean $\{A\} \cup V(\mathcal{M})$.

We also split F into F_a and F_b. If the root r_i of T_i has an odd distance from the root r of T, we put T_i into F_a. Otherwise we put T_i into F_b. By moving the root of

T to a neighbour of r if necessary[1] we may assume that $|F_a| \geq |F_b|$. As mentioned before, we intend to embed the roots of F into A and B. More specifically, the roots of F_a will be embedded into A, and the roots of F_b into B. It is for this reason that we required that any close roots were either at even distance or at distance 1.

For if r_i and r_j are close roots, where $j < i$, then $P(r_j) \in T_i$. If for example $r_i \in F_a$, then $T_i - r_i$ will be embedded into some regular pair (X, Y) which intersects the neighbourhood of A in H, but may not intersect the neighbourhood of B. Then if $r_j \in F_a$, we will embed r_j into A, which will be possible because $P(r_j) \in F_a$ will be embedded into (X, Y) which intersects $N_H(A)$. On the other hand, if $r_j \in F_b$, then we will want to embed $P(r_j)$ into a cluster which is adjacent to B in H, which may not be the case for X or Y. But since r_j must be at odd distance from r_i, with our additional assumption we know that in fact $P(r_j) = r_i$. Therefore $P(r_j)$ has already been embedded into A, which will be a neighbour of B in H as required.

Observe from (3.1) that $t \ll dM$. Note therefore that if we have embedded a parent p_i of a root $r_i \in F_a$ and if p_i has been embedded onto a vertex x in a cluster D adjacent to A, then provided x is typical with respect to A we have at least $(d-\varepsilon)M - t \geq dM/2$ neighbours of x still available for the embedding of r_i. In fact, we will embed roots into a subset $A' \subseteq A$ of size at least $\sqrt{d}M$, which will be defined later. Provided x is typical with respect to A', at least $d^{3/2}M/2$ neighbours of x in A' will be available. Furthermore at most $\sqrt{\varepsilon}M$ vertices of D will not be typical with respect to A', and since removing these vertices will not affect any calculations significantly, we may demand that all vertices of F_a are embedded onto vertices of G'' which are typical with respect to A'. Similarly, we will assume that vertices of F_b are embedded onto vertices of G'' which are typical with respect to a subset $B' \subseteq B$ of size at least $\sqrt{d}M$.

[1] When moving the root we may have to re-order some of the trees of F to ensure that if $r_j \in T(r_i)$ then $j \leq i$.

Meanwhile, vertices of A and B onto which we embed roots may need to be typical with respect to some clusters of H, or some subsets of these clusters. These subsets of clusters will always have size at least $\sqrt{d}M$, and since $d \gg \varepsilon$, as observed when we defined typical vertices, at most $\varepsilon^{1/3}M$ vertices of A or B will not be typical with respect to such subsets as we require. On the other hand, a parent vertex whose child should be a root in A will be typical with respect to A, and thus have at least $dM/2$ available neighbours in A, and thus at least $dM/3$ available and typical neighbours. Thus we will always have appropriate unused neighbours remaining. By an identical argument, the same is true for roots to be embedded into B. Thus we will be able to perform any embedding of roots greedily, and we need only concentrate on the embedding of the remainder of F. From now on, and for the rest of the chapter, we will assume implicitly that the roots of F can always be embedded appropriately.

Let $R = \{r_1, \ldots, r_t\}$ denote the set of roots of F. We now define $Level_i(F)$, for any integer $i \geq 0$, to be the set of vertices at distance i from a root in F. Thus $Level_0(F)$ is exactly R, $Level_1(F) = N_F(R)$ etc.

3.5.4 Proof of Theorems 3.8 and 3.9

As mentioned before, most of the proofs of Theorems 3.8 and 3.9 will be presented together. It is only at the end of the argument, when the two proofs become significantly different, that we distinguish between them.

Let us first observe that under the conditions of Theorems 3.8 and 3.9, $S := V(G')\backslash L$ contains few edges. To see this we will give the argument under the conditions of Theorem 3.8; the other case is similar. For observe that since $e(V(G'), V(G)\backslash V(G')) \leq \alpha_2^2 k^2$, at most $\alpha_2 k$ vertices in $V(G')$ have at least $\alpha_2 k$ neighbours in $V(G)\backslash V(G')$, and so at most $\alpha_2 k$ vertices lie both in $L(G)$ and in S.

Thus $e(S) \le \alpha_2 kn \le \sqrt{\alpha_2}k^2$. But then, since G' does not satisfy $EC(\alpha_1)$, even if we remove all edges within S, G' does not satisfy $EC(\alpha_1 - \sqrt{\alpha_2})$. Since $\alpha_2 \ll \alpha_1$ the $\sqrt{\alpha_2}$ error term will not affect calculations significantly, and so we will assume that S is an independent set. We will find it convenient to prove Theorems 3.8 and 3.9 by contradiction. Thus we assume that we have some fixed tree $T \in \mathcal{T}_k$ such that $T \nsubseteq G$. From this assumption we will go on to prove certain properties that the tree T and the graph G must satisfy, and eventually derive a contradiction.

We begin with a claim which corresponds to Claim 5.14 in [72]. Recall that the *ratio* of a tree T' is defined to be $ratio(T') := |T'_{odd}|/|T'|$. Let $\xi := 12c$ and let $F^2 := \{T' \in F : c < ratio(T') < 1 - c\}$. In other words, F^2 is the set of balanced trees in F.

Claim 3.14 *$|V(F^2)| > ck$.*

Proof. Suppose $|V(F^2)| \le ck$. Let $F^1 := F - F^2 = \{T' \in F : ratio(T) \notin (c, 1-c)\}$. Then $|V(F^1)| \ge (1-c)k$.

For each $T' \in F^1$, either $|T'_{odd}| - |T'_{even}| \ge (1-2c)|T'|$ or $|T'_{odd}| - |T'_{even}| \le -(1-2c)|T'|$. In either case by Fact 3.12 T' contains at least $(1-2c)|T'|$ leaves. Thus F^1 contains at least $(1-2c)(1-c)k = (1-3c)k + 2c^2k$ leaves. Now by (3.1), F contains at most $f(\varepsilon)$ trees, so F has at most $2f(\varepsilon)$ more leaves than T. Since $2c^2k > 2f(\varepsilon) + 1$, T contains at least $(1-3c)k + 1$ leaves, and at most $3ck = \xi k/4$ non-leaf vertices.

Since in both Theorems we have $1/k \ll \alpha_2 \ll \nu$ we may apply Lemma 3.11 to G' with $\nu'' = \alpha_2$. Thus we may assume that SC does not hold in G', and observing that $\nu > C\xi$, we have $e(L) \ge \nu k^2 > \xi C k^2$, and so $d(G'[L]) \ge 2\xi k$. Therefore there is an induced subgraph G^* of G' with $V(G^*) \subseteq L$ and with $\delta(G^*) \ge \xi k$. We can embed the non-leaf vertices of T into G^* greedily, and since each vertex embedded is large in G, we can embed the leaves greedily, proving that $T \subseteq G$, which contradicts our initial assumption. $\square$

Claim 3.14 states that a reasonable proportion of the vertices of F are contained in balanced trees.

Let us now consider some properties of the reduced graph H. Let

$$\mathcal{L} := \{A \in V(H) : |A \cap L| \geq \sqrt{d}M\}.$$

From our comments immediately after the statement of the regularity lemma, all vertices in L still have degree at least $(1-2d)k$ in G''. Thus any cluster A of $\mathcal{L}$ contains at least $\sqrt{d}M/2$ typical (with respect to $V(H) - A$) vertices of degree at least $(1-2d)k$ in G''. We pick one such vertex, x. Then for all but $\sqrt{\varepsilon}N$ clusters $B \in V(H) - A$ we have $\mathbf{d}_H(A, B) \geq d_{G''}(x, B) - \varepsilon M$, and therefore the weighted degree of A in H is

$$\begin{aligned}\mathbf{d}_H(A) &\geq d_{G''}(x) - (\varepsilon M)N - (\sqrt{\varepsilon}N)M \\ &\geq (1-2d)k - 2\sqrt{\varepsilon}n' \\ &\geq (1-3d)k.\end{aligned}$$

In other words, the vertices of $\mathcal{L}$ are in some sense large in H (or equivalently are large clusters in G''). It is also easy to see that at least $(1-2\alpha_2)N/2$ vertices of H are in $\mathcal{L}$, for otherwise

$$\begin{aligned}|L| &\leq (1-2\alpha_2)MN/2 + \sqrt{d}MN + \varepsilon n' \\ &< (1-\alpha_2)n'/2\end{aligned}$$

which is a contradiction. For the proof of Theorem 3.9 a similar calculation shows that in this case $|\mathcal{L}| \geq (1+\sqrt{\nu'})N/2$. Finally, a cluster $A \notin \mathcal{L}$ has at most $\sqrt{d}M$ large vertices, and so most of its vertices will have degree less than k in G (and therefore also in G'). We therefore have at least $M/2$ typical (with respect to $V(H) - A$) vertices in A of degree less than k in G', and so $\mathbf{d}_H(A) \leq k + (\varepsilon M)N + (\sqrt{\varepsilon}N)M \leq (1+d)k$.

We note here that when we embed the roots of the forest F, we will embed them onto vertices of A and B which are not only typical with respect to $V(H)\backslash\{A, B\}$, but also typical with respect to the sets of large vertices in clusters of $\mathcal{L}$. More precisely, for $V_i \in \mathcal{L}$, let $L_i := L \cap V_i$. Then let $\mathcal{B} := \{L_i : V_i \in \mathcal{L}\}$. We will demand that roots of F are embedded onto vertices of A and B which are typical with respect to $\mathcal{B}$. Since $|L_i| \geq \sqrt{d}M$ for each $V_i \in \mathcal{L}$, as observed in Section 3.5.1, almost all vertices of A and B are typical with respect to $\mathcal{B}$, and so making this restriction will not affect calculations significantly.

Let $\mathcal{S} := V(H)\backslash\mathcal{L}$. Note that two clusters A and B of $\mathcal{S}$ each have subsets A', B' of size greater than $M/2$ which consist entirely of vertices from S, and so have no edges between them. Thus $d_{G''}(A', B') = 0$, and therefore $d_{G''}(A, B) = 0$, which means that A and B are non-adjacent in H. Thus $\mathcal{S}$ is an independent set in H.

Note also that if there is a set $\mathcal{S}' \subseteq \mathcal{S}$ of size at least $k/(3M)$ such that $|N_H(\mathcal{S}')| \leq 7|\mathcal{S}'|/20$, then the small vertices in the clusters of $\mathcal{S}'$ give a set S' of size at least $(M - \sqrt{d}M)|\mathcal{S}'| \geq k/4$ such that the neighbourhood of S' in G'' is contained in the clusters belonging to $N_H(\mathcal{S}')$ together with the large vertices of $\mathcal{S}'$. Thus $|N_{G''}(S')| \leq 7M|\mathcal{S}'|/20 + \sqrt{d}M|\mathcal{S}'| + |V_0| \leq 2|S'|/5$. This gives us a set $L' \subseteq L$, and for the proof of Theorem 3.8 we have

$$e_G(S', L\backslash L') = e_{G'}(S', L\backslash L') \leq d|S'||L\backslash L'| \leq dn^2 \leq \tau k^2$$

which leads us as before to the conditions of Theorem 3.3. Thus unless we are in the case when L is substantially larger than $n'/2$, i.e. in the proof of Theorem 3.9, we may assume that no such set $\mathcal{S}'$ exists.

Using these properties of $\mathcal{L}$ and $\mathcal{S}$ we will find an appropriate structure in H into which we will be able to embed T.

We will make use of the Gallai-Edmonds decomposition (see for example [22]). We say that a graph G^* is 1-*factor-critical* if for any $x \in V(G^*)$, $G^* - x$ has a perfect matching.

Theorem 3.15 *Every graph H contains a set $U \subseteq V(H)$ such that each component of $H - U$ is 1-factor-critical, and such that there is a matching which covers U and which matches the vertices of U to different components of $H - U$.* □

Using this theorem, we obtain the following lemma which will give us the appropriate structure in H. The lemma and its proof are very similar to Lemma 7 in [61].

Lemma 3.16 (Structure Lemma) *Let H be a weighted graph on N vertices, in which $\mathbf{d}(A, B) \leq M$ for all pairs of distinct vertices (A, B). Let $k \in \mathbb{N}$, and let d, α_2, η, ν' be positive real numbers satisfying $0 < d \ll \alpha_2 \ll \eta \ll \nu' \ll k/(MN)$. Suppose there is a set $\mathcal{L} \subseteq V(H)$ such that*

- *for all $x \in \mathcal{L}$, $\mathbf{d}(x) \geq (1 - 3d)k$*
- *for all $x \notin \mathcal{L}, \mathbf{d}(x) \leq (1 + d)k$*
- *$|\mathcal{L}| \geq (1 - 2\alpha_2)N/2$*
- *$\mathcal{S} = V(H)\backslash\mathcal{L}$ is independent*
- *$e(\mathcal{L}) > 0$.*

Then there are two adjacent vertices $A, B \in \mathcal{L}$ and a matching $\mathcal{M}$ in H such that one of the following holds:

1. $\mathcal{M}$ covers $N(A) \cup N(B)$ except for at most $5\alpha_2 N$ vertices.

2. $\mathcal{M}$ covers $N(A)$ and $\mathbf{d}_{\mathcal{L}\cup\mathcal{M}}(B) \geq (1-\eta)k/2$. Moreover, each edge in $\mathcal{M}$ has at most one endvertex in $N(A)$. Furthermore, if in fact $|\mathcal{L}| \geq (1+\sqrt{\nu'})N/2$, then we even have $\mathbf{d}_{\mathcal{L}\cup\mathcal{M}}(B) \geq (1+\nu')k/2$.

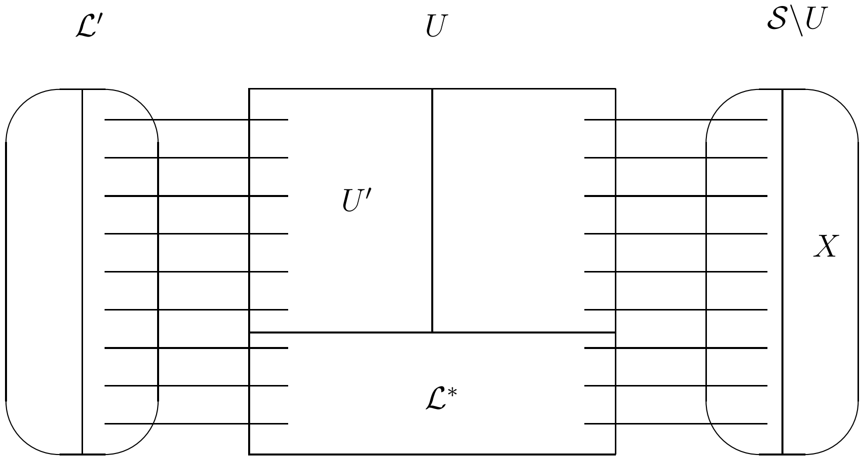
The "furthermore" in case 2 will allow us to prove Theorem 3.9, and therefore also Theorem 3.3. Because the additional assumption in this case gives us a strictly stronger condition in the conclusion, for most of the proof of Theorems 3.8 and 3.9 we will ignore it and use only the weaker bound for case 2. This will allow us to go through both proofs together. We will use the stronger bound only towards the end of the proof, when we need to distinguish the proofs of Theorems 3.8 and 3.9.

Proof. We apply Theorem 3.15 to (the unweighted version of) H to find a set U and a matching $\mathcal{M}'$. We fix U and choose $\mathcal{M}'$ to contain the maximal number of vertices of $\mathcal{S}$. Let $\mathcal{M}$ consist of $\mathcal{M}'$ together with a maximal matching of $V(H)-\mathcal{M}'$. Now let $\mathcal{L}' := \mathcal{L}\backslash U$. If there are adjacent vertices $A, B \in \mathcal{L}'$, then they are in the same component of $H-U$, and since this component is 1-factor-critical, at most one vertex of it is not covered by $\mathcal{M}$. Since all of U is covered by $\mathcal{M}$, at most one vertex of $N(A) \cup N(B)$ is not covered by $\mathcal{M}$, so Case 1 holds.

We may therefore assume that $\mathcal{L}'$ is independent. Since $\mathcal{S}\backslash U$ is also independent, every component of $H-U$ is bipartite. But then since every component is also 1-factor-critical, each component is in fact a single vertex, and we have $\mathcal{M} = \mathcal{M}'$.

Now let $\mathcal{L}^* := N(\mathcal{L}') \cap \mathcal{L} \subseteq U$. Suppose first that $\mathcal{L}^* = \emptyset$. Then either $\mathcal{L}' = \emptyset$, in which case $U = \mathcal{L}$ (and so $|U| \geq (1-2\alpha_2)N/2$) [1], or else for every $A \in \mathcal{L}'$, $N(A) \subseteq U \cap \mathcal{S}$. (See Figure 3.3.)

[1] Note that $U \cap \mathcal{S} = \emptyset$, for otherwise, $\mathcal{M}'$ would match each vertex $A \in U \cap \mathcal{S}$ to a vertex outside U. But vertices outside U are also in $\mathcal{S}$, and $\mathcal{S}$ is independent, which is a contradiction.



$$\mathcal{L}' := \mathcal{L}\backslash U \qquad X := \mathcal{S}\backslash\mathcal{M}$$
$$U' := U \cap \mathcal{S} \qquad \mathcal{L}^* := N(\mathcal{L}') \cap U$$

Figure 3.3: The structure of H.

In this latter case $\mathbf{e}(\mathcal{L}', U\cap\mathcal{S}) \geq (1-3d)k|\mathcal{L}'|$ and $\mathbf{e}(U\cap\mathcal{S}, \mathcal{L}') \leq (1+d)k|U\cap\mathcal{S}|$. Combining the two gives $|\mathcal{L}'| \leq \frac{1+d}{1-3d}|U\cap\mathcal{S}|$. Thus

$$|\mathcal{L}'| - |U\cap\mathcal{S}| \leq \left(\frac{1+d}{1-3d} - 1\right)|U\cap\mathcal{S}| \leq \frac{4d}{1-3d}N \leq 5dN.$$

Also $|\mathcal{S}| - |\mathcal{L}| \leq 4\alpha_2 N$, and thus

$$|\mathcal{S}\backslash U| - |\mathcal{L}\cap U| = |\mathcal{S}| - |\mathcal{L}| + |\mathcal{L}'| - |U\cap\mathcal{S}| \leq 4\alpha_2 N + 5dN \leq 5\alpha_2 N.$$

In both cases we have $\mathcal{L}\cap U$ matched with $\mathcal{S}\backslash U$, and any vertices of $N(\mathcal{L})$ uncovered by $\mathcal{M}$ are in $\mathcal{S}\backslash U$. Since $|\mathcal{S}\backslash U| - |\mathcal{L}\cap U| \leq 5\alpha_2 N$, at most $5\alpha_2 N$ vertices in $N(\mathcal{L})$ are uncovered. Since $e(\mathcal{L}) > 0$ by assumption, $\mathcal{L}\cap U$ contains an edge AB and the end-vertices A and B together with the matching $\mathcal{M}$ will satisfy Case 1.

Therefore we may assume that $\mathcal{L}^* \neq \emptyset$. Let $X := \mathcal{S}\backslash V(\mathcal{M})$. Now if there exists

$B \in \mathcal{L}^*$ such that $\mathbf{d}_{H-X}(B) \geq (1-\eta)k/2$, then B together with any neighbour $A \in \mathcal{L}'$ satisfy Case 2 without the "furthermore" part, which will be proved at the end.

So we assume that $\mathbf{d}_{H-X}(B) < (1-\eta)k/2$ for every $B \in \mathcal{L}^*$. So $\mathbf{d}_X(B) \geq (1+\eta-3d)k/2$ for all $B \in \mathcal{L}^*$, and therefore

$$\mathbf{e}(\mathcal{L}^*, X) \geq (1+\eta-3d)(k/2)|\mathcal{L}^*|.$$

On the other hand

$$\mathbf{e}(X, \mathcal{L}^*) \leq (1+d)k|X|$$

and thus $|\mathcal{L}^*| \leq \frac{2(1+d)}{1+\eta-3d}|X|$.

Let $U' := U \cap \mathcal{S}$. Then $\mathbf{e}(\mathcal{L}^* \cup U', \mathcal{L}') < (1-\eta)k|\mathcal{L}^*|/2 + (1+d)k|U'|$. But for all $A \in \mathcal{L}'$, $\mathbf{d}_{\mathcal{L}^* \cup U'}(A) = \mathbf{d}_H(A) \geq (1-3d)k$. So

$$\mathbf{e}(\mathcal{L}', \mathcal{L}^* \cup U') \geq (1-3d)k|\mathcal{L}'|.$$

Thus $|\mathcal{L}'| < \frac{1-\eta}{1-3d}\frac{|\mathcal{L}^*|}{2} + \frac{1+d}{1-3d}|U'|$.

$\mathcal{S}$ is an independent set, and so $\mathcal{M}$ matches $U' \subseteq \mathcal{S}$ to $\mathcal{L}'$. Thus $|\mathcal{L}'| \geq |U'| + |\mathcal{L} \backslash \mathcal{M}|$. Therefore

$$|U'| + |\mathcal{L}\backslash\mathcal{M}| \leq \frac{1-\eta}{1-3d}\frac{|\mathcal{L}^*|}{2} + \frac{1+d}{1-3d}|U'| \leq \frac{1-\eta}{1-3d}\frac{1+d}{1+\eta-3d}|X| + \frac{1+d}{1-3d}|U'|.$$

So

$$|\mathcal{L}\backslash\mathcal{M}| \leq \frac{1+d}{(1+\eta-3d)(1-3d)}(1-\eta)|X| + \frac{4d}{1-3d}|U'| \leq (1-\eta)|X| + 5d|\mathcal{S} \cap \mathcal{M}|,$$

which we express as

$$|\mathcal{L}| - |\mathcal{L} \cap \mathcal{M}| \leq (1-\eta)(|\mathcal{S}| - |\mathcal{S} \cap \mathcal{M}|) + 5d|\mathcal{S} \cap \mathcal{M}|.$$

Thus

$$\begin{aligned} |\mathcal{L} \cap \mathcal{M}| - |\mathcal{S} \cap \mathcal{M}| &\geq |\mathcal{L}| - (1-\eta)|\mathcal{S}| - 5d|\mathcal{S} \cap \mathcal{M}| - \eta|\mathcal{S} \cap \mathcal{M}| \\ &\geq |\mathcal{L}| - (1-\eta^2)|\mathcal{S}|. \end{aligned}$$

To see the last line, observe that $|\mathcal{S}| - |\mathcal{S} \cap \mathcal{M}| = |X| \geq \mathbf{d}_X(B)/M \geq k/(2M)$ for $B \in \mathcal{L}^*$. Thus $|\mathcal{S}| - |\mathcal{S} \cap \mathcal{M}| \geq 2\eta N \geq 2\eta|\mathcal{S}|$.

Now since $|\mathcal{L}| \geq (1-2\alpha_2)N/2$ and $|\mathcal{S}| \leq (1+2\alpha_2)N/2$, we have $|\mathcal{S}| \leq \frac{1+2\alpha_2}{1-2\alpha_2}|\mathcal{L}|$. So

$$|\mathcal{L} \cap \mathcal{M}| - |\mathcal{S} \cap \mathcal{M}| \geq |\mathcal{L}| - (1-\eta^2)\frac{1+2\alpha_2}{1-2\alpha_2}|\mathcal{L}| > 0.$$

Thus $\mathcal{M}$ must contain two adjacent vertices of $\mathcal{L}$, A and B say. Assume without loss of generality that $A \in \mathcal{L}', B \in \mathcal{L}^*$. Now B has a neighbour $D \in X$. But then replacing AB with BD in $\mathcal{M}$ gives a matching covering more vertices of $\mathcal{S}$ than $\mathcal{M} = \mathcal{M}'$ does, contradicting the choice of $\mathcal{M}'$.

To see the "furthermore" in case 2, suppose that $|\mathcal{L}| \geq (1+\sqrt{\nu'})N/2$, and suppose that $d_{H-X}(B) < (1+\nu')k/2$ for every $B \in \mathcal{L}^*$ (otherwise the conditions would be satisfied immediately). The argument is similar to the previous argument in the case when $\mathcal{L}^* \neq \emptyset$, and we simply alter the calculations from that case.

Let $X = \mathcal{S} \backslash V(\mathcal{M})$ and let $U' = U \cap \mathcal{S}$ be defined as before. We have $\mathbf{d}_X(B) \geq (1-\nu'-3d)k/2$ for all $B \in \mathcal{L}^*$, and therefore

$$\mathbf{e}(\mathcal{L}^*, X) \geq (1-\nu'-3d)(k/2)|\mathcal{L}^*|.$$

On the other hand

$$\mathbf{e}(X,\mathcal{L}^*) \le (1+d)k|X|$$

and thus $|\mathcal{L}^*| \le \frac{2(1+d)}{1-\nu'-3d}|X|$.

Now $\mathbf{e}(\mathcal{L}^*\cup U',\mathcal{L}') < (1+\nu')k|\mathcal{L}^*|/2+(1+d)k|U'|$. But for all $A \in \mathcal{L}'$, $\mathbf{d}_{\mathcal{L}^*\cup U'}(A) = \mathbf{d}_H(A) \ge (1-3d)k$, and so

$$\mathbf{e}(\mathcal{L}',\mathcal{L}^*\cup U') \ge (1-3d)k|\mathcal{L}'|.$$

Thus $|\mathcal{L}'| < \frac{1+\nu'}{1-3d}\frac{|\mathcal{L}^*|}{2} + \frac{1+d}{1-3d}|U'|$.

$\mathcal{S}$ is an independent set, and so $\mathcal{M}$ matches $U' \subseteq \mathcal{S}$ to $\mathcal{L}'$. Thus $|\mathcal{L}'| \ge |U'| + |\mathcal{L}\backslash\mathcal{M}|$. Therefore

$$|U'| + |\mathcal{L}\backslash\mathcal{M}| \le \frac{1+\nu'}{1-3d}\frac{|\mathcal{L}^*|}{2} + \frac{1+d}{1-3d}|U'| \le \frac{1+\nu'}{1-\nu'-3d}\frac{1+d}{1-3d}|X| + \frac{1+d}{1-3d}|U'|.$$

So

$$|\mathcal{L}\backslash\mathcal{M}| \le \frac{1+\nu'}{1-\nu'-3d}\frac{1+d}{1-3d}|X| + \frac{4d}{1-3d}|U'| \le (1+3\nu')|X| + 5d||\mathcal{S}\cap\mathcal{M}|,$$

which we express as

$$|\mathcal{L}| - |\mathcal{L}\cap\mathcal{M}| \le (1+3\nu')(|\mathcal{S}| - |\mathcal{S}\cap\mathcal{M}|) + 5d|\mathcal{S}\cap\mathcal{M}|.$$

Thus

$$\begin{aligned} ||\mathcal{L}\cap\mathcal{M}| - |\mathcal{S}\cap\mathcal{M}| &\ge |\mathcal{L}| - (1+3\nu')|\mathcal{S}| - 5d|\mathcal{S}\cap\mathcal{M}| + 3\nu'|\mathcal{S}\cap\mathcal{M}| \\ &\ge |\mathcal{L}| - (1+\sqrt{\nu'})|\mathcal{S}|. \end{aligned}$$

Now since $|\mathcal{L}| \geq (1+\sqrt{\nu'})N/2$ we have $|\mathcal{L}| - |\mathcal{S}| \geq \sqrt{\nu'}N > \sqrt{\nu'}|\mathcal{S}|$, and so

$$|\mathcal{L} \cap \mathcal{M}| - |\mathcal{S} \cap \mathcal{M}| > 0.$$

Thus $\mathcal{M}$ must contain two adjacent vertices of $\mathcal{L}$, A and B say. Assume without loss of generality that $A \in \mathcal{L}', B \in \mathcal{L}^*$. Now B has a neighbour $D \in X$. But then replacing AB with BD in $\mathcal{M}$ gives a matching covering more vertices of $\mathcal{S}$ than $\mathcal{M} = \mathcal{M}'$ does, contradicting the choice of $\mathcal{M}'$. This therefore completes the proof of the "furthermore" part of case 2, and therefore also completes the proof of Lemma 3.16. □

We now make a remark based on the proof of Lemma 3.16 which will be required later on. We define $\mathcal{M}^2(A)$ to be the set of edges of $\mathcal{M}$ with both end-vertices lying in $N(A)$.

Remark 3.17 *Case 1 can arise in one of three ways:*

- *A: If $\mathcal{L}'$ is not independent, then A lies in some component of $H - U$ and at most one vertex of this component is uncovered by $\mathcal{M}$. Furthermore at most one vertex lying in $\mathcal{M}^2(A)$ does not lie in the same component as A. In particular, all but at most one of the large vertices of $\mathcal{M}^2(A)$ could be used as a vertex for B.*
- *B: If $\mathcal{L}'$ is independent and $\mathcal{L}^* = \emptyset$, then the conclusion of A holds.*
- *C: If $\mathcal{L}'$ is independent and $\mathcal{L}^* = \emptyset$, then $|\mathcal{M}| \geq (1/2 - 10\alpha_2)N$ and any two adjacent clusters of $\mathcal{L}$ can play the same roles as A and B.*

Lemma 3.16 gives us two possible cases. We will deal with these cases separately.

3.5.5 Case 1

We now have adjacent clusters A and B and a cluster matching $\mathcal{M}$ in $G'' - \{A, B\}$ satisfying

$$\mathbf{d}_{\mathcal{M}}(A), \mathbf{d}_{\mathcal{M}}(B) \geq (1 - 3d)k - 5\alpha_2 NM - 4M \geq (1 - \eta)k$$

(the $-4M$ term appears because we may have to delete edges from $\mathcal{M}$ incident to A and B), and furthermore $A, B \in \mathcal{L}$, so they contain many vertices of L. The proof in this case will be very similar to the proof in [72] for $k = n/2$. We will assume that $\mathbf{d}_{\mathcal{M}}(A) = \mathbf{d}_{\mathcal{M}}(B) = (1 - \eta)k$.[1]

Recall that we split our tree T into εM-trees to obtain forests F_a, F_b, and parent vertices $p_1, \ldots, p_t$ (not necessarily distinct) where $t \leq f(\varepsilon) \ll M$. Let $f_a := |F_a|$, $f_b := |F_b|$ and recall that we assume without loss of generality that $f_a \geq f_b$. Since $f_a + f_b \leq k + 1$, $f_b \leq (k + 1)/2$.

We quote two important embedding results from [72]. The first is a simple consequence of Corollary 5.7 and Lemma 5.9 Part 1 in that paper, while the second appears as Lemma 5.11.

Lemma 3.18 *Let A and B be two adjacent clusters in H. If there are disjoint cluster matchings $\mathcal{M}_a$ and $\mathcal{M}_b$ in $H - \{A, B\}$ such that*

$$f_a \leq \mathbf{d}(A, \mathcal{M}_a) - 5\sqrt{\varepsilon}n \quad \text{and} \quad f_b \leq \mathbf{d}(B, \mathcal{M}_b) - 5\sqrt{\varepsilon}n \tag{3.2}$$

then T can be embedded with $F_a \to A \cup \mathcal{M}_a$, $F_b \to B \cup \mathcal{M}_b$. □

[1] We can ensure that $(1-\eta)k \leq \mathbf{d}_{\mathcal{M}}(A), \mathbf{d}_{\mathcal{M}}(B) \leq (1-\eta)k+M$ simply by deleting some regular bipartite graphs between A or B and $V(\mathcal{M})$. Since M is comparatively small, the error term will not affect calculations significantly.

Lemma 3.19 *Let A and B be two adjacent clusters, and let $\mathcal{M}$ be a cluster matching in $H - \{A, B\}$. If a tree T satisfies*

$$|T| \leq min\{\mathbf{d}(A, \mathcal{M}), \mathbf{d}(B, \mathcal{M})\} - 12\sqrt{\varepsilon}n$$

then $T \rightarrow \{A, B\} \cup \mathcal{M}$. □

Using Lemma 3.19 we can embed into $A \cup B \cup \mathcal{M}$ a subtree T' of T of size $(1 - 2\eta)k$. Our aim now is to show that we can do slightly better than this, and embed the whole tree T.

Roughly speaking, Lemma 3.19 is proved from Lemma 3.18 in [72] simply by splitting the matching $\mathcal{M}$ into $\mathcal{M}_a$ and $\mathcal{M}_b$, where an edge e of $\mathcal{M}$ will generally be placed into $\mathcal{M}_a$ if A has a greater neighbourhood within this edge than B, i.e. if $\mathbf{d}_e(A) \geq \mathbf{d}_e(B)$. (Here $\mathbf{d}_e(A) = \mathbf{d}(A, X) + \mathbf{d}(A, Y)$, where $e = XY$.) Using this construction, we only lose in $N(B)$ at most what we need in $N(A)$, and in particular $\mathbf{d}(B, \mathcal{M}_b) = \mathbf{d}(B, \mathcal{M}) - \mathbf{d}(B, \mathcal{M}_a)$ is almost large enough to embed F_b.

However, if $\mathbf{d}_e(B)$ is substantially less than $\mathbf{d}_e(A)$, then we do not lose as much as we assumed. If this happens in many edges, then we may gain enough room to embed F_b. This is formalised in the following claim (c.f. Claim 5.15 in [72]).

Claim 3.20 *If $f_b > \eta^{1/3}k$, then $\sum_{e \in \mathcal{M}} |\mathbf{d}(A, e) - \mathbf{d}(B, e)| < \eta^{1/3}k$.*

Proof. Suppose not. We will partition $\mathcal{M}$ into $\mathcal{M}_a$ and $\mathcal{M}_b$ such that (3.2) holds. Then by Lemma 3.18, $T \rightarrow G$ which is a contradiction. We define

$$\mathcal{M}^1 := \{e \in M : \mathbf{d}(A, e) \geq \mathbf{d}(B, e)\};$$

$$\mathcal{M}^2 := \mathcal{M} - \mathcal{M}^1;$$

$$a^{(i)} := \mathbf{d}(A, \mathcal{M}^i) \text{ for } i = 1, 2;$$

$$b^{(i)} := \mathbf{d}(B, \mathcal{M}^i) \text{ for } i = 1, 2.$$

Since $a^{(1)} + a^{(2)} = b^{(1)} + b^{(2)} = (1-\eta)k$, we have that

$$a^{(1)} - b^{(1)} = b^{(2)} - a^{(2)} = \frac{1}{2}\sum_{e\in\mathcal{M}} |\mathbf{d}(A,e) - \mathbf{d}(B,e)| \geq \eta^{1/3}k/2.$$

Without loss of generality[1] we assume that $a^{(1)} \leq b^{(2)}$. Then

$$b^{(2)} - b^{(1)} = (b^{(2)} - a^{(1)}) + (a^{(1)} - b^{(1)}) \geq 0 + \eta^{1/3}k/2.$$

Also $b^{(2)} + b^{(1)} = (1-\eta)k$, and so

$$b^{(2)} \geq \frac{1}{2}((1-\eta)k + \eta^{1/3}k/2) \geq f_b + 5\sqrt{\varepsilon}n$$

where the second inequality follows since $f_b \leq (k+1)/2$. Now there exists $\mathcal{M}_b \subseteq \mathcal{M}^2$ such that

$$f_b + 5\sqrt{\varepsilon}n \leq \mathbf{d}(B,\mathcal{M}_b) \leq f_b + 5\sqrt{\varepsilon}n + 2M$$

and furthermore

$$\frac{\mathbf{d}(B,\mathcal{M}_b) - \mathbf{d}(A,\mathcal{M}_b)}{\mathbf{d}(B,\mathcal{M}_b)} \geq \frac{\mathbf{d}(B,\mathcal{M}^2) - \mathbf{d}(A,\mathcal{M}^2)}{\mathbf{d}(B,\mathcal{M}^2)} \geq \frac{\eta^{1/3}k}{2b^{(2)}}.$$

This is because we can order the edges of $\mathcal{M}^2$ as $e_1, e_2, \ldots, e_s$ in such a way that, setting $a_i := \mathbf{d}(B,e_i) - \mathbf{d}(A,e_i)$ and $b_i := \mathbf{d}(B,e_i)$ we have $a_i/b_i \geq a_{i+1}/b_{i+1}$ for $1 \leq i \leq s-1$. Then for any $s' \leq s$ we have $\frac{\sum_{i=1}^{s'} a_i}{\sum_{i=1}^{s'} b_i} \geq \frac{\sum_{i=1}^{s} a_i}{\sum_{i=1}^{s} b_i}$. We then simply pick s' to be minimum such that $\sum_{i=1}^{s'} b_i = \mathbf{d}(B,\mathcal{M}_b) \geq f_b + 5\sqrt{\varepsilon}n$.

[1] Strictly speaking this is not completely without loss of generality, because edges to which A and B had the same density were put into $\mathcal{M}^1$. However, such edges do not affect the relevant calculations.

Now $\mathcal{M}_b$ certainly satisfies (3.2). Also

$$\mathbf{d}(B,\mathcal{M}_b) - \mathbf{d}(A,\mathcal{M}_b) \geq \mathbf{d}(B,\mathcal{M}_b)\frac{\eta^{1/3}k}{2b^{(2)}} > f_b\frac{\eta^{1/3}k}{2k} > \eta^{2/3}k/2.$$

Now let $\mathcal{M}_a := \mathcal{M} - \mathcal{M}_b$. Then

$$\begin{aligned}
\mathbf{d}(A,\mathcal{M}_a) &= \mathbf{d}(A,\mathcal{M}) - \mathbf{d}(A,\mathcal{M}_b) \\
&= \mathbf{d}(A,\mathcal{M}) + (\mathbf{d}(B,\mathcal{M}_b) - \mathbf{d}(A,\mathcal{M}_b)) - \mathbf{d}(B,\mathcal{M}_b) \\
&\geq (1-\eta)k + \eta^{2/3}k/2 - (f_b + 5\sqrt{\varepsilon}n + 2M) \\
&\geq f_a + 5\sqrt{\varepsilon}n.
\end{aligned}$$

Thus $\mathcal{M}_a$ and $\mathcal{M}_b$ satisfy (3.2), as required. □

Corollary 3.21 *If $f_b \geq \eta^{1/3}k$, then except for at most $\eta^{1/6}k/M$ edges, all edges in $\mathcal{M}$ satisfy $|\mathbf{d}(A,e) - \mathbf{d}(B,e)| < \eta^{1/6}M$.* □

We now define a new constant β such that $\eta \ll \beta \ll \alpha_1$. The following claim (c.f. Claim 5.16 in [72]) is the crucial point of the argument. It says that if somewhere within the matching we can embed slightly more of the forest F than we expected to, then we gain enough room that we can embed the rest of F into the rest of the matching.

Claim 3.22 *Let $\mathcal{M}_0 \subseteq \mathcal{M}$ be a matching of size at most $k/(4M)$. Suppose that we have a subforest $\tilde{F}_a \subseteq F_a$ with $|\tilde{F}_a| \geq \mathbf{d}(A,\mathcal{M}_0) + \beta k$, and suppose that for any embedding of Root($\tilde{F}_a$), the set of roots of $\tilde{F}_a$, to vertices of A typical with respect to $\mathcal{M}_0$, we can extend this to an embedding of $\tilde{F}_a$ into $\mathcal{M}_0 \cup (V(H)\backslash(\mathcal{M} \cup B))$. Then $T \to G$. The corresponding result holds for a subforest $\tilde{F}_b \subseteq F_b$.*

Since both the result and the proof are essentially the same as the corresponding claim and proof in [72] we will only give an outline of the proof here.

We will partition $\mathcal{M}$ into $\mathcal{M}_a$ and $\mathcal{M}_b$ in such a way that $\mathcal{M}_0 \subseteq \mathcal{M}_a$, $\mathbf{d}(A, \mathcal{M}_a \backslash \mathcal{M}_0)$ will be slightly greater than $|F_a \backslash \tilde{F}_a|$, and $\mathbf{d}(B, \mathcal{M}_b)$ will be slightly greater than f_b. The fact that $|\tilde{F}_a|$ is greater than $\mathbf{d}(A, \mathcal{M}_0)$, together with Corollary 3.21, will ensure that we can find such $\mathcal{M}_a \backslash \mathcal{M}_0$ and $\mathcal{M}_b$. These will then satisfy Lemma 3.18, and so we can embed $T \backslash \tilde{F}_a$ into $\mathcal{M}_a \backslash \mathcal{M}_0$ and $\mathcal{M}_b$ using this lemma.

In the case when $f_b < \eta^{1/3}k$ we must be slightly more careful because Corollary 3.21 does not apply. However, in this case f_b is small enough that we can find any appropriate $\mathcal{M}_b$ similarly to the method in the proof of Claim 3.20, and removing this $\mathcal{M}_b$ will not subtract too much from $\mathbf{d}(A, \mathcal{M})$, and so we will still be able to embed $F_a \backslash \tilde{F}_a$ into $\mathcal{M}_a \backslash (\mathcal{M}_0 \cup \mathcal{M}_b)$ as required.

From now on we will assume that such $\mathcal{M}_0$ and $\tilde{F}_a$ or $\tilde{F}_b$ do not exist. Note also that when we perform our embedding, we can embed $R = Root(F)$ into large typical vertices of A and B, and so any adjacent leaf can be embedded greedily at the end. So apart from at most t parent vertices, we may assume that all vertices in $Level_1(F)$ have at least one child. Thus almost every tree in $F \backslash R$ contains at least two vertices.

We now define some notation. Suppose that we have a graph H' which we want to embed into G''. Suppose that we also have an assignment of the vertices of H' to clusters of the reduced graph H, i.e. for each vertex of H' we have already determined into which cluster we would like to embed it. We say $H' \xrightarrow{q} G''$ if there is an embedding algorithm which embeds H' into G'' one vertex at a time, which respects the pre-determined assignment, and in which we always have at least q choices in G'' for where to embed each vertex of H'. We also write $H' \xrightarrow{-q} G''$ if such an embedding exists in which for each vertex of H' we can pick all but q of the vertices in the appropriate cluster of G'' which have not yet been used in the embedding. More generally for a subset $S \subseteq V(G'')$, we write $H' \xrightarrow{q} S$ to mean $H' \xrightarrow{q} G''[S]$, and similarly for $H' \xrightarrow{-q} S$.

Let ε' be a new constant such that $\varepsilon \ll \varepsilon' \ll \delta$. We quote another result from [72] (c.f. Lemma 5.3 in that paper). In fact, Part 1 was proved in [3].

Lemma 3.23 *Let (X, Y) be an ε-regular pair with $|X| = |Y| = M$ and $d(X, Y) \geq d$. Let A be a third cluster and let $d_x := d(A, X)$, $d_y := d(A, Y)$. Let F be an ordered forest consisting of εM-trees, and with at most εM roots.*

1. If $|F| \leq (d_x + d_y - \varepsilon')M$, then there is an embedding algorithm with $x \xrightarrow{-2\varepsilon M - |R|} A$ for $x \in R = Root(F)$ and $x \xrightarrow{\varepsilon M+1} X \cup Y$ for $x \notin R$.

2. Suppose furthermore that every tree in F has ratio between c and $1 - c$ for some $0 < c \leq 1/2$ and that $d_x \leq d_y$. Then the conclusion of 1 holds provided $|F| \leq (2d_x - \varepsilon')M + \frac{1}{1-c}(d_y - d_x)M$.

3. Suppose that for some $0 \leq \lambda \leq 1/2$ we have $\lambda \leq d_x, d_y \leq (1 - \lambda)$, and that every tree in $F - Root(F)$ contains at least 2 vertices. Then the conclusion of 1 holds provided that $|F| \leq (d_x + d_y + \lambda - \varepsilon')M$. □

Note that $|F|$ denotes the number of vertices in F, and not the number of trees. Note also that since $t \leq f(\varepsilon) \ll M$, F does indeed have at most εM roots. For part 2, observe that $\frac{1}{1-c}(d_y - d_x) + 2d_x - \varepsilon' \geq d_x + d_y + c(d_y - d_x) - \varepsilon'$. We also observe that since F consists of εM-trees, we can find a subforest F' such that

$$(d_x + d_y + c(d_y - d_x) - \varepsilon' - \varepsilon)M \leq |F'| \leq (d_x + d_y + c(d_y - d_x) - \varepsilon')M.$$

Therefore provided we can find F' such that it also consists of balanced trees, then by Lemma 3.23 part 2 we will be able to embed a subforest of size at least $(d_x + d_y + c(d_y - d_x) - \varepsilon' - \varepsilon)M$.

We define $\mathcal{M}_1 = \mathcal{M}_1(A) := \{(X,Y) \in \mathcal{M} : d_{G''}(A,X) \in [\beta^{1/3}, 1-\beta^{1/3}]$ or $d_{G''}(A,Y) \in [\beta^{1/3}, 1-\beta^{1/3}]\}$. The following result roughly corresponds to Claim 5.18 in [72].

Claim 3.24 $|\mathcal{M}_1| < 2\sqrt{\beta}k/M$.

Proof. Suppose $|\mathcal{M}_1| \geq 2\sqrt{\beta}k/M$. Let $\mathcal{M}_0 \subseteq \mathcal{M}_1$ be a matching of size $2\sqrt{\beta}k/M$. Now for almost every edge $e = (X,Y) \in \mathcal{M}_0$ we can apply Lemma 3.23 to embed a subforest of $F_a \backslash R$ as large as possible. We assume without loss of generality that $d_x := d_{G''}(A,X) \leq d_{G''}(A,Y) =: d_y$. Recall that by Claim 3.14 $|V(F^2)| \geq ck$. Therefore we also assume that $|F_a \cap F^2| \geq ck/2$. We may do this without loss of generality here because we will not need the fact that $|F_a| \geq |F_b|$.

If $d_y - d_x > \beta^{1/3}/2$, we will apply Lemma 3.23 Part 2. We therefore set

$$\ell := d_x + d_y + c(d_y - d_x) - \varepsilon' \geq \mathbf{d}_e(A)/M + c\beta^{1/3}/2 - \varepsilon' \geq \mathbf{d}_e(A)/M + \sqrt{\beta}.$$

Otherwise note that $\beta^{1/3}/2 \leq d_x \leq d_y \leq 1 - \beta^{1/3}/2$, and we will apply Lemma 3.23 Part 3 with $\lambda = \beta^{1/3}/2$. In this case we set

$$\ell := d_x + d_y + \beta^{1/3}/2 - \varepsilon' \geq \mathbf{d}_e(A)/M + \sqrt{\beta}.$$

In either case we can find a subforest F_e (which consists of trees in F_a, i.e. we do not split up the trees of F_a) of size $\ell M - \varepsilon M \leq |F_e| \leq \ell M$, and so $|F_e| \geq \mathbf{d}_e(A) + 2\sqrt{\beta}M/3$. For the former case we also choose F_e from trees of $F_a \cap F^2$, and then we can apply the appropriate part of Lemma 3.23 to embed F_e into $A \cup e$. We can do this in all but at most $\sqrt{\varepsilon}|\mathcal{M}_0|$ edges if $Root(F)$ is mapped to large vertices in A which are typical with respect to $\mathcal{M}_0$. Then in total we have embedded a union of subforests $\tilde{F}_a$ with

$$\begin{aligned}|\tilde{F}_a| &\geq \mathbf{d}(A, \mathcal{M}_0) + |\mathcal{M}_0|(2\sqrt{\beta}M/3) - 2M\sqrt{\varepsilon}|\mathcal{M}_0| \\ &\geq \mathbf{d}(A, \mathcal{M}_0) + (2\sqrt{\beta}k/M)(2\sqrt{\beta}M/3) - (2M\sqrt{\varepsilon})(2\sqrt{\beta}k/M) \\ &= \mathbf{d}(A, \mathcal{M}_0) + 4\beta k/3 - 4\sqrt{\varepsilon\beta}k \\ &\geq \mathbf{d}(A, \mathcal{M}_0) + \beta k.\end{aligned}$$

So $\mathcal{M}_0$ and $\tilde{F}_a$ satisfy the conditions of Claim 3.22, which contradicts our assumption that no such $\mathcal{M}_0$ and $\tilde{F}_a$ exist. □

Similarly we define $\mathcal{M}_2 = \mathcal{M}_2(A) := \{(X, Y) \in \mathcal{M}\backslash\mathcal{M}_1 : d_{G''}(A, X) < \beta^{1/3}$ and $d_{G''}(A, Y) > (1 - \beta^{1/3})\}$.

Claim 3.25 $|\mathcal{M}_2| < \sqrt{\beta}k/M$.

Proof. Suppose instead that there is some $\mathcal{M}_0 \subseteq \mathcal{M}_2$ of size $\sqrt{\beta}k/M$. Recall that $F^2 = \{T \in F : \mathit{Ratio}(T) \in [c, 1-c]\}$, and that by Claim 3.14, $|V(F^2)| \geq ck$. Let $F_a^2 := F^2 \cap F_a$ and $F_b^2 := F^2 \cap F_b$. We will assume that $|F_a^2| \geq ck/2$; the other case (when $|F_b^2| \geq ck/2$) is similar. Suppose $\mathit{Root}(F_a)$ has been mapped into large vertices of A typical with respect to $\mathcal{M}_0$. For all but at most $\sqrt{\varepsilon}|\mathcal{M}_0|$ edges $e \in \mathcal{M}_0$ we can apply Lemma 3.23 Part 2 to embed at least

$$\mathbf{d}(A, e) + c(d_y - d_x)M - \varepsilon' M - \varepsilon M \geq \mathbf{d}(A, e) + cM/2$$

vertices of F_a^2 into e. Thus in $\mathcal{M}_0$ we embed a forest $\tilde{F}_a$ of size

$$\begin{aligned}|\tilde{F}_a| &\geq \mathbf{d}(A, \mathcal{M}_0) + |\mathcal{M}_0|cM/2 - 2M\sqrt{\varepsilon}|\mathcal{M}_0| \\ &\geq \mathbf{d}(A, \mathcal{M}_0) + (c/2)\sqrt{\beta}k - 2\sqrt{\varepsilon\beta}k \\ &\geq \mathbf{d}(A, \mathcal{M}_0) + \beta k.\end{aligned}$$

So $\mathcal{M}_0$ and $\tilde{F}_a$ satisfy the conditions of Claim 3.22, which is a contradiction once again. □

We now define $\mathcal{M}_3 = \mathcal{M}_3(A) := \{e \in \mathcal{M}\backslash(\mathcal{M}_1 \cup \mathcal{M}_2) : \mathbf{d}(A,e) < 2\beta^{1/3}M\}$ and set $\mathcal{M}' := \mathcal{M} - \mathcal{M}_1 - \mathcal{M}_2 - \mathcal{M}_3$, i.e.

$$\mathcal{M}' = \{(X,Y) \in \mathcal{M} : d_{G''}(A,X), d_{G''}(A,Y) > 1 - \beta^{1/3}\}.$$

By Claims 3.24 and 3.25, $|\mathcal{M}_1| + |\mathcal{M}_2| \leq 3\sqrt{\beta}k/M$, and since A has low degree to vertices in $\mathcal{M}_3$, $\mathcal{M}'$ carries most of the weight of $\mathbf{d}(A)$. More precisely, note that by the definition of $\mathcal{M}_3$, $\mathbf{d}(A, \mathcal{M}_3) < \beta^{1/3}MN$. Thus

$$\mathbf{d}(A, \mathcal{M}') > (1-\eta)k - 3\sqrt{\beta}(k/M)2M - \beta^{1/3}MN > (1-\beta^{2/7})k$$

and so $|\mathcal{M}'| > (1-\beta^{2/7})k/(2M)$. Furthermore since for any edge $e = (X,Y) \in \mathcal{M}'$ we have $\mathbf{d}(A,X), \mathbf{d}(A,Y) \geq (1-\beta^{1/3})M$, we also have

$$(2 - 2\beta^{1/3})M|\mathcal{M}'| \leq \mathbf{d}(A, \mathcal{M}') = (1-\eta)k$$

and so

$$|\mathcal{M}'| \leq \frac{1-\eta}{1-\beta^{1/3}}\frac{k}{2M} \leq (1+\beta^{2/7})k/(2M).$$

Recall that $\mathcal{M}^2(A) = \{(X,Y) \in \mathcal{M} : X, Y \in N(A)\}$. Thus $\mathcal{M}' \subseteq \mathcal{M}^2(A)$. For any real number $\alpha \in [0,1]$, let $N_\alpha(J) := \{X \in V(H) : \mathbf{d}(J,X) \geq \alpha\}$.

Remark 3.26 *If two clusters J and K can play the same roles as A and B, then all but at most $3\sqrt{\beta}k/M$ vertices of $N_{\beta^{1/3}}(J)$ in $\mathcal{M}$ are contained in $\mathcal{M}^2(J)$.*

This is because we can simply follow all of the above arguments with A and B replaced by J and K. We obtain sets $\mathcal{M}_1(J)$, $\mathcal{M}_2(J)$ and $\mathcal{M}_3(J)$, but observe that $\mathcal{M}_3(J) \cap N_{\beta^{1/3}}(J) = \emptyset$. Then every edge of $\mathcal{M}$ which contains a vertex of $N_{\beta^{1/3}}(J)$ but does not lie in $\mathcal{M}^2(J)$ must lie in $\mathcal{M}_1(J) \cup \mathcal{M}_2(J)$. But we also have $|\mathcal{M}_1(J)| + |\mathcal{M}_2(J)| \leq 3\sqrt{\beta}k/M$ as required.

Now if $f_b \geq \eta^{1/3}k$, then we let $\mathcal{M}_{in} = \mathcal{M}_{in}(A, B) := \{e = (X, Y) \in \mathcal{M}' : |\mathbf{d}(A, e) - \mathbf{d}(B, e)| \leq \beta M\}$ and $\mathcal{M}_{out} := \mathcal{M} - \mathcal{M}_{in}$. Let $\mathcal{V}_1 := V(\mathcal{M}_{in})$ and $\mathcal{V}_2 := V(H) - \mathcal{V}_1$. By Corollary 3.21,

$$\mathbf{d}(A, \mathcal{M}_{in}) \geq \mathbf{d}(A, \mathcal{M}') - 2M\eta^{1/6}k/M \geq (1 - 2\beta^{2/7})k$$

and so $|\mathcal{V}_1| \geq (1 - 2\beta^{2/7})k/M$. Also

$$\mathbf{d}(B, \mathcal{M}_{in}) \geq \mathbf{d}(A, \mathcal{M}') - 2M\eta^{1/6}k/M - \beta MN \geq (1 - 2\beta^{2/7})k.$$

On the other hand, if $f_b < \eta^{1/3}k$, then we observe that since we are in Case 1, we could without loss of generality have switched A and B at the start of the argument. Then we would have obtained a submatching $\mathcal{M}'(B) \subseteq \mathcal{M}$ of size at least $(1 - \beta^{2/7})k/M$. We pick a further submatching $\mathcal{M}_0 \subseteq \mathcal{M}'(B)$ of size $\eta^{1/3}k/M$. Then

$$\mathbf{d}(B, \mathcal{M}_0) \geq (2 - 2\beta^{1/3})M\eta^{1/3}k/M \geq \eta^{1/3}k + 5\sqrt{\varepsilon}n \geq f_b + 5\sqrt{\varepsilon}n.$$

Therefore we can embed F_b into $B \cup \mathcal{M}_0$ by Lemma 3.18. We now set $\mathcal{M}_{in} = \mathcal{M}_{in}(A) := \mathcal{M}'(A) \backslash \mathcal{M}_0$. Note that

$$\mathbf{d}(A, \mathcal{M}_{in}) \geq (1 - \beta^{1/3})k - 2M|\mathcal{M}_0| \geq (1 - 2\beta^{2/7})k.$$

As before we set $\mathcal{M}_{out} := \mathcal{M} - \mathcal{M}_{in}$ and $\mathcal{V}_1 := V(\mathcal{M}_{in})$, $\mathcal{V}_2 := V(H) - \mathcal{V}_1$. Thus in either case we have $|\mathcal{V}_1| \geq (1 - 2\beta^{2/7})k/M$, and

$$\mathbf{d}(A, \mathcal{M}_{in}) \geq (1 - 2\beta^{2/7})k. \tag{3.3}$$

In the first case we also have

$$\mathbf{d}(B, \mathcal{M}_{in}) \geq (1 - 2\beta^{2/7})k. \tag{3.4}$$

In the second case we have already embedded F_b outside $\mathcal{M}_{in}$. In order to go through the proof of both cases together, we will sometimes refer to embedding some subforest $F_0 \subseteq F$ in $\mathcal{M}_{in}$. It should be understood that some of these vertices may already have been embedded outside $\mathcal{M}_{in}$ in the case when F_b is small, and we do not attempt to rearrange this embedding. Rather, we embed only $F_0 \backslash F_b$ in $\mathcal{M}_{in}$.

In both cases we also have $|\mathcal{V}_1| \leq 2|\mathcal{M}'| \leq (1 + \beta^{2/7})k/M$. We define a new constant ρ such that $\beta \ll \rho \ll \alpha_1$. Recall that α_1 is the constant used in EC (i.e. we assume that $EC(\alpha_1)$ does not hold).

We first remove all edges between regular pairs which run between $\mathcal{V}_1$ and $\mathcal{V}_2$ with density less than $\beta^{1/3}$, and denote by H' the (unweighted) graph which we obtain from H by deleting the corresponding edges. Let W_i denote the set of vertices of G contained in the clusters of $\mathcal{V}_i$ for $i = 1, 2$ (and we also put V_0 into W_2). Suppose first that $e_{H'}(\mathcal{V}_1, \mathcal{V}_2) \le \rho(k/M)^2$. Then

$$\begin{aligned} e_G(W_1, W_2) &\le e_{G''}(W_1, W_2 \backslash V_0) + \varepsilon n^2 + |V_0| n \\ &\le \sum_{X \in \mathcal{V}_1, Y \in \mathcal{V}_2} d_H(X, Y) M^2 + 2\varepsilon n^2 \\ &\le e_{H'}(\mathcal{V}_1, \mathcal{V}_2) M^2 + \beta^{1/3} M^2 |\mathcal{V}_1||\mathcal{V}_2| + 2\varepsilon n^2 \\ &< \rho k^2 + \beta^{1/3} M^2 N^2 + 2\varepsilon n^2 \\ &< (\rho + \beta^{1/4}) k^2. \end{aligned}$$

So $e_G(W_1, W_2) < 2\rho k^2$, and even after moving a few vertices to ensure that $|W_1| = k$, we have $e_G(W_1, W_2) < 3\rho k^2 < \alpha_1 k^2$. But this would imply that $EC(\alpha_1)$ holds, which is a contradiction. Thus we may assume that $e_{H'}(\mathcal{V}_1, \mathcal{V}_2) > \rho(k/M)^2$.

We need to quote one more result from [72] (c.f. Lemma 5.8 part 2 in that paper). We call a forest consisting of εM-trees an *εM-forest.*

Lemma 3.27 *Given a cluster matching $\mathcal{M}$, a cluster set $\mathcal{C}$ outside $V(\mathcal{M})$ and a cluster $A \notin V(\mathcal{M}) \cup \mathcal{C}$, let $\delta_1 := \min_{C \in \mathcal{C}} |\{(X, Y) \in \mathcal{M} : d(C, X) > 0$ or $d(C, Y) > 0\}|$.*

If F is an εM-forest with $|V(F)| \le (1 - \varepsilon')\delta_1 M$ and $|Level_1(F)| \le \mathbf{d}(A, \mathcal{C}) - 2\sqrt{\varepsilon} M|\mathcal{C}|$, then F can be embedded (in any order of the trees) into $A \cup \mathcal{C} \cup \mathcal{M}$ with $Root(F) \to A$ and $Level_1(F) \to \mathcal{C}$. □

Note that the result in [72] actually requires $\delta_1' := \min_{C \in \mathcal{C}} |\{(X, Y) \in \mathcal{M} : d(C, X) > 0$ **and** $d(C, Y) > 0\}|$. However, the proof in that paper does not use this stronger assumption, and we require the result in the stated form.

We define $F_3 := \{T \in F_a \backslash Root(F_a) : |V(T)| \ge 3\}$. The following claim corresponds to Claim 5.19 in [72].

Claim 3.28 *$|F_3| < 16\sqrt{\beta} k$.*

Proof. Suppose instead that $|F_3| \geq 16\sqrt{\beta}k$. Now since $e_{H'}(\mathcal{V}_1, \mathcal{V}_2) > \rho(k/M)^2 > 24\sqrt{\beta}N^2$, $\mathcal{V}_1$ contains at least $8\sqrt{\beta}N$ clusters which have at least $16\sqrt{\beta}N$ neighbours in $\mathcal{V}_2$.

We now claim that there is a set of at least $7\sqrt{\beta}N$ clusters in $\mathcal{V}_1$ which have at least $14\sqrt{\beta}N$ neighbours in $\mathcal{M}_{out}$. This comes from any of the three cases in Remark 3.17.

In case C of that remark, when $\mathcal{M}$ covers all but at most $10\alpha_2 N$ vertices of H, this is trivial, since each one of the $8\sqrt{\beta}N$ vertices we have already chosen has at least $16\sqrt{\beta}N$ neighbours in $\mathcal{V}_2$, at most $10\alpha_2 N$ of which are not covered by M. (Recall that $\alpha_2 \ll \beta$.)

In cases A and B we need to be a bit more careful. However, we observe that each of our $8\sqrt{\beta}N$ clusters lies in $V(\mathcal{M}^2(A))$, and so all but at most one of them lies in the same component as A. Therefore for all but one of these clusters, all but one of its neighbours lies in $\mathcal{M}$, and the result follows.

From the set of at least $7\sqrt{\beta}N$ clusters, we pick a set of $3\beta N$ clusters which lie in different edges of $\mathcal{M}$, and we call this set $\mathcal{C}$. Let $\mathcal{M}_0 := \{(X,Y) \in \mathcal{M}_{in} : \{X,Y\} \cap \mathcal{C} \neq \emptyset\}$. Then $|\mathcal{M}_0| = |\mathcal{C}|$ and so $\mathbf{d}(A, \mathcal{M}_0) \leq 2M3\sqrt{\beta}N$. Observe also that since each vertex $C \in \mathcal{C}$ has at least $14\sqrt{\beta}N$ neighbours in $\mathcal{M}_{out}$ there are at least $7\sqrt{\beta}N$ edges $e = (X,Y)$ such that $d(C,X) > 0$ or $d(C,Y) > 0$.

Now let $\tilde{F_a} \subseteq F_a$ with $\tilde{F_a} \backslash Root(F_a) \subseteq F_3$ be the largest subset of the trees of F_a which we can embed into $A \cup \mathcal{C} \cup \mathcal{M}_{out}$ with $Root(\tilde{F_a}) \to A$, $Level_1(\tilde{F_a}) \to \mathcal{C}$. By Lemma 3.27 with $\delta_1 = 7\sqrt{\beta}N$, either $|V(\tilde{F_a})| \geq (1-\varepsilon')7\sqrt{\beta}NM > \mathbf{d}(A, \mathcal{M}_0) + \beta k$, or else $|Level_1(\tilde{F_a})| \geq \mathbf{d}(A, \mathcal{C}) - 2\sqrt{\varepsilon}|\mathcal{C}|M$. In the latter case, since each tree in F_3 has at least three vertices, we have $|\tilde{F_a}| \geq 3|Level_1(\tilde{F_a})| \geq 3\mathbf{d}(A, \mathcal{C}) - 6\sqrt{\varepsilon}|\mathcal{C}|M > \mathbf{d}(A, \mathcal{M}_0) + \beta k$.

In either case, we have $\tilde{F_a}$ and $\mathcal{M}_0$ satisfying the conditions of Claim 3.22, which is a contradiction. □

Recall that R denotes the set of roots of F. Thus we may assume that most vertices in $F_a\backslash R$ are contained in trees with at most 2 vertices, and since we already assumed that (apart from a few parent vertices) all are contained in trees with at least two vertices, we may in fact assume that almost all vertices of F_a are covered by root-2-paths, where a *root-2-path* is a path of length two with one endvertex in R. Furthermore, these root-2-paths are disjoint except for the vertices in R.

We define $\mathcal{S}_1 := \{Y : \exists X \in \mathcal{L},\ (X,Y) \in \mathcal{M}_{in}\}$, and $\mathcal{L}_1 := \mathcal{V}_1\backslash\mathcal{S}_1$. Note that all small clusters of $\mathcal{V}_1$ are contained in $\mathcal{S}_1$ and that $\mathcal{L}_1 \subseteq \mathcal{L}$. We will aim to bound both $e_{H'}(\mathcal{S}_1, \mathcal{V}_2)$ and $e_{H'}(\mathcal{L}_1, \mathcal{V}_2)$ from above and thus obtain a contradiction.

Claim 3.29 $e_{H'}(\mathcal{S}_1, \mathcal{V}_2) < 16\beta^{1/4}N^2$.

Proof. Suppose not. By Claim 3.28 we can pick $3\beta^{1/4}k$ root-2-paths in F_a which contain no parent vertices. We denote the set of non-root vertices in these paths by Z, so $|Z| = 6\beta^{1/4}k$. Note that because Z contains no parent vertices, it can be embedded at any time.

From our assumption it is easy to see that there are at least $8\beta^{1/4}N$ clusters in $\mathcal{S}_1$ each with at least $8\beta^{1/4}N$ neighbours in $\mathcal{V}_2$. For suppose not. Then $e_H(\mathcal{S}_1, \mathcal{V}_2) < 8\beta^{1/4}N \cdot N + (1 - 8\beta^{1/4})N \cdot 8\beta^{1/4}N < 16\beta^{1/4}N^2$, which is a contradiction. Pick $4\beta^{1/4}k/M$ such clusters which belong to different edges of $\mathcal{M}_{in}$. Denote this set by $\mathcal{S}_0$ and the submatching containing it by $\mathcal{M}_0$. Let $\mathcal{L}_0 := V(\mathcal{M}_0)\backslash\mathcal{S}_0 \subseteq \mathcal{L}$.

As $d_{\mathcal{V}_2}(J) \geq 8\beta^{1/4}N \geq |\mathcal{S}_0|$ for all $J \in \mathcal{S}_0$, we can form a new matching of size $\mathcal{S}_0$ between $\mathcal{S}_0$ and $\mathcal{V}_2$, and replace $\mathcal{M}_0$ in $\mathcal{M}_{in}$ by this new matching. Now by Lemma 3.19 we can embed $F\backslash Z$ into $\mathcal{M}_{in}$ since $|V(F)\backslash Z| \leq (1 - 6\beta^{1/4})k/M$ and by (3.3), even after our rearrangement of $\mathcal{M}_{in}$ we have

$$\mathbf{d}(A, \mathcal{M}_{in}) \geq (1 - 2\beta^{2/7})k - 4\beta^{1/4}(k/M)M \geq (1 - 6\beta^{1/4})k + 12\varepsilon n$$

and similarly for $\mathbf{d}(B, \mathcal{M}_{in})$.

The clusters in $\mathcal{L}_0$ have not been used yet. Each such cluster J has at least $\sqrt{d}M$ large vertices, and at least $(1-\sqrt{\varepsilon})M$ vertices typical with respect to $V(H)-J$, which have degree at least $(1-5d)k$ in G''. Furthermore, if we have already embedded a parent vertex in A (recall that $Z \subseteq F_a$), then at most $2\beta^{1/3}M$ are not in the neighbourhood of this vertex of A. We map the midpoints of paths of Z to typical vertices or large vertices in clusters of $\mathcal{L}_0$, using large vertices wherever possible. Note that $(1-\varepsilon-2\beta^{1/3})M4\beta^{1/4}k/M > 3\beta^{1/4}k$, so we always have available vertices. Note also that since $\mathcal{L}_0 \subseteq \mathcal{L}$, at least $\sqrt{d}M|\mathcal{L}_0| = \sqrt{d}4\beta^{1/4}k$ vertices of $\mathcal{L}_0$ are large. Furthermore, since the roots of the paths in Z have been embedded onto vertices typical with respect to the subsets of large vertices in $\mathcal{L}_0$, each such vertex has at least $(1-\beta^{1/3}-\varepsilon)\sqrt{d}M$ large neighbours in all but $\sqrt{\varepsilon}|\mathcal{L}_0|$ clusters of $\mathcal{L}_0$. Thus by a simple greedy argument, we can use at least

$$(1-\sqrt{\varepsilon})|\mathcal{L}_0|(1-\beta^{1/3}-\varepsilon)\sqrt{d}M \geq d^{2/3}k$$

large vertices for midpoints of Z.

We now pick $6dk$ such large midpoints, set these aside and consider the remaining $(3\beta^{1/4}-6d)k$ midpoints. Since they are either large or typical, they all have degree at least $(1-5d)k$ in G'', and since $6dk$ endpoints have been kept aside and have not yet been embedded, we can greedily find neighbours in G'' onto which to embed the endpoints of these $(3\beta^{1/4}-6d)k$ midpoints.

We now have just $6dk$ midpoints remaining, each of which is embedded onto a large vertex of G. Thus we can greedily find neighbours of these vertices for the endpoints of these root-2-paths, and thus complete the embedding of T. □

Let $\beta_1 := \beta^{1/16}$. We have the following final claim to complete Case 1 (c.f.

Claim 5.21 in [72]).

Claim 3.30 $e_{H'}(\mathcal{L}_1, \mathcal{V}_2) < 16\beta_1 N^2$.

Proof. Suppose not. From the definition the clusters in $\mathcal{L}_1$ are large and belong to different edges in $\mathcal{M}$, the other end of each edge being a small cluster. Choose $\mathcal{L}_0 \subseteq \mathcal{L}_1$ of size $8\beta_1 N$ such that $d_{H'}(X, \mathcal{V}_2) \geq 8\beta_1 N$ for each $X \in \mathcal{L}_0$. This is possible for otherwise $e_{H'}(\mathcal{L}_1, \mathcal{V}_2) \leq (8\beta_1 N)|\mathcal{V}_2| + |\mathcal{L}_1|(8\beta_1|\mathcal{V}_2|) \leq 16\beta_1 N^2$ which contradicts our initial assumption. Let $\mathcal{S}_0 := \{Y : (X,Y) \in \mathcal{M}_{in}, X \in \mathcal{L}_0\}$, so $\mathcal{S}_0 \subseteq \mathcal{S}_1$. We will show that $e_{H'}(\mathcal{S}_0, \mathcal{V}_2) \geq 16\beta^{1/4}N^2$, contradicting Claim 3.29.

Consider $X \in \mathcal{L}_0$. As in Remark 3.17, since X is large and $X \in \mathcal{M}^2(A)$, unless X is a vertex in U matched to the component of $H - U$ to which A belongs (which can only be the case for at most one X, by the initial construction of $\mathcal{M}$), X and A can play the roles of A and B respectively.

Let us now delete any regular pairs which still have density less than $\beta^{1/3}$. Recall that we had already deleted such regular pairs between $\mathcal{V}_1$ and $\mathcal{V}_2$, and so this deletion will not affect $d_{H'}(X, \mathcal{V}_2)$ or $e_{H'}(\mathcal{S}_0, \mathcal{V}_2)$ at all.

Since we have deleted regular pairs of density less than $\beta^{1/3}$, Remark 3.26 implies that all but at most $3\sqrt{\beta}k/M$ neighbours of X are contained in $\mathcal{M}^2(X)$, and so make up edges of $\mathcal{M}$. We pick one large cluster from each of these edges in $\mathcal{M}_{out}$ to form a set $\tilde{N}(X)$ of size at least $(8\beta_1 N - 3\sqrt{\beta}k/M)/2 > 3\beta_1 k/M$. Now since X also lies in the component of $H - U$ containing A, by Remark 3.17 it is still true that for all but at most one $Y \in \tilde{N}(X)$, Y and X can play the roles of A and B respectively. Thus all but at most $3\sqrt{\beta}k/M$ neighbours of Y in $\mathcal{V}_1$ make up edges of $\mathcal{M}$. So $|d(Y, \mathcal{L}_0) - d(Y, \mathcal{S}_0)| \leq 3\sqrt{\beta}k/M$. (Note that these degrees are unweighted.)

Let $\mathcal{N} := \cup_{X \in \mathcal{L}_0} \tilde{N}(X)$. Consider the (unweighted) bipartite subgraph $H'' \subseteq H'$ induced on $(\mathcal{L}_0, \mathcal{N})$. Let $\mathcal{N}_0$ consist of those vertices of degree at least $e(H'')/(2|\mathcal{N}|)$. Then $e(H'') \leq \frac{e(H'')}{2|\mathcal{N}|}(|\mathcal{N}| - |\mathcal{N}_0|) + |\mathcal{N}_0||\mathcal{L}_0| \leq e(H'')/2 + |\mathcal{N}_0||\mathcal{L}_0|$. Thus $|\mathcal{N}_0||\mathcal{L}_0| \geq$

$e(H'')/2$. So

$$|\mathcal{N}_0| \geq \frac{e(H'')}{2|\mathcal{L}_0|} \geq \frac{\delta_{H''}(\mathcal{L}_0, \mathcal{N})}{2} \geq \frac{3}{2}\beta_1 k/M$$

and for all $Y \in \mathcal{N}_0$,

$$\begin{aligned} d_{\mathcal{L}_0}(Y) &\geq \frac{e(H'')}{2|\mathcal{N}|} \geq \frac{1}{2|\mathcal{N}|} 8\beta_1(k/M)3\beta_1(k/M) \\ &= \frac{1}{|\mathcal{N}|} 12\beta_1^2(k/M)^2 \geq \beta_1^{5/2} k/M. \end{aligned}$$

Thus $d_{\mathcal{S}_0}(Y) \geq \beta_1^{5/2}k/M - 3\sqrt{\beta}k/M > \frac{1}{2}\beta_1^{5/2}k/M$ for all $Y \in \mathcal{N}_0$. Therefore

$$\begin{aligned} e(\mathcal{N}_0, \mathcal{S}_0) &> \frac{3}{2}\beta_1(k/M)\frac{1}{2}\beta_1^{5/2}(k/M) = \frac{3}{4}\beta_1^{7/2}(k/M)^2 \\ &= \frac{3}{4}\beta^{7/32}(k/M)^2 > 16\beta^{1/4}N^2. \end{aligned}$$

Which is a contradiction, as required. □

Now Claims 3.29 and 3.30 together show that

$$e_{H'}(\mathcal{V}_1, \mathcal{V}_2) = e_{H'}(\mathcal{S}_1, \mathcal{V}_2) + e_{H'}(\mathcal{L}_1, \mathcal{V}_2) < 16\beta^{1/4}N^2 + 16\beta_1 N^2 < \rho(k/M)^2.$$

But we already assumed that $e_{H'}(\mathcal{V}_1, \mathcal{V}_2) > \rho(k/M)^2$, which is a contradiction. This therefore completes the proof of the non-extremal theorems in Case 1.

3.5.6 Case 2

Recall that we have adjacent vertices $A, B \in \mathcal{L}$ and a matching $\mathcal{M}$ such that $\mathcal{M}$ covers $N(A)$ and $\mathbf{d}_{\mathcal{L}\cup\mathcal{M}}(B) \geq (1-\eta)k/2$. In the case when $|\mathcal{L}| \geq (1+\sqrt{\nu'})N/2$, i.e. in the proof of Theorem 3.9 we even have $\mathbf{d}_{\mathcal{L}\cup\mathcal{M}}(B) \geq (1+\nu')k/2$. Moreover, in both cases each edge in $\mathcal{M}$ has at most one endpoint in $N(A)$. Recall also that by Claim 3.14, $|V(F^2)| > ck$, where $F^2 = \{T \in F : c < ratio(T) < 1-c\}$.

Roughly speaking, we will attempt to embed F (and therefore T) as follows. Split F_b into $F_b^{(\mathcal{M})}$, of size approximately $\mathbf{d}_{\mathcal{M}}(B)$, and $F_b^{(\mathcal{L})}$. We embed $F_b^{(\mathcal{M})}$ into $\mathcal{M}_b$, an appropriate sub-matching of $\mathcal{M}$ intersecting $N(B)$. For each vertex $J \in N(B) \cap \mathcal{L}$, pick a neighbour to form a *fractional matching* (which we will define later) $\mathcal{M}_{\mathcal{L}}$, and embed $F_b^{(\mathcal{L})}$ into $\mathcal{M}_{\mathcal{L}}$. Finally, we embed F_a into $\mathcal{M}_a$, which will consist of the unused part of $\mathcal{M}$.

Of course, we cannot necessarily do this immediately, since $\mathbf{d}_{\mathcal{M}}(A)$ and $\mathbf{d}_{\mathcal{L}\cup\mathcal{M}}(B)$ are not quite large enough.

We use the same hierarchy of constants as we had in Case 1, so in particular we have $\alpha_2 \ll \eta \ll \beta \ll \alpha_1$. (We will not need the constant ρ for this case.) We split the proof further into two cases.

- Case a: $|F_b| \leq (1-\beta)k/2$
- Case b: $|F_b| > (1-\beta)k/2$.

In fact, almost all of the same problems that arise in Case a will also arise in Case b, but we concentrate first on the easier Case a for the sake of clarity.

In both cases we will assume that $\mathbf{d}_{\mathcal{M}}(A) = (1-\eta)k$ and $\mathbf{d}_{\mathcal{L}\cup\mathcal{M}}(B) = (1-\eta)k/2$, or else $\mathbf{d}_{\mathcal{L}\cup\mathcal{M}}(B) = (1+\nu')k/2$ if $|\mathcal{L}| \geq (1+\sqrt{\nu'})N/2$, (i.e. for the proof of Theorem 3.9).[1] We will not need this stronger assumption in Case a.

Case a

Now if $\mathbf{d}_{\mathcal{M}}(B) \geq f_b + 5\sqrt{\varepsilon}n$, then by Lemma 3.18 (in a degenerate form, since we ignore $\mathcal{M}_a$ and F_a), we can embed F_b into $B \cup \mathcal{M}$ easily.

[1] As in Case 1, we can ensure that the true values are within M by deleting some regular pairs as appropriate. As in that case, M is comparatively small, and so will not affect the relevant calculations significantly.

Otherwise we find a subforest $F_b^{(\mathcal{M})}$ such that

$$\mathbf{d}_{\mathcal{M}}(B) - 5\sqrt{\varepsilon}n - \varepsilon M \leq |F_b^{(\mathcal{M})}| \leq \mathbf{d}_{\mathcal{M}}(B) - 5\sqrt{\varepsilon}n.$$

We can do this because F consists of εM-trees. Now $F_b^{(\mathcal{M})}$ and $\mathcal{M}$ satisfy the degenerate conditions of Lemma 3.18, and so we can embed $F_b^{(\mathcal{M})}$ into $B \cup \mathcal{M}$. Then if $F_b^{(\mathcal{L})} := F_b \backslash F_b^{(\mathcal{M})}$, we have

$$\begin{aligned}\mathbf{d}_{\mathcal{L}\backslash\mathcal{M}}(B) &= (1-\eta)k/2 - \mathbf{d}_{\mathcal{M}}(B)\\ &\geq (1-\eta)k/2 - |F_b^{(\mathcal{M})}| - 5\sqrt{\varepsilon}n - \varepsilon M\\ &\geq |F_b^{(\mathcal{L})}| + \beta^2 k.\end{aligned}$$

We define a *fractional matching* to be a set of edges, each with a positive weight, such that the sum of the weights of the edges incident to any vertex is at most 1. (Thus a matching is just a fractional matching in which every edge has weight 1.) For our purposes we will also allow loops in a fractional matching. Our convention is that when calculating the weighted degree of a vertex with a loop attached to it, the weight of the loop is counted only once. We will define a fractional matching which will prescribe where we embed the remainder of F_b. The weight of an edge will indicate approximately how many vertices of $F_b^{(\mathcal{L})}$ we will embed into that edge.

For any $D \in N(B) \cap (\mathcal{L}\backslash\mathcal{M})$, we have $\mathbf{d}(D) \geq (1-\eta)k$. We now define the fractional matching $\mathcal{M}_{\mathcal{L}}$ into which we intend to embed $F_b^{(\mathcal{L})}$. Firstly for every cluster K of $\mathcal{M}$, if M' vertices of $F_b^{(\mathcal{M})}$ have been embedded into K we add a loop of weight M'/M to K. We do this to take account of vertices which have already been chosen for the embedding of $F_b^{(\mathcal{M})}$, and are therefore forbidden for $F_b^{(\mathcal{L})}$.

We would like to end up with a fractional matching in which the total weight of the edges is at least $(1-\beta^2)k/(2M)$. We first delete any edges between B and

$D \in N(B) \cap (\mathcal{L} \backslash \mathcal{M})$ which satisfy $\mathbf{d}(B, D) \leq \eta^2 M$. Note that we still have

$$\mathbf{d}_{\mathcal{M} \cup \mathcal{L}}(B) \geq (1-\eta)k/2 - \eta^2 MN \geq (1-2\eta)k/2.$$

Let $N := N(B) \cap (\mathcal{L} \backslash \mathcal{M})$. For each cluster $D \in N$ we also temporarily add in a loop of weight $\mathbf{d}(B, D)$. These loops will ensure that we do not match the clusters of N together. This is not strictly necessary for Case a, but we will want to use the same construction in Case b later, and so we prove the existence of a stronger structure than we need at the moment.

Now for each cluster D of N in turn we will delete the loop attached to it and find neighbours $D_1, D_2, \ldots, D_s$ of D in $H - \{A, B\}$ such that each D_i has a weight of at most $1-\eta^2$ in the fractional matching so far, and such that the total weight of all the D_i is at most $s - \mathbf{d}(B, D)/M$. We will assume that s is minimal such that these properties hold. Then we add edges (D, D_i) to the fractional matching such that the weight of each D_i (except possibly D_s) is 1, and the total weight of these new edges is $\mathbf{d}(B, D)/M$. We continue doing this until we reach a D for which we can no longer find the appropriate D_i. Suppose that at this stage the total weight of edges in the fractional matching, not including the loops of N, is less than $(1-\beta^2)k/(2M)$. Then since $\mathbf{d}_{\mathcal{L} \cup \mathcal{M}}(B) \geq (1-2\eta)k/2$, and since the set of loops in $\mathcal{M}$ carried a total weight of $|F_b^{(\mathcal{M})}|/M \geq \mathbf{d}_{\mathcal{M}}(B)/M - 6\sqrt{\varepsilon}n/M$, we have a total weight on the non-loop edges attached to N of at most $(1-\beta^2)k/(2M) - \mathbf{d}_{\mathcal{M}}(B)/M + 6\sqrt{\varepsilon}n/M \leq \mathbf{d}_N(B)/M - \beta^2 k/(4M)$. Thus in particular, there must be some $D \in N$ which is not yet used, and this must be because we could not find the appropriate D_i. However, since $D \in N \subseteq \mathcal{L}$ we have $\mathbf{d}_{H-\{A,B\}}(D) \geq (1-3d)k - 2M \geq (1-\eta)k$. The total weight of edges in the fractional matching, now including the loops of N, is

$$|F_b^{(\mathcal{M})}|/M + \mathbf{d}_N(B)/M \leq \mathbf{d}_{\mathcal{M}}(B)/M - 5\sqrt{\varepsilon}n/M + \mathbf{d}_N(B)/M \leq (1-\eta)k/(2M).$$

Thus the total weight of all the vertices in the fractional matching is at most $(1-\eta)k/(2M)+(1-\beta^2)k/(2M)$. But since we could not find the appropriate D_i for D we have

$$\begin{aligned}\mathbf{d}_{H-\{A,B\}}(D)/M &\le (1-\eta)k/(2M)+(1-\beta^2)k/(2M)+\eta^2N+\mathbf{d}(B,D)/M\\ &\le (1-\eta)k/M-\beta^2k/(2M)+\eta^2N+1\\ &< (1-\eta)k/M.\end{aligned}$$

But this is clearly a contradiction. Therefore the process of replacing loops by matching edges does not stop until we have a weight of at least $(1-\beta^2)k/(2M)$, not including loops of N. Now let $\mathcal{M}_{\mathcal{L}}$ be the resulting fractional matching obtained by removing all loops of both $\mathcal{M}$ and N.

For each edge $e=(X,Y)$ in the fractional matching we choose subsets $X_e \subseteq X$ and $Y_e \subseteq Y$ of size $w(e)M$, where $w(e)$ is the weight of e in the fractional matching. If X is incident to more than one edge e, we choose the subsets X_e to be disjoint, and to avoid any vertices of $F_b^{(\mathcal{M})}$ that have already been embedded. This is possible since with the loops of $\mathcal{M}$ which we initially included, we had a fractional matching and so the total weight of any cluster was not more than 1. We now note that since the weight of any edge is at least η^2, each of these subsets has size at least η^2M. By standard regularity arguments it is easy to see that each edge therefore still corresponds to an (ε/η^2)-regular pair, and $\varepsilon/\eta^2 \le \sqrt{\varepsilon}$, so we may say that each edge of $\mathcal{M}_{\mathcal{L}}$ represents a $\sqrt{\varepsilon}$-regular pair. Since $|F_b^{(\mathcal{L})}| \le \mathbf{d}_{\mathcal{L}\setminus\mathcal{M}}(B)-\beta^2k \le \mathbf{d}_{\mathcal{M}_{\mathcal{L}}}(B)-5\varepsilon^{1/4}n$, we can embed $F_b^{(\mathcal{L})}$ into $\mathcal{M}_{\mathcal{L}}$ by Lemma 3.18.

We now aim to embed F_a in $\mathcal{M}$, while avoiding vertices which have already been used for the embedding of F_b. We will define a new matching $\mathcal{M}_a$: For each edge of $\mathcal{M}$ we choose subsets of the two clusters which have equal size, and where the subsets are chosen to be as large as possible without including any previously embedded vertices. However, if this size is less than $\eta^2 M$, then we will ignore the edge entirely. This leaves us with a matching $\mathcal{M}_a$, in which every cluster has size at least $\eta^2 M$. Thus each edge is still (ε/η^2)-regular, and therefore also $\sqrt{\varepsilon}$-regular.

Now from the definition of $\mathcal{M}_a$, $\mathbf{d}_{\mathcal{M}_a}(A) \geq (1-\eta)k - f_b - \eta^2 MN \geq f_a - 2\eta k$. However, this is not quite enough to embed F_a, and so we will either need to gain some extra room while embedding F_a, or else show that we have already gained room during the embedding of F_b, and thus we have a better bound on $\mathbf{d}_{\mathcal{M}_a}(A)$ than the one above. This leads to a case distinction based on whether we have a reasonably large number of balanced trees in F_a or in F_b (recall that by Claim 3.14, we have a reasonably large number of balanced trees in total). Roughly, if F_a contains many balanced trees, then since each edge of $\mathcal{M}_a \subseteq \mathcal{M}$ has only one cluster in $N(A)$, we will be able to apply Lemma 3.23 part 2 to embed F_a. On the other hand, if F_b contains many balanced trees then whenever we embed a balanced tree T', at most a $(1-c)$-proportion of the vertices of T' will be embedded into a cluster $D \in V(\mathcal{M})$. The remaining vertices will be embedded either in the partner of D in $\mathcal{M}$ or outside $\mathcal{M}$. When we come to define $\mathcal{M}_a$ we consider subsets of the clusters such that the endclusters of each edge still contain the same number of vertices. Since we have often embedded vertices either outside $\mathcal{M}$ or into vertices in partner clusters, we will need to remove from $\mathcal{M}$ significantly less than $2|F_b|$ vertices of G''. In particular this will mean that $\mathbf{d}_{\mathcal{M}_a}(A) > \mathbf{d}_{\mathcal{M}}(A) - |F_b|$, and indeed we will gain an extra term which will be enough to allow us to embed F_a into $\mathcal{M}_a$.

More precisely, since $|V(F^2)| > ck$, either $|V(F_a^2)| > ck/2$ or $|V(F_b^2)| > ck/2$ (or both). We first assume the former.

Suppose therefore that $|F_a^2| \geq ck/2$. Then since every edge in $\mathcal{M}$ has at most one endvertex in $N(A)$, when we come to embed the trees in F_a^2 we will embed at least a c proportion of them into clusters not lying in $N(A)$. In particular, by Lemma 3.23

part 2 (with $d_x = 0$), we will be able to embed in such an edge $e = (X, Y)$ a subforest of size $(1+c)\mathbf{d}_e(A) - \varepsilon' M - \varepsilon M$. Thus overall we will be able to find a sub-matching $\tilde{\mathcal{M}_a}$ and a subforest $\tilde{F_a} \subseteq F_a^2$ such that $\tilde{F_a} \to A \cup \tilde{\mathcal{M}_a}$ and $|\tilde{F_a}| \geq \mathbf{d}_{\tilde{\mathcal{M}_a}}(A) + c^3 k$. So

$$\begin{aligned}
\mathbf{d}_{\mathcal{M}_a \setminus \tilde{\mathcal{M}_a}}(A) &= \mathbf{d}_{\mathcal{M}_a}(A) - \mathbf{d}_{\tilde{\mathcal{M}_a}}(A) \\
&\geq f_a - 2\eta k - \mathbf{d}_{\tilde{\mathcal{M}_a}}(A) \\
&\geq |F_a| - 2\eta k - |\tilde{F_a}| + c^3 k \\
&\geq |F_a \setminus \tilde{F_a}| + 5\varepsilon^{1/4} n
\end{aligned}$$

and thus by Lemma 3.18 we can embed $F_a \setminus \tilde{F_a}$ into $A \cup (\mathcal{M} \setminus \tilde{\mathcal{M}_a})$ as required.

Suppose instead that $|F_b^2| \geq ck/2$. Again, since every edge in $\mathcal{M}$ has at most one endvertex in $N(A)$, when we embedded a tree T' of F_b^2 we embedded at most $(1-c)|T'|$ vertices into a cluster $D \in N(A)$, and also at most $(1-c)|T'|$ vertices onto its partner. We modify the clusters by taking away at most $(1-c)|T'|$ vertices from both classes, including all embedded vertices of T', and keeping the sizes the same. Repeating this for every $T' \in F_b^2$, and then deleting any clusters which now have size at most $\eta^2 M$, we obtain a matching $\mathcal{M}_a$ with

$$\begin{aligned}
\mathbf{d}_{\mathcal{M}_a}(A) &\geq (1-\eta)k - |F_b^2|(1-c) - (f_b - |F_b^2|) - \eta^2 MN \\
&= f_a + c^2 k/2 - \eta k - \eta^2 n' \geq f_a + 5\varepsilon^{1/4} n.
\end{aligned}$$

Thus by Lemma 3.18 we will be able to embed F_a into $\mathcal{M}_a$ as required. So we may assume that we are in Case b.

Case b

Recall that in Case b, $|F_b| \geq (1-\beta)k/2$. We now need to make ourselves some extra room for F_b as well as for F_a. However, the extra room for F_a will be gained similarly as in Case a, so we will not repeat the argument here, focussing instead only on embedding F_b in a similar way to before. We can therefore observe that for the proof of Theorem 3.9, i.e. if we are in the case where $|\mathcal{L}| \geq (1+\sqrt{\nu'})n'/2$, we have $\mathbf{d}_{\mathcal{L}\cup\mathcal{M}}(B) = (1+\nu')k/2$, and since $\nu' \gg \beta$, the extra weighted degree that we have attached to B will allow us to complete the embedding in the same way as in Case a. More precisely, we will embed within $\mathcal{M}_b \subseteq \mathcal{M}$ a subforest $F_b^{(\mathcal{M})}$ of F_b of size at least $\mathbf{d}_{\mathcal{M}}(B) - 5\sqrt{\varepsilon}n - \varepsilon M$. We will then find a fractional matching $\mathcal{M}_{\mathcal{L}}$ attached to $N(B) \cap (\mathcal{L}\backslash\mathcal{M})$ of weight at least

$$\begin{aligned}\mathbf{d}_{\mathcal{L}\backslash\mathcal{M}}(B) - \beta^2 k &\geq \mathbf{d}_{\mathcal{L}\cup\mathcal{M}}(B) - |F_b^{(\mathcal{M})}| - 5\sqrt{\varepsilon}n - \varepsilon M - \beta^2 k \\ &\geq |F_b \backslash F_b^{(\mathcal{M})}| + \nu' k/2 - 5\sqrt{\varepsilon}n - \varepsilon M - \beta^2 k \geq |F_b \backslash F_b^{(\mathcal{M})}| + \beta n\end{aligned}$$

and since all the edges in the fractional matching will be $\sqrt{\varepsilon}$-regular, we will be able to embed $F_b \backslash F_b^{\mathcal{M}}$ into $\mathcal{M}_{\mathcal{L}}$ by Lemma 3.18. Thus the proof of Theorem 3.9 is complete, and we turn our attention to the proof of Theorem 3.8. We may assume that $|\mathcal{L}| \leq (1+\sqrt{\nu'})n'/2$, which as observed in the paragraph before Theorem 3.15 means that we may assume that there are no sets $\mathcal{S}' \subseteq \mathcal{S}$ and $\mathcal{L}' \subseteq \mathcal{L}$ such that $k/(10M), |\mathcal{L}'| \leq (7/20)|\mathcal{S}'|$ and $N(\mathcal{S}') \subseteq \mathcal{L}'$. This will be important later on.

Note that the condition for Case b means that F_a and F_b have approximately the same size. In some cases it will be convenient to switch them around in order to complete the embedding. To ensure that we lose no generality doing this switching, we will assume for the rest of this proof that $f_a = f_b = (1+\beta)k/2$. We can ensure this simply by adding some extra leaves adjacent to roots of trees in F_a and F_b. Note that this does not affect the fact that the trees of F are εM-trees. It may affect whether trees are balanced, but by choosing to add the new vertices in such a way that they are all adjacent to just one root of F_a or one root of F_b, only at most 2 trees can become unbalanced, and this will not affect calculations significantly.

We may therefore assume without loss of generality that $|V(F_b^2)| \geq ck/2$, for if not we simply switch F_a and F_b.

Now suppose first that $\mathbf{d}_{\mathcal{L}\backslash\mathcal{M}}(B) \geq \beta^{1/3}k$. Then as in Case a we embed a subforest $F_b^{(\mathcal{M})} \subseteq F_b$ into $\mathcal{M}$, where $\mathbf{d}_{\mathcal{M}}(B) - 5\sqrt{\varepsilon}n - \varepsilon M \leq |F_b^{(\mathcal{M})}| \leq \mathbf{d}_{\mathcal{M}}(B) - 5\sqrt{\varepsilon}n$. We also choose $F_b^{(\mathcal{M})}$ in such a way that $F_b^{(\mathcal{L})} := F_b \backslash F_b^{(\mathcal{M})}$ contains as many balanced trees as possible. Now

$$|F_b^{(\mathcal{L})}| \geq |F_b| - \mathbf{d}_{\mathcal{M}}(B) + 5\sqrt{\varepsilon}n = \mathbf{d}_{\mathcal{M}\cup\mathcal{L}}(B) + \eta k/2 + \beta k/2 - \mathbf{d}_{\mathcal{M}}(B) + 5\sqrt{\varepsilon}n \geq \beta^{1/3}k.$$

The equality holds since we assumed that $|F_b| = (1+\beta)k/2$ and that $\mathbf{d}_{\mathcal{M}\cup\mathcal{L}}(B) = (1-\eta)k/2$. Since $\beta \ll c$ we may assume that $F_b^{(\mathcal{L})}$ contains at least $\beta^{1/3}k$ vertices which lie in balanced trees. We also have $|F_b^{(\mathcal{L})}| \leq \mathbf{d}_{\mathcal{L}\backslash\mathcal{M}}(B) + \beta k/2 + \eta k/2 + 5\sqrt{\varepsilon}n + \varepsilon M \leq \mathbf{d}_{\mathcal{L}\backslash\mathcal{M}}(B) + \beta k$. Since each edge e of $\mathcal{M}_{\mathcal{L}}$ is $\sqrt{\varepsilon}$-regular, and since only one endvertex of e lies in $N(B)$, we may use Lemma 3.23 part 2 to embed at least $(1+c)\mathbf{d}_e(B) - \varepsilon' M - \varepsilon M \geq (1+c/2)\mathbf{d}_e(B)$ vertices into $B \cup e$. Thus taking a submatching $\mathcal{M}_0 \subseteq \mathcal{M}_{\mathcal{L}}$ such that $\mathbf{d}(B, \mathcal{M}_0) = \beta^{1/3}k/2$, we can embed into $B \cup \mathcal{M}_0$ a subforest $\tilde{F}_b \subseteq F_b^{(\mathcal{L})}$ of size at least $(1+c/2)\beta^{1/3}k/2 \geq \mathbf{d}(B, \mathcal{M}_0) + \beta^{1/2}k$. Thus

$$\begin{aligned} |F_b^{(\mathcal{L})} \backslash \tilde{F}_b| &\geq \mathbf{d}(B, \mathcal{M}_{\mathcal{L}} \backslash \mathcal{M}_0) - \beta k + \beta^{1/2}k \\ &\geq \mathbf{d}(B, \mathcal{M}_{\mathcal{L}} \backslash \mathcal{M}_0) + \beta n \end{aligned}$$

and therefore we can embed $F_b^{\mathcal{L}} \backslash \tilde{F}_b$ into $\mathcal{M}_{\mathcal{L}} \backslash \mathcal{M}_0$ using Lemma 3.18.

We now note also that when embedding $\tilde{F}_b$, at least $c|\tilde{F}_b|$ vertices were embedded into those endvertices of $\mathcal{M}_0$ which lie in $N(B) \cap (\mathcal{L}\backslash\mathcal{M})$, and so in particular at least $c(1+c)\beta^{1/3}k/3 \geq \beta^{1/2}k$ vertices of F_b have been embedded outside $\mathcal{M}$. Thus when we come to define $\mathcal{M}_a$ (including removing those edges where the clusters now have size less than $\eta^2 M$) we now have

$$\mathbf{d}(A, \mathcal{M}_a) \geq (1-\eta)k - |F_b| + \beta^{1/2}k - \eta^2 MN \geq |F_a| + \beta n$$

and since the pairs in $\mathcal{M}_a$ are still $\sqrt{\varepsilon}$-regular, we can embed F_a into $\mathcal{M}_a$ by Lemma 3.18 as required.

So we may assume that $\mathbf{d}_{\mathcal{L}\setminus\mathcal{M}}(B) < \beta^{1/3}k$. Thus $\mathbf{d}_{\mathcal{M}}(B) \geq (1-\beta^{1/4})k/2$. The following claim is similar to Claim 3.22.

Claim 3.31 *Let $\mathcal{M}_0 \subseteq \mathcal{M}$ be a matching of size at most $k/(4M)$. Suppose $\tilde{F}_b \subseteq F_b$ with $|\tilde{F}_b| \geq \mathbf{d}(B, \mathcal{M}_0) + \beta^{1/5}k$ can be embedded into $V(\mathcal{M}_0) \cup (V(H)\setminus(V(\mathcal{M}) \cup A))$ after we map Root($\tilde{F}_b$) to any vertices of B typical with respect to $\mathcal{M}_0$. Then $T \to G$.*

Proof. The proof is essentially similar to that of Claim 3.22, and so we only sketch it here. Similarly to that proof, since we have embedded more than we expected into $\mathcal{M}_0$, $F_b\setminus\tilde{F}_b$ is now small enough that we can embed it into $\mathcal{M}\setminus\mathcal{M}_0$. Note also that since many trees in F_b are balanced, when we define $\mathcal{M}_a$ we gain room for F_a automatically: For if we have embedded balanced trees into some edge $e = CD$, where $C \in N(A)$, and if in total we have used up $c'M$ vertices in $C \cup D$ (where $c' \in [0,2]$) then we have used at most $(1-c)c'M$ in both C and D, which means that what we have lost for $\mathcal{M}_a$ in each cluster has size at most $(1-c)c'M$. Note in particular that this is significantly better than the worst case scenario, in which everything would be embedded into one of the clusters and we would have had to delete just as many vertices in the other cluster to maintain equal size. Then we would have deleted $c'M$ vertices from both clusters, whereas in this case we only have to delete at most $(1-c)c'M$ vertices from each cluster. It is this that allows us to gain the extra room we need.

Summing up over all the balanced trees in F_b, we find that $\mathbf{d}(A, \mathcal{M}) - \mathbf{d}(A, \mathcal{M}_a)$

is significantly smaller than $|F_b|$, and thus we also gain enough room to embed F_a, even after deleting edges in which the clusters now have size at most $\eta^2 M$. The details are very similar to those given for the case when $\mathbf{d}(B, \mathcal{L}\backslash\mathcal{M}) \geq \beta^{1/3}k$ before the Claim, and we do not repeat them here. □

Similarly to Case 1, we define $\mathcal{M}_1 := \{(X,Y) \in \mathcal{M} : d_{G''}(B,X) \in [\beta^{1/12}, 1-\beta^{1/12}]$ or $d_{G''}(B,Y) \in [\beta^{1/12}, 1-\beta^{1/12}]\}$. The following claim is similar to Claim 3.24.

Claim 3.32 *$|\mathcal{M}_1| < \beta^{1/12}k/M$.*

Proof. Suppose not, and let $\mathcal{M}_0 \subseteq \mathcal{M}_1$ be a matching of size $\beta^{1/12}k/M$. For almost every edge $e = XY \in \mathcal{M}_0$ we apply Lemma 3.23 Part 3 to embed a subforest of F_b with at least $\mathbf{d}(B,e) - \varepsilon' M + \beta^{1/12}M - \varepsilon M$ vertices into $X \cup Y$. We can do this in all but $\sqrt{\varepsilon}N$ edges of $\mathcal{M}_0$, since the roots of F_b have been embedded into vertices of B typical with respect to $V(H) - B$. We denote the union of such subforests by $\tilde{F}_b$, and observe that

$$\begin{aligned}
|\tilde{F}_b| &\geq \mathbf{d}(B,\mathcal{M}_0) - 2M\sqrt{\varepsilon}N + (\beta^{1/12}M - \varepsilon' M - \varepsilon M)(|\mathcal{M}_0| - \sqrt{\varepsilon}N) \\
&\geq \mathbf{d}(B,\mathcal{M}_0) - \varepsilon^{1/3}k + \beta^{1/11}M\beta^{1/11}k/M \\
&\geq \mathbf{d}(B,\mathcal{M}_0) + \beta^{1/5}k.
\end{aligned}$$

So $\tilde{F}_b$ and $\mathcal{M}_0$ satisfy the conditions of Claim 3.31. So $T \to G$, which is a contradiction. □

We also define $\mathcal{M}_2 := \{(X,Y) \in \mathcal{M}\backslash\mathcal{M}_1 : d_{G''}(B,X) < \beta^{1/12}$ and $d_{G''}(B,Y) > (1-\beta^{1/12})\}$, and the following claim corresponds to Claim 3.25 in Case 1.

Claim 3.33 *$|\mathcal{M}_2| < \beta^{1/12}k/M$.*

Proof. Suppose not, and let $\mathcal{M}_0 \subseteq \mathcal{M}_2$ be a matching of size $\beta^{1/12}k/M$. For all but at most $\sqrt{\varepsilon}N$ edges e in $\mathcal{M}_0$ we can apply Lemma 3.23 part 2 to embed at least

$\mathbf{d}_e(B) + c(1 - 2\beta^{1/12})M - \varepsilon' M - \varepsilon M \geq \mathbf{d}_e(B) + cM/2$ vertices of F_b^2. Thus in $\mathcal{M}_0$ we embed a forest $\tilde{F}_b$ with

$$\begin{aligned}|\tilde{F}_b| &\geq \mathbf{d}(B, \mathcal{M}_0) - 2M\sqrt{\varepsilon}N + (|\mathcal{M}_0| - \sqrt{\varepsilon}N)cM/2 \\ &\geq \mathbf{d}(B, \mathcal{M}_0) + (cM/2)|\mathcal{M}_0| - \sqrt{\varepsilon}MN(2 + c/2) \\ &\geq \mathbf{d}(B, \mathcal{M}_0) + \beta^{1/12}(c/2)k - \varepsilon^{1/3}k \\ &\geq \mathbf{d}(B, \mathcal{M}_0) + \beta^{1/5}k.\end{aligned}$$

So $\tilde{F}_b$ and $\mathcal{M}_0$ satisfy the conditions of Claim 3.31, and therefore $T \to G$, which is a contradiction. □

Let $\mathcal{M}_3 := \{(X,Y) \in \mathcal{M} : d_{G''}(B,X), d_{G''}(B,Y) < \beta^{1/12}\}$, and let $\mathcal{M}' := \mathcal{M} - \mathcal{M}_1 - \mathcal{M}_2 - \mathcal{M}_3$. Then

$$\begin{aligned}\mathbf{d}(B, \mathcal{M}') &\geq (1-\eta)k/2 - \mathbf{d}_{\mathcal{L}\setminus\mathcal{M}}(B) - 2\beta^{1/12}(k/M)M \\ &\quad -2\beta^{1/12}(k/M)M - \beta^{1/12}MN \\ &\geq (1 - \beta^{1/15})k/2.\end{aligned}$$

Thus $|\mathcal{M}'| \geq (1 - \beta^{1/15})k/(4M)$. Furthermore, since $\mathbf{d}(B, e) \geq (1 - \beta^{1/12})2M$ for any $e \in \mathcal{M}'$, and since $\mathbf{d}(B, \mathcal{M}') \leq \mathbf{d}(B, \mathcal{M}) \leq \mathbf{d}(B, \mathcal{L} \cup \mathcal{M}) = (1 - \eta)k/2$, we have

$$|\mathcal{M}'| \leq \frac{(1-\eta)k/2}{2(1-\beta^{1/12})M} \leq (1 + \beta^{1/15})k/(4M).$$

Now let $\mathcal{S}_0 := \{Y : (X,Y) \in \mathcal{M}',\ X \in \mathcal{L}\}$, and let $\mathcal{L}_0 := \{Y : (X,Y) \in \mathcal{M}',\ X \in \mathcal{S}\} = V(\mathcal{M}') - \mathcal{S}_0$. Observe that $\mathcal{L}_0 \subseteq \mathcal{L}$, since the vertices of $\mathcal{L}_0$ are matched to vertices of $\mathcal{S}$ which is independent, but it is not necessarily true that $\mathcal{S}_0 \subseteq \mathcal{S}$.

Lemma 3.34 *There is no matching between $\mathcal{S}_0$ and $V(H)\backslash V(\mathcal{M}')$ of size $\beta^{1/30}k/(2M)$.*

Proof. Suppose there is such a matching, and let $\mathcal{S}_0'$ be the intersection of this matching with $\mathcal{S}_0$. Now let $\mathcal{L}_0'$ be the partners of $\mathcal{S}_0'$ in $\mathcal{M}'$ which do not themselves lie in $\mathcal{S}_0'$. Then since the vertices of $\mathcal{L}_0'$ are large, $\delta(\mathcal{L}_0', V(H)\backslash V(\mathcal{M}')) \geq (1-\eta)k/M - (1+\beta^{1/15})k/(2M) \geq k/(3M)$, and so we can also match $\mathcal{L}_0'$ to a subset of $V(H)\backslash V(\mathcal{M}')$ avoiding the previous matching. We call these two matchings together $\mathcal{M}_0$.

But now replacing $\mathcal{M}_0' := \{(X,Y) \in \mathcal{M}' : X \in \mathcal{S}_0' \text{ or } Y \in \mathcal{S}_0'\}$ with $\mathcal{M}_0$ in $\mathcal{M}'$, we can embed F_b. For recall that $|V(F_b^2)| > ck/2$, so by Lemma 3.23 part 2 we can embed a subforest of size $(1+c)(1-\beta^{1/12})M|\mathcal{M}_0| - \varepsilon M|\mathcal{M}_0| \geq (1+c/2)\beta^{1/30}k \geq \mathbf{d}(B, \mathcal{M}_0') + c\beta^{1/30}k/2$ into $\mathcal{M}_0$. Let $\tilde{F}_b$ be this subforest. Then

$$\begin{aligned}
\mathbf{d}(B, \mathcal{M}'\backslash\mathcal{M}_0) &= (1-\eta)k/2 - \mathbf{d}(B, \mathcal{M}_0') - \mathbf{d}_{\mathcal{L}\backslash\mathcal{M}}(B) \\
&\geq |F_b| - \beta k - \eta k/2 - \mathbf{d}(B, \mathcal{M}_0') - \beta^{1/3}k \\
&\geq |F_b\backslash\tilde{F}_b| - \beta k - \eta k/2 + c\beta^{1/30}k - \beta^{1/3}k \\
&\geq |F_b\backslash\tilde{F}_b| + 5\sqrt{\varepsilon}n
\end{aligned}$$

and so we can apply Lemma 3.18 to embed $F_b\backslash\tilde{F}_b$. Note that we still have plenty of room for F_a, and so we can embed T into G, which is a contradiction. □

Corollary 3.35 *At most $|\mathcal{M}'| + \beta^{1/30}k/(2M)$ clusters of $\mathcal{M}'$ lie in $\mathcal{L}$.*

Proof. If not, then at least $\beta^{1/30}k/(2M)$ vertices of $\mathcal{S}_0$ are large, and therefore we could easily find a matching between these vertices and $V(H)\backslash V(\mathcal{M}')$, contradicting Lemma 3.34 □

Now let $\mathcal{R}_0 := \{X \in N(B)\backslash\mathcal{M}' : \mathbf{d}(B, X) \geq \beta^{1/60}M\}$.

Lemma 3.36 *There is no matching between $\mathcal{R}_0$ and $V(H)\backslash(V(\mathcal{M}')\cup\mathcal{R}_0)$ of size $\beta^{1/30}k/(2M)$.*

Proof. Suppose there were such a matching $\mathcal{M}^*$, and let $\mathcal{R}_0'$ be the intersection of this matching with $\mathcal{R}_0$. Since $\mathbf{d}(B,e)\geq\beta^{1/60}M$ for each edge e of this matching, we have $\mathbf{d}(B,\mathcal{M}^*)\geq\beta^{1/60}M\beta^{1/30}k/(2M)=\beta^{1/20}k$. Thus

$$\mathbf{d}(B,\mathcal{M}'\cup\mathcal{M}^*)\geq(1-\beta^{1/15})k/2+\beta^{1/20}k\geq|F_b|+5\sqrt{\varepsilon}n$$

and so by Lemma 3.18 we can embed F_b into $B\cup\mathcal{M}'\cup\mathcal{M}^*$. Once again, we still have plenty of room for F_a, and so $T\to G$, which is a contradiction. □

Lemmas 3.34 and 3.36 together give the following.

Lemma 3.37 *Suppose $\mathcal{M}''$ is a matching from $\mathcal{S}_0\cup\mathcal{R}_0$ into $V(H)\backslash\mathcal{V}_0$, where $\mathcal{V}_0:=V(\mathcal{M}')\cup\mathcal{R}_0$. Then $|\mathcal{M}''|\leq 2\beta^{1/30}k/M$.* □

However, recall that $\mathbf{d}_{\mathcal{L}\backslash\mathcal{M}}(B)<\beta^{1/3}k$, that $|\mathcal{M}_1|,|\mathcal{M}_2|<\beta^{1/12}k/M$, and that $\mathbf{d}(B,\mathcal{M}_3)<\beta^{1/12}MN$. Therefore

$$\mathbf{d}(B,\mathcal{V}_0)>(1-\eta)k/2-2\beta^{1/12}k-2\beta^{1/12}k-\beta^{1/12}n-\beta^{1/3}k>(1-c)k/2.$$

Recall that $|\mathcal{M}'|\leq(1+\beta^{1/15})k/(4M)$ and thus $|\mathcal{L}_0|\leq(1+\beta^{1/15})k/(8M)$, therefore

$$\begin{aligned}|\mathcal{S}_0\cup\mathcal{R}_0|&\geq(1-c)k/(2M)-(1+\beta^{1/15})k/(8M)\\&\geq(3/8-2c)k/M.\end{aligned}$$

By Lemma 3.37 there is a set $\mathcal{S}_1\subseteq\mathcal{S}_0\cup\mathcal{R}_0$ of size $|\mathcal{S}_0\cup\mathcal{R}_0|-2\beta^{1/30}k/M$ whose neighbourhood outside $\mathcal{V}_0$ has size at most $2\beta^{1/30}k/M$. For consider a maximum matching $\mathcal{M}''$ between $\mathcal{S}_0\cup\mathcal{R}_0$ and $V(H)\backslash\mathcal{V}_0$, and let $\mathcal{S}_1:=\mathcal{S}_0\cup\mathcal{R}_0\backslash V(\mathcal{M}'')$. Then the neighbourhood of $\mathcal{S}_1$ outside $\mathcal{V}_0$ lies within $V(\mathcal{M}'')\backslash\mathcal{V}_0$, and therefore has size at most $|V(\mathcal{M}'')| \leq \beta^{1/30}k/M$

Note also that Lemma 3.37 implies that $|(\mathcal{S}_0 \cup \mathcal{R}_0) \cap \mathcal{L}| \leq 2\beta^{1/30}k/M$, and so there is a set $\mathcal{S}_1' \subseteq \mathcal{S}_1$ of size $|\mathcal{S}_0 \cup \mathcal{R}_0| - 4\beta^{1/30}k/M$ consisting only of small clusters. Now

$$\begin{aligned}
|N(\mathcal{S}_1')| &\leq |V(\mathcal{M}') \cap \mathcal{L}| + |\mathcal{R}_0 \cap \mathcal{L}| + 2\beta^{1/30}k/(2M) \\
&\leq |\mathcal{M}'| + \beta^{1/30}k/(2M) + \beta^{1/30}k/(2M) + 2\beta^{1/30}k/(2M) \\
&\leq (1 + \beta^{1/40})k/(4M) \leq 7|\mathcal{S}_1'|/20
\end{aligned}$$

where the last line follows since $|\mathcal{S}_1'| \geq |\mathcal{M}'| + |\mathcal{R}_0| - \beta^{1/40}k/M \geq (3/4 - \beta^{1/50})k/M$. But observing that $\mathcal{S}_1' \subseteq \mathcal{S}$, this contradicts our assumption at the start of the Case b that no such set exists, which gives a contradiction. This completes the proof of Theorem 3.8, and since the proof of Theorem 3.9 was completed earlier, this also completes the non-extremal case. □

3.6 The Extremal Case

3.6.1 Outline and main results

As in the non-extremal case, we will prove the extremal case by contradiction, i.e. we will assume that there is some tree T on $k+1$ vertices that is not contained as a subgraph of G, and show that this leads to a contradiction. We therefore consider the tree T to be fixed. Since the proof holds for any choice of T, this contradiction then shows that $\mathcal{T}_k \subseteq G$. In this section, we will present statements of the main results for the extremal case without proof. This will give an extended outline of the main ideas. The results will then be proved in Section 3.6.3. Before this, though, in Section 3.6.2 we will complete the proof of Theorem 3.3. The proof uses some of the results of this section, and includes some of the ideas needed for the proof of the extremal case of Theorem 3.1 while being considerably shorter and easier. Thus it serves as a useful introduction to the main proof. It is here that we use the constants $\theta_i^\dagger$ instead of θ_i.

The results in this section often take the form of saying that if some property P holds in G (or in T) then $T \subseteq G$. Since we assumed that $T \nsubseteq G$, in context this amounts to saying that P does not hold.

Recall that we have constants

$$0 < \theta_1 \ll \theta_2 \ll \ldots \ll \theta_{\lfloor n/k \rfloor + 4} \ll 1/C.$$

Let j be maximal such that there are pairwise disjoint vertex sets $V_1, \ldots, V_j$ in G satisfying $|V_1| = \ldots = |V_j| = k$, and $e(V_i, V(G) \backslash V_i) \leq \theta_j k^2$ for $1 \leq i \leq j$. (In Section 3.6.2 for the proof of Theorem 3.3 we will use a similar condition with θ_j replaced by $\theta_j^\dagger$.) Let $V_0 := V(G) \backslash (\bigcup_{i=1}^{j} V_i)$, and for each i we define $L_i := V_i \cap L$ and $S_i := V_i \cap S$. With j defined to be maximal in this way, we say that we are in EC_j. Throughout this section we assume that G satisfies EC_j for some $j \geq 1$. We also define a slightly stronger condition with parameter α, which we call EC'.

$EC'(\alpha)$: There are pairwise disjoint sets $V_0, V_1, \ldots, V_j \subseteq V(G)$, where $j = \lfloor n/k \rfloor$ such that

- $|V_i| = k$ *for* $i = 1, \ldots, j$,
- $(1 - \alpha)k/2 \leq |L_i| \leq (1 + \alpha)k/2$ *for* $i = 1, \ldots, j$,
- $e(V_i, V(G) \backslash V_i) \leq \alpha k^2$ *for* $i = 0, 1, \ldots, j$ *and*
- *for all* $x \in V_i \cap L$ *we have* $d(x, V(G) \backslash V_i) \leq \alpha k$ *for* $i = 0, \ldots, j$.

Our proof of the extremal case will proceed as follows. We will first show that either we can continue applying the non-extremal case to split the vertex set of G essentially completely into "almost components" of size k, i.e. that $j = \lfloor n/k \rfloor$, or we find some "almost component" of size k with significantly more than $k/2$ large vertices. In the latter case we will prove that $\mathcal{T}_k \subseteq G$ directly using the following proposition. Let $\theta_i \ll \mu_i \ll \theta_{i+1}$ for each i.

Proposition 3.38 *Suppose $j < \lfloor n/k \rfloor$. Then there is some $1 \leq i \leq j$ such that $|L_i| \geq (1/2 + \mu_j)k$.*

This proposition will be proved by contradiction; if the conclusion does not hold then V_0 will satisfy the conditions of Theorem 3.8, and so $T \subseteq G$, which we already assumed was not the case. We will also have a similar argument in Section 3.6.2 for the proof of Theorem 3.3, but we will not need to present it as a separate proposition in that case.

Once we know that there such an i, we assume without loss of generality that it is $i = 1$, and obtain $T \subseteq G$ by the following lemma.

Lemma 3.39 *Suppose we have a set $V_1 \subseteq V(G)$ of size k such that $e(V_1, V(G)\backslash V_1) \leq \alpha_1 k^2$ and $|L_1| \geq (1/2 + \alpha_2)k$, where $0 < \alpha_1 \ll \alpha_2 \ll 1$. Then $T \subseteq G$.*

We will also need a result very similar to this for the proof of Theorem 3.3.

Lemma 3.40 *Suppose we have constants satisfying*

$$0 < 1/k \ll \tau \ll \tau' \ll \alpha_1 \ll \alpha_2, \nu \ll 1/C$$

and suppose we have subgraphs $G^\dagger \subseteq G^* \subseteq G$ as in Theorem 3.3. Suppose furthermore that we have a set $V_1 \subseteq V(G^\dagger)$ of size k such that $e_{G^*}(V_1, V(G^*)\backslash V_1) \leq \alpha_1 k^2$ and $|L_1| \geq (1/2 + \alpha_2)k$, where $L_1 = L^* \cap V_1$. Then $T \subseteq G$.

Let us note that in the case when we have $G^{\dagger}$ as in Theorem 3.3, we will obtain such a set V_1 even if $j = \lfloor n/k \rfloor$, and therefore the proof of Theorem 3.3 will be complete. A more precise argument for this is given in Section 3.6.2. We will prove Lemma 3.39 in Section 3.6.3 and then note that Lemma 3.40 can be proved in an almost identical way.

Lemma 3.39 and Proposition 3.38 together give the following.

Theorem 3.41 *If $j < \lfloor n/k \rfloor$, then $T \subseteq G$.* □

The proof of Lemma 3.39 and Lemma 3.40 is rather involved, and requires some preliminary results (Claims 3.42 and 3.43 and Lemma 3.44). We define a *path segment* of the tree T to be a subgraph of T which forms a path, and furthermore each vertex of this subgraph has degree 2 in T (thus the only neighbours of the internal vertices of such a path also lie on the path).

Claim 3.42 *Let q be a positive integer and let γ_1, γ_2 be real numbers such that $1/k < \gamma_1, \gamma_2 \ll 1/q$. Suppose T contains at most $\gamma_1 k$ leaves. Then T contains at least $\gamma_2 k$ vertex-disjoint path-segments on q vertices.*

Observe that we do not require any relation between γ_1 and γ_2.

Claim 3.43 *Let $A \subseteq V(G)$ and suppose that $|A \cap L| \geq k/2$ and $|A| \geq (1-\alpha)k$, for some $0 < \alpha \ll 1/C$. Consider in G a maximal set $\mathcal{P}'$ of vertex-disjoint paths of length between 2 and 6 which have their end-points in $A \cap L$ and their internal vertices in $V(G) \backslash A$. Then $|A \cup V(\mathcal{P}')| \geq k-1$.*

These two claims are designed to complement each other. One guarantees paths in T, and the other guarantees paths in G. Naturally, we will aim to embed the paths of T onto the paths in G. This, and much more, is the aim of the following technical embedding lemma.

Lemma 3.44 *Let $0 < \gamma_1 \ll \gamma_2 \ll 1$. Suppose that we have a tree T' on at most $k-1$ vertices such that $V(T') = U_1 + U_2$ where*

- *U_1 and U_2 are independent sets;*
- *There exists a set $\mathcal{P}$ of $\gamma_2 k$ disjoint path-segments of length 8 in T'.*

Suppose also that we have a bipartite graph G' with vertex classes L_1 and S_1, and further partitions $S_1 = C_1 + D_1$ and $L_1 = L_1'' + L_1'$ such that the following conditions hold:

- *$|L_1| = |U_1|$, $|S_1| = |U_2|$;*
- *$\delta(L_1'', S_1) \geq |S_1| - \gamma_1 k$;*
- *$\delta(C_1, L_1) \geq |L_1| - \gamma_1 k$;*
- *$|D_1| \leq \gamma_1 k$;*
- *S_1 is independent;*
- *G' contains a set $\mathcal{P}'$ of disjoint paths of even length between 2 and 6 covering D_1, L_1' and $2|\mathcal{P}'|$ vertices of L_1'', disjoint from C_1 and whose endpoints lie in L_1''.*

Then T' can be embedded into G', with U_1 embedded in L_1 and U_2 embedded in S_1.

Roughly speaking, we prove Lemma 3.39 and Lemma 3.40 by first discarding any large vertices which do not have almost all of V_1 as neighbours. We then discard any small vertices which do not have almost all of L_1 as neighbours. We may now have $|V_1| < k$, but we will still have $|L_1| \geq k/2$ and so we apply Claim 3.43 to find paths with which to extend V_1. Then if T has few leaves, Claim 3.42 will give us the paths in T which will complete the conditions of Lemma 3.44. On the other hand, if T has many leaves then we can show that in fact U_2 contains many leaves, where U_2 is the larger of the two bipartition classes of T. Deleting these leaves will give us sets that are sufficiently small that we could embed them into L_1 and C_1 greedily. Furthermore, neighbours of deleted leaves will be embedded onto large vertices and so we can add the remaining leaves greedily at the end.

In the case when $j = \lfloor n/k \rfloor$, the following two theorems will prove Theorem 3.1.

Theorem 3.45 *$EC_{\lfloor n/k \rfloor} \Rightarrow$ either $EC'(\theta^3_{\lfloor n/k \rfloor+2})$ holds or $\mathcal{T}_k \subseteq G$.*

Theorem 3.46 *$EC'(\theta^3_{\lfloor n/k \rfloor+2}) \Rightarrow \mathcal{T}_k \subseteq G$.*

Theorem 3.45 is proved using two main propositions. The first implies that we may assume large vertices have almost all of their neighbours in one class. It is very similar to Proposition 6.12 in [72].

Proposition 3.47 *If some $v_0 \in L$ has at least $\theta^{1/4}_{j+1}k$ neighbours in both V_{i_1} and V_{i_2} ($1 \leq i_1 < i_2 \leq j$), then $\mathcal{T}_k \subseteq G$.*

We can then show that in fact almost all large vertices are already in the class in which they have most of their neighbours. For those few that remain, we move them into the appropriate class. This tidies up the large vertices so that they have the properties required for EC'. Thus we have the following proposition.

Proposition 3.48 *If the condition of Proposition 3.47 does not hold, then we can rearrange the sets V_i to ensure that $EC'(\theta^3_{j+2})$ holds.*

This will complete the proof of Theorem 3.45. Theorem 3.46 is harder to prove. We first rearrange any small vertices which do not have appropriately high degree in their own class such that they now belong to the class in which they have the most neighbours. After this rearrangement, we no longer have sets of size exactly

k. However, we can now prove the following proposition, which is very similar to Proposition 3.47. In Section 3.6.3 we will define what we mean by "good" and "bad" vertices. Roughly speaking, a vertex x in S_i is good if it has almost all of L_i as neighbours, and bad otherwise. We define good and bad vertices of L_i similarly, although by the time we need such a definition all large vertices will be good.

Proposition 3.49 *No small vertex has more than $(1/2 + 2\theta_{j+3})k$ neighbours. In particular, no good small vertex in V_i has more than $3\theta_{j+3}k$ neighbours outside V_i, and no bad small vertex in V_i has more than $(1/4 + \theta_{j+3})k$ neighbours in any $V_{i'}$ for $i' \neq i$.*

The "in particular" will follow very easily from the first statement later on, once we have properly defined what it means for a vertex to be good or bad.

This proposition will be required for the proof of the following lemma. Since at least half of the vertices of G are large, there must be some i such that $|L_i| \geq |V_i|/2$. Without loss of generality we will assume that this holds for $i = 1$, and we will do most of our embedding in V_1. For each $i = 0, \dots, j$ let $m_i := k - |V_i|$.

Lemma 3.50 *Let $q_1, q_2, \dots, q_s$ be positive integers such that $q := \sum_{i=1}^{s} q_i \leq 2(m_1 + 1)/3$, and let $C_1, \dots, C_s \subseteq L_1$ be (not necessarily distinct or disjoint) sets of size $(1/2 - 2\theta_{j+4})k$. Then there are*

- *q disjoint $(1/\theta_{j+4})$-stars in $V \backslash V_1$ with midpoints $y_1, \dots, y_q$;*
- *distinct vertices $x_1, \dots, x_s \in L_1$ and*
- *a partition of the set of stars into $Q_1, \dots, Q_s$*

such that for each $i = 1, \dots, s$ we have

- *$x_i \in C_i$;*

- $|Q_i| = q_i$ *and*

- $M(Q_i) \subseteq N(x_i)$

where $M(Q_i)$ *denotes the set of midpoints of the stars in* Q_i. *Furthermore, all endpoints of stars are good vertices.*

Let us note here that if we set $c_1 := k/2 - |L_1|$ then by assumption we have $c_1 \leq m_1/2$, and therefore

$$2(m_1+1)/3 \geq \lfloor 2(m_1+1)/3 \rfloor \geq m_1/2 \geq c_1$$

for all non-negative integers m_1.

The idea of this lemma is that we will be able to use a midpoint of a star to embed a vertex y of T, and the endpoints of the star to embed the children of y. Then the remainder of $T(y)$ will be embedded outside V_1. This enables us to gain room within V_1 if V_1 has size less than k, and is therefore not large enough to contain T.

The proof of Theorem 3.46 will be split into two cases.

Case 1: T contains at least $36\theta_{\lfloor n/k \rfloor + 4} k$ leaves.

Recall that a skew-partition of T is an ordered partition $V(T) = U_1 + U_2$ such that $|U_1| \leq |U_2|$ and U_2 is an independent set. We say that a skew-partition is *ideal* if both U_1 and U_2 contain at least $5\theta_{\lfloor n/k \rfloor + 4} k$ leaves. The following two propositions prove Theorem 3.46 in Case 1. (In both propositions we implicitly assume that $EC'(\theta^3_{\lfloor n/k \rfloor + 2})$ holds.)

Proposition 3.51 *If* T *has an ideal skew-partition, then* $T \subseteq G$.

Proposition 3.52 *If* T *contains at least* $36\theta_{\lfloor n/k \rfloor + 4} k$ *leaves, then either* $T \subseteq G$ *or* T *has an ideal skew-partition.*

Case 2: T contains at most $36\theta_{\lfloor n/k \rfloor+4}k$ leaves.

In this case we will apply Claim 3.42 to find path segments in T. We will also apply Lemma 3.50 to find 2-paths in G with one endpoint in L_1 and the other being a good endpoint outside V_1. We will then join up these endpoints using bounded length paths to create a situation in which we can apply Lemma 3.44 to obtain $T \subseteq G$, as required. This will complete the proof of the extremal case.

3.6.2 Proof of Theorem 3.3

We can now complete the proof of Theorem 3.3. We assume for now that Lemma 3.40 holds, although we will delay the proof of this fact until Section 3.6.3, so that we can combine it with the proof of Lemma 3.39.

In $G^\dagger$ we have $L^* = L^*(G^\dagger, G^*) := \{v \in V(G^\dagger) : d_{G^\dagger}(v) \geq (1-\tau')k$ and $d_{G^*}(v) \geq k\}$, and $|L^*| \geq (1+\nu)|G^\dagger|/2$.

We define j to be maximal such that there exist disjoint sets $V_1, \ldots, V_j \subseteq V(G^\dagger)$ of size k and such that $e(V_i, V(G^\dagger)\backslash V_i) \leq \theta_j^\dagger k^2$ for $i = 1, \ldots, j$. If for some i we have $|L^* \cap V_i| \geq (1+\nu/2)k/2$, we assume without loss of generality that this holds for $i = 1$. Then we recall that $e_{G^*}(V(G^\dagger), V\backslash V(G^\dagger)) \leq \tau k^2$, and so $e_{G^*}(V_i, V\backslash V_i) \leq (\tau + \theta_j^\dagger)k^2 \leq 2\theta_j^\dagger k^2$, and so we can apply Lemma 3.40 with $\alpha_1 = 2\theta_j^\dagger$ and $\alpha_2 = \nu/2$ to obtain $T \subseteq G$.

But if this does not hold for any i, then let $V_0 := V(G^\dagger)\backslash(\bigcup_{i=1}^j V_i)$, and observe that $|L^* \cap V_0| \geq (1+\nu)|V_0|/2$. We also have

$$e_{G^*}(V_0, V\backslash V_0) \leq j\theta_j^\dagger k^2 + \tau k^2 \leq \sqrt{\theta_j^\dagger}k^2.$$

Therefore at most $(\theta_j^\dagger)^{1/4}k$ vertices have more than $(\theta_j^\dagger)^{1/4}k$ neighbours outside V_0. Thus in $G[V_0]$ at least $(1+\nu/2)|V_0|/2$ vertices have degree at least $(1-(\tau' + (\theta_j^\dagger)^{1/4}))k \geq (1-2(\theta_j^\dagger)^{1/4})k$. Therefore if $|V_0| \geq k$, we can apply Theorem 3.9

with $\alpha_1 = \theta^\dagger_{j+1}$ and $\alpha_2 = 2(\theta^\dagger_j)^{1/4}$ to find $T \subseteq G$, which is a contradiction.

We must therefore have $|V_0| < k$. Suppose that $\theta^\dagger_{j+1}k \leq |V_0| \leq (1-\theta^\dagger_{j+1})k$. Then $|L^* \cap V_0| \geq \theta^\dagger_{j+1}k/2$, and each vertex of $L^* \cap V_0$ must have at least $k - |V_0| \geq \theta^\dagger_{j+1}k$ neighbours in G^* which lie outside V_0, and so $e_{G^*}(V_0, V\backslash V_0) \geq ((\theta^\dagger_{j+1})^2/2)k^2 \geq \sqrt{\theta^\dagger_j}k^2$, which is also a contradiction. So either $|V_0| \geq (1-\theta^\dagger_{j+1})k$ or $|V_0| \leq \theta^\dagger_{j+1}k$.

Now if $|V_0| \leq \theta^\dagger_{j+1}k$, then trivially $|L^* \cap V_0| \leq \theta^\dagger_{j+1}k$, and so

$$|L^*| \leq j(1+\nu/2)k/2 + \theta^\dagger_{j+1}k < j(1+\nu)k/2 < (1+\nu)|G^\dagger|/2 \leq |L^*|$$

which is clearly a contradiction.

On the other hand, if $|V_0| \geq (1-\theta^\dagger_{j+1})k$, then $|L^* \cap V_0| \geq (1+\nu)(1-\theta^\dagger_{j+1})k/2 \geq (1+\nu/2)k/2$. By moving at most $\theta^\dagger_{j+1}k$ small vertices[1] into V_0, we can ensure that $|V_0| = k$, and we now have $e_{G^*}(V_0, V\backslash V_0) \leq \sqrt{\theta^\dagger_j}k^2 + k\theta^\dagger_{j+1}k \leq 2\theta^\dagger_{j+1}k^2$, and we can apply Lemma 3.40 once again with $\alpha_1 = 2\theta^\dagger_{j+1}$ and $\alpha_2 = \nu/2$ to obtain $T \subseteq G$. This completes the proof of Theorem 3.3, except for Lemma 3.40, which will be proved along with Lemma 3.39 later.

3.6.3 Proof of Theorem 3.1

Proof of Theorem 3.41

We assume that $j < \lfloor n/k \rfloor$. Recall that Theorem 3.41 followed immediately from Proposition 3.38 and from Lemma 3.39, which in turn required Claims 3.42 and 3.43 and Lemma 3.44. Recall that we define a new constant μ_j such that $\theta_j \ll \mu_j \ll \theta_{j+1}$.

Proof of Proposition 3.38. Suppose the conclusion of Proposition 3.38 does not hold. Recall that $V_0 = V(G)\backslash(\bigcup_{i=1}^j V_i)$ and so we have $|V_0| > k$ and $|L_0| \geq n/2 - j(1/2+\mu_j)k = |V_0|/2 - j\mu_j k \geq (1-\sqrt{\mu_j})|V_0|/2$. Now since $e(V_0, V(G)\backslash V_0) \leq$

[1]These small vertices exist in $G^\dagger$ since $|L_i| \leq (1+\nu/2)k/2$ and therefore $|S_i| \geq (1-\nu/2)k/2$ for $i = 1, \ldots, j$.

$j\theta_j k^2 \leq \sqrt{\theta_j}k^2$, at most $\mu_j k$ vertices of L_0 have at least $\mu_j k$ neighbours outside V_0, and so at least $(1 - 2\sqrt{\mu_j})|V_0|/2$ vertices of L_0 have at least $(1 - \mu_j)k$ neighbours in V_0. Since we also chose j to be maximal, the conditions of Theorem 3.8 hold with $G' = G[V_0]$ and with $\alpha_1 = \theta_{j+1}$ and $\alpha_2 = 2\sqrt{\mu_j}$, and so $T \subseteq G$, which is a contradiction. $\square$

Proof of Claim 3.42. Let $L(T)$ denote the set of leaves in T and let $\ell(T) := |L(T)|$. Note that $\sum_{x \in V(T)}(d(x) - 2) = 2e(T) - 2|T| = -2$. Thus

$$\sum_{x \in V(T),\, d(x)>2} (d(x) - 2) = \ell(T) - 2 \leq \gamma_1 k.$$

and so there are at most $\gamma_1 k$ vertices of degree more than 2. We remove vertices which have degree more than 2 and all leaves. Removing a vertex x from T increases the number of components by at most $d(x) - 1$, and so we obtain a forest F' with at most $\sum_{x \in V(T),\, d(x)>2}(d(x) - 1) \leq 2\gamma_1 k$ components. Now the remaining components are in fact path-segments, as are any sub-paths. Let $\mathcal{P}$ be a maximal set of vertex-disjoint paths on q vertices in F'. Suppose $|\mathcal{P}| < \gamma_2 k$. Then $\mathcal{P}$ covers at most $q\gamma_2 k$ vertices. Let F'' be the graph obtained by removing $\mathcal{P}$. Then F'' has at most $2\gamma_1 k + 1 + \gamma_2 k$ components, and at least $(1 - 2\gamma_1 - q\gamma_2)k$ vertices. But then some component of F'' contains at least $\frac{1-2\gamma_1-q\gamma_2}{2\gamma_1+\gamma_2+1/k} \geq q$ vertices, and so we can find another path on q vertices in $F'' \subseteq F'$, contradicting the maximality of $\mathcal{P}$. Thus $|\mathcal{P}| \geq \gamma_2 k$. $\square$

Proof of Claim 3.43. Suppose that $|A \cup V(\mathcal{P}')| \leq k - 2$. Let $L_A := A \cap L$, let $L'_A = L_A \cap V(\mathcal{P}')$, the set of endpoints of $\mathcal{P}'$, and $L''_A := L_A \backslash V(\mathcal{P}')$. We will try to find an additional path with end-vertices in L''_A, thus contradicting the maximality of $\mathcal{P}'$. Note that $|V(\mathcal{P}') \backslash L_A| \leq k - 2 - |A| \leq \alpha k$. Thus, since each path in $\mathcal{P}'$ has exactly two vertices in L'_A and at least one vertex in $V(\mathcal{P}') \backslash L'_A$, we have

$|L'_A| \leq 2|V(\mathcal{P}')\backslash L_A| \leq 2\alpha k$, and so $|L''_A| \geq k/2 - 2\alpha k$. Let $A' = A \cup V(\mathcal{P}')$. Then since $|A'| \leq k-2$, each vertex of L''_A has at least three neighbours outside A'. Furthermore, by the maximality of $\mathcal{P}'$, no such neighbour is adjacent to more than one vertex of L''_A. We thus obtain disjoint sets N_1, N_2 and N_3, each of size $|L''_A|$, such that there is a perfect matching between L''_A and N_i for each i.

Now observe that for any two vertices in N_i their neighbourhoods outside A' are disjoint, by the maximality of $\mathcal{P}'$. So suppose that there is a set of αk large vertices in N_i. Each one has at most one neighbour in L''_A, and $|A'\backslash L''_A| \leq k - (k/2 - 2\alpha k)$, and so certainly each one of these vertices has at least $k/3$ neighbours outside A'. But these are all distinct, since otherwise we would have a path of length 4 which contradicts the maximality of $\mathcal{P}'$, and so we have a set of at least $\alpha k(k/3) > n$ distinct vertices in G, which is impossible. Thus at most αk vertices of N_i are large for each i, and so we obtain sets $N'_i \subseteq N_i$ of size $|L''_A| - \alpha k \geq k/2 - 4\alpha k$ consisting entirely of small vertices.

Now consider a maximal matching M_i between N'_i and $V(G)\backslash A'$. Observe that since the N'_i consist of small vertices, N'_i is an independent set, and furthermore any neighbours of a vertex in N'_i are large. Suppose $|M_i| \geq \alpha k$. Then $V(M_i)\backslash N'_i$ is a set of at least αk large vertices, each of which has at most one neighbour in L''_A (otherwise we have a path of length 2, contradicting the choice of $\mathcal{P}'$) and at most $k/2 + 3\alpha k$ neighbours in $A'\backslash L''_A$, and whose neighbourhoods outside A' are disjoint (otherwise we have a path of length 6, contradicting the choice of $\mathcal{P}'$). Thus as before we obtain a set of at least $\alpha k(k/3) > n$ distinct vertices in G, which is impossible and therefore $|M_i| \leq \alpha k$.

Let $N^*_i := V(M_i) \cap N'_i$, $Q_i := V(M_i)\backslash N'_i$ and $N''_i := N'_i\backslash N^*_i$. Then $|N''_i| \geq |N'_i| - \alpha k \geq |L''_A| - 2\alpha k$. Furthermore, $|Q_i| \leq \alpha k$ and N''_i has neighbours only in $Q_i \cup A'$, by the maximality of the matching M_i. Thus since N''_i consists of small vertices, N_i'' has no neighbours in $S \cap A$, and so has neighbours only in Q_i, L_A'' and $V(\mathcal{P}')$. Now let $S' := N_1'' \cup N_2'' \cup N_3''$. Then S' is a set of small vertices, and $N(S') \subseteq Q_1 \cup Q_2 \cup Q_3 \cup L_A'' \cup V(\mathcal{P}')$. Thus $|N(S')| \leq 3\alpha k + |L_A''| + \alpha k$. Also $|S'| \geq 3(|L_A''| - 2\alpha k)$. Altogether,

$$|S'|/|N(S')| \geq \frac{3(|L_A''| - 2\alpha k)}{|L_A''| + 4\alpha k} \geq 2$$

since $|L_A''| \geq (1/2 - 2\alpha)k$. But initially in Section 3.2 we assumed that there was no such set $S' \subseteq S$. This is a contradiction, and completes the proof of the claim. □

Proof of Lemma 3.44. We have a set $\mathcal{P}$ of $\gamma_2 k$ vertex-disjoint path-segments of length 8 in T' and a set $\mathcal{P}'$ of paths in G' which cover D_1 and L_1'. For each path P' in $\mathcal{P}'$, we pick a path P of $\mathcal{P}$. This is possible because $|\mathcal{P}'| \leq \gamma_1 k \leq \gamma_2 k = |\mathcal{P}|$. Thus we can pick a distinct P for each P'. We then pick a subpath P^* of P such that $|P^*| = |P'|$ and P^* has both its endpoints in U_1. This is possible since $|P| \geq |P'|+1$, and because the vertices of P alternate between U_1 and U_2. We can also pick the subpath P^* so that it does not include the vertex of P nearest the root. We then embed P^* onto P' in the obvious way. Let us observe that because the endpoints of P^* are in U_1 while the endpoints of P' are in L_1, and because L_1 and S_1 are independent sets, we have embedded vertices of $P^* \cap U_1$ into L_1 and vertices of $P^* \cap U_2$ into S_1.

Note also that because we avoided the vertices of P closest to the root of T, the vertices of the paths P^* nearest the root are always at distance at least 3 from each other, and so do not have a common neighbour. This will be important later on.

We also have sufficiently many paths in $\mathcal{P}$ left over to find $\gamma_2 k/3$ paths of length 2 with midpoints in U_1 and $\gamma_2 k/3$ paths of length 2 with midpoints in U_2. Indeed we can also choose these paths sufficiently far from the ends of paths in $\mathcal{P}$ that the endpoints have distance at least three from any other paths. We construct a forest T^* from T' by deleting the midpoints of these paths. Observe that we now have $U_1^* = U_1 \cap V(T^*)$ and $U_2^* \cap V(T^*)$ with

$$|U_1^*| = |U_1| - \gamma_2 k/3 \leq \delta(C_1, L_1)$$

$$|U_2^*| = |U_2| - \gamma_2 k/3 \leq \delta(L_1, C_1)$$

Now let $x_1, \ldots, x_\ell$ be the endpoints closest to the root of the paths of $\mathcal{P}$ chosen to cover $\mathcal{P}'$, and let $y_i := P(x_i)$, $z_i := P(y_i)$. (Recall that $P(x)$ denotes the parent of x, i.e. the vertex directly above x in the rooted tree T.) Note that because the x_i all have distance at least 3 from each other, the y_i are distinct. We embed the remainder of T^* greedily, starting at the root of T and placing vertices of U_1 into L_1 and vertices of U_2 into S_1. To see that we can do this, observe that when x_i has been embedded to u_i and z_i to w_i, then $u_i, w_i \in L_1$ and we have

$$|N(u_i) \cap N(w_i) \cap C_1| \geq |C_1| - 2\gamma_1 k \geq |U_2^*|$$

and so there is always a free common neighbour v_i available for y_i. It is important here that the y_i are distinct, although the z_i may not be. Therefore when embedding this parent, we only need to find a common neighbourhood of two vertices that have already been embedded (the vertex of the path and its grandparent).

It remains only to embed the midpoints of the paths of length 2 which were deleted. We begin with those midpoints which were in U_1. Note that $\gamma_2 k/3$ vertices of L_1'' remain free. Let b_i, c_i denote the vertices of C_1 which were chosen as endpoints of such paths.

We construct a bipartite auxiliary graph H. One class of H will consist of pairs (b_i, c_i) and the other class will consist of the $\gamma_2 k/3$ free vertices in L_1''. Such a vertex will be joined to a pair (b_i, c_i) if it is adjacent to both of these vertices in G'. Note that $d_H((b_i, c_i)) \geq \gamma_2 k/3 - 2\gamma_1 k > \gamma_2 k/4$, since $\delta(C_1, L_1) \geq |L_1| - \gamma_1 k$, while for $x \in L_1''$, at most $\gamma_1 k$ pairs (b_i, c_i) are not adjacent to x, since $\delta(L_1'', S_1) \geq |S_1| - \gamma_1 k$. Thus $d_H(x) \geq \gamma_2 k/3 - \gamma_1 k > \gamma_2 k/4$. Thus $\delta(H) > \gamma_2 k/4$. It is now easy to see that Hall's condition is satisfied. For we have a bipartite graph on classes A and B, where $|A| = |B|$ the minimum degree is at least $|A|/2 = |B|/2$. Suppose that there exists a set $S \subseteq A$ such that Hall's condition is violated, i.e. $|N(S)| < |S|$. Then since $|N(S)| \geq \delta(H) \geq |A|/2$ we have $|S| > |A|/2$. Let $S' = B \backslash N(S)$. Then $S' \neq \emptyset$ and therefore $|N(S')| \geq \delta(H) \geq |A/2|$. On the other hand $N(S') \subseteq A \backslash S$ and so $|N(S')| \leq |A| - |S| < |A|/2$, which is a contradiction. Thus Hall's condition is satisfied, and so a perfect matching exists. This corresponds to finding suitable vertices of L_1 onto which to embed the U_1-midpoints of the paths of length 2.

By an identical argument we can also find a perfect matching in a bipartite auxiliary graph, with one vertex class consisting of unused vertices in C_1 and the other consisting of pairs of vertices in L_1 onto which the endpoints of some 2-path have been embedded. This allows us to embed the midpoints which lie in U_2. Thus we can embed the whole of T' as required. □

Proof of Lemma 3.39. We will first tidy up the set V_1 to ensure that all vertices have an appropriately high minimum degree. Note that at most $\sqrt{\alpha_1}k$ vertices in L_1 have at least $\sqrt{\alpha_1}k$ neighbours outside V_1. We remove these vertices from L_1. Now by relocating some vertices of L_1 to S_1, and by removing some vertices from S_1, we may assume that $|L_1| = (1/2 + \alpha_2/2)k$, and that $|S_1| = (1/2 - 3\alpha_2/4)k$. It is still true, however, that each large vertex has at least $|V_1| - \sqrt{\alpha_1}k$ neighbours in V_1. Thus

$$e(L_1, S_1) \geq |L_1|(|S_1| - \sqrt{\alpha_1}k) \geq |L_1||S_1| - \sqrt{\alpha_1}k^2.$$

Thus at most $\alpha_1^{1/4}k$ vertices of S_1 have fewer than $|L_1| - \alpha_1^{1/4}k$ neighbours in L_1.

We remove these vertices from S_1 (and thus also from V_1). We also remove a few more vertices from S_1 to obtain $|S_1| = (1/2 - \alpha_2)k$.

We still denote the sets thus obtained by L_1, S_1 and V_1. We now have that $|V_1| = (1/2 - \alpha_2/2)k$, that every vertex in L_1 has at least $|V_1| - \sqrt{\alpha_1}k$ neighbours in V_1 (and thus at least $|L_1| - \sqrt{\alpha_1}k$ neighbours in L_1 and at least $|S_1| - \sqrt{\alpha_1}k$ neighbours in S_1) and that every vertex in S_1 has at least $|L_1| - \alpha_1^{1/4}k$ neighbours in L_1. Note that we may treat a vertex of L_1 as a vertex of S_1 if necessary. This is why it is useful that L_1 is larger than actually required. However, as we will see V_1 may now not be large enough, which will present some technical difficulties.

We consider a bipartition of the tree T into independent sets U_1 and U_2, where $|U_1| \leq |U_2|$. If $|U_1| = |U_2|$, we will choose U_2 to be the set with the greater number of leaves. Now by Fact 3.12, U_2 contains at least $|U_2| - |U_1| + 1$ leaves. So U_2 contains at least two leaves except when $|U_1| = |U_2|$ and T contains only two leaves in total, i.e. T is a path. In this special case, we will move the leaf in U_1 into U_2, and move its parent into U_1. Now U_1 contains exactly one edge. Since we will usually be embedding U_1 into L_1, this will not be a problem. Indeed, this case will be so similar to the more general case when U_1 and U_2 are independent that we will not mention it any further, noting only that the proof can be trivially adapted to resolve it.

Thus we assume that U_2 contains at least two leaves, and all leaves in U_2 are adjacent to vertices in U_1. Since we will be embedding U_1 into L_1, which consists of large vertices, we may embed any leaves in U_2 greedily at the end of the embedding process. So we delete any leaves from U_2. We still denote this set by U_2, and the tree by T. Now $|T| \leq k - 1$. Suppose in fact that $|T| \leq k - 2\alpha_2 k/3 = |V_1| - \alpha_2 k/6$. Since $|U_1| \leq (k+1)/2 \leq |L_1|$, we move vertices from L_1 to S_1 to ensure that $|L_1| = |U_1| + \alpha_2 k/12$. Then we also have $|S_1| \geq |U_2| + \alpha_2 k/12$. The minimum degree conditions between L_1 and S_1 ensure that we can complete the embedding greedily.

Now suppose instead that $|T| > k - 2\alpha_2 k/3$. This means that originally U_2 contained at most $2\alpha_2 k/3+1$ leaves, and thus by Fact 3.12, originally $|U_2|-|U_1|+1 \leq 2\alpha_2 k/3 + 1$. Since also $|U_1| + |U_2| = k + 1$, we have $|U_2| \leq k/2 + \alpha_2 k/3 + 1/2 \leq |L_1| - \alpha_2 k/7$. Thus if U_1 contains at least $\alpha_2 k$ leaves we can perform the same process as before, now removing leaves of U_1 and embedding U_2 into L_1.

Thus in total we may assume that T contains at most $2\alpha_2 k$ leaves. We may therefore apply Claim 3.42 with $\gamma_1 = 2\alpha_2$ and $\gamma_2 = \alpha_3$, where $\alpha_2 \ll \alpha_3 \ll 1$, to find a set of $\alpha_3 k$ vertex-disjoint path-segments of length 8 in T. We will use these to help us embed T by using some extra vertices which are not in V_1.

Now by Claim 3.43 applied to V_1 we have a set of vertex-disjoint paths $\mathcal{P}'$ of length at most 6, each of which has its endpoints in L_1 and its internal vertices, of which there is at least one for each path, outside V_1 and such that $V_1' := V_1 \cup V(\mathcal{P}')$ has size at least $k - 1 \geq |T|$ (recall that we have removed all leaves from U_2). If $|V_1'| \geq |T| + 5$, we simply remove some paths from $\mathcal{P}'$ to ensure that V_1' is not substantially bigger than we need it to be, i.e. that $|V_1'| \leq |T|+4$. This ensures that $|V(\mathcal{P}')| \leq 3\alpha_2 k$. If we still have $|V_1'| > |T|$, we simply discard some small vertices (not in $V(\mathcal{P}')$), or large vertices if no small vertices are left, to ensure that $|V_1'| = |T|$. Now for each path in $\mathcal{P}'$ of odd length, we find a neighbour in $L_1'' := L_1 \backslash V(\mathcal{P}')$ of one of the end-vertices, and add this to the path. This ensures that all paths in $\mathcal{P}'$ have even length, and means that we will be able to use the endpoints for vertices in U_1 while still respecting the bipartition of T. We rearrange the vertices of the paths of $\mathcal{P}'$ to ensure that they alternate between L_1 and S_1, with the endpoints lying in L_1.

We now note that by moving some vertices of L_1 into S_1 (and deleting any edges which now lie within S_1 and L_1), the conditions of Lemma 3.44 are satisfied, where $D_1 \cup L'_1$ consists of the internal vertices of paths in $\mathcal{P}'$, C_1 consists of any remaining vertices of S_1 and L''_1 consists of any remaining vertices of L_1, and where $\gamma_1 = \max(|\mathcal{P}'|/k, \sqrt{\alpha_1}) \leq \alpha_2$, and $\gamma_2 = \alpha_3$. So we can apply that lemma to embed the tree T into V_1 and since $|U_1|$ vertices will be embedded into $L_1 \subseteq L$, we can also embed the remaining leaves of T greedily, and thus $T \subseteq G$, as required. □

Proof of Lemma 3.40. It is in this proof that we finally need to use the full strength of the conditions on G rather than G^*.

At first we remain in G^* and we go through exactly the same proof as the one for Lemma 3.39 to obtain $V_1 = L_1 + S_1$ such that $|L_1| = (1/2 + \alpha_2/2)k$ and $|S_1| = (1/2 - \alpha_2)k$. We also have that each vertex in L_1 has at least $|V_1| - \sqrt{\alpha_1}k$ neighbours in V_1 and each vertex in S_1 has at least $|L_1| - (\alpha_1)^{1/4}k$ neighbours in L_1.

We now transfer to G and continue to go through the same proof as the one for Lemma 3.39. Note that we needed to apply Claim 3.43 for that proof. In order to see that this is still permissible, we must observe that $L^* \subseteq L$. Note also that in the proof of Claim 3.43 we needed to use the fact that G is edge-minimal subject to satisfying the conditions of Theorem 3.1, and in particular that there is no set $S' \subseteq S$ such that $|N(S')| \leq |S'|/2$. This is not necessarily true in G^* and this is the reason that we need to use G instead. □

Note that with the proof of Lemma 3.40 we have finally completed the proof of Theorem 3.3, and therefore also the proof of Theorem 3.8.

Now Proposition 3.38 and Lemma 3.39 (with $\alpha_1 = \theta_j$ and $\alpha_2 = \mu_j$) prove Theorem 3.41, and so we move on to proving Theorems 3.45 and 3.46.

3.6.4 Proof of Theorem 3.45

We now know that $j = \lfloor n/k \rfloor$. Let us define θ'_j and θ''_j such that $\theta_j \ll \theta'_j \ll \theta''_j \ll \theta_{j+1}$. We also know that for each $i > 0$, $|L_i| \leq (1/2 + \theta'_j)k$ (from Lemma 3.39). We begin the proof of Theorem 3.45 with two simple claims.

Claim 3.53 *Either* $|V_0| \leq \theta''_j k$ *or* $|V_0| \geq (1-\theta''_j)k$.

Note that although we did not state it as a numbered result, a similar argument appeared in Section 3.6.2 for the proof of Theorem 3.3.

Proof. Suppose instead that $|V_0| \in (\theta''_j k, (1-\theta''_j)k)$. Then

$$\begin{aligned}|V_0 \cap L| &\geq n/2 - \sum_{i=1}^{j}(1+\theta'_j)k/2 = (n-jk)/2 - j\theta'_j k/2 \\ &= |V_0|/2 - j\theta'_j k/2 \geq \theta''_j k/4.\end{aligned}$$

But $e(V_0, V(G)\backslash V_0) \leq \sum_{i=1}^{j} \theta_j k^2 \leq \theta_j^{2/3} k^2$. So at most $\theta_j^{1/3} k$ vertices of V_0 have degree at least $\theta_j^{1/3} k$ outside V_0. But this does not cover all the large vertices in V_0, and so there must still be some vertices with degree at least $(1-\theta_j^{1/3})k$ in V_0. So $|V_0| \geq (1-\theta_j^{1/3})k \geq (1-\theta''_j)k$, which is a contradiction. □

Note in particular that if $|L_0| \geq (1/2+\theta_{j+1})k$, then we could move a few small vertices into V_0 to ensure that $|V_0| = k$. Then we have

$$e(V_0, V(G)\backslash V_0) \leq \sum_{i=1}^{j} \theta_j k^2 + \theta''_j k^2 \leq 2\theta''_j k^2$$

and we can apply Lemma 3.39 with $\alpha_1 = 2\theta''_j$ and $\alpha_2 = \theta_{j+1}$ to obtain $T \subseteq G$. Thus we may assume that $|L_0| \leq (1/2+\theta_{j+1})k$. Thus for every i, including $i = 0$, we have $|L_i| \leq (1/2+\theta_{j+1})k$.

Claim 3.54 *For all* $i \neq 0$, $|L_i| > (1/2 - \sqrt{\theta_{j+1}})k$.

Proof. Suppose not. Then since no vertex class has substantially more than half its vertices being large (including V_0) there cannot possibly be a total of $n/2$ large vertices in the classes of G. More precisely

$$\begin{aligned}|L| &\leq |L_0| + (1 - \sqrt{\theta_{j+1}})k/2 + (j-1)(1+\theta_{j+1})k/2 \\ &\leq |L_0| - 2\sqrt{\theta_{j+1}}k/2 + (n - |V_0|)/2\end{aligned}$$

and so

$$|L_0| \geq |V_0|/2 + 2\sqrt{\theta_{j+1}}k/2.$$

In particular $|V_0| \geq (1-\theta_j'')k$, and so $|L_0| > (1/2+\theta_{j+1})k$. But we already assumed that this is not the case, which is a contradiction □

Note that by a similar argument we have either $|V_0| < \theta_j''k$ or $|L_0| > (1/2 - \sqrt{\theta_{j+1}})k$.

In general, if $|V_0| < \theta_j''k$ we can ignore it, and if $|V_0| \geq (1-\theta_j'')k$ we can, if necessary depending on whether we intend to embed into V_0, add in a few vertices from some other class to increase the size to k. This doesn't affect calculations significantly, so for the remainder of the proof of Theorem 3.45 we will assume for simplicity that $V_0 = \emptyset$, or $|V_0| = k$, in which case we will call it V_{j+1} and increase j. In either case we now have that $e(V_i, V \backslash V_i) \leq \theta_j k^2$ for $i = 1, \ldots, j$. In the proof of Theorem 3.46 we will need to consider the case when n is not divisible by k (and therefore $0 < |V_0| < k$) and deal with it more carefully.

Let us observe that since we are in EC_j, we have $e(V_i, V\backslash V_i) \leq \theta_j k^2$ for each $1 \leq i \leq j$. Thus at most $\sqrt{\theta_j}k$ vertices of L_i have more than $\sqrt{\theta_j}k$ neighbours outside V_i. Thus at least $|L_i| - \sqrt{\theta_j}k$ vertices of L_1 have at least $(1-\sqrt{\theta_j})k$ neighbours in V_i, and so in particular have at least $|S_i| - \sqrt{\theta_j}k$ neighbours in S_i. Thus $e(L_i, S_i) \geq (|L_i| - \sqrt{\theta_j}k)(|S_i| - \sqrt{\theta_j}k) \geq |S_i|(|L_i| - \theta_{j+1}k)$. This last inequality holds since $(1/2 - \sqrt{\theta_{j+1}})k \leq |L_i|, |S_i| \leq (1/2 + \sqrt{\theta_{j+1}})k$. Therefore at most $\sqrt{\theta_{j+1}}k$ vertices of S_1 have fewer than $|L_i| - \sqrt{\theta_{j+1}}k$ neighbours in L_i. Thus we obtain sets

$L'_i \subseteq L_i$ and $S'_i \subseteq S_i$ such that

$$\delta(L'_i, L'_i), \delta(L'_i, S'_i), \delta(S'_i, L'_i) \geq (1/2 - 2\sqrt{\theta_{j+1}})k. \tag{3.5}$$

We call the vertices of $S_i \backslash S'_i$ *bad* vertices, and denote this set by B_i.

We now aim to prove Proposition 3.47. In the proof we will use two simple facts from [72] (Fact 6.2 Part 1 and Fact 6.8 in that paper).

Fact 3.55 *Suppose G' is a graph with $V(G') = C + D$, and T' a tree. Suppose also that $V(T') = U_1 + U_2$, where $|U_1| \leq |U_2|$ and U_2 is an independent set, and that G' satisfies $\delta(C, C), \delta(D, C) \geq |U_1|$ and $\delta(C, D) \geq |U_2|$. Then $T' \subseteq G'$ with U_1 embedded into C and U_2 into D.* □

Recall that $\ell(T)$ denotes the number of leaves of T.

Fact 3.56 *1. For any positive integer $q \leq k+1$ there is a vertex x of T, and some children $y_1, y_2, \ldots, y_t$ of x such that $q/2 \leq |T(x) \backslash (\bigcup_{i=1}^t T(y_i))| < q$.*

2. For any positive integer $q \leq \ell(T)$ there is a vertex x of T, and some children $y_1, y_2, \ldots, y_t$ of x such that $T(x) \backslash (\bigcup_{i=1}^t T(y_i))$ contains $[q/2, q)$ leaves of T. □

We also need the following simple result, which is very similar to Claim 6.6 in [72]. Recall that a skew-partition is an ordered vertex partition $V(T) = U_1 + U_2$ such that $|U_1| \leq |U_2|$ and U_2 is an independent set. Recall also that we call $g(U_1, U_2) := |U_2| - |U_1|$ the *gap* of the partition, and that $g(T) := g(T_{odd}, T_{even})$ is the *gap of* T.

Remark 3.57 *If T has a skew-partition with $g(U_1, U_2) \geq 5\sqrt{\theta_{j+1}}k$, then $T \subseteq G$.*

Proof. By Fact 3.12, U_2 contains at least $|U_2| - |U_1| + 1$ leaves. Deleting these leaves gives two vertex sets U_1, U'_2 each of size at most $(k+1)/2 - 5\sqrt{\theta_{j+1}}k/2 \leq (1/2 - 2\sqrt{\theta_{j+1}})k$. Now the minimum degree conditions of (3.5) ensure that for any i we can embed U_1 into L'_i and U'_2 into S'_i greedily, starting at the root. Now the remaining vertices of U_2 are leaves adjacent to vertices of U_1. Since vertices of U_1 were embedded onto large vertices, we can embed these remaining leaves greedily.

□

Given a partition U_1, U_2 of the vertices of T and a subtree T', *flipping* T' means moving the vertices of T' that lie in U_1 into U_2 and vice versa.

Corollary 3.58 *$k/2 - \theta_{j+2}k \leq |T_{odd}|, |T_{even}| \leq k/2 + \theta_{j+2}k$. Furthermore, for any subtree $T' \subseteq T$, of the form $T' = T(x) \setminus \bigcup_{i=1}^{s} T(y_i)$, where $x \in V(T)$ and $y_1, \dots, y_s$ are children of x, we have*

$$|T' \cap T_{odd}| - |T' \cap T_{even}| \in (-3\theta_{j+2}k, 3\theta_{j+2}k)$$

or else $T \subseteq G$.

Proof. The first part is immediate from Remark 3.57. For the second, suppose that $|T' \cap T_{odd}| - |T' \cap T_{even}| \geq 3\theta_{j+2}k$. Then we start with $U_1 = T_{odd}, U_2 = T_{even}$ and we flip T' (except for x if x lies in U_1). Together with the bound in the first part, this gives a skew-partition with a gap of size at least $3\theta_{j+2}k - 1 - 2\theta_{j+2} \geq 5\sqrt{\theta_{j+1}}$, and thus by Remark 3.57 we could embed T in G. A similar argument shows that $|T' \cap T_{odd}| - |T' \cap T_{even}| \geq -3\theta_{j+2}k$. □

Proof of Proposition 3.47. Let $I := \{i : d(v_0, V_i) \geq \theta_{j+1}^{1/4}k\}$ and note that by assumption $|I| \geq 2$. By relabelling if necessary, we may assume that $I = \{1, \dots, s\}$. We also assume without loss of generality that for $1 \leq i_1 < i_2 \leq s$,

$$d(v_0, V_{i_1}) \leq d(v_0, V_{i_2}). \tag{3.6}$$

We will follow the proof of Proposition 6.12 in [72] with some minor modifications.

Recall that we have sets $L_i' \subseteq L_i$ and $S_i' \subseteq S_i$ satisfying (3.5). By Remark 3.57 we may assume that T has no skew-partition with a large gap.

Now by Fact 3.56 we can find a vertex x in T and some children $y_1, \dots, y_t$ such that setting $T = T(x)\backslash(\bigcup_{i=1}^t T(y))$ we have $|T'| \in [\theta_{j+1}^{1/4}k/4, \theta_{j+1}^{1/4}k/2)$.

Now $d_{T'}(x) \leq \theta_{j+1}^{1/4}k/2$, and $d_{L_1' \cup S_1'}(v_0) \geq \theta_{j+1}^{1/4}k - 4\sqrt{\theta_{j+1}}k \geq \theta_{j+1}^{1/4}k/2$. So we embed x onto v_0, and T' into $L_1' \cup S_1'$ greedily (note that $\delta(L_1' \cup S_1') \geq |T'|$).

Now let $F := T\backslash T'$ be the forest consisting of $T\backslash T(x), T(y_1), T(y_2), \dots, T(y_t)$. Since any isolated vertices of F are neighbours of x, and since $v_0 \in L$, we can embed these greedily at the end. So we assume that F contains no isolated vertices. Thus the number of roots in F is at most $|F|/2 \leq (1 - \theta_{j+1}^{1/4}/4)k/2$. Let $V_I := \bigcup_{i \in I}(L_i' \cup S_i')$. Then

$$d(v_0, V_I \backslash V_1) \geq k - j\theta_{j+1}^{1/4}k - 4j\sqrt{\theta_{j+1}}k - d(v_0, L_1' \cup S_1').$$

If $s \geq 3$ then by (3.6) we have $d(v_0, L_1' \cup S_1') \leq \frac{1}{3}d(v_0, V_I)$, and so $d(v_0, V_I \backslash V_1) \geq k/2$, which will be enough for our purposes. If $s = 2$, however, we need to be more careful. (The following argument also works for $s \geq 3$, although in that case it is substantially more complicated than necessary, as indicated by the easy argument above.) We have

$$d(v_0, \bigcup_{i=2}^{j} L_i' \cup S_i') \geq k/2 - 4j\sqrt{\theta_{j+1}}k \geq \frac{1}{2}(1 - \theta_{j+1}^{1/4}/4)k$$

so we can embed the roots of F greedily into $\bigcup_{i=2}^{j} L_i' \cup S_i'$.

For $2 \leq i \leq j$, let F_i^a, F_i^b denote the subforests of F with roots in L_i', S_i' respectively. Let $U_i^{(1)} := (F_i^a)_{even} \cup (F_i^b)_{odd}$, and $U_i^{(2)} := (F_i^a)_{odd} \cup (F_i^b)_{even}$. If $|U_i^{(1)}|, |U_i^{(2)}| \leq (1/2 - 2\sqrt{\theta_{j+1}})k$ for each i, then the conditions of (3.5) ensure that we can complete the embedding using the greedy algorithm. (We embed $U_i^{(1)}$ into L_i' and $U_i^{(2)}$ into S_i'.) Note also that since $[\theta_{j+1}^{1/4}k/4, \theta_{j+1}^{1/4}k/2)$ vertices have already been embedded into V_1, only at most one i can fail to satisfy this condition. We will show in this case that T has a skew-partition (U_1, U_2) with gap at least $5\sqrt{\theta_{j+1}}k$.

If $|U_i^{(2)}| > (1/2 - 2\sqrt{\theta_{j+1}})k$, we put $U_i^{(2)}$ into U_2. We place x into U_1 along with the smaller half of the bipartition of T'. We place the larger part of T' into U_2, and any remaining vertices of T into U_1. Then U_2 is indeed an independent set, and

$$|U_2| \geq (1/2 - 2\sqrt{\theta_{j+1}})k + \theta_{j+1}^{1/4}k/8 - 1 \geq (1/2 + \theta_{j+1}^{1/4}/10)k.$$

Thus $|U_1| \leq (1/2 - \theta_{j+1}^{1/4}/10)k$, and so $|U_2| - |U_1| \geq \theta_{j+1}^{1/4}k/5 \geq 5\sqrt{\theta_{j+1}}k$, and U_1, U_2 is a skew-partition with a large gap, which by Remark 3.57 we assumed earlier was not the case, and so we have a contradiction.

If on the other hand $|U_i^{(1)}| > (1/2 - 2\sqrt{\theta_{j+1}})k$, the process is similar. Now $U_i^{(1)}$ goes into U_2 along with the larger part of T', except for x. The rest of T, including x, goes into U_1. The calculations are the same and again yield the desired contradiction. This completes the proof of Proposition 3.47. □

Proof of Proposition 3.48. For each $i \in [1, j]$, let $L_i^* := \{x \in L : d(x, V_i) \geq \theta_{j+1}^{1/4}k\}$. Since the conditions of Proposition 3.47 do not hold, the L_i^* are pairwise disjoint. Furthermore, for $x \in L_i^*$, $d(x, V_i) \geq k - j\theta_{j+1}^{1/4}k \geq (1 - \theta_{j+1}^{1/5})k$. Since $e(V_i, V(G)\backslash V_i) \leq \theta_j k^2$, we have $|L_i^*\backslash V_i| \leq \sqrt{\theta_j}k$.

We move each L_i^* into V_i, and move some small vertices to rebalance the sizes of the V_i. We have moved at most $2j\sqrt{\theta_j}k$ vertices, and thus[1]

$$e(V_i, V(G)\backslash V_i) \leq \theta_j k^2 + (\sqrt{\theta_j}k)(j\theta_{j+1}^{1/4}k) + (j\sqrt{\theta_j}k)(\theta_{j+1}^{1/4}k) + (j\sqrt{\theta_j}k)k \leq \theta_{j+2}^3 k^2.$$

[1]The four terms in the central expression come from the original number of edges, the edges coming off those vertices of L_i^* which we had to move into V_i, those edges coming off large vertices moved out of V_i and those edges coming off small vertices which were moved either into or out of V_i.

Furthermore, each vertex in L_i still has at least

$$(1-\theta_{j+1}^{1/4})k - 2j\sqrt{\theta_j}k^2 \geq (1-\theta_{j+2}^3)k$$

neighbours in V_i, and at most $\theta_{j+1}^{1/4}k + 2j\sqrt{\theta_j}k \leq \theta_{j+2}^3 k$ neighbours outside. Thus the conditions of $EC'(\theta_{j+2}^3)$ are now satisfied. □

This therefore also completes the proof of Theorem 3.45.

3.6.5 Proof of Theorem 3.46

We have now tidied up the large vertices in each of the classes to ensure that $EC'(\theta_{j+2}^3)$ holds. Note that every class has approximately half its vertices in L. We would like to say now that there is some class which has at least $k/2$ large vertices. However, this may not be the case if $(1-\theta_j^{1/4})k \leq |V_0| < k$. Then V_0 may have more than half its vertices lying in L, but nevertheless $|L_0| < k/2$. Therefore if all the remaining vertices have very slightly less than half their vertices lying in L, we have no class for which $|L_i| \geq k/2$. This will cause some difficulty later on.[1]

The proof in the case when there is some i with $|L_i| \geq k/2$ is substantially easier. However, since the harder case would rely on many very similar results, we prove the two together. This involves stating and proving certain results in considerably more generality than we would require for the easier case.

We begin by rearranging some small vertices. Recall that for each $i = 1, \ldots, j$ we have $|L_i| \geq (1/2 - \theta_{j+2}^3)k$, and that for each $x \in L_i$, $d_{V_i}(x) \geq (1-\theta_{j+2}^3)k$. This means that $e(L_i, S_i) \geq |L_i|(|S_i| - \theta_{j+2}^3 k) \geq |S_i|(|L_i| - 2\theta_{j+2}^3 k)$. (Note that this also holds for $i = 0$ in the case when $|V_0| \geq (1-\theta_j'')k$.) Thus at most $\sqrt{2}\theta_{j+2}^{3/2}|S_i|$ vertices

[1]We don't have the same problem if $|V_0| \leq \theta_j^{1/4}$, since then by Proposition 3.47 we could move any large vertex of V_0 to the (unique) class in which it has almost k neighbours, effectively leaving $L_0 = \emptyset$.

of S_i have fewer than $|L_i| - \sqrt{2}\theta_{j+2}^{3/2}k$ neighbours in L_i. We call such vertices *bad*, and we move them into the class V_i in which they have most neighbours. Note that since all bad vertices are small, and because the set of small vertices is independent, this rearrangement is well-defined. We still call those vertices bad, and denote by B_i the set of bad vertices in V_i. Because we are moving fewer than $3\theta_{j+2}^{3/2}n$ vertices in total, we still have sets V_i with the following properties:

- $(1 - \theta_{j+2})k \leq |V_i| \leq (1 + \theta_{j+2})k$;
- $(1/2 - \theta_{j+2})k \leq |L_i| \leq (1/2 + \theta_{j+2})k$;
- $e(V_i, V(G)\backslash V_i) \leq \theta_{j+2}k^2$;
- For each $x \in L_i$, $d(x, V(G)\backslash V_i) \leq \theta_{j+2}k$;
- $|B_i| \leq \theta_{j+2}k$, and $\delta(S_i\backslash B_i, L_i) \geq (1/2 - \theta_{j+2})k$.

Recall that in the proof of Theorem 3.45 we ignored the set V_0 if its size was small, or called it V_{j+1} and increased j if it had size almost k. It did not affect calculations significantly to assume that either $|V_0| = 0$ or $|V_0| = k$. Now, however, we no longer need to make such assumptions, because we no longer demand that $|V_i| = k$. If originally V_0 was small (i.e. $|V_0| \leq \theta_j''k \leq \theta_{j+1}k$), then its vertices will now be distributed among the other V_i, each vertex being placed into the class in which it has most neighbours. Since $\theta_{j+1} \ll \theta_{j+2}$, this does not affect the above conditions significantly. If it had size almost k, then because we treated it as a set of size k we have already performed all the same rearrangements as we have performed for all the other classes, and therefore V_0 will satisfy the above conditions just like all the other V_i.

We note also that for $0 \leq i \leq j$ we have $|S_i| \geq (1 - \theta_{j+2})k - (1/2 + \theta_{j+2})k \geq (1/2 - 2\theta_{j+2})k$, and each $x \in L_i$ has at least $k - \theta_{j+2}k \geq |V_i| - 2\theta_{j+2}k$ neighbours in

V_i, so (with the bound for $\delta(S_i\backslash B_i, L_i)$ which we had before) we have

$$\begin{aligned} \delta(L_i, S_i\backslash B_i) &\geq (1/2-2\theta_{j+2})k - \theta_{j+2}k - 2\theta_{j+2}k = (1/2-5\theta_{j+2})k \\ \delta(L_i, L_i) &\geq (1/2-\theta_{j+2})k - 2\theta_{j+2}k = (1/2-3\theta_{j+2})k \qquad (3.7) \\ \text{and} \quad \delta(S_i\backslash B_i, L_i) &\geq (1/2-\theta_{j+2})k. \end{aligned}$$

We can now prove Proposition 3.49.

Proof of Proposition 3.49 First of all, to see how the "in particular" follows from the first statement, recall that a good small vertex in V_i has at least $(1/2-\theta_{j+3})k$ neighbours in V_i, and so has at most $(1/2+2\theta_{j+3})k - (1/2-\theta_{j+3})k = 3\theta_{j+3}k$ neighbours outside V_i. On the other hand a bad small vertex in V_i has at least as many neighbours in V_i as in any other $V_{i'}$. Therefore if it has more than $(1/4+\theta_{j+3})k$ neighbours in $V_{i'}$, it also has at least $(1/4+\theta_{j+3})k$ neighbours in V_i, and so has at least $(1/2+2\theta_{j+3})k$ neighbours in total, which is a contradiction.

Therefore we need only show that any small vertex has at most $(1/2+2\theta_{j+3})k$ neighbours. Suppose instead that we have some vertex $v_0 \in S_i$ with at least $(1/2+2\theta_{j+3})k$ neighbours. Then since $|L_i| \leq (1/2+\theta_{j+2})k$, v has at least $\theta_{j+3}k$ neighbours outside V_i, and since v_0 is small these neighbours must be large vertices.

By Corollary 3.58 we may assume that $|T_{odd}|, |T_{even}| \leq (1/2+\theta_{j+2})k$. Thus $d(v_0) \geq (1/2+2\theta_{j+3})k \geq |T_{odd}|, |T_{even}|$. In particular, $d(v_0) \geq \Delta(T)$.

As in the proof of Proposition 3.47, by Fact 3.56 we can find a vertex x and some children $y_1, \ldots, y_t$ in T such that $T' := T(x)\backslash(\bigcup_{i=1}^t T(y_i))$ satisfies $|T'| \in [\theta_{j+3}k/2, \theta_{j+3}k)$.

Note that by Corollary 3.58 we have $|T' \cap T_{even}|, |T' \cap T_{odd}| \geq \theta_{j+3}k/4 - 3\theta_{j+2}k \geq \theta_{j+3}k/5$. Thus

$$|T_{odd}\backslash T'|, |T_{even}\backslash T'| \leq (1/2+\theta_{j+2})k - \theta_{j+3}k/5 \leq (1/2-\theta_{j+3}/6)k.$$

We now embed x onto v_0 and $N(x) \cap T'$ greedily onto neighbours of v_0 in $V \backslash V_i$. This is possible since $|T'| \leq \theta_{j+3}k \leq d(v_0, V \backslash V_i)$. Now since v_0 is small its neighbours must be large, and the minimum degree conditions of (3.7) (applied with i' instead of i) ensure that we can easily embed the remainder of T' greedily into $V \backslash (V_i \cup B)$, where $B = \bigcup_{i'=0}^{j} B_{i'}$.

We now embed $T \backslash T'$ into V greedily with x as the root. If v_0 is good, then it has at least $(1/2 - 5\theta_{j+2})k \geq |T_{odd} \backslash T'|, |T_{even} \backslash T'|$ neighbours in its own class, and so we can embed $T \backslash T'$ into $L_i \cup (S_i \backslash B_i)$ greedily, using the minimum degree of at least $(1/2 - 5\theta_{j+2})k$ between these two sets.

On the other hand, if v_0 is bad then it contains at least as many neighbours in its own class as in any other. We order the remaining neighbours $y_{t+1}, \ldots, y_{t'}$ of x in such a way that $|T(y_{t+1})| \geq |T(y_{t+2})| \geq \ldots \geq |T(y_{t'})|$. Then if we begin by embedding the y_m in order, first embedding as many as possible into V_i, then we will eventually embed at least $(1 - \theta_{j+3})k/C$ vertices into V_i. Since we have already embedded at least $\theta_{j+3}k/2$ vertices into $V \backslash V_i$, this ensures that we never attempt to embed too many vertices (i.e. $(1 - 20\theta_{j+2})k$) into any one class. Note that $T'' = T(x) \backslash \left(\bigcup_{i=1}^{m-1} T(y_i) \cup \bigcup_{i=m'+1}^{s'} T(y_i) \right)$ is balanced by Corollary 3.58, and so $T'' - x$, which we intend to embed in $V_{i'}$, is also balanced. Thus the minimum degree conditions of (3.7) between the $L_{i'}$ and the $S_{i'} \backslash B_{i'}$ ensure that we can do the remainder of the embedding greedily. But since we assumed that T cannot be embedded into G this is a contradiction, as required. $\square$

We now note that since in total half the vertices of G are large, there must be some set V_i for which $|L_i| \geq |S_i|$. Without loss of generality, we will assume that this set is V_1. We will do most of the embedding in V_1, although it may be too small to embed all of T, and a few vertices will be embedded into other classes. This is the purpose of Lemma 3.50, which we will prove shortly.

Recall that T is rooted at r. We may assume without loss of generality that $|T_{even}| \leq |T_{odd}|$ (otherwise we move the root to one of its neighbours). Recall that we define the gap of T to be $g(T) = |T_{odd}| - |T_{even}|$. Note in particular that by Remark 3.57 we may assume that $g(T) \leq 5\sqrt{\theta_{j+1}}k$, and so $k/2 - \theta_{j+2}k \leq |T_{even}| \leq |T_{odd}| \leq k/2 + \theta_{j+2}k$. We split the proof further into two cases:

- *Case 1: T has at least $36\theta_{j+4}k$ leaves,*
- *Case 2: T has fewer than $36\theta_{j+4}k$ leaves.*

Recall that $m_i := k - |V_i|$. If a vertex is not bad, then we call it *good*. Note that this includes all large vertices. Recall Lemma 3.50 which we will need in both cases, although its full strength is only needed in Case 1.

Lemma 3.50. *Let $q_1, q_2, \ldots, q_s$ be positive integers such that $q := \sum_{i=1}^{s} q_i \leq 2(m_1 + 1)/3$, and let $C_1, \ldots, C_s \subseteq L_1$ be (not necessarily distinct or disjoint) sets of size $(1/2 - 2\theta_{j+4})k$. Then there are*

- *q disjoint $(1/\theta_{j+4})$-stars in $V \backslash V_1$ with midpoints $y_1, \ldots, y_q$;*
- *distinct vertices $x_1, \ldots, x_s \in L_1$ and*
- *a partition of the set of stars into $Q_1, \ldots, Q_s$*

such that for each $i = 1, \ldots, s$ we have

- *$x_i \in C_i$;*
- *$|Q_i| = q_i$ and*
- *$M(Q_i) \subseteq N(x_i)$*

where $M(Q_i)$ denotes the set of midpoints of the stars in Q_i. Furthermore, all endpoints of stars are good vertices.

Proof. We prove the lemma inductively on s. For $s = 0$, there is nothing to prove, so we assume that $s \geq 1$ and that we have a set Q of appropriate stars for integers $q_1, \ldots, q_{s-1}$, along with vertices $x_1, \ldots, x_{s-1}$.

Now let P be the set of vertices outside V_1 with fewer than $(m_1 + 1)/\theta_{j+4}$ neighbours in their own class and let $C'_s := C_s \backslash \{x_1, \ldots, x_{s-1}\}$. Now if any vertex in C'_s has at least q_s neighbours in $V \backslash (V_1 \cup P \cup V(Q))$, then we call this vertex x_s and we can greedily pick q_s neighbours $y_1, \ldots, y_{q_s}$ and $(1/\theta_{j+4})$ further neighbours for each y_i to find the required further stars to form Q_s, and so the proof is complete. (Note that if y_i is bad, then it is small and all its neighbours are large and therefore good. On the other hand, if y_i is good, then by (3.7) there are plenty of good neighbours to choose from. So we can ensure that all the endpoints are good.) Thus we may assume that any vertex in C'_s has fewer than q_s neighbours in $V \backslash (V_1 \cup P \cup V(Q))$. This means that

$$e(C'_s, V \backslash (V_1 \cup P \cup V(Q))) \leq q_s |C'_s|. \tag{3.8}$$

Note also that since any vertex in $End(Q) := V(Q) \backslash M(Q)$ (the set of endpoints of the stars in Q) is good, it has at most $3\theta_{j+3}k$ neighbours in C'_s by Proposition 3.49. Thus we have

$$e(C'_s, End(Q)) \leq 3\theta_{j+3}k|End(Q)| = 3\theta_{j+3}k(q - q_s)/\theta_{j+4}. \tag{3.9}$$

Now we also have that any vertex in $M(Q)$ has at most $(1/4 + \theta_{j+3})k$ neighbours outside its own class (by Proposition 3.49) and so

$$e(C'_s, M(Q)) \leq (1/4 + \theta_{j+3})k|M(Q)| = (1/4 + \theta_{j+3})k(q - q_s). \tag{3.10}$$

Note also that the vertices of P are bad, and so have at most as many neighbours in V_1 as they have in their own class, and therefore certainly at most as many neighbours as they have in $V \backslash V_1$, which itself is at most $(m_1+1)/\theta_{j+4}$ neighbours. Thus, since $|P| \leq |B| \leq \theta_{j+2} n$,

$$e(C'_s, P) \leq ((m_1+1)/\theta_{j+4})|P| \leq ((m_1+1)/\theta_{j+4})\theta_{j+2} n \leq (m_1+1)\theta_{j+3} k. \quad (3.11)$$

Finally we note that each vertex in C'_s has at least $k - |V_1| + 1 = m_1 + 1$ neighbours in $V \backslash V_1$, and so

$$e(C'_s, V \backslash V_1) \geq (m_1+1)|C'_s|. \quad (3.12)$$

Now (3.8), (3.9), (3.10) and (3.11) together give

$$\begin{aligned} e(C'_s, V \backslash V_1) &\leq q_s|C'_s| + 3(q-q_s)k\theta_{j+3}/\theta_{j+4} \\ &\quad +(1/4+\theta_{j+3})k(q-q_s) + (m_1+1)\theta_{j+3}k \\ &\leq q_s|C'_s| + k(q-q_s)/3 + (m_1+1)\theta_{j+3}k \\ &\leq q_s|C'_s| + (q-q_s)|C'_s| + 3(m_1+1)\theta_{j+3}|C'_s| \\ &\leq (2(m_1+1)/3)|C'_s| + 3(m_1+1)\theta_{j+3}|C'_s| \\ &< (m_1+1)|C'_s| \end{aligned}$$

which contradicts (3.12). □

We will need a slightly different version of Lemma 3.50 later on, for which we make the following remark.

Remark 3.59 *In the proof of Lemma 3.50, we did not need any properties of $L_1 \backslash C_i$. In particular, we did not use the fact that vertices of $L_1 \backslash C_i$ had degree at least k. Therefore the Lemma remains true even if we had already deleted some set of vertices $S'_1 \subseteq S_1$ with $N(S'_1) \subseteq L_1 \backslash \left(\bigcup_{i=1}^{s} C_i\right)$.*

Case 1*: T contains at least $36\theta_{j+4}k$ leaves.*

Recall that we call a skew-partition ideal if both U_1 and U_2 contain at least $5\theta_{j+4}k$ leaves. Instead of Proposition 3.51, which states that if T has an ideal skew-partition then it can be embedded into G, we will prove a very slightly stronger result which also allows for the possibility that $|U_1| = k/2 + 1$ (so $|U_2| = k/2$). We will need this extra possibility later on.

Proposition 3.60 *Let $V(T) = U_1 + U_2$ where U_2 is an independent set and $|U_1| \leq k/2 + 1$ if k is even and $|U_1| \leq (k+1)/2$ if k is odd. Suppose that both U_1 and U_2 contain at least $5\theta_{j+4}k$ leaves of T. Then $T \subseteq G$.*

Proof. Let W_i be the set of leaves in U_i. Let $\hat{W}_1$ be the set of leaves in U_1 whose parent is in U_2. (Note that the corresponding $\hat{W}_2$ would just be W_2.) If $\hat{W}_1 \leq 4\theta_{j+4}k$, we can move at least $\theta_{j+4}k$ leaves from U_1 to U_2, thus giving a new skew-partition with gap at least $2\theta_{j+4}k \geq 12\theta_{j+2}k$, and we can apply Remark 3.57. So we assume that $|\hat{W}_1| > 4\theta_{j+4}k$.

Let $W_1' := \{v \in \hat{W}_1 : v$ is the only leaf among the children of $P(v)\}$. Note that $|P(W_1')| = |W_1'|$.

Case 1.1: $|W_1'| < 2\theta_{j+4}k$.

Let $W_1'' := \hat{W}_1 \backslash W_1'$, and flip the subforest on $P(W_1'') \cup W_1''$. Note that any new edges within a class come from the vertices of $P(W_1'')$, which now lie in U_1, so U_2 is still independent. Also, since $|P(W_1'')| \leq |W_1''|/2$ and $|W_1''| > 2\theta_{j+4}k$, we now have

$$|U_1| \leq k/2 + 1 - (|W_1''| - |P(W_1'')|) < k/2 - \theta_{j+4}k/2 < (1/2 - \theta_{j+2})k$$

and therefore we can apply Remark 3.57 to obtain $T \subseteq G$.

Case 1.2: $|W_1'| \geq 2\theta_{j+4}k$.

This case is considerably harder. Recall that if $|L_1| < k/2$, then $m_1 \geq 2(k/2 - |L_1|) >$

0. We pick a set $W_1'' \subseteq W_1'$ of size $2\theta_{j+4}k$. Let $W_1'' = \{c_1, c_2, \ldots, c_t\}$, let $P(W_1'') = \{b_1, b_2, \ldots, b_t\}$ and let $P(P(W_1'')) = \{a_1, a_2, \ldots, a_{t'}\}$ (where $t' \le t = 2\theta_{j+4}k$).

Now we would like to use Lemma 3.50 to find a set of at least $2(m_1+1)/3$ stars with $(1/\theta_{j+4})$ endpoints outside V_1 and vertices $x_1, \ldots, x_s$ in L_1 such that we can embed some of the a_i onto x_i, the b_i onto the midpoints of stars (which we call y_i) and c_i and any remaining neighbours onto the endpoints of stars. Since these endpoints are good, we could embed whatever remains of T below these neighbours into the appropriate classes greedily using the minimum degree conditions of (3.7). We would then have embedded at least $2(m_1+1)/3$ vertices of U_1 (namely the c_i) outside V_1, and this would give us enough room in V_1 to embed the remainder of the tree greedily.

However, performing this process naïvely may fail for any one of three reasons.

(1) When attempting to embed the trees $T(b_i)$ outside of V_1, we may inadvertently end up attempting to embed almost all of the tree in some other V_i, thus merely moving our problems to a different class.

(2) Some of the b_i may have degree greater than $1/\theta_{j+4}$, and so the stars guaranteed by Lemma 3.50 are not large enough to fit in all of the neighbours as we would wish to.

(3) Some of the a_i may be at distance two from each other. Thus when we attempt to embed what remains of the tree into V_1 greedily, we may be looking for common neighbours of a large number of vertices. The minimum degree conditions may not be sufficient to guarantee that we can find this.

The first two problems are easy to deal with, but the third is harder, and it is to solve this problem that we introduced the candidate sets C_i in Lemma 3.50.

We define the weight of a vertex x to be $|T(x)|$ (and r has weight $|T| = k+1$). To deal with problem (1), we first note that if some child r' of the the root r of T

satisfies $|T(r')| \geq (k+2)/2$, then we can move the root to r'. Since this can only happen for at most one r', and since with this new root we have $|T(r)| \leq k/2$, we can continue with this process until no child of the root r carries more than half the weight of the tree. In order to ensure that we have not switched T_{odd} and T_{even} we then move the root to a neighbour once more arbitrarily if necessary. Now at most one child and no grandchild of the root carries at least half the weight of the tree.

Thus in particular, unless b_i is the root or this one special child (which can only happen for at most two b_i) $|T(b_i)| \leq (k+1)/2$. By removing at most two b_i from consideration, we assume that every b_i satisfies this property. Since $|W_1''|$ is large, removing two vertices will not affect matters significantly. Now if $|\bigcup_i T(b_i)| \geq 3k/4$, we will simply take a subset of the b_i such that together they carry a weight of between $k/4$ and $3k/4$ (successively remove vertices b_i from consideration until the combined weight is at most $3k/4$, and since the last vertex to be removed had weight at most $(k+1)/2$, the remaining weight is at least $k/4$).

To deal with problem (2) we note that $|P(W_1'')| = |W_1''| = 2\theta_{j+4}k$, and if at least $\theta_{j+4}k+1$ of these vertices have degree more than $1/\theta_{j+4}$, then $|T| \geq (\theta_{j+4}k + 1)/\theta_{j+4} > k+1$ which is a contradiction. So we can take a set $W_1''' \subseteq W_1''$ such that $|W_1'''| = \theta_{j+4}k$ and the vertices of $P(W_1''')$ all have degree at most $1/\theta_{j+4}$. Without loss of generality we will assume that $W_1''' = \{c_1, \ldots, c_{t/2}\}$, $P(W_1''') = \{b_1, \ldots, b_{t/2}\}$ and $P(P(W_1''')) = \{a_1, \ldots, a_{t''}\}$.

We now turn our attention to problem (3). Instead of embedding the a_i, b_i and c_i straight away, we will first embed some preliminary vertices. Let $P_1 \subseteq V(T)$ be the set of vertices which are parents of more than one a_i. Inductively we then define P_i to be the set of vertices which are parents of more than one vertex of P_{i-1}. We observe that $|P_i| \leq |P_{i-1}|/2$ (where we may define P_0 to be $P(P(W_1'''))$, the set of a_i), and so $P := \bigcup_{i\geq 1} P_i$ satisfies $|P| \leq |P(P(W_1'''))| \leq \theta_{j+4}k$. In particular, the process must terminate at some $i = p$, say. We now greedily embed P into V_1 starting with P_p and embedding each P_i in order of decreasing i. Furthermore, if a vertex is in U_1 we will embed it into L_1, and if it is in U_2 we will embed it into $S_1 \backslash B_1$. The fact that $|P| \leq \theta_{j+4}k$ means that the minimum degree conditions of (3.7) applied with $i = 1$ will be more than sufficient. Let $\tilde{P}$ denote the set in V_1 onto which P is embedded.

We now show that we can apply Lemma 3.50. If a vertex a_i of $P(P(W_1'''))$ is not a child of any vertex of P_1, then the candidate set C_i for the corresponding x_i will be $L_1 \backslash \tilde{P}$. If on the other hand a_i is a child of a vertex d_i in P_1, then let $\tilde{d}_i$ be the vertex in $V_1 \backslash B_1$ onto which d_i is embedded. The candidate set C_i in this case will be $(L_1 \cap N(\tilde{d}_i)) \backslash \tilde{P}$. Observe that the minimum degree condition of (3.7) ensures that $|C_i| \geq (1/2 - 2\theta_{j+2})k - |\tilde{P}| \geq (1/2 - 2\theta_{j+4})k$. Thus we may apply Lemma 3.50 to find appropriate x_i and y_i onto which to embed the a_i and b_i. Since the b_i have degree at most $1/\theta_{j+4}$, we may embed the children of the b_i onto endpoints of the stars. Since $|\bigcup_i T(b_i)| \leq 3k/4$ and $\bigcup_i T(b_i)$ is "well-balanced" (by Corollary 3.58 applied to $T' = \{x\} \cup \bigcup_i T(b_i)$ it has a gap of size at most $4\theta_{j+2}k + 1$), and since the endpoints of stars were good, we may then embed the remainder of the $T(b_i)$ greedily outside V_1.

We now embed what is left of T into V_1. Observe that at least $2(m_1 + 1)/3 > m_1/2 \geq k/2 - |L_1|$ vertices of U_1 have been embedded outside L_1. This ensures that L_1 is now big enough to hold the remainder of U_1 (even in the case when $|U_1| = k/2 + 1$ since then k is even and at least $k/2 - |L_1| + 1$ vertices of U_1 have been embedded outside L_1).

Since no vertices of P_i (including $P_0 = P(P(W_1''')))$ are leaves, any leaves that have been embedded have been embedded outside V_1. We delete any remaining leaves from U_2 and from W_1'', and since we have at most $(1/2 - \theta_{j+4})k$ vertices remaining in each of U_1 and U_2 (because $|W_1''| \geq 2\theta_{j+4}k$ and because U_2 had at least $5\theta_{j+4}k$ leaves, but $g(U_1, U_2) \leq 12\theta_{j+2}k \leq \theta_{j+4}k$), we may embed U_1 into L_1 and U_2 into $S_1 \backslash B_1$ greedily. To do this we start at the root and work down the tree, observing that any already embedded vertices have different parents, and so we will only ever have to find an image vertex in the common neighbourhood of at most two vertices during the embedding process (one vertex embedded before the greedy algorithm began, and its grandparent). The minimum degree conditions of (3.7) ensure that the common neighbourhood of two good vertices has size at least $(1/2 - 7\theta_{j+2})k$, which is larger than the number of vertices already embedded, so we can always find an appropriate vertex for the final embedding.

It now remains only to embed those leaves which we deleted. We begin with the leaves deleted from W_1'', which we call U_1', and observe that there are at least as many unused vertices in L_1 as there are leaves in U_1'. In fact, there are also at least $\theta_{j+4}k/2$ unused vertices of L_1, although there may be fewer leaves. We take a subset X of size $t := \max(\theta_{j+4}k/2, |U_1'|)$ of unused vertices in L_1. We also take a subset Y of t vertices from $S_1 \backslash B_1$ consisting of the parents of unembedded leaves in U_1' and some extra vertices chosen arbitrarily if necessary (i.e. if $|U_1'| < \theta_{j+4}k/2$).

We now consider the bipartite subgraph between X and Y, and observe that it has minimum degree at least $|X| - \theta_{j+2}k = |Y| - \theta_{j+2}k \geq |X|/2 = |Y|/2$. So Hall's condition holds, and therefore we can find a perfect matching between X and Y. In particular we can find a matching between the vertices chosen for the parents of U_1' and unused vertices of L_1. This allows us to embed U_1' into L_1, as required.

Finally recall that the leaves deleted from U_2 were adjacent to vertices of U_1, which have been embedded into L_1. Since these vertices are large, we may embed the remaining leaves greedily. This completes the proof of Case 1.2. □

Proof of Proposition 3.52. Recall that we want to prove that if T contains at least $36\theta_{j+4}k$ leaves then either $T \subseteq G$ or T has an ideal skew-partition. Let $g = |T_{odd}| - |T_{even}| \geq 0$ without loss of generality. We start with the skew-partition given by $U_1 = T_{even}$ and $U_2 = T_{odd}$. Note that by Remark 3.57 we may assume that $g \leq 2\theta_{j+4}k$.

Let W_o be the set of leaves of T in T_{odd}, and similarly let W_e be the set of leaves of T in T_{even}. Let $w_o := |W_o|$ and $w_e := |W_e|$. Thus by the assumption of Case 1, $w_o + w_e \geq 36\theta_{j+4}k$. We now split into three further cases:

- Case A: $w_o, w_e \geq 5\theta_{j+4}k$;
- Case B: $w_o < 5\theta_{j+4}k$;
- Case C: $w_e < 5\theta_{j+4}k$.

Case A: $w_o, w_e \geq 5\theta_{j+4}k$. In this case, (T_{even}, T_{odd}) is already an ideal skew-partition, as required.

Case B: $w_o < 5\theta_{j+4}k$. So $w_e \geq 31\theta_{j+4}k$.

Case B (i): $|P(W_e)| \leq 15\theta_{j+4}k$. We flip $P(W_e)$ and W_e. Since we originally chose $U_1 = T_{even}$, $U_2 = T_{odd}$, then with this flip $|U_2|$ increases by at least $31\theta_{j+4}k - 15\theta_{j+4}k$, and so the gap of the partition increases by at least $16\theta_{j+4}k$. Thus we can apply Remark 3.57 to obtain $T \subseteq G$.

Case B (ii): $|P(W_e)| \geq 15\theta_{j+4}k$. We choose $5\theta_{j+4}k$ vertices of $P(W_e)$ (in U_2) and flip these along with those children which are leaves. There are at least $5\theta_{j+4}k$ such children, so now U_2 has at least $5\theta_{j+4}k$ leaves. But at least $10\theta_{j+4}k$ vertices of $P(W_e)$ remained unflipped, and so at least $10\theta_{j+4}k$ leaves of W_e remain in U_1. Note also that $|U_1| \leq |U_2|$, and so we have an ideal skew-partition as required.

Case C: $w_e < 5\theta_{j+4}k$. So $w_o > 31\theta_{j+4}k$.

We apply Fact 3.56 to find a subtree T' of T which contains $[12\theta_{j+4}k, 24\theta_{j+4}k)$ leaves of T, rooted at a vertex x. Let $d = |V(T') \cap T_{odd}| - |V(T') \cap T_{even}|$.

Case C(i): $d \geq g/2$, $x \in T_{even}$.

Let $U_2 = T_{even}$, $U_1 = T_{odd}$ and then flip T'. Since $d \geq g/2$, we now have $|U_1| \leq |U_2|$. Since $x \in T_{even}$ and x was flipped, U_2 is independent, and U_1 has only edges coming from x (to $P(x)$ and $y_1, \dots, y_s$).

Now U_2 contains at least $12\theta_{j+4}k - 5\theta_{j+4}k = 7\theta_{j+4}k$ leaves, and U_1 contains at least $31\theta_{j+4}k - 24\theta_{j+4}k = 7\theta_{j+4}k$ leaves. Thus (U_1, U_2) is an ideal skew-partition.

Case C(ii): $d \leq g/2$ and $x \in T_{odd}$.

Let $U_1 = T_{even}$, $U_2 = T_{odd}$ and flip T'. Similarly to case C(i) this gives an ideal skew-partition.

Case C(iii): $d \leq g/2 - 1$ and $x \in T_{even}$.

Let $U_1 = T_{even}$, $U_2 = T_{odd}$ and flip $T' \backslash \{x\}$ to obtain an ideal skew-partition.

Case C(iv): $d \geq g/2 + 1$ and $x \in T_{odd}$.

Let $U_1 = T_{odd}$, $U_2 = T_{even}$ and flip $T' \backslash \{x\}$ to obtain an ideal skew-partition.

We now only have two special cases left, and these only when g (and therefore $|T| = k + 1$) is odd.

Case C(v): $d = (g - 1)/2$ and $x \in T_{even}$.

Case C(vi): $d = (g + 1)/2$ and $x \in T_{odd}$.

In either case we start with $U_1 = T_{odd}, U_2 = T_{even}$. In Case (v) we flip T' and in Case (vi) we flip $T' \backslash \{x\}$. In either case we obtain (letting W_i denote the set of leaves in U_i)

- $|U_1| = k/2 + 1$, $|U_2| = k/2$;
- U_1, U_2 are independent except for some edges in U_1;
- $|W_2 \cap T'| > 12\theta_{j+4}k - 5\theta_{j+4}k = 7\theta_{j+4}k$;

- $|W_1 \cap (T \backslash T')| > 31\theta_{j+4}k - 24\theta_{j+4}k = 7\theta_{j+4}k$.

But then (U_1, U_2) satisfies the conditions of Proposition 3.60, and so $T \subseteq G$. □

Case 2: T contains at most $36\theta_{j+4}k$ leaves.

Our aim is to "cover" the bad vertices of V_1 first so that we can then apply Lemma 3.44. Consider a maximal set $\mathcal{P}$ of disjoint 2-paths in G with their midpoints in B_1 and their endpoints in L_1. Now if any vertices of B_1 remain uncovered by these paths, then each is adjacent to at most one uncovered vertex of L_1. We delete such vertices from G. Now let L_1' consist of the uncovered neighbours of such deleted vertices in L_1 together with the vertices of L_1 already used as endpoints of 2-paths. The fact that each deleted vertex had only at most one uncovered neighbour in L_1 means that $|L_1'| \leq 2|B_1| \leq 2\theta_{j+2}k$. Therefore setting $L_1'' := L_1 \backslash L_1$ we have $|L_1''| \geq |L_1| - \theta_{j+2}k \geq (1/2 - \theta_{j+4})k$. Thus L_1'' is certainly large enough to be used as a set C_i in Lemma 3.50, which we will want to apply. Note that since we may have deleted some vertices from B_1, some vertices in L_1 may now have degree less than k. However, all such vertices lie in L_1', and it is for this reason that we made Remark 3.59 after the proof of Lemma 3.50.

We split Case 2 further into two subcases:

Case (a): $m_1 \geq 6C$. Thus $m_1/2 + C \leq 2m_1/3$.

In this case we use Claim 3.42 to find a set of at least $m_1\theta_{j+3}/\theta_{j+2}$ disjoint path-segments of length 10 in T. (Note that $m_1 \leq \theta_{j+2}k$, and so $m_1\theta_{j+3}/\theta_{j+2} \leq \theta_{j+3}k$.) In particular, we can take sub-paths of length 8 and ensure that both endpoints of any path-segment are in U_1, and furthermore such that we avoid the vertices of the paths closest to the root. As before, this ensures that for the paths which we choose, the vertices closest to the root do not have a common neighbour.

Now in G, using Proposition 3.50 we can find a set $\mathcal{P}'$ of at least $\lfloor 2(m_1+1)/3 \rfloor \geq 2m_1/3$ disjoint paths of length 2 each with one endvertex in L_1'', the other endvertex being a good vertex of $V \backslash V_1$ and with midpoint also outside V_1. Our candidate sets C_i will all just be L_1''. Whenever we have two paths of $\mathcal{P}'$ with endpoints in the same V_i, we can use the fact that these endpoints are good, together with the minimum degree conditions within V_i, to join them together using vertices in V_i to create a path of length 6 with endpoints in L_1. As long as we have at least C such paths available, we can always find two with their endpoints in the same class. Since $2m_1/3 \geq m_1/2 + C$ we can connect at least $m_1/2$ of the paths and we obtain a set of $m_1/4$ paths of length 6 in G whose endpoints lie in L_1, and for which the rest of the vertices lie outside L_1. Now if $|S_1| < |U_2|$, then we move some (unused) vertices of L_1 to S_1 to ensure that $|S_1| = |U_2|$. We also delete some (unused) vertices from L_1 if necessary to ensure that $|L_1| = |U_1|$. For those original 2-paths whose midpoints were in B_1, we move the endpoints from L_1' into L_1''. Finally we delete all edges within S_1 and L_1. It is then simple to check that the conditions of Lemma 3.44 hold with $\gamma_1 = m_1/(4k)$ and $\gamma_2 = m_1\theta_{j+3}/(\theta_{j+2}k) \gg \gamma_1$, and so we have $T \subseteq G$.

Finally, we consider:

Case (b): $m_1 \leq 6C$.

In this case we use Claim 3.42 to find $\theta_{j+4}k$ disjoint paths on $20C$ vertices in T. We first consider one such path and by taking a subpath P of length $15C$, we may assume that both endpoints lie in U_1. Removing the internal vertices of this path splits the tree T into T_1 and T_2. Without loss of generality we assume that $|T_1| \geq |T_2|$. Let $v = V(P) \cap T_1 \in U_1$, and let w be the neighbour of v on P.

Now any vertex $x \in L_1$ has a neighbour y in some other class V_i (wlog in V_2). We embed v onto x and w onto y. Now if y is small it must have at least one neighbour z in V_2 as well, and we embed the other neighbour of w onto z, which must be large.

We now note that both T_1 and T_2 are "well-balanced" by Corollary 3.58, i.e. $|T_2 \cap U_1|, |T_2 \cap U_2| \leq |T_2|/2 + 2\theta_{j+2}k \leq k/3$. Thus we can easily embed the rest of the path P and T_2 into V_2 using the minimum degree conditions between L_2 and $S_2 \backslash B_2$.

Also, we have now embedded all but one vertex of P outside V_1, and so we have embedded at least $7C > m_1$ vertices of both U_1 and U_2 outside V_1. Together with the 2-paths in $\mathcal{P}$ which we found earlier to cover B_1, this ensures that the conditions of Lemma 3.44 hold, and so we can embed T into G as required. This completes the proof of Theorem 3.46, and therefore also completes the proof of Theorem 3.1.

CHAPTER 4

HYPERGRAPH EMBEDDINGS

4.1 Introduction

The main aim of this chapter is to prove the following.

Theorem 4.1 *For all $\Delta, k \in \mathbb{N}$ there exists a constant $C = C(\Delta, k)$ such that all k-uniform hypergraphs $\mathcal{H}$ of maximum degree at most Δ satisfy $R(\mathcal{H}) \leq C|\mathcal{H}|$.*

The proof given in this chapter also appeared in [18], although here I have added some extra details which were omitted in that paper.

This chapter is organised as follows. In Section 4.2 we give an overview of the proof of Theorem 4.1 and we state the embedding theorem (Theorem 4.2) mentioned above. Our proof of Theorem 4.2 relies on a more general version (Lemma 4.4) of the well-known counting lemma for hypergraphs as well as an 'extension lemma' (Lemma 4.5), whose proofs are postponed until Sections 4.7 and 4.8. We introduce these lemmas, along with further tools, in Section 4.3. We then prove a strengthened version (Theorem 4.3) of Theorem 4.2 in Section 4.4. The regularity lemma for k-uniform hypergraphs is introduced in Section 4.5. In Section 4.6 we deduce Theorem 4.1 from the regularity lemma and Theorem 4.2. In Section 4.7 we derive our version of the counting lemma (Lemma 4.4) from that in [65]. Finally, in Section 4.8

we use it to deduce the extension lemma (Lemma 4.5).

4.2 Overview of the proof of Theorem 4.1 and statement of the embedding theorem

4.2.1 Overview of the proof of Theorem 4.1

The proof in [12] that graphs of bounded degree have linear Ramsey numbers proceeds roughly as follows: Let H be a graph of maximum degree Δ. Take a complete graph K_n, where n is a sufficiently large integer. Colour the edges of K_n with red and blue, and apply the graph regularity lemma to the denser of the two monochromatic graphs, G_{red} say, to obtain a partition of the vertex set into a bounded number of clusters. Since almost all pairs of clusters are regular or 'quasi-random', by Turán's theorem there will be a set of r clusters, where $r := R(K_{\Delta+1})$, in which each pair of clusters is regular. A pair of clusters will be coloured red if its density in G_{red} is at least $1/2$, and blue otherwise. By the definition of r, there must be a set of $\Delta + 1$ clusters such that all the pairs have the same colour. If this colour is red, then one can apply the so-called embedding or key lemma for graphs to find a (red) copy of H in the subgraph of G_{red} spanned by these $\Delta + 1$ clusters. This is possible since $\chi(H) \leq \Delta + 1$. If all the pairs of clusters are coloured blue we apply the embedding theorem in the blue subgraph G_{blue} of K_n to find a blue copy of H. It turns out that in this proof we only needed $n \geq C|H|$, where C is a constant dependent only on Δ. Thus $R(H) \leq C|H|$.

We will generalise this approach to k-uniform hypergraphs. As mentioned in Section 4.1, the main obstacle is the proof of an embedding theorem for k-uniform hypergraphs (Theorem 4.2 below), which allows us to embed a k-uniform hypergraph $\mathcal{H}$ within a suitable 'quasi-random' k-uniform hypergraph $\mathcal{G}$, where the order of $\mathcal{H}$ might be linear in the order of $\mathcal{G}$. Our proof uses ideas from [17].

4.2.2 Notation and statement of the embedding theorem

Before we can state the embedding theorem, we first have to say what we mean by a regular or 'quasi-random' hypergraph. In the setup below, this will involve the relationship between certain i-uniform hypergraphs and $(i-1)$-uniform hypergraphs on the same vertex set. Given a hypergraph $\mathcal{G}$, we write $E(\mathcal{G})$ for the set of its hyperedges and define $e(\mathcal{G}) := |E(\mathcal{G})|$. We write $K_i^{(j)}$ for the complete j-uniform hypergraph on i vertices. Given a j-uniform hypergraph $\mathcal{G}$ and $j \leq i$, we write $\mathcal{K}_i(\mathcal{G})$ for the set of i-sets of vertices of $\mathcal{G}$ which form a copy of $K_i^{(j)}$ in $\mathcal{G}$. Given an i-partite i-uniform hypergraph $\mathcal{G}_i$, and an i-partite $(i-1)$-uniform hypergraph $\mathcal{G}_{i-1}$ on the same vertex set, we define the *density of* $\mathcal{G}_i$ *with respect to* $\mathcal{G}_{i-1}$ to be

$$d(\mathcal{G}_i|\mathcal{G}_{i-1}) := \frac{|\mathcal{K}_i(\mathcal{G}_{i-1}) \cap E(\mathcal{G}_i)|}{|\mathcal{K}_i(\mathcal{G}_{i-1})|}$$

if $|\mathcal{K}_i(\mathcal{G}_{i-1})| > 0$, and $d(\mathcal{G}_i|\mathcal{G}_{i-1}) := 0$ otherwise. More generally, if $\mathbf{Q} := (Q(1), Q(2), \dots, Q(r))$ is a collection of r subhypergraphs of $\mathcal{G}_{i-1}$, we define $\mathcal{K}_i(\mathbf{Q}) := \bigcup_{j=1}^{r} \mathcal{K}_i(Q(j))$ and

$$d(\mathcal{G}_i|\mathbf{Q}) := \frac{|\mathcal{K}_i(\mathbf{Q}) \cap E(\mathcal{G}_i)|}{|\mathcal{K}_i(\mathbf{Q})|}$$

if $|\mathcal{K}_i(\mathbf{Q})| > 0$, and $d(\mathcal{G}_i|\mathbf{Q}) := 0$ otherwise. We sometimes write $|K_i^{(i-1)}|_{\mathbf{Q}}$ instead of $|\mathcal{K}_i(\mathbf{Q})|$.

We say that $\mathcal{G}_i$ is (d_i, δ, r)*-regular with respect to* $\mathcal{G}_{i-1}$ if every r-tuple $\mathbf{Q}$ with $|\mathcal{K}_i(\mathbf{Q})| > \delta|\mathcal{K}_i(\mathcal{G}_{i-1})|$ satisfies

$$d(\mathcal{G}_i|\mathbf{Q}) = d_i \pm \delta.$$

Given $\ell \geq i \geq 3$, an ℓ-partite i-uniform hypergraph $\mathcal{G}_i$ and an ℓ-partite $(i-1)$-uniform hypergraph $\mathcal{G}_{i-1}$ on the same vertex set, we say that $\mathcal{G}_i$ is (d_i, δ, r)-regular with respect to $\mathcal{G}_{i-1}$ if for every i-tuple K of vertex classes, either $\mathcal{G}_i[K]$ is (d_i, δ, r)-regular with respect to $\mathcal{G}_{i-1}[K]$ or $d(\mathcal{G}_i[K]|\mathcal{G}_{i-1}[K]) = 0$ (but the latter should not hold for all K). Instead of $(d_i, \delta, 1)$-regularity we sometimes refer to (d_i, δ)-*regularity.*

Recall from Chapter 1 that the density of a bipartite graph G with vertex classes A and B is defined by $d(A, B) := e(A, B)/|A||B|$ and G is (d, δ)-regular if for all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \delta|A|$ and $|Y| \geq \delta|B|$ we have $d(X, Y) = d \pm \delta$. We say that an ℓ-partite graph $\mathcal{G}_2$ is (d_2, δ)-*regular* if each of the $\binom{\ell}{2}$ bipartite subgraphs forming it is either (d_2, δ)-regular or has density 0 (and if for at least one of them the former holds).

Suppose that we have $\ell \geq k$ vertex classes $V_1, \dots, V_\ell$, and that for each $i = 2, \dots, k$ we are given an ℓ-partite i-uniform hypergraph $\mathcal{G}_i$ with these vertex classes. Suppose also that $\mathcal{H}$ is an ℓ-partite k-uniform hypergraph with vertex classes $X_1, \dots, X_\ell$. We will aim to embed $\mathcal{H}$ into $\mathcal{G}_k$, and in particular to embed X_j into V_j for each $j = 1, \dots, \ell$. So we make the following definition: We say that $(\mathcal{G}_k, \dots, \mathcal{G}_2)$ *respects the partition of* $\mathcal{H}$ if whenever $\mathcal{H}$ contains a hyperedge with vertices in $X_{j_1}, \dots, X_{j_k}$, then there is a hyperedge of $\mathcal{G}_k$ with vertices in $V_{j_1}, \dots, V_{j_k}$ which also forms a copy of $K_k^{(i)}$ in $\mathcal{G}_i$ for each $i = 2, \dots, k-1$.

The following is our main tool in the proof of Theorem 4.1. It states that provided an ℓ-partite hypergraph $\mathcal{G}$ satisfies certain suitable regularity conditions, then we can embed in $\mathcal{G}$ any bounded degree ℓ-partite hypergraph $\mathcal{H}$, provided that the partition classes of $\mathcal{H}$ are not too large compared with those of $\mathcal{G}$. In particular, we allow for the classes of $\mathcal{H}$ to have up to a small linear size compared with those of $\mathcal{G}$.

Theorem 4.2 (Embedding theorem for hypergraphs) *Let* Δ, k, ℓ, r, n_0 *be pos-*

itive integers with $k \le \ell$ *and let* $c, d_2, d_3, \dots, d_k, \delta, \delta_k$ *be positive constants such that* $1/d_i \in \mathbb{N}$,

$$1/n_0 \ll 1/r, \delta \ll \min\{\delta_k, d_2, \dots, d_{k-1}\} \le \delta_k \ll d_k, 1/\Delta, 1/\ell$$

and

$$c \ll d_2, \dots, d_k, 1/\Delta, 1/\ell.$$

Then the following holds for all integers $n \ge n_0$. *Suppose that* $\mathcal{H}$ *is an* ℓ*-partite* k*-uniform hypergraph of maximum degree at most* Δ *with vertex classes* $X_1, \dots, X_\ell$ *such that* $|X_i| \le cn$ *for all* $i = 1, \dots, \ell$. *Suppose that for each* $i = 2, \dots, k$, $\mathcal{G}_i$ *is an* ℓ*-partite* i*-uniform hypergraph with vertex classes* $V_1, \dots, V_\ell$, *which all have size* n. *Suppose also that* $\mathcal{G}_k$ *is* (d_k, δ_k, r)*-regular with respect to* $\mathcal{G}_{k-1}$, *that for each* $i = 3, \dots, k-1$, $\mathcal{G}_i$ *is* (d_i, δ)*-regular with respect to* $\mathcal{G}_{i-1}$, *that* $\mathcal{G}_2$ *is* (d_2, δ)*-regular, and that* $(\mathcal{G}_k, \dots, \mathcal{G}_2)$ *respects the partition of* $\mathcal{H}$. *Then* $\mathcal{G}_k$ *contains a copy of* $\mathcal{H}$.

4.3 Further notation and tools

4.3.1 Embedding theorem for complexes

Instead of Theorem 4.2, we will prove a considerably stronger version which appears as Theorem 4.3 below. It allows the embedding of hypergraphs which are not necessarily uniform and gives a lower bound on the number of such embeddings. This enables us to prove the lemma by induction on $|\mathcal{H}|$. Before we can state Theorem 4.3, we need to make the following definitions.

A *complex* $\mathcal{H}$ on a vertex set V is a collection of subsets of V, each of size at least 2, such that if $B \in \mathcal{H}$, and if $A \subseteq B$ has size at least 2, then $A \in \mathcal{H}$. (So if we add each vertex in V as a singleton into a complex, we obtain a downset.) A

k-complex is a complex in which no member has size greater than k. The members of size $i \geq 2$ are called the *i-edges* of $\mathcal{H}$ and the elements of V are called the *vertices* of $\mathcal{H}$. We write $E_i(\mathcal{H})$ for the set of all i-edges of $\mathcal{H}$ and set $e_i(\mathcal{H}) := |E_i(\mathcal{H})|$. We also write $|\mathcal{H}| := |V|$ for the *order* of $\mathcal{H}$. Note that a k-uniform hypergraph can be turned into a k-complex by making every hyperedge into a complete i-uniform hypergraph $K_k^{(i)}$, for each $2 \leq i \leq k$. (In a more general k-complex we may have i-edges which do not lie within an $(i+1)$-edge.) Given $k \leq \ell$, a *(k, ℓ)-complex* is an ℓ-partite k-complex. Given a k-complex $\mathcal{H}$, for each $i = 2, \dots, k$ we denote by $\mathcal{H}_i$ the *underlying i-uniform hypergraph of* $\mathcal{H}$. So the vertices of $\mathcal{H}_i$ are those of $\mathcal{H}$ and the hyperedges of $\mathcal{H}_i$ are the i-edges of $\mathcal{H}$.

Two vertices x and y in a k-complex are *neighbours* if they are joined by a 2-edge. (Note that if x and y lie in a common i-edge for some $2 \leq i \leq k$, then they are joined by a 2-edge.) The *degree* $d(x)$ of a vertex x is the maximum (over $2 \leq i \leq k$) of the number of i-edges containing x. Thus x has at most $d(x)$ neighbours. The *maximum degree* of the complex $\mathcal{H}$ is the greatest degree of any vertex. Note that if $\mathcal{H}$ is a k-uniform hypergraph of maximum degree Δ, the maximum degree of the corresponding k-complex is crudely at most $\Delta 2^k$. The *distance* between two vertices x and y in a k-complex $\mathcal{H}$ is the length of the shortest path between x and y in the underlying 2-graph $\mathcal{H}_2$ of $\mathcal{H}$. The *components* of $\mathcal{H}$ are the subcomplexes induced by the components of $\mathcal{H}_2$.

We say that a k-complex $\mathcal{G}$ is $(d_k, \dots, d_2, \delta_k, \delta, r)$-regular if $\mathcal{G}_k$ is (d_k, δ_k, r)-regular with respect to $\mathcal{G}_{k-1}$, if $\mathcal{G}_i$ is (d_i, δ)-regular with respect to $\mathcal{G}_{i-1}$ for each $i = 3, \dots, k-1$, and if $\mathcal{G}_2$ is (d_2, δ)-regular. We denote $(d_k, \dots, d_2)$ by $\mathbf{d}$ and refer to $(\mathbf{d}, \delta_k, \delta, r)$-regularity.

Suppose that $\mathcal{G}$ is a (k, ℓ)-complex with vertex classes $V_1, \dots, V_\ell$, which all have size n. Suppose also that $\mathcal{H}$ is a (k, ℓ)-complex with vertex classes $X_1, \dots, X_\ell$ of size at most n. Similarly as for hypergraphs we say that $\mathcal{G}$ *respects the partition of* $\mathcal{H}$ if whenever $\mathcal{H}$ contains an i-edge with vertices in $X_{j_1}, \ldots, X_{j_i}$, then there is an i-edge of $\mathcal{G}$ with vertices in $V_{j_1}, \ldots, V_{j_i}$. On the other hand, we say that a labelled copy of $\mathcal{H}$ in $\mathcal{G}$ is *partition-respecting* if for each $i = 1, \ldots, \ell$ the vertices corresponding to those in X_i lie within V_i. We denote by $|\mathcal{H}|_{\mathcal{G}}$ the number of labelled, partition-respecting copies of $\mathcal{H}$ in $\mathcal{G}$.

Theorem 4.3 (Embedding theorem for complexes) *Let Δ, k, ℓ, r, n_0 be positive integers and let $c, \alpha, d_2, \ldots, d_k, \delta, \delta_k$ be positive constants such that $1/d_i \in \mathbb{N}$,*

$$1/n_0 \ll 1/r, \delta \ll \min\{\delta_k, d_2, \ldots, d_{k-1}\} \leq \delta_k \ll \alpha \ll d_k, 1/\Delta, 1/\ell$$

and

$$c \ll \alpha, d_2, \ldots, d_k.$$

Then the following holds for all integers $n \geq n_0$. Suppose that $\mathcal{H}$ is a (k, ℓ)-complex of maximum degree at most Δ with vertex classes $X_1, \ldots, X_\ell$ such that $|X_i| \leq cn$ for all $i = 1, \ldots, \ell$. Suppose also that $\mathcal{G}$ is a $(\mathbf{d}, \delta_k, \delta, r)$-regular (k, ℓ)-complex with vertex classes $V_1, \ldots, V_\ell$, all of size n, which respects the partition of $\mathcal{H}$. Then for every vertex h of $\mathcal{H}$ we have that

$$|\mathcal{H}|_{\mathcal{G}} \geq (1-\alpha)n \left(\prod_{i=2}^{k} d_i^{e_i(\mathcal{H})-e_i(\mathcal{H}_h)} \right) |\mathcal{H}_h|_{\mathcal{G}},$$

where $\mathcal{H}_h$ denotes the induced subcomplex of $\mathcal{H}$ obtained by removing h. In particular, $\mathcal{G}$ contains at least $((1-\alpha)n)^{|\mathcal{H}|} \prod_{i=2}^{k} d_i^{e_i(\mathcal{H})}$ labelled partition-respecting copies of $\mathcal{H}$.

As discussed in the next subsection, Theorem 4.3 is a generalization of the hypergraph counting lemma (which counts subcomplexes $\mathcal{H}$ of bounded order) to subcomplexes $\mathcal{H}$ of bounded degree and linear order. Note that the bound relating $|\mathcal{H}|_{\mathcal{G}}$ to $|\mathcal{H}_h|_{\mathcal{G}}$ in Theorem 4.3 is close to what one would get with high probability if $\mathcal{G}$ were a random complex[1]. This also shows that the bound is close to best possible. Theorem 4.3 will be proved in Section 4.4. In the proof we will need two lemmas on embeddings of complexes of bounded order, which are stated in the next subsection.

Recall that if the maximum degree of a k-uniform hypergraph $\mathcal{H}$ is at most Δ then the maximum degree of the corresponding k-complex is at most $\Delta 2^k$. So it is easy to see that Theorem 4.3 does indeed imply Theorem 4.2.

4.3.2 Counting lemma and extension lemma

We will need a variant (Lemma 4.4) of the counting lemma for k-unifom hypergraphs due to Rödl and Schacht [65, Thm 9]. (A similar result was proved earlier by Gowers [31] as well as Nagle, Rödl and Schacht [60].) It states that if $|\mathcal{H}|$ is bounded and $\mathcal{G}$ is suitably regular, then the number of copies of $\mathcal{H}$ in $\mathcal{G}$ is as large as one would expect if $\mathcal{G}$ were random. The main difference to the result in [65] is that Lemma 4.4 counts copies of k-complexes $\mathcal{H}$ instead of copies of k-uniform hypergraphs $\mathcal{H}$ and also includes an upper bound on the number of these copies. We will derive Lemma 4.4 from the result in [65] in Section 4.7.

Lemma 4.4 (Counting lemma) *Let k, ℓ, r, t, n_0 be positive integers and let $\varepsilon, d_2, \dots, d_k, \delta, \delta_k$ be positive constants such that $1/d_i \in \mathbb{N}$ and*

$$1/n_0 \ll 1/r, \delta \ll \min\{\delta_k, d_2, \dots, d_{k-1}\} \leq \delta_k \ll \varepsilon, d_k, 1/\ell, 1/t.$$

Then the following holds for all integers $n \geq n_0$. Suppose that $\mathcal{H}$ is a (k, ℓ)-complex on t vertices with vertex classes $X_1, \dots, X_\ell$. Suppose also that $\mathcal{G}$ is a $(\mathbf{d}, \delta_k, \delta, r)$-

[1] That is, $\mathcal{G}_2$ is an ℓ-partite random graph with density d_2, each triangle of $\mathcal{G}_2$ is an edge of $\mathcal{G}_3$ with probability d_3 etc.

regular (k,ℓ)-complex with vertex classes $V_1,\dots,V_\ell$, all of size n, which respects the partition of $\mathcal{H}$. Then

$$|\mathcal{H}|_{\mathcal{G}} = (1\pm\varepsilon)n^t\prod_{i=2}^{k} d_i^{e_i(\mathcal{H})}.$$

The main difference between the counting lemma and Theorem 4.3 is that the counting lemma only allows for complexes $\mathcal{H}$ of bounded order. We will apply the counting lemma to embed complexes of order $\leq f(\Delta,k)$ for some appropriate function f. Note that the upper and lower bounds of the counting lemma imply Theorem 4.3 for the case when $|\mathcal{H}|$ is bounded. A formal proof of this (which settles the base case for the induction in the proof of Theorem 4.3) can be found at the beginning of Section 4.4.

In the induction step of the proof of Theorem 4.3 we will also need the following extension lemma, which states that if $\mathcal{H}'$ is a complex of bounded order, $\mathcal{H}\subseteq\mathcal{H}'$ is an induced subcomplex and $\mathcal{G}$ is suitably regular, then almost all copies of $\mathcal{H}$ in $\mathcal{G}$ can be extended to about the 'right' number of copies of $\mathcal{H}'$, where the 'right' number is the number one would expect if $\mathcal{G}$ were random. We will derive Lemma 4.5 from Lemma 4.4 in Section 4.8.

Lemma 4.5 (Extension lemma) *Let k,ℓ,r,t,t',n_0 be positive integers, where $t<t'$, and let $\beta,d_2,\dots,d_k,\delta,\delta_k$ be positive constants such that $1/d_i\in\mathbb{N}$ and*

$$1/n_0 \ll 1/r,\delta \ll \min\{\delta_k,d_2,\dots,d_{k-1}\} \leq \delta_k \ll \beta,d_k,1/\ell,1/t'.$$

Then the following holds for all integers $n\geq n_0$. Suppose that $\mathcal{H}'$ is a (k,ℓ)-complex on t' vertices with vertex classes $X_1,\dots,X_\ell$ and let $\mathcal{H}$ be an induced subcomplex of $\mathcal{H}'$ on t vertices. Suppose also that $\mathcal{G}$ is a $(\mathbf{d},\delta_k,\delta,r)$-regular (k,ℓ)-complex with vertex classes $V_1,\dots,V_\ell$, all of size n, which respects the partition of $\mathcal{H}'$. Then all *but at most $\beta|\mathcal{H}|_{\mathcal{G}}$ labelled partition-respecting copies of $\mathcal{H}$ in $\mathcal{G}$ are extendible to*

$$(1 \pm \beta) n^{t'-t} \prod_{i=2}^{k} d_i^{e_i(\mathcal{H}')-e_i(\mathcal{H})}$$

labelled partition-respecting copies of $\mathcal{H}'$ in $\mathcal{G}$.

As well as these versions of the counting lemma and extension lemma, we will need to be able to apply versions of these lemmas to underlying $(k-1)$-complexes. In this case, we have that the regularity constant δ is much smaller than all the densities $d_2, \dots, d_{k-1}$, but on the other hand we have no r in the highest level and thus we cannot apply Lemmas 4.4 and 4.5. So instead of Lemma 4.4 we will use the following variant of a result of Kohayakawa, Rödl and Skokan [44, Cor. 6.11].

Lemma 4.6 (Dense counting lemma) *Let k, ℓ, t, n_0 be positive integers and let $\varepsilon, d_2, \dots, d_{k-1}, \delta$ be positive constants such that*

$$1/n_0 \ll \delta \ll \varepsilon \ll d_2, \dots, d_{k-1}, 1/\ell, 1/t.$$

Then the following holds for all integers $n \geq n_0$. Suppose that $\mathcal{H}$ is a $(k-1, \ell)$-complex on t vertices with vertex classes $X_1, \dots, X_\ell$. Suppose also that $\mathcal{G}$ is a $(d_{k-1}, \dots, d_2, \delta, \delta, 1)$-regular $(k-1, \ell)$-complex with vertex classes $V_1, \dots, V_\ell$, all of size n, which respects the partition of $\mathcal{H}$. Then

$$|\mathcal{H}|_{\mathcal{G}} = (1 \pm \varepsilon) n^t \prod_{i=2}^{k-1} d_i^{e_i(\mathcal{H})}.$$

In Section 4.7 we will show how Lemma 4.6 can be deduced from the result in [44]. The following dense version of the extension lemma can be deduced from the dense counting lemma (see Section 4.8).

Lemma 4.7 (Dense extension lemma) *Let k, ℓ, t, t', n_0 be positive integers and let $\beta, d_2, \ldots, d_{k-1}, \delta$ be positive constants such that*

$$1/n_0 \ll \delta \ll \beta \ll d_2, \ldots, d_{k-1}, 1/\ell, 1/t'.$$

Then the following holds for all integers $n \geq n_0$. Suppose that $\mathcal{H}'$ is a $(k-1, \ell)$-complex on t' vertices with vertex classes $X_1, \ldots, X_\ell$ and let $\mathcal{H}$ be an induced subcomplex of $\mathcal{H}'$ on t vertices. Suppose also that $\mathcal{G}$ is a $(d_{k-1}, \ldots, d_2, \delta, \delta, 1)$-regular $(k-1, \ell)$-complex with vertex classes $V_1, \ldots, V_\ell$, all of size n, which respects the partition of $\mathcal{H}'$. Then all but at most $\beta|\mathcal{H}|_{\mathcal{G}}$ labelled partition-respecting copies of $\mathcal{H}$ in $\mathcal{G}$ can be extended into

$$(1 \pm \beta) n^{|\mathcal{H}'| - |\mathcal{H}|} \prod_{i=2}^{k-1} d_i^{e_i(\mathcal{H}') - e_i(\mathcal{H})}$$

labelled partition-respecting copies of $\mathcal{H}'$ in $\mathcal{G}$.

An overview of how all these lemmas are used in the proof of Theorem 4.1 is shown in Figure 4.1.

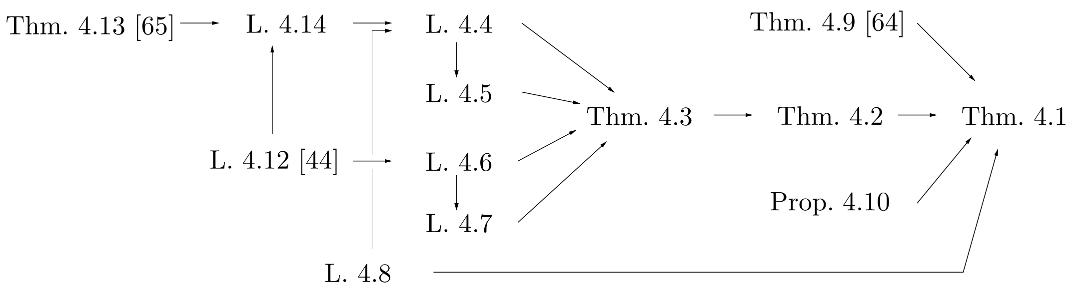


Figure 4.1: Proof of Theorem 4.1 - Flowchart

Another auxiliary result that we will use in the proof of Lemma 4.4 as well as in the proof of Theorem 4.1 is the slicing lemma. Roughly speaking, this says that in a regular complex $\mathcal{G}$, we can partition the edge set $E_j(\mathcal{G})$ of the jth level into an arbitrary number of parts so that each part is still regular with respect to $\mathcal{G}_{j-1}$ with the appropriate density, at the expense of a larger regularity constant. This can be proved using a simple application of a Chernoff bound.

Lemma 4.8 (Slicing lemma [64]) *Let $j \geq 2$ and $s_0, r \geq 1$ be integers and let δ_0, d_0 and p_0 be positive real numbers. Then there is an integer $n_0 = n_0(j, s_0, r, \delta_0, d_0, p_0)$ such that the following holds. Let $n \geq n_0$ and let $\mathcal{G}_j$ be a j-partite j-uniform hypergraph with vertex classes $V_1, \dots, V_j$ which all have size n. Also let $\mathcal{G}_{j-1}$ be a j-partite $(j-1)$-uniform hypergraph with the same vertex classes and assume that each j-set of vertices that spans a hyperedge in $\mathcal{G}_j$ also spans a $K_j^{(j-1)}$ in $\mathcal{G}_{j-1}$. Suppose that*

1. $|K_j^{(j-1)}(\mathcal{G}_{j-1})| > n^j / \ln n$ *and*

2. $\mathcal{G}_j$ *is (d, δ, r)-regular with respect to $\mathcal{G}_{j-1}$, where $d \geq d_0 \geq 2\delta \geq 2\delta_0$.*

Then for any positive integer $s \leq s_0$ and all positive reals $p_1, \dots, p_s \geq p_0$ with $\sum_{i=1}^s p_i \leq 1$ there exists a partition of $E(\mathcal{G}_j)$ into $s+1$ parts $E^{(0)}(\mathcal{G}_j), E^{(1)}(\mathcal{G}_j), \dots, E^{(s)}(\mathcal{G}_j)$ such that if $\mathcal{G}_j(i)$ denotes the spanning subhypergraph of $\mathcal{G}_j$ whose edge set is $E^{(i)}(\mathcal{G}_j)$, then $\mathcal{G}_j(i)$ is $(p_i d, 3\delta, r)$-regular with respect to $\mathcal{G}_{j-1}$ for every $i = 1, \dots, s$.

Moreover, $\mathcal{G}_j(0)$ is $((1 - \sum_{i=1}^s p_i)\, d, 3\delta, r)$-regular with respect to $\mathcal{G}_{j-1}$ and $E^{(0)}(\mathcal{G}_j) = \emptyset$ if $\sum_{i=1}^s p_i = 1$.

4.4 Proof of the embedding theorem for complexes (Theorem 4.3)

Throughout the rest of the chapter, whenever we talk about a copy of a complex $\mathcal{H}$ in $\mathcal{G}$ we mean that this copy is labelled and partition-respecting, without mentioning this explicitly. We prove Theorem 4.3 by induction on $|\mathcal{H}|$. [17] contains a sketch of the argument for the graph case which gives a good idea of the proof. We first suppose that the connected component of $\mathcal{H}$ which contains the vertex h has order less than Δ^5. In this case we will use the counting lemma to prove the embedding theorem. So let $\mathcal{C}$ be the component of $\mathcal{H}$ containing h, and let $\mathcal{D} := \mathcal{H} - \mathcal{C}$. Also, let $\mathcal{C}_h := \mathcal{C} - h$. We may assume that both $\mathcal{C}_h$ and $\mathcal{D}$ are non-empty. (If $\mathcal{D}$ is empty then the result follows from Lemma 4.5, and if $\mathcal{C}_h$ is empty then h is an isolated vertex and the result is trivial.) Note that a copy of $\mathcal{H}$ consists of disjoint copies of $\mathcal{C}$ and $\mathcal{D}$, while $\mathcal{H}_h$ consists of disjoint copies of $\mathcal{C}_h$ and $\mathcal{D}$. Copies of these complexes in $\mathcal{G}$ will be denoted by C, D and C_h.

Choose a new constant β such that $c, \delta_k \ll \beta \ll \alpha$. Now note that $|\mathcal{H}|_{\mathcal{G}} = \sum_{D \subseteq \mathcal{G}} |\mathcal{C}|_{\mathcal{G}-D}$, and by applying the upper and lower bounds of the counting lemma (Lemma 4.4) to copies of $\mathcal{C}$ in $\mathcal{G}$ and $\mathcal{G} - D$ respectively, we obtain $|\mathcal{C}|_{\mathcal{G}-D} \geq \frac{(1-c)^{\Delta^5}(1-\beta)}{(1+\beta)} |\mathcal{C}|_{\mathcal{G}} \geq (1-3\beta)|\mathcal{C}|_{\mathcal{G}}$. So

$$|\mathcal{H}|_{\mathcal{G}} \geq \sum_{D \subseteq \mathcal{G}} (1-3\beta)|\mathcal{C}|_{\mathcal{G}} = (1-3\beta)|\mathcal{C}|_{\mathcal{G}}|\mathcal{D}|_{\mathcal{G}}. \tag{4.1}$$

On the other hand, by a similar argument using the upper and lower bounds from the counting lemma in $\mathcal{G}$ for $\mathcal{C}_h$ and $\mathcal{C}$ respectively,

$$|\mathcal{H}_h|_{\mathcal{G}} \leq |\mathcal{C}_h|_{\mathcal{G}}|\mathcal{D}|_{\mathcal{G}} \leq \frac{1+\beta}{1-\beta} \frac{|\mathcal{C}|_{\mathcal{G}}|\mathcal{D}|_{\mathcal{G}}}{n \prod_{i=2}^{k} d_i^{e_i(\mathcal{C})-e_i(\mathcal{C}_h)}}. \tag{4.2}$$

Combining (4.1) and (4.2) gives the desired result.

Thus we may assume that the component of $\mathcal{H}$ containing h has order at least Δ^5. This deals with the base case of the inductive argument, and it also means that the fourth neighbourhood of h in $\mathcal{H}$ will be non-empty, which will be convenient later on in the proof as we will only be counting complexes which are non-empty.

We pick new constants ε_k and ε_{k-1} satisfying the following hierarchies:

$$\delta \ll \varepsilon_{k-1} \ll d_2, d_3, \ldots, d_k, 1/\Delta,$$

$$c, \delta_k, \varepsilon_{k-1} \ll \varepsilon_k \ll \alpha.$$

Let $\mathcal{N}_h$ be the subcomplex of $\mathcal{H}$ induced by the neighbours of h, and let $\mathcal{B}$ be the subcomplex of $\mathcal{H}$ induced by h and the neighbours of h. Then any copy of $\mathcal{H}$ in $\mathcal{G}$ extending a copy N_h of $\mathcal{N}_h$ can be obtained by first extending N_h into a copy of $\mathcal{H}_h$ and then extending N_h into a copy of $\mathcal{B}$, where the vertex chosen for h has to be distinct from all the vertices chosen for $\mathcal{H}_h$. In particular,

$$|\mathcal{H}|_{\mathcal{G}} \geq \sum_{N_h \subseteq \mathcal{G}} |N_h \to \mathcal{H}_h|(|N_h \to \mathcal{B}| - cn). \tag{4.3}$$

We now introduce some more notation. Given k-complexes $\mathcal{H}' \subseteq \mathcal{H}''$ such that $\mathcal{H}'$ is induced, and a copy H' of $\mathcal{H}'$ in $\mathcal{G}$, we define $|H' \to \mathcal{H}''|$ to be the number of ways in which H' can be extended to a copy of $\mathcal{H}''$ in $\mathcal{G}$. We also define

$$\overline{|\mathcal{H}' \to \mathcal{H}''|} := n^{|\mathcal{H}''|-|\mathcal{H}'|} \prod_{i=2}^{k} d_i^{e_i(\mathcal{H}'')-e_i(\mathcal{H}')}.$$

Thus $\overline{|\mathcal{H}' \to \mathcal{H}''|}$ is roughly the expected number of ways H' could be extended to a copy of $\mathcal{H}''$ if $\mathcal{G}$ were a random complex.

We define a copy N_h of $\mathcal{N}_h$ to be *typical* if it has about the correct number of extensions into $\mathcal{B}$, i.e. if $|N_h \to \mathcal{B}| = (1 \pm \varepsilon_k)\overline{|\mathcal{N}_h \to \mathcal{B}|}$. An application of the extension lemma (Lemma 4.5) shows that at most $\varepsilon_k |\mathcal{N}_h|_{\mathcal{G}}$ copies of $\mathcal{N}_h$ in $\mathcal{G}$ are not typical. We denote the set of typical copies of $\mathcal{N}_h$ by `typ`, and the set of all atypical copies by `atyp`.

Now observe that if all of the copies of $\mathcal{N}_h$ were typical, the proof would be complete, since then

$$\begin{aligned}|\mathcal{H}|_{\mathcal{G}} &\overset{4.3}{\geq} \sum_{N_h \subseteq \mathcal{G}} |N_h \to \mathcal{H}_h|(|N_h \to \mathcal{B}| - cn) \\ &\geq \Big((1-\varepsilon_k)\overline{|\mathcal{N}_h \to \mathcal{B}|} - cn\Big) \sum_{N_h \subseteq \mathcal{G}} |N_h \to \mathcal{H}_h| \\ &\geq (1-\alpha)\overline{|\mathcal{N}_h \to \mathcal{B}|}|\mathcal{H}_h|_{\mathcal{G}} = (1-\alpha)\overline{|\mathcal{H}_h \to \mathcal{H}|}|\mathcal{H}_h|_{\mathcal{G}}. \end{aligned} \tag{4.4}$$

The third inequality follows since $c \ll \alpha, d_2, \dots, d_k$, and $\varepsilon_k \ll \alpha$.

However, we also need to take account of the atypical copies of $\mathcal{N}_h$. The proportion of these is about ε_k, which may be larger than some d_i. It will turn out that this is too large for our purposes, and so we will need to consider the atypical copies more carefully.

We define, instead of $|H' \to \mathcal{H}''|$, the expression $|H' \overset{k-1}{\to} \mathcal{H}''|$, where $\mathcal{H}' \subseteq \mathcal{H}''$ are induced subcomplexes of $\mathcal{H}$ and H' is a copy of $\mathcal{H}'$ in $\mathcal{G}$. We consider the underlying $(k-1)$-complexes in each case, and define $|H' \overset{k-1}{\to} \mathcal{H}''|$ to be the number of ways in which the underlying $(k-1)$-complex of H' can be extended to the underlying $(k-1)$-complex of $\mathcal{H}''$ within (the underlying $(k-1)$-complex of) $\mathcal{G}$. Clearly $|H' \overset{k-1}{\to} \mathcal{H}''| \geq |H' \to \mathcal{H}''|$. We also define

$$\overline{|\mathcal{H}' \overset{k-1}{\to} \mathcal{H}''|} := n^{|\mathcal{H}''|-|\mathcal{H}'|} \prod_{i=2}^{k-1} d_i^{e_i(\mathcal{H}'')-e_i(\mathcal{H}')}.$$

Thus $\overline{|\mathcal{H}' \overset{k-1}{\to} \mathcal{H}''|}$ is roughly the expected value of $|H' \overset{k-1}{\to} \mathcal{H}''|$ if $\mathcal{G}$ were a random complex. Also,

$$\overline{|\mathcal{H}' \overset{k-1}{\to} \mathcal{H}''|} = \overline{|\mathcal{H}' \to \mathcal{H}''|}/d_k^{e_k(\mathcal{H}'')-e_k(\mathcal{H}')} \geq \overline{|\mathcal{H}' \to \mathcal{H}''|}.$$

We define $\mathcal{N}_h^*$ to be the subcomplex of $\mathcal{H}$ induced by the vertices at distance 3 from h. We also define $\mathcal{F}$ to be the subcomplex of $\mathcal{H}$ induced by the vertices at distance $1, 2$ or 3 from h, i.e. the subcomplex induced by $\mathcal{N}_h$, $\mathcal{N}_h^*$ and the vertices in between (see Figure 4.2).

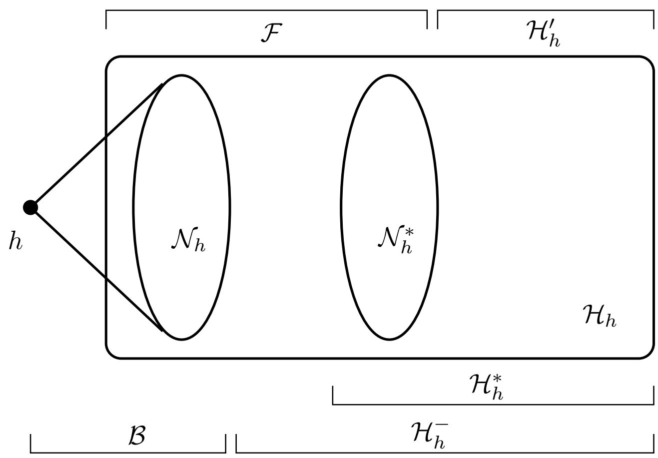


Figure 4.2: The complex $\mathcal{H}$

Given copies N_h of $\mathcal{N}_h$ and N_h^* of $\mathcal{N}_h^*$, we say that the pair N_h, N_h^* is *useful* if N_h and N_h^* are disjoint and if the pair has about the expected number of extensions into copies of $\mathcal{F}$ as $(k-1)$-complexes, i.e. if

$$|N_h \cup N_h^* \stackrel{k-1}{\rightarrow} \mathcal{F}| = (1 \pm \varepsilon_{k-1})\overline{|\mathcal{N}_h \cup \mathcal{N}_h^* \stackrel{k-1}{\rightarrow} \mathcal{F}|}.$$

We use Lemmas 4.4, 4.6 and 4.7 applied to $\mathcal{N}_h \cup \mathcal{N}_h^*$ to show that at most $\sqrt{\varepsilon_{k-1}}|\mathcal{N}_h|_{\mathcal{G}}|\mathcal{N}_h^*|_{\mathcal{G}}$ disjoint pairs N_h, N_h^* are not useful. Let $|\mathcal{N}_h \cup \mathcal{N}_h^*|_{\mathcal{G}}^{(k-1)}$ denote the number of copies of the underlying $(k-1)$-complex of $\mathcal{N}_h \cup \mathcal{N}_h^*$ in $\mathcal{G}$. Then Lemmas 4.4 and 4.6 together imply that $|\mathcal{N}_h \cup \mathcal{N}_h^*|_{\mathcal{G}}^{(k-1)} \le (1+2\varepsilon_k)|\mathcal{N}_h \cup \mathcal{N}_h^*|_{\mathcal{G}}/d_k^{e_k(\mathcal{N}_h \cup \mathcal{N}_h^*)}$. Moreover, the dense extension lemma (Lemma 4.7) shows that all but at most $\varepsilon_{k-1}|\mathcal{N}_h \cup \mathcal{N}_h^*|_{\mathcal{G}}^{(k-1)}$ copies of the underlying $(k-1)$-complex of $\mathcal{N}_h \cup \mathcal{N}_h^*$ in $\mathcal{G}$ are useful. Altogether this shows that all but at most

$$\varepsilon_{k-1}(1+2\varepsilon_k)|\mathcal{N}_h \cup \mathcal{N}_h^*|_{\mathcal{G}}/d_k^{e_k(\mathcal{N}_h \cup \mathcal{N}_h^*)} \le \sqrt{\varepsilon_{k-1}}|\mathcal{N}_h \cup \mathcal{N}_h^*|_{\mathcal{G}} \le \sqrt{\varepsilon_{k-1}}|\mathcal{N}_h|_{\mathcal{G}}|\mathcal{N}_h^*|_{\mathcal{G}} \quad (4.5)$$

disjoint pairs of copies of $\mathcal{N}_h$ and $\mathcal{N}_h^*$ are useful. Note that if we had chosen $\mathcal{N}_h^*$ to be the subcomplex of $\mathcal{H}$ induced by the vertices at distance 2 from h (instead of 3), then we could not have applied Lemma 4.6, since $\mathcal{N}_h \cup \mathcal{N}_h^*$ would not be an induced subcomplex of $\mathcal{F}$. Together with the fact that only comparatively few of the pairs N_h, N_h^* will intersect, this shows that at most $2\sqrt{\varepsilon_{k-1}}|\mathcal{N}_h|_\mathcal{G}|\mathcal{N}_h^*|_\mathcal{G}$ pairs N_h, N_h^* are not useful. Hence at most $\varepsilon_{k-1}^{1/4}|\mathcal{N}_h|_\mathcal{G}$ copies of $\mathcal{N}_h$ form a non-useful pair together with more than $2\varepsilon_{k-1}^{1/4}|\mathcal{N}_h^*|_\mathcal{G}$ copies of $\mathcal{N}_h^*$. We call all other copies of $\mathcal{N}_h$ *useful* and let `Usef` denote the set of all these copies. Then

$$|\mathcal{N}_h|_\mathcal{G} - |\texttt{Usef}| \leq \varepsilon_{k-1}^{1/4}|\mathcal{N}_h|_\mathcal{G}. \tag{4.6}$$

We denote by $\texttt{Usef}^*(N_h)$ the set of all N_h^* which form a useful pair together with N_h.

Claim. *Any useful copy N_h of $\mathcal{N}_h$ satisfies*

$$|N_h \to \mathcal{H}_h| \leq \frac{10}{d_k^{\Delta^3}} \frac{|\mathcal{H}_h|_\mathcal{G}}{|\mathcal{N}_h|_\mathcal{G}}.$$

Note that $\sum_{N_h} |N_h \to \mathcal{H}_h| = |\mathcal{H}_h|_\mathcal{G}$, so $|\mathcal{H}_h|_\mathcal{G}/|\mathcal{N}_h|_\mathcal{G}$ is the average value of $|N_h \to \mathcal{H}_h|$ over all copies N_h of $\mathcal{N}_h$. Later on, we will apply the claim to show that only a small fraction of copies of $\mathcal{H}$ contain a useful but atypical copy of $\mathcal{N}_h$.

Proof of Claim. Fix a useful copy N_h of $\mathcal{N}_h$. Put $\mathcal{H}_h^* := \mathcal{H}_h - (\mathcal{F} - \mathcal{N}_h^*)$. We aim to extend N_h to a copy of $\mathcal{H}_h$ by first picking a copy N_h^* of $\mathcal{N}_h^*$, then extending this to a copy of $\mathcal{H}_h^*$ and also extending $N_h \cup N_h^*$ to a copy of $\mathcal{F}$. We must also make sure that no vertices are used more than once. However, since we are only looking for an upper bound on $|N_h \to \mathcal{H}_h|$, and ignoring this restriction can only increase the number of extensions we find, we may ignore this difficulty. Thus

$$|N_h \to \mathcal{H}_h| \leq \sum_{N_h^* \in \mathtt{Usef}^*(N_h)} |N_h \cup N_h^* \to \mathcal{F}||N_h^* \to \mathcal{H}_h^*| + \sum_{N_h^* \notin \mathtt{Usef}^*(N_h)} |N_h \cup N_h^* \to \mathcal{F}||N_h^* \to \mathcal{H}_h^*|. \tag{4.7}$$

We bound the two sums separately. To bound the first sum, we need to bound $|N_h \cup N_h^* \to \mathcal{F}|$ in the case when the pair N_h, N_h^* is useful. But clearly $|N_h \cup N_h^* \to \mathcal{F}| \leq |N_h \cup N_h^* \stackrel{k-1}{\to} \mathcal{F}|$, and

$$|N_h \cup N_h^* \stackrel{k-1}{\to} \mathcal{F}| \leq (1+\varepsilon_{k-1})\overline{|\mathcal{N}_h \cup \mathcal{N}_h^* \stackrel{k-1}{\to} \mathcal{F}|} = \frac{(1+\varepsilon_{k-1})\overline{|\mathcal{N}_h \cup \mathcal{N}_h^* \to \mathcal{F}|}}{d_k^{e_k(\mathcal{F})-e_k(\mathcal{N}_h)-e_k(\mathcal{N}_h^*)}}$$

whenever $N_h^* \in \mathtt{Usef}^*(N_h)$. So the first sum in (4.7) is bounded by

$$\frac{1+\varepsilon_{k-1}}{d_k^{e_k(\mathcal{F})-e_k(\mathcal{N}_h)-e_k(\mathcal{N}_h^*)}}\overline{|\mathcal{N}_h \cup \mathcal{N}_h^* \to \mathcal{F}|}|\mathcal{H}_h^*|_{\mathcal{G}} \leq \frac{2}{d_k^{\Delta^3}}\overline{|\mathcal{N}_h \cup \mathcal{N}_h^* \to \mathcal{F}|}|\mathcal{H}_h^*|_{\mathcal{G}}. \tag{4.8}$$

To see the bound of Δ^3 on the number of k-edges which we used in the final inequality, note that $|\mathcal{F} - \mathcal{N}_h - \mathcal{N}_h^*| \leq \Delta^2$ and that the number of k-edges each of these vertices lies in is at most Δ. We now want to express the bound in (4.8) in terms of $|\mathcal{H}_h^-|_{\mathcal{G}}$, where $\mathcal{H}_h^- := \mathcal{H}_h - \mathcal{N}_h$. By the induction hypothesis applied $|H_h^-| - |H_h^*|$ times,

$$\begin{aligned}|\mathcal{H}_h^*|_{\mathcal{G}} &\leq ((1-\alpha)n)^{-(|\mathcal{H}_h^-|-|\mathcal{H}_h^*|)}\left(\prod_{i=2}^{k} d_i^{-(e_i(\mathcal{H}_h^-)-e_i(\mathcal{H}_h^*))}\right)|\mathcal{H}_h^-|_{\mathcal{G}} \\ &\leq 2\frac{\prod_{i=2}^k d_i^{e_i(\mathcal{H}_h)-e_i(\mathcal{H}_h^-)-e_i(\mathcal{N}_h)}}{\overline{|\mathcal{N}_h \cup \mathcal{N}_h^* \to \mathcal{F}|}}|\mathcal{H}_h^-|_{\mathcal{G}}.\end{aligned}$$

In the last line we used that $e_i(\mathcal{H}_h) = e_i(\mathcal{H}_h^*) + e_i(\mathcal{F}) - e_i(\mathcal{N}_h^*)$ and $|\mathcal{F}| - |\mathcal{N}_h| - |\mathcal{N}_h^*| =$

$|\mathcal{H}_h^-|-|\mathcal{H}_h^*|$ (see Figure 4.2). We also used that $(1-\alpha)^{-(|\mathcal{H}_h^-|-|\mathcal{H}_h^*|)} \leq 2$. So we obtain

$$\sum_{N_h^* \in \texttt{Usef}^*(N_h)} |N_h \cup N_h^* \to \mathcal{F}||N_h^* \to \mathcal{H}_h^*| \leq \frac{4\prod_{i=2}^k d_i^{e_i(\mathcal{H}_h)-e_i(\mathcal{H}_h^-)-e_i(\mathcal{N}_h)}}{d_k^{\Delta^3}} |\mathcal{H}_h^-|_{\mathcal{G}}. \quad (4.9)$$

To bound the second sum in (4.7), we define $\mathcal{H}_h' := \mathcal{H}_h^* - \mathcal{N}_h^*$, and observe that trivially any copy N_h^* of $\mathcal{N}_h^*$ satisfies $|N_h^* \to \mathcal{H}_h^*| \leq |\mathcal{H}_h'|_{\mathcal{G}}$. Note that $\mathcal{H}_h'$ is nonempty. On the other hand, by the induction hypothesis applied $|H_h^-| - |H_h'|$ times,

$$\begin{aligned} |\mathcal{H}_h'|_{\mathcal{G}} &\leq ((1-\alpha)n)^{|\mathcal{H}_h'|-|\mathcal{H}_h^-|} \left(\prod_{i=2}^k d_i^{e_i(\mathcal{H}_h')-e_i(\mathcal{H}_h^-)} \right) |\mathcal{H}_h^-|_{\mathcal{G}} \\ &\leq \frac{2|\mathcal{H}_h^-|_{\mathcal{G}}}{\left(\prod_{i=2}^k d_i\right)^{2\Delta^4} n^{|\mathcal{H}_h^-|-|\mathcal{H}_h'|}}. \end{aligned}$$

Since at most $2\varepsilon_{k-1}^{1/4}|\mathcal{N}_h^*|_{\mathcal{G}} \leq 2\varepsilon_{k-1}^{1/4} n^{|\mathcal{N}_h^*|}$ copies of $\mathcal{N}_h^*$ do not lie in $\texttt{Usef}(N_h)$, the second sum in (4.7) is bounded by

$$\begin{aligned} \sum_{N_h^* \notin \texttt{Usef}(N_h)} &|N_h \cup N_h^* \to \mathcal{F}||N_h^* \to \mathcal{H}_h^*| \\ &\leq 2\varepsilon_{k-1}^{1/4} n^{|\mathcal{N}_h^*|} n^{|\mathcal{F}|-|\mathcal{N}_h|-|\mathcal{N}_h^*|} \frac{2|\mathcal{H}_h^-|_{\mathcal{G}}}{\left(\prod_{i=2}^k d_i\right)^{2\Delta^4} n^{|\mathcal{H}_h^-|-|\mathcal{H}_h'|}} \\ &= 2\varepsilon_{k-1}^{1/4} \frac{2|\mathcal{H}_h^-|_{\mathcal{G}}}{\left(\prod_{i=2}^k d_i\right)^{2\Delta^4}} \\ &\leq \left(\prod_{i=2}^k d_i^{e_i(\mathcal{H}_h)-e_i(\mathcal{H}_h^-)-e_i(\mathcal{N}_h)}\right) |\mathcal{H}_h^-|_{\mathcal{G}}. \end{aligned} \quad (4.10)$$

The last inequality follows since $\varepsilon_{k-1} \ll d_2, d_3, \ldots, d_k, 1/\Delta$. Substituting (4.9) and

(4.10) into (4.7) we obtain

$$\begin{aligned}
|N_h \to \mathcal{H}_h| &\le \left(1+\frac{4}{d_k^{\Delta^3}}\right)\left(\prod_{i=2}^{k} d_i^{e_i(\mathcal{H}_h)-e_i(\mathcal{H}_h^-)-e_i(\mathcal{N}_h)}\right)|\mathcal{H}_h^-|_{\mathcal{G}} \\
&\le \frac{5\left(\prod_{i=2}^{k} d_i^{e_i(\mathcal{H}_h)-e_i(\mathcal{H}_h^-)-e_i(\mathcal{N}_h)}\right)}{d_k^{\Delta^3}}|\mathcal{H}_h^-|_{\mathcal{G}}. \qquad (4.11)
\end{aligned}$$

It now remains only to relate $|\mathcal{H}_h^-|_{\mathcal{G}}$ to $|\mathcal{H}_h|_{\mathcal{G}}/|\mathcal{N}_h|_{\mathcal{G}}$. Once again we apply the induction hypothesis $|H_h| - |H_h^-|$ times to obtain

$$|\mathcal{H}_h|_{\mathcal{G}} \ge ((1-\alpha)n)^{|\mathcal{H}_h|-|\mathcal{H}_h^-|}\prod_{i=2}^{k} d_i^{e_i(\mathcal{H}_h)-e_i(\mathcal{H}_h^-)}|\mathcal{H}_h^-|_{\mathcal{G}}.$$

On the other hand, the counting lemma implies that

$$|\mathcal{N}_h|_{\mathcal{G}} \le (1+\alpha)\left(\prod_{i=2}^{k} d_i^{e_i(\mathcal{N}_h)}\right)n^{|\mathcal{N}_h|}.$$

Putting these two bounds together, we obtain

$$\begin{aligned}
\frac{|\mathcal{H}_h|_{\mathcal{G}}}{|\mathcal{N}_h|_{\mathcal{G}}} &\ge \frac{((1-\alpha)n)^{|\mathcal{H}_h|-|\mathcal{H}_h^-|}\left(\prod_{i=2}^{k} d_i^{e_i(\mathcal{H}_h)-e_i(\mathcal{H}_h^-)}\right)|\mathcal{H}_h^-|}{(1+\alpha)\left(\prod_{i=2}^{k} d_i^{e_i(\mathcal{N}_h)}\right)n^{|\mathcal{N}_h|}} \\
&\ge \frac{1}{2}\left(\prod_{i=2}^{k} d_i^{e_i(\mathcal{H}_h)-e_i(\mathcal{H}_h^-)-e_i(\mathcal{N}_h)}\right)|\mathcal{H}_h^-|_{\mathcal{G}}. \qquad (4.12)
\end{aligned}$$

Together with (4.11), this shows that

$$|N_h \to \mathcal{H}_h| \le \frac{5\cdot 2}{d_k^{\Delta^3}}\frac{|\mathcal{H}_h|_{\mathcal{G}}}{|\mathcal{N}_h|_{\mathcal{G}}},$$

which completes the proof of the claim. □

Using the claim we now go on to prove the induction step. Given a copy H_h of $\mathcal{H}_h$, we denote by $N_h(H_h)$ the induced copy of $\mathcal{N}_h$. We have

$$\begin{aligned}
|\mathcal{H}|_{\mathcal{G}} &= \sum_{H_h \subseteq \mathcal{G}} |H_h \to \mathcal{H}| \\
&\geq \sum_{H_h \subseteq \mathcal{G}} (|N_h(H_h) \to \mathcal{B}| - cn) \\
&= \sum_{N_h \subseteq \mathcal{G}} |N_h \to \mathcal{H}_h||N_h \to \mathcal{B}| - cn|\mathcal{H}_h|_{\mathcal{G}} \\
&\geq (1-\varepsilon_k)\overline{|\mathcal{N}_h \to \mathcal{B}|}\left(\sum_{N_h \subseteq \mathcal{G}} |N_h \to \mathcal{H}_h| - \sum_{N_h \notin \texttt{typ}} |N_h \to \mathcal{H}_h|\right) - cn|\mathcal{H}_h|_{\mathcal{G}}.
\end{aligned} \tag{4.13}$$

We want to show that the term in this expression which comes from the atypical copies of $\mathcal{N}_h$ does not affect the calculations too much, and so we aim to bound the contribution from atypical copies of $\mathcal{N}_h$. We have

$$\sum_{N_h \notin \texttt{typ}} |N_h \to \mathcal{H}_h| = \sum_{N_h \notin \texttt{typ}, N_h \in \texttt{Usef}} |N_h \to \mathcal{H}_h| + \sum_{N_h \notin \texttt{typ}, N_h \notin \texttt{Usef}} |N_h \to \mathcal{H}_h|. \tag{4.14}$$

Now the claim implies that we can bound the first sum in (4.14) by

$$\begin{aligned}
\sum_{N_h \notin \texttt{typ}, N_h \in \texttt{Usef}} |N_h \to \mathcal{H}_h| &\leq \sum_{N_h \notin \texttt{typ}, N_h \in \texttt{Usef}} \frac{10}{d_k^{\Delta^3}} \frac{|\mathcal{H}_h|_{\mathcal{G}}}{|\mathcal{N}_h|_{\mathcal{G}}} \\
&\leq |\texttt{atyp}| \frac{10}{d_k^{\Delta^3}} \frac{|\mathcal{H}_h|_{\mathcal{G}}}{|\mathcal{N}_h|_{\mathcal{G}}} \leq \sqrt{\varepsilon_k}|\mathcal{H}_h|_{\mathcal{G}}.
\end{aligned} \tag{4.15}$$

Meanwhile we can also bound the second sum by

$$\begin{aligned}
\sum_{N_h \notin \texttt{typ}, N_h \notin \texttt{Usef}} |N_h \to \mathcal{H}_h| &\leq \sum_{N_h \notin \texttt{typ}, N_h \notin \texttt{Usef}} |\mathcal{H}_h^-|_{\mathcal{G}} \\
&\overset{(4.12)}{\leq} (|\mathcal{N}_h|_{\mathcal{G}} - |\texttt{Usef}|) \frac{2}{\prod_{i=1}^k d_i^{\Delta^2}} \frac{|\mathcal{H}_h|_{\mathcal{G}}}{|\mathcal{N}_h|_{\mathcal{G}}} \\
&\overset{(4.6)}{\leq} \varepsilon_{k-1}^{1/5}|\mathcal{H}_h|_{\mathcal{G}}.
\end{aligned} \tag{4.16}$$

Combining (4.14), (4.15) and (4.16), we have

$$\sum_{N_h \notin \texttt{typ}} |N_h \to \mathcal{H}_h| \leq 2\sqrt{\varepsilon_k}|\mathcal{H}_h|_{\mathcal{G}}$$

and combining this with (4.13), we obtain

$$\begin{aligned}
|\mathcal{H}|_{\mathcal{G}} &\geq (1-\varepsilon_k)\overline{|\mathcal{N}_h \to \mathcal{B}|}\left(|\mathcal{H}_h|_{\mathcal{G}} - 2\sqrt{\varepsilon_k}|\mathcal{H}_h|_{\mathcal{G}}\right) - cn|\mathcal{H}_h|_{\mathcal{G}} \\
&= (1-\varepsilon_k)n\left(\prod_{i=2}^{k} d_i^{e_i(\mathcal{B})-e_i(\mathcal{N}_h)}\right)(1-2\sqrt{\varepsilon_k})|\mathcal{H}_h|_{\mathcal{G}} - cn|\mathcal{H}_h|_{\mathcal{G}} \\
&\geq (1-\alpha)n\left(\prod_{i=2}^{k} d_i^{e_i(\mathcal{H})-e_i(\mathcal{H}_h)}\right)|\mathcal{H}_h|_{\mathcal{G}},
\end{aligned}$$

as required. This completes the proof of Theorem 4.3.

4.5 The regularity lemma for k-uniform hypergraphs

4.5.1 Preliminary definitions and statement

In this section we state the version of the regularity lemma for k-uniform hypergraphs due to Rödl and Schacht [64], which we use in the proof of Theorem 4.1 in the next section. To prepare for this we will first need some notation. We follow [64]. Given a finite set V of vertices, we will define a family $\mathcal{P} = \{\mathcal{P}^{(1)}, \dots, \mathcal{P}^{(k-1)}\}$ where each $\mathcal{P}^{(j)}$ is a partition of certain j-subsets of V. These partitions will satisfy properties which we will describe below. We denote by $[V]^j$ the set of all j-subsets of V. Suppose that we are given a partition $\mathcal{P}^{(1)} = \{V_1, \dots, V_{|\mathcal{P}^{(1)}|}\}$ of $[V]^1 = V$. We will call the V_i *clusters*. We denote by $\text{Cross}_j = \text{Cross}_j(\mathcal{P}^{(1)})$ the set of all those j-subsets of V that meet each part of $\mathcal{P}^{(1)}$ in at most 1 element. Each $\mathcal{P}^{(j)}$ will be a partition of Cross_j.

Moreover, any two j-sets that belong to the same part of $\mathcal{P}^{(j)}$ will meet the same j clusters. This means that each part of $\mathcal{P}^{(j)}$ can be viewed as a j-partite j-uniform hypergraph whose vertex classes are these clusters. In particular, the parts of $\mathcal{P}^{(2)}$ can be thought of as bipartite subgraphs between two of the clusters. Moreover, for each part A of $\mathcal{P}^{(3)}$ there will be 3 clusters and 3 bipartite graphs belonging to $\mathcal{P}^{(2)}$ between these clusters such that all the 3-sets in A form triangles in the union of these 3 bipartite graphs.

More generally, suppose that we have already defined partitions $\mathcal{P}^{(1)}, \dots, \mathcal{P}^{(j-1)}$ and are about to define $\mathcal{P}^{(j)}$. Given $i < j$ and $I \in \mathrm{Cross}_i$, we let $P^{(i)}(I)$ denote the part of $\mathcal{P}^{(i)}$ the set I belongs to. Given $J \in \mathrm{Cross}_j$, the *polyad* $\hat{P}^{(j-1)}(J)$ *of* J is defined by

$$\hat{P}^{(j-1)}(J) := \bigcup \{P^{(j-1)}(I) : \ I \in [J]^{j-1}\}.$$

Thus $\hat{P}^{(j-1)}(J)$ is the unique collection of j parts of $\mathcal{P}^{(j-1)}$ in which J spans a copy of the complete $(j-1)$-uniform hypergraph $K_j^{(j-1)}$ on j vertices. Moreover, note that $\hat{P}^{(j-1)}(J)$ can be viewed as a j-partite $(j-1)$-uniform hypergraph whose vertex classes are the j clusters containing the vertices of J. We set

$$\hat{\mathcal{P}}^{(j-1)} := \{\hat{P}^{(j-1)}(J) : \ J \in \mathrm{Cross}_j\}.$$

Note that the polyads $\hat{P}^{(j-1)}(J)$ and $\hat{P}^{(j-1)}(J')$ need not be distinct for different $J, J' \in [V]^j$. However, if these polyads are distinct then $\mathcal{K}_j(\hat{P}^{(j-1)}(J)) \cap \mathcal{K}_j(\hat{P}^{(j-1)}(J')) = \emptyset$. (Recall that $\mathcal{K}_j(\hat{P}^{(j-1)}(J))$ is the set of all j-sets of vertices which form a $K_j^{(j-1)}$ in $\hat{P}^{(j-1)}(J)$. So in particular, $\mathcal{K}_j(\hat{P}^{(j-1)}(J))$ contains J.) This implies that $\{\mathcal{K}_j(\hat{P}^{(j-1)}) : \hat{P}^{(j-1)} \in \hat{\mathcal{P}}^{(j-1)}\}$ is a partition of Cross_j. The property of $\mathcal{P}^{(j)}$ which we require is that it refines $\{\mathcal{K}_j(\hat{P}^{(j-1)}) : \ \hat{P}^{(j-1)} \in \hat{\mathcal{P}}^{(j-1)}\}$, i.e. each part of $\mathcal{P}^{(j)}$ has to be contained in some $\mathcal{K}_j(\hat{P}^{(j-1)})$.

We also need a notion which generalises that of a polyad: given $J \in \mathrm{Cross}_j$ and $i < j$ we set

$$\hat{P}^{(i)}(J) := \bigcup\{P^{(i)}(I) : \ I \in [J]^i\}.$$

Then the properties of our partitions imply that $\bigcup_{i=1}^{j-1} \hat{P}^{(i)}(J)$ is a $(j-1, j)$-complex.

Altogether, given $\boldsymbol{a} = (a_1, \ldots, a_{k-1}) \in \mathbb{N}^{k-1}$ we say that $\mathcal{P}(k-1, \boldsymbol{a}) = \{\mathcal{P}^{(1)}, \ldots, \mathcal{P}^{(k-1)}\}$ is a *family of partitions on V* if

1. $\mathcal{P}^{(1)}$ is a partition of V into a_1 clusters.
2. For all $j = 2, \ldots, k-1$, $\mathcal{P}^{(j)}$ is a partition of Cross_j such that for each part there is a polyad $\hat{P}^{(j-1)} \in \hat{\mathcal{P}}^{(j-1)}$ so that the part is contained in $\mathcal{K}_j(\hat{P}^{(j-1)})$. Moreover, for each polyad $\hat{P}^{(j-1)} \in \hat{\mathcal{P}}^{(j-1)}$, the set $\mathcal{K}_j(\hat{P}^{(j-1)})$ is the union of a_j parts of $\mathcal{P}^{(j)}$.

We say that $\mathcal{P} = \mathcal{P}(k-1, \boldsymbol{a})$ is *t-bounded* if $a_1, \ldots, a_{k-1} \leq t$. Suppose that a_1 divides $|V|$. Then $\mathcal{P} = \mathcal{P}(k-1, \boldsymbol{a})$ is called *$(\eta, \delta, \boldsymbol{a})$-equitable* if

1. $\mathcal{P}^{(1)}$ is a partition of V into a_1 clusters of equal size;
2. $|[V]^k \setminus \mathrm{Cross}_k| \leq \eta\binom{|V|}{k}$;
3. for every $K \in \mathrm{Cross}_k$, the $(k-1, k)$-complex $\bigcup_{i=1}^{k-1} \hat{P}^{(i)}(K)$ is $(\mathbf{d}, \delta, \delta, 1)$-regular, where $\mathbf{d} = (1/a_{k-1}, \ldots, 1/a_2)$.

In particular, the second condition implies that $1/a_1$ is small compared to η.

Let $\delta_k > 0$ and $r \in \mathbb{N}$. Suppose that $\mathcal{G}$ is a k-uniform hypergraph on V and $\mathcal{P} = \mathcal{P}(k-1, \boldsymbol{a})$ is a family of partitions on V. Recall that we can view each polyad $\hat{P}^{(k-1)} \in \hat{\mathcal{P}}^{(k-1)}$ as a $(k-1)$-uniform k-partite hypergraph. $\mathcal{G}$ is called *(δ_k, r)-regular with respect to $\hat{P}^{(k-1)}$* if $\mathcal{G}$ is (d, δ_k, r)-regular with respect to $\hat{P}^{(k-1)}$

for some d. Setting

$$\text{Irreg}(\mathcal{G}) := \{\hat{P}^{(k-1)} \in \hat{\mathcal{P}}^{(k-1)} : \mathcal{G} \text{ is not } (\delta_k, r)\text{-regular with respect to } \hat{P}^{(k-1)}\},$$

we say that $\mathcal{G}$ is (δ_k, r)*-regular with respect to* $\mathcal{P}$ if

$$\left| \bigcup_{\hat{P}^{(k-1)} \in \text{Irreg}(\mathcal{G})} \mathcal{K}_k(\hat{P}^{(k-1)}) \right| \leq \delta_k |V|^k.$$

This means that not much more than a δ_k-fraction of the k-subsets of V form a $K_k^{(k-1)}$ that lies within a polyad with respect to which $\mathcal{G}$ is not regular.

Now, we are ready to state the regularity lemma, which we are going to use in the proof of Theorem 4.1.

Theorem 4.9 (Rödl and Schacht [64]) *Let $k \geq 2$ be a fixed integer. For all positive constants η and δ_k and all functions $r : \mathbb{N}^{k-1} \to \mathbb{N}$ and $\delta : \mathbb{N}^{k-1} \to (0, 1]$, there are integers t and m_0 such that the following holds for all $m \geq m_0$ which are divisible by $t!$. Suppose that $\mathcal{G}$ is a k-uniform hypergraph of order m. Then there exists an $\boldsymbol{a} \in \mathbb{N}^{k-1}$ and a family of partitions $\mathcal{P} = \mathcal{P}(k-1, \boldsymbol{a})$ of the vertex set V of $\mathcal{G}$ such that*

1. *$\mathcal{P}$ is $(\eta, \delta(\boldsymbol{a}), \boldsymbol{a})$-equitable and t-bounded and*

2. *$\mathcal{G}$ is $(\delta_k, r(\boldsymbol{a}))$-regular with respect to $\mathcal{P}$.*

The advantage of this regularity lemma compared to the one proved earlier by Rödl and Skokan [66] is that it uses only two regularity constants δ and δ_k instead of $k-1$ different ones. The regularity constants $\delta_2, \ldots, \delta_k$ produced by the regularity lemma in [66] might satisfy $\delta_2 \ll 1/a_2 \ll \delta_3 \ll 1/a_3 \ll \cdots \ll 1/a_{k-1} \ll \delta_k$, which would make the proof of the corresponding embedding theorem more technical in appearance.

Note that the constants in Theorem 4.9 can be chosen such that they satisfy the following hierarchy:

$$\frac{1}{m_0} \ll \frac{1}{r} = \frac{1}{r(\boldsymbol{a})}, \delta = \delta(\boldsymbol{a}) \ll \min\{\delta_k, \eta, 1/a_1, 1/a_2, \ldots, 1/a_{k-1}\}. \tag{4.17}$$

4.5.2 The reduced hypergraph

In the proof of Theorem 4.1 that follows in the next section, we will use the so-called reduced hypergraph. If $\mathcal{P} = \{\mathcal{P}^{(1)}, \ldots, \mathcal{P}^{(k-1)}\}$ is the partition of the vertex set of $\mathcal{G}$ given by the regularity lemma, the *reduced hypergraph* $\mathcal{R} = \mathcal{R}(\mathcal{G}, \mathcal{P})$ is a k-uniform hypergraph whose vertices are the clusters, i.e. the parts of $\mathcal{P}^{(1)}$. To define the set of hyperedges we need the following notion. We say that a k-tuple of clusters is *fruitful* if $\mathcal{G}$ is (δ_k, r)-regular with respect to all but at most a $\sqrt{\delta_k}$-fraction of all those polyads $\hat{P}^{(k-1)}$ which are induced on these k clusters. The set of hyperedges of $\mathcal{R}$ consists of precisely those k-tuples that are fruitful. In the proof of Theorem 4.1, we shall need an estimate on the number of these hyperedges. In particular, we need to show that $\mathcal{R}$ is very dense. This is conveyed in the following proposition.

Proposition 4.10 *All but at most $2\sqrt{\delta_k}a_1^k$ of the k-tuples of clusters are fruitful.*

Proof. By the dense counting lemma (Lemma 4.6) each polyad in $\hat{\mathcal{P}}^{(k-1)}$ contains at least

$$f(m, \boldsymbol{a}) := \frac{1}{2}\left(\frac{m}{a_1}\right)^k \prod_{i=2}^{k-1}\left(\frac{1}{a_i}\right)^{\binom{k}{i}}$$

copies of $K_k^{(k-1)}$. Since $\mathcal{G}$ is (δ_k, r)-regular with respect to $\mathcal{P}$, the number of polyads in $\hat{\mathcal{P}}^{(k-1)}$ with respect to which $\mathcal{G}$ is not (δ_k, r)-regular is at most

$$\frac{\delta_k m^k}{f(m, \boldsymbol{a})} = \frac{2\prod_{i=1}^{k-1} a_i^{\binom{k}{i}}}{m^k}\delta_k m^k = 2\delta_k \prod_{i=1}^{k-1} a_i^{\binom{k}{i}}. \tag{4.18}$$

We call these polyads *bad*. Now, each k-tuple of clusters induces $\prod_{i=2}^{k-1} a_i^{\binom{k}{i}}$ polyads in $\hat{\mathcal{P}}^{(k-1)}$. Thus if there were more than $2\sqrt{\delta_k}a_1^k$ k-tuples of clusters each inducing more than $\sqrt{\delta_k}\prod_{i=2}^{k-1} a_i^{\binom{k}{i}}$ bad polyads, the total number of bad polyads would exceed the bound given in (4.18), yielding a contradiction. □

4.6 Proof of Theorem 4.1

We now give a brief outline of the proof of Theorem 4.1: consider any red/blue colouring of the hyperedges of $K_m^{(k)}$, where $m = C|\mathcal{H}|$ and C is a large constant depending only on k and the maximum degree of $\mathcal{H}$. We apply the hypergraph regularity lemma to the red subhypergraph $\mathcal{G}_{red}$ to obtain a reduced hypergraph $\mathcal{R}$ which is very dense. Thus the following fact will show that $\mathcal{R}$ contains a copy of $K_\ell^{(k)}$ with $\ell := R(K_{k\Delta}^{(k)})$.

Fact 4.11 *For all $\ell, k \in \mathbb{N}$ with $\ell \geq k$, every k-uniform hypergraph $\mathcal{R}$ on $t \geq \ell$ vertices with $e(\mathcal{R}) > \left(1 - \binom{\ell}{k}^{-1}\right)\binom{t}{k}$ contains a copy of $K_\ell^{(k)}$.*

Proof. Let $\mathcal{R}$ be as in the statement of the fact. Assume for the sake of contradiction that $\mathcal{R}$ is $K_\ell^{(k)}$-free. Then for each ℓ-subset S of $V(\mathcal{R})$, we have $e(\mathcal{R}[S]) \leq \binom{\ell}{k} - 1$. But note that

$$e(\mathcal{R}) = \binom{t-k}{\ell-k}^{-1} \sum_S e(\mathcal{R}[S]).$$

Thus $e(\mathcal{R}) \leq \binom{t-k}{\ell-k}^{-1}\binom{t}{\ell}\left(\binom{\ell}{k} - 1\right)$. Now the observation that $\binom{t-k}{\ell-k}^{-1}\binom{t}{\ell}\binom{\ell}{k} = \binom{t}{k}$ yields the required contradiction. □

The copy of $K_\ell^{(k)}$ in $\mathcal{R}$ involves ℓ clusters and for each k-tuple of them the red hypergraph $\mathcal{G}_{red}$ is regular with respect to almost all of the polyads induced on it.

We will then show that we can find a $(k-1,\ell)$-complex $\mathcal{S}$ on these clusters such that for each $j=2,\dots,k-1$ the restriction of its underlying j-uniform hypergraph $\mathcal{S}_j$ to any $(j+1)$-tuple of clusters is a polyad. Moreover, $\mathcal{G}_{red}$ will be regular with respect to $\mathcal{S}_{k-1}$. By combining $E(\mathcal{G}_{red})\cap\mathcal{K}_k(\mathcal{S}_{k-1})$ with $\mathcal{S}$, we will obtain a regular k-complex $\mathcal{S}_{red}$. Similarly we obtain a k-complex $\mathcal{S}_{blue}$ which also turns out to be regular. We then consider the following red/blue colouring of $K_\ell^{(k)}$. We colour a hyperedge red if $\mathcal{G}_{red}$ has density at least $1/2$ with respect to the corresponding polyad in $\mathcal{S}_{k-1}$ and blue otherwise. By the definition of ℓ, we can find a monochromatic $K_{k\Delta}^{(k)}$. If it is red, then we can apply the embedding lemma to $\mathcal{S}_{red}$ to find a red copy of $\mathcal{H}$. This can be done since $\Delta(\mathcal{H})\le\Delta$ implies that the chromatic number of $\mathcal{H}$ is at most $(k-1)\Delta+1\le k\Delta$. If our monochromatic copy of $K_{k\Delta}^{(k)}$ is blue, then we can apply the embedding theorem to $\mathcal{S}_{blue}$ and obtain a blue copy of $\mathcal{H}$.

Proof of Theorem 4.1. Given Δ and k, we choose C to be a sufficiently large constant. We will describe the bounds that C has to satisfy at the end of the proof. Let $m:=C|\mathcal{H}|$ and consider any red/blue colouring of the hyperedges of $K_m^{(k)}$. Let $\mathcal{G}_{red}$ be the red and $\mathcal{G}_{blue}$ be the blue subhypergraph on $V=V(K_m^{(k)})$. We may assume without loss of generality that $e(\mathcal{G}_{red})\ge e(\mathcal{G}_{blue})$. We apply the hypergraph regularity lemma to $\mathcal{G}_{red}$ with constants $\eta,\delta_k\ll 1/\Delta,1/k$ as well as functions r and δ satisfying the hierarchy in (4.17). This gives us clusters $V_1,\dots,V_{a_1}$, each of size n say, together with a t-bounded $(\eta,\delta,\boldsymbol{a})$-equitable family of partitions $\mathcal{P}=\mathcal{P}(k-1,\boldsymbol{a})$ on V where $\boldsymbol{a}=(a_1,\dots,a_{k-1})$. (Note that by deleting some vertices of $\mathcal{G}_{red}$ if necessary we may assume that $m=|\mathcal{G}_{red}|$ is divisible by $t!$.) Since $\eta\ll 1/\Delta,1/k$, condition (2) in the definition of an $(\eta,\delta,\boldsymbol{a})$-equitable family of partitions implies that the a_1 which we obtain from the regularity lemma satisfies

$$a_1\ge R(K_{k\Delta}^{(k)})=:\ell.$$

Note that the definition of ℓ involves a hypergraph Ramsey number whose value is unknown. However, for the argument below all we need is that this number exists.

Let $\mathcal{R}$ denote the reduced hypergraph, defined in the previous section. Proposition 4.10 implies that $\mathcal{R}$ has at least $(1-\varepsilon)\binom{a_1}{k}$ hyperedges, where $\varepsilon := 4\sqrt{\delta_k}k!$. Since $\delta_k \ll 1/\Delta, 1/k$, we may assume that $e(\mathcal{R}) \geq (1-\varepsilon)\binom{|\mathcal{R}|}{k} > \left(1 - \binom{\ell}{k}^{-1}\right)\binom{|\mathcal{R}|}{k}$. Since $|\mathcal{R}| = a_1 \geq \ell$, this means that we can apply Fact 4.11 to $\mathcal{R}$ to obtain a copy of $K_\ell^{(k)}$ in $\mathcal{R}$. Without loss of generality we may assume that the vertices of this copy are the clusters $V_1, \dots, V_\ell$.

As mentioned above, we now want to find a $(k-1,\ell)$-complex $\mathcal{S}$ on these clusters such that for each $j = 2, \dots, k-1$ its underlying j-uniform hypergraph $\mathcal{S}_j$ is a union of parts of $\mathcal{P}^{(j)}$ and $\mathcal{G}_{red}$ is regular with respect to $\mathcal{S}_{k-1}$. We construct $\mathcal{S}$ inductively starting from the lower levels. To begin with, for each pair (V_i, V_j) $(1 \leq i < j \leq \ell)$ independently, we choose with probability $1/a_2$ one of the parts of $\mathcal{P}^{(2)}$ induced on (V_i, V_j). $\mathcal{S}_2$ will be the union of these parts. Now suppose that we have chosen $\mathcal{S}_{j-1}$ such that its restriction to any j-tuple of clusters forms a polyad (clearly this is the case for $\mathcal{S}_2$). Now, if $\hat{P}^{(j-1)}$ is such a polyad, we choose a part of $\mathcal{P}^{(j)}$ uniformly at random among the a_j parts of $\mathcal{P}^{(j)}$ that form $\mathcal{K}_j(\hat{P}^{(j-1)})$, independently for each j-tuple of clusters. We let $\mathcal{S}$ be the $(k-1,\ell)$-complex thus obtained.

We will show that there is some choice of $\mathcal{S}$ such that for every k-tuple among the clusters $V_1, \dots, V_\ell$ the hypergraph $\mathcal{G}_{red}$ is (δ_k, r)-regular with respect to the restriction of $\mathcal{S}_{k-1}$ to this k-tuple. Note that $\mathcal{S}_{k-1}$ restricted to any particular k-tuple of clusters is in fact a polyad selected uniformly at random among all polyads $\hat{P}^{(k-1)}$ induced by these k clusters. Therefore, since all the k-tuples of clusters are fruitful, the definition of a fruitful k-tuple implies that the probability that $\mathcal{G}_{red}$ has the necessary regularity is at least

$$1 - \sqrt{\delta_k}\binom{\ell}{k} > \frac{1}{2}.$$

The final inequality holds since we may assume that δ_k is sufficiently small compared to $1/\ell$. This shows the existence of a $(k-1,\ell)$-complex $\mathcal{S}$ with the required properties. In what follows, $P_{\mathcal{S}}$ will always denote a $(k-1)$-uniform subhypergraph of $\mathcal{S}$ induced by k of the clusters $V_1, \dots, V_\ell$. So each such $P_{\mathcal{S}}$ is a polyad and to each hyperedge of the subhypergraph of $\mathcal{R}$ induced by the clusters $V_1, \dots, V_\ell$ there corresponds such a polyad $P_{\mathcal{S}}$.

We now use the densities of $\mathcal{G}_{red}$ with respect to $\mathcal{S}_{k-1}$ to define a red/blue colouring of the $K_\ell^{(k)}$ which we found in $\mathcal{R}$: we colour a hyperedge of this $K_\ell^{(k)}$ red if the polyad $P_{\mathcal{S}}$ corresponding to this hyperedge satisfies $d(\mathcal{G}_{red}|P_{\mathcal{S}}) \geq 1/2$; otherwise we colour it blue. Since $\ell = R(K_{k\Delta}^{(k)})$, we find a monochromatic copy K of $K_{k\Delta}^{(k)}$ in our $K_\ell^{(k)}$. We now greedily assign the vertices of $\mathcal{H}$ to the clusters that form the vertex set of K in such a way that if k vertices of $\mathcal{H}$ form a hyperedge, then they are assigned to k different clusters. (We may think of this as a $(k\Delta)$-vertex-colouring of $\mathcal{H}$.) We now need to show that with this assignment we can apply the embedding lemma to find a monochromatic copy of $\mathcal{H}$ in either the subhypergraph of $\mathcal{G}_{red}$ induced by the $k\Delta$ clusters in K or the subhypergraph of $\mathcal{G}_{blue}$ induced by these clusters.

First suppose that K is red, so we want to apply the embedding theorem to the k-complex formed by $\mathcal{G}_{red}$ and $\mathcal{S}$ (induced on the $k\Delta$ clusters in K). However, the embedding theorem requires all the densities involved to be equal and of the from $1/a$ for $a \in \mathbb{N}$, whereas all we know is that for every polyad $P_{\mathcal{S}}$ corresponding to a hyperedge of K, we have $d(\mathcal{G}_{red}|P_{\mathcal{S}}) \geq 1/2$. This minor obstacle can be overcome by choosing a subhypergraph $\mathcal{G}'_{red} \subseteq \mathcal{G}_{red}$ such that $\mathcal{G}'_{red}$ is $(1/2, 3\delta_k, r)$-regular with respect to each polyad $P_{\mathcal{S}}$. The existence of such a $\mathcal{G}'_{red}$ follows immediately from the slicing lemma (Lemma 4.8). We then add $E(\mathcal{G}'_{red}) \cap \mathcal{K}_k(\mathcal{S}_{k-1})$ to the subcomplex of $\mathcal{S}$ induced by the clusters in K to obtain a regular $(k, k\Delta)$-complex $\mathcal{S}_{red}$ and we apply the embedding theorem (Theorem 4.2) there to find a copy of $\mathcal{H}$ in $\mathcal{G}'_{red}$, and therefore also in $\mathcal{G}_{red}$.

On the other hand, if K is blue, we need to prove that $\mathcal{G}_{blue}$ is regular with respect to all chosen polyads $P_{\mathcal{S}}$. So suppose $\mathbf{Q} = (Q(1), \ldots, Q(r))$ is an r-tuple of subhypergraphs of one of these polyads $P_{\mathcal{S}}$, satisfying $|\mathcal{K}_k(\mathbf{Q})| > \delta_k |\mathcal{K}_k(P_{\mathcal{S}})|$. Let d be such that $\mathcal{G}_{red}$ is (d, δ_k, r)-regular with respect to $P_{\mathcal{S}}$. Then

$$|(1-d) - d(\mathcal{G}_{blue}|\mathbf{Q})| = |d - (1 - d(\mathcal{G}_{blue}|\mathbf{Q}))| = |d - d(\mathcal{G}_{red}|\mathbf{Q})| < \delta_k.$$

Thus $\mathcal{G}_{blue}$ is $(1-d, \delta_k, r)$-regular with respect to $P_{\mathcal{S}}$ (note that $\delta_k \ll 1/2 \leq 1-d$). Following the same argument as in the previous case, we add $E(\mathcal{G}'_{blue}) \cap \mathcal{K}_k(\mathcal{S}_{k-1})$ to the subcomplex of $\mathcal{S}$ induced by the clusters in K to derive the regular $(k, k\Delta)$-complex $\mathcal{S}_{blue}$ to which we can apply the embedding theorem to obtain a copy of $\mathcal{H}$ in $\mathcal{G}_{blue}$.

It remains to check that we can choose C to be a constant depending only on Δ and k. Note that the constants and functions η, δ_k, r and δ we defined at the beginning of the proof all depend only on Δ and k. So this is also true for the integers m_0 and t and the vector $\boldsymbol{a} = (a_1, \ldots, a_{k-1})$ which we then obtained from the regularity lemma. Note that in order to be able to apply the regularity lemma to $\mathcal{G}_{red}$ we needed $m \geq m_0$, where $m = C|\mathcal{H}|$. This is certainly true if we set $C \geq m_0$. The embedding theorem allows us to embed subcomplexes of size at most cn, where n is the cluster size and where c satisfies $c \ll 1/a_2, , \ldots, 1/a_{k-1}, d_k, 1/(k\Delta)$ (recall that $d_k = 1/2$ and $d_i = 1/a_i$ for all $i = 2, \ldots, k-1$). Thus c too depends only on Δ and k. In order to apply the embedding theorem we needed that $n \geq n_0$, where n_0 as defined in the embedding theorem depends only on Δ and k. Since the number of clusters is at most t, this is satisfied if $m \geq tn_0$, which in turn is certainly true if $C \geq tn_0$. When we applied the embedding lemma to $\mathcal{H}$, we needed that $|\mathcal{H}| \leq cn$. Since $n = m/a_1 = C|\mathcal{H}|/a_1 \geq C|\mathcal{H}|/t$, it suffices to choose $C \geq t/c$ for this. Altogether, this shows that we can define the constant C in Theorem 4.1 by $C := \max\{m_0, tn_0, t/c\}$. □

4.7 Deriving Lemmas 4.4 and 4.6 from earlier work

First, we deduce Lemma 4.6 from [44, Cor. 6.11]. The difference between the two is that the latter result only counts complete hypergraphs but on the other hand it allows for different densities within each level. We need a few definitions that make this notion precise. Let $\mathcal{G}$ be a (k,t)-complex. Recall that $\mathcal{G}_i$ denotes the underlying i-uniform hypergraph of $\mathcal{G}$. For each $3 \leq i < k$, we say that $\mathcal{G}_i$ *is* $(\geq d_i, \delta_i)$*-regular with respect to* $\mathcal{G}_{i-1}$, if for every i-tuple Λ_i of vertex classes of $\mathcal{G}$ the induced hypergraph $\mathcal{G}_i[\Lambda_i]$ is $(d_{\Lambda_i}, \delta_i)$-regular with respect to $\mathcal{G}_{i-1}[\Lambda_i]$, for some $d_{\Lambda_i} \geq d_i$. Similarly we define when $\mathcal{G}_k$ *is* $(\geq d_k, \delta_k, r)$*-regular with respect to* $\mathcal{G}_{k-1}$ and when $\mathcal{G}_2$ *is* $(\geq d_2, \delta_2)$*-regular.* Let $\mathbf{d} := (d_k, \dots, d_2)$. We say that a (k,t)-complex $\mathcal{G}$ is $(\geq \mathbf{d}, \delta_k, \delta, r)$*-regular* if

- $\mathcal{G}_k$ is $(\geq d_k, \delta_k, r)$-regular with respect to $\mathcal{G}_{k-1}$;
- $\mathcal{G}_i$ is $(\geq d_i, \delta)$-regular with respect to $\mathcal{G}_{i-1}$ for each $3 \leq i < k$;
- $\mathcal{G}_2$ is $(\geq d_2, \delta)$-regular.

Lemma 4.12 (Dense counting lemma for complete complexes [44]) *Let k, t, n_0 be positive integers and let $\varepsilon, d_2, \dots, d_{k-1}, \delta$ be positive constants such that*

$$1/n_0 \ll \delta \ll \varepsilon \ll d_2, \dots, d_{k-1}, 1/t.$$

Then the following holds for all integers $n \geq n_0$. Suppose that $\mathcal{G}$ is a $(\geq (d_{k-1}, \dots, d_2), \delta, \delta, 1)$-regular $(k-1, t)$-complex with vertex classes $V_1, \dots, V_t$, all of size n. Then

$$|K_t^{(k-1)}|_{\mathcal{G}} = (1 \pm \varepsilon) n^t \prod_{i=2}^{k-1} \prod_{\Lambda_i} d_{\Lambda_i},$$

where the second product is taken over all i-tuples Λ_i of vertex classes of $\mathcal{G}$.

We now show how to deduce Lemma 4.6 from this.

Proof of Lemma 4.6. First we prove the lemma for the case when $\ell = t$, i.e. when each of the vertex classes $X_1, \dots, X_t$ of $\mathcal{H}$ consists of exactly one vertex, say $X_i := \{h_i\}$. Given such an $\mathcal{H}$ and a complex $\mathcal{G}$ as in Lemma 4.6, we construct a complex $\mathcal{G}'$ from $\mathcal{G}$ as follows: Starting with $i = 2$, for all i with $2 \leq i \leq k-1$ in turn, we successively consider each i-tuple $\Lambda_i = (V_{j_1}, \dots, V_{j_i})$ of vertex classes of $\mathcal{G}$. If $h_{j_1}, \dots, h_{j_i}$ forms an i-edge of $\mathcal{H}$ we let $\mathcal{G}'_i[\Lambda_i] = \mathcal{G}_i[\Lambda_i]$. If $h_{j_1}, \dots, h_{j_i}$ does not form an i-edge we make each copy of $K_i^{(i-1)}$ in $\mathcal{G}'_{i-1}[\Lambda_i]$ into an i-edge of $\mathcal{G}'_i$. Thus in the latter case the density of $\mathcal{G}'_i[\Lambda_i]$ with respect to $\mathcal{G}'_{i-1}[\Lambda_i]$ will be 1. (If $i = 2$, this means that we let $\mathcal{G}'_i[\Lambda_i]$ be the complete bipartite graph with vertex classes V_{j_1} and V_{j_2}.) Using that $\mathcal{H}$ is a complex, it is easy to see that $\mathcal{G}'$ is also $(\geq (d_{k-1}, \dots, d_2), \delta, \delta, 1)$-regular. Clearly, there is a bijection between the copies of $\mathcal{H}$ in $\mathcal{G}$ and the copies of $K_t^{(k-1)}$ in $\mathcal{G}'$. So $|\mathcal{H}|_{\mathcal{G}} = |K_t^{(k-1)}|_{\mathcal{G}'}$. The result now follows if we apply Lemma 4.12 to $\mathcal{G}'$.

It now remains to deduce Lemma 4.6 for arbitrary ℓ-partite complexes $\mathcal{H}$ from the result for the above case. For this, we use a simple argument that was also used in [17] to obtain Lemma 4.4 in the case $k = 3$. We define a complex $\mathcal{G}^*$ from $\mathcal{G}$ by making $|X_i|$ copies $V_i^1, \dots, V_i^{|X_i|}$ of each vertex class V_i in such a way that for any selection of indices $i_1, \dots, i_t$ the complex $\mathcal{G}^*[V_1^{i_1}, \dots, V_t^{i_t}]$ is isomorphic to $\mathcal{G}$. Note that $\mathcal{G}^*$ is $|\mathcal{H}|$-partite. Also, we can turn $\mathcal{H}$ into an $|\mathcal{H}|$-partite complex $\mathcal{H}^*$ by viewing each vertex as a single vertex class. Note that different copies of $\mathcal{H}$ in $\mathcal{G}$ give rise to different copies of $\mathcal{H}^*$ in $\mathcal{G}^*$. Thus $|\mathcal{H}|_{\mathcal{G}} \leq |\mathcal{H}^*|_{\mathcal{G}^*}$. Conversely, the only case where a copy of $\mathcal{H}^*$ in $\mathcal{G}^*$ does not correspond to a copy of $\mathcal{H}$ in $\mathcal{G}$ is when there is some i and indices $j_1 \neq j_2$ such that the vertices that are used by $\mathcal{H}^*$ in $V_i^{j_1}$ and $V_i^{j_2}$ correspond to the same vertex of V_i. It is easy to see that the number of such copies is comparatively small. Thus the desired bounds on $|\mathcal{H}|_{\mathcal{G}}$ immediately follow from the bounds on $|\mathcal{H}^*|_{\mathcal{G}^*}$ which we obtained in the previous paragraph. □

We now prove Lemma 4.4. Its proof is based on the following version of the counting lemma that accompanies the hypergraph regularity lemma (Theorem 4.9) from [64]. Theorem 4.13 gives a lower bound on the number of complete complexes $K_t^{(k)}$ in a regular (k,t)-complex $\mathcal{G}$, under less restrictive assumptions on the regularity constants than those in Lemma 4.12.

Theorem 4.13 (Counting lemma for complete complexes [65]) *Let* $k, r, t,$ n_0 *be positive integers and let* $\varepsilon, d_2, \dots, d_k, \delta, \delta_k$ *be positive constants such that* $1/d_i \in \mathbb{N}$ *for* $i = 2, \dots, k-1$ *and*

$$1/n_0 \ll 1/r, \delta \ll \min\{\delta_k, d_2, \dots, d_{k-1}\} \leq \delta_k \ll \varepsilon, d_k, 1/t.$$

Then the following holds for all integers $n \geq n_0$. *Suppose that* $\mathcal{G}$ *is a* $(\mathbf{d}, \delta_k, \delta, r)$*-regular* (k,t)*-complex with vertex classes* $V_1, \dots, V_t$*, all of size* n*, which respects the partition of* $K_t^{(k)}$. *Then*

$$|K_t^{(k)}|_{\mathcal{G}} \geq (1-\varepsilon) n^t \prod_{i=2}^{k} d_i^{\binom{k}{i}}.$$

Lemma 4.4 is more general in the sense that it counts copies of complexes that may not be complete, and also gives an upper bound on their number. We will deduce

Lemma 4.4 from Theorem 4.13 in several steps. The first (and main) step is to deduce a counting lemma which gives the number of copies of complete complexes, but now in a (k, t)-complex $\mathcal{G}$ where the density of $\mathcal{G}_i[\Lambda_i]$ with respect to $\mathcal{G}_{i-1}[\Lambda_i]$ might be different for different i-tuples Λ_i of vertex classes of $\mathcal{G}$.

Lemma 4.14 (Counting lemma for complete complexes – different densities)

Let k, r, t, n_0 be positive integers and let $\varepsilon, d_2, \ldots, d_k, \delta, \delta_k$ be positive constants such that

$$1/n_0 \ll 1/r, \delta \ll \min\{\delta_k, d_2, \ldots, d_{k-1}\} \le \delta_k \ll \varepsilon, d_k, 1/t.$$

Then the following holds for all integers $n \ge n_0$. Suppose $\mathcal{G}$ is a $(\ge \mathbf{d}, \delta_k, \delta, r)$-regular (k, t)-complex with vertex classes $V_1, \ldots, V_t$, all of size n, such that for all $2 < i < k$ and all i-tuples Λ_i of vertex classes of $\mathcal{G}$ the hypergraph $\mathcal{G}_i[\Lambda_i]$ is (d_{Λ_i}, δ)-regular with respect to $\mathcal{G}_{i-1}[\Lambda_i]$ where d_{Λ_i} can be written as $d_{\Lambda_i} = p_{\Lambda_i}/q_{\Lambda_i}$ such that $p_{\Lambda_i}, q_{\Lambda_i} \in \mathbb{N}$ and $1/q_{\Lambda_i} \ge d_i$. Suppose that the analogue holds for all the d_{Λ_2} and all the d_{Λ_k}. Then

$$|K_t^{(k)}|_{\mathcal{G}} \ge (1-\varepsilon) n^t \prod_{i=2}^{k} \prod_{\Lambda_i} d_{\Lambda_i},$$

where the second product is taken over all i-tuples Λ_i of vertex classes of $\mathcal{G}$.

Proof. We will prove this lemma by an inductive argument, in which we allow for different densities in the top levels but not in the lower levels, and show that we can always move down another level, until we allow different densities in all levels. This leads to the following definition. For any $2 < j \le k$, we say that a complex $\mathcal{G}$ is $(\ge d_k, \ldots, \ge d_j, d_{j-1}, \ldots, d_2, \delta_k, \delta, r)$-*regular* if

- $\mathcal{G}_k$ is $(\ge d_k, \delta_k, r)$-regular with respect to $\mathcal{G}_{k-1}$;
- $\mathcal{G}_i$ is $(\ge d_i, \delta)$-regular with respect to $\mathcal{G}_{i-1}$ for each $j \le i \le k-1$;

- $\mathcal{G}_i$ is (d_i, δ)-regular with respect to $\mathcal{G}_{i-1}$ for each $3 \le i \le j-1$;
- $\mathcal{G}_2$ is (d_2, δ)-regular.

Choose new constants $\eta_i, \xi_i, \varepsilon_i$ and integers r_i satisfying

$$1/n_0 \ll \delta = \xi_2 \ll \cdots \ll \xi_k \ll \min\{\delta_k, d_2, \dots, d_{k-1}\} \le \delta_k = \eta_2 \ll \cdots \ll \eta_{k+1} \\ \ll \varepsilon_k \ll \cdots \ll \varepsilon_2 = \varepsilon, d_k, 1/t$$

and $1/n_0 \ll 1/r = 1/r_2 \ll \cdots \ll 1/r_k \ll \min\{\delta_k, d_2, \dots, d_{k-1}\}$. Then the following claim immediately implies the lemma:

Claim. *Let $2 \le j \le k$. Suppose that $\mathcal{G}$ satisfies the conditions of Lemma 4.14 but is $(\ge d_k, \dots, \ge d_j, d_{j-1}, \dots, d_2, \eta_j, \xi_j, r_j)$-regular instead of $(\ge \mathbf{d}, \delta_k, \delta, r)$-regular if $j > 2$, where $1/d_i \in \mathbb{N}$ for all $i = 2, \dots, j-1$. Then*

$$|K_t^{(k)}|_{\mathcal{G}} \ge (1-\varepsilon_j) n^t \left(\prod_{i=2}^{j-1} d_i^{\binom{t}{i}} \right) \prod_{i=j}^{k} \prod_{\Lambda_i} d_{\Lambda_i}.$$

We prove this claim by backward induction on j as follows: given a t-partite complex $\mathcal{G}$ which is $(\ge d_k, \dots, \ge d_j, d_{j-1}, \dots, d_2, \eta_j, \xi_j, r_j)$-regular, we will partition the hyperedges of $\mathcal{G}_j$ to obtain several $(\ge d_k, \dots, \ge d_{j+1}, d'_j, d_{j-1}, \dots, d_2, \eta_{j+1}, \xi_{j+1}, r_{j+1})$-regular complexes for some d'_j. We will then apply the lower bound from the induction hypothesis to each of these complexes. Summing over all of them will give the lower bound in the claim.

We first consider the case $j = k$. We will apply the slicing lemma (Lemma 4.8) to split the kth level $\mathcal{G}_k$ of the complex $\mathcal{G}$ to obtain regular complexes whose densities within the kth level are the same. Set $d'_k := 1/\prod_{\Lambda_k} q_{\Lambda_k}$. The slicing lemma implies that for all Λ_k there is a partition $P(\Lambda_k)$ of the set $E(\mathcal{G}_k[\Lambda_k])$ of k-edges induced on Λ_k such that each part is (d'_k, η_{k+1}, r_k)-regular with respect to $\mathcal{G}_{k-1}[\Lambda_k]$. So for each Λ_k, $P(\Lambda_k)$ has d_{Λ_k}/d'_k parts. Now for each Λ_k, choose one part from $P(\Lambda_k)$ and let $\mathcal{C}_k$ denote the resulting k-uniform t-partite hypergraph. Let $\mathcal{G}^{\mathcal{C}_k}$ denote the k-complex obtained from $\mathcal{G}$ by replacing $\mathcal{G}_k$ with $\mathcal{C}_k$. Then

$$|K_t^{(k)}|_{\mathcal{G}} = \sum_{\mathcal{C}_k} |K_t^{(k)}|_{\mathcal{G}^{\mathcal{C}_k}}.$$

Here the summation is over all possible choices of parts from each of the $\binom{t}{k}$ partitions $P(\Lambda_k)$. So the number of summands is $\prod_{\Lambda_k} d_{\Lambda_k}/d'_k = {d'}_k^{-\binom{t}{k}} \prod_{\Lambda_k} d_{\Lambda_k}$. Moreover, by Theorem 4.13 each summand in the above sum can be bounded below:

$$|K_t^{(k)}|_{\mathcal{G}^{\mathcal{C}_k}} \geq (1-\varepsilon_k)n^t \left(\prod_{i=2}^{k-1} d_i^{\binom{t}{i}}\right) {d'}_k^{\binom{t}{k}}.$$

Altogether, this implies the claim for $j = k$.

Now suppose that $j < k$ and that the claim holds for $j+1$. To apply the induction hypothesis, we now need to get equal densities in the jth level. We will achieve this by applying the slicing lemma (Lemma 4.8) to this level. Set $d'_j := 1/\prod_{\Lambda_j} q_{\Lambda_j}$. So $1/d'_j \in \mathbb{N}$. The slicing lemma implies that for every j-tuple Λ_j of vertex classes of $\mathcal{G}$ there is a partition $P(\Lambda_j)$ of the set $E(\mathcal{G}_j[\Lambda_j])$ of j-edges induced on Λ_j such that each part is (d'_j, ξ_{j+1})-regular with respect to $\mathcal{G}_{j-1}[\Lambda_j]$. For each Λ_j, the corresponding partition $P(\Lambda_j)$ will have $a_{\Lambda_j} := d_{\Lambda_j}/d'_j$ parts. Now for each Λ_j, choose one part from $P(\Lambda_j)$ and let $\mathcal{C}_j$ denote the resulting j-uniform t-partite hypergraph. We let $\mathcal{G}^{\mathcal{C}_j}$ denote the (k,t)-complex obtained from $\mathcal{G}$ as follows: we replace $\mathcal{G}_j$ by $\mathcal{C}_j$ and for each $j < i \leq k$ we replace $\mathcal{G}_i$ with the subhypergraph whose i-edges are all those i-sets of vertices that span a $K_i^{(j)}$ in $\mathcal{C}_j$. Thus $\mathcal{G}_j^{\mathcal{C}_j}$ is (d'_j, ξ_{j+1})-regular with respect to $\mathcal{G}_{j-1} = \mathcal{G}_{j-1}^{\mathcal{C}_j}$. However, to apply the induction hypothesis this is not enough. We also need to prove the following more general assertion.

For all $i = j, \dots, k$ and any Λ_i the following holds. If $i = j$ then $\mathcal{G}_i^{\mathcal{C}_j}[\Lambda_i]$ is (d'_j, ξ_{j+1})-regular with respect to $\mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]$. If $j < i < k$ then $\mathcal{G}_i^{\mathcal{C}_j}[\Lambda_i]$ is $(d_{\Lambda_i}, \xi_{j+1})$-regular with respect to $\mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]$. If $i = k$ then $\mathcal{G}_i^{\mathcal{C}_j}[\Lambda_i]$ is $(d_{\Lambda_i}, \eta_{j+1}, r_{j+1})$-regular with respect to $\mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]$ for all but at most $\sqrt{\eta_{j+1}}\prod_{\Lambda_j} a_{\Lambda_j}$ hypergraphs $\mathcal{C}_j$. $(*)$

We will prove $(*)$ by induction on i. If $i = j$ then we already know that the assertion is true. So suppose that $i > j$ and that the claim holds for $i - 1$. We will first consider the case when $i < k$. The induction hypothesis together with the dense counting lemma for complete complexes (Lemma 4.12) implies that

$$|K_i^{(i-1)}|_{\mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]} \geq \frac{1}{2} n^i \left(\prod_{\ell=2}^{j-1} d_\ell^{\binom{i}{\ell}} \right) d'^{\binom{i}{j}}_j \prod_{s=j+1}^{i-1} \prod_{\Lambda_s \subseteq \Lambda_i} d_{\Lambda_s}. \tag{4.19}$$

Similarly, the assumptions on $\mathcal{G}$ in the claim together with Lemma 4.12 imply

$$|K_i^{(i-1)}|_{\mathcal{G}_{i-1}[\Lambda_i]} \leq 2n^i \left(\prod_{\ell=2}^{j-1} d_\ell^{\binom{i}{\ell}} \right) \prod_{s=j}^{i-1} \prod_{\Lambda_s \subseteq \Lambda_i} d_{\Lambda_s}. \tag{4.20}$$

If we combine these inequalities and use the fact that $\xi_j \ll \xi_{j+1} \ll d_j, 1/k$, we obtain

$$|K_i^{(i-1)}|_{\mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]} \geq \sqrt{\xi_{j+1}} |K_i^{(i-1)}|_{\mathcal{G}_{i-1}[\Lambda_i]} \geq \frac{\xi_j}{\xi_{j+1}} |K_i^{(i-1)}|_{\mathcal{G}_{i-1}[\Lambda_i]}. \tag{4.21}$$

In other words, a ξ_{j+1}-proportion of copies of $K_i^{(i-1)}$ in $\mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]$ gives rise to a ξ_j-proportion of copies in $\mathcal{G}_{i-1}[\Lambda_i]$. Moreover, $\mathcal{K}_i(\mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]) \cap E(\mathcal{G}_i^{\mathcal{C}_j}[\Lambda_i]) = \mathcal{K}_i(\mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]) \cap E(\mathcal{G}_i[\Lambda_i])$ by the definition of $\mathcal{G}^{\mathcal{C}_j}$ and so $d(\mathcal{G}_i^{\mathcal{C}_j}[\Lambda_i] | \mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]) = d(\mathcal{G}_i[\Lambda_i] | \mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]) = d_{\Lambda_i} \pm \xi_j$ by (4.21) and the (d_{Λ_i}, ξ_j)-regularity of $\mathcal{G}_i[\Lambda_i]$ with respect to $\mathcal{G}_{i-1}[\Lambda_i]$. Thus the $(d_{\Lambda_i}, \xi_{j+1})$-regularity of $\mathcal{G}_i^{\mathcal{C}_j}[\Lambda_i]$ with respect to $\mathcal{G}_{i-1}^{\mathcal{C}_j}[\Lambda_i]$ follows from the (d_{Λ_i}, ξ_j)-regularity of $\mathcal{G}_i[\Lambda_i]$ with respect to $\mathcal{G}_{i-1}[\Lambda_i]$.

But if $i = k$, this might not be true, as η_{j+1} may not be small compared to d_j. However, given a k-tuple Λ_k of vertex classes of $\mathcal{G}$, it is true for most complexes $\mathcal{G}^{\mathcal{C}_j}[\Lambda_k]$. To see this, given Λ_k, let $\mathcal{B}$ be a (k,k)-complex obtained as follows: For each $\Lambda_j \subset \Lambda_k$, choose one part from $P(\Lambda_j)$ and let $\mathcal{B}_j$ denote the resulting j-uniform k-partite hypergraph. To obtain $\mathcal{B}$ from $\mathcal{G}[\Lambda_k]$, we replace $\mathcal{G}_j[\Lambda_k]$ by $\mathcal{B}_j$ and for each $j < i \leq k$ we replace $\mathcal{G}_i[\Lambda_k]$ with the subhypergraph whose i-edges are all those i-sets of vertices which span a $K_i^{(j)}$ in $\mathcal{B}_j$. Thus there are $\prod_{\Lambda_j \subset \Lambda_k} a_{\Lambda_j} =: A_{\Lambda_k}$ such complexes $\mathcal{B}$. (Recall that $a_{\Lambda_j} = d_{\Lambda_j}/d'_j$ was the number of parts of the partition $P(\Lambda_j)$.) Using that $(*)$ holds for all $i < k$, similarly as in (4.19)–(4.21) one can show that

$$|K_k^{(k-1)}|_{\mathcal{B}_{k-1}} \geq \frac{d_j'^{\binom{k}{j}}}{4\prod_{\Lambda_j \subset \Lambda_k} d_{\Lambda_j}} |K_k^{(k-1)}|_{\mathcal{G}_{k-1}[\Lambda_k]} = \frac{|K_k^{(k-1)}|_{\mathcal{G}_{k-1}[\Lambda_k]}}{4A_{\Lambda_k}}. \tag{4.22}$$

We will now prove the following:

The underlying k-uniform hypergraph $\mathcal{B}_k$ is not $(d_{\Lambda_k}, \eta_{j+1}, r_{j+1})$-regular with respect to $\mathcal{B}_{k-1}$ for less than $\eta_{j+1}A_{\Lambda_k}$ of the complexes $\mathcal{B}$. $(**)$

If $(**)$ is false then we can find $T := \eta_{j+1}A_{\Lambda_k}/2$ such complexes $\mathcal{B}^1, \dots, \mathcal{B}^T$, such that each $\mathcal{B}^\ell$ has a $\mathbf{Q}^\ell = (Q_1^\ell, \dots, Q_{r_{j+1}}^\ell)$ satisfying $Q_s^\ell \subseteq \mathcal{B}_{k-1}^\ell$ for all $s = 1, \dots, r_{j+1}$ and $|K_k^{(k-1)}|_{\mathbf{Q}^\ell} \geq \eta_{j+1}|K_k^{(k-1)}|_{\mathcal{B}_{k-1}^\ell}$, but either $d(\mathcal{B}_k^\ell|\mathbf{Q}^\ell) > d_{\Lambda_k} + \eta_{j+1}$ for each ℓ or $d(\mathcal{B}_k^\ell|\mathbf{Q}^\ell) < d_{\Lambda_k} - \eta_{j+1}$ for each ℓ. We will assume the latter – the proof in the former case is similar. But then let $\mathbf{Q} = (\mathbf{Q}^1, \mathbf{Q}^2, \dots, \mathbf{Q}^T)$. Thus $\mathbf{Q}$ is a Tr_{j+1}-tuple and

$$|K_k^{(k-1)}|_{\mathbf{Q}} \geq \sum_{\ell=1}^{T} \eta_{j+1}|K_k^{(k-1)}|_{\mathcal{B}_{k-1}^\ell} \overset{(4.22)}{\geq} \eta_j |K_k^{(k-1)}|_{\mathcal{G}_{k-1}[\Lambda_k]}.$$

Since we may assume that $Tr_{j+1} \leq r_j$ our assumption on the regularity of $\mathcal{G}_k[\Lambda_k]$ with respect to $\mathcal{G}_{k-1}[\Lambda_k]$ implies that $d(\mathcal{G}_k[\Lambda_k]|\mathbf{Q}) \geq d_{\Lambda_k} - \eta_j$. On the other hand, the definition of $\mathcal{B}$ implies that $d(\mathcal{B}_k^\ell|\mathbf{Q}^\ell) = d(\mathcal{G}_k[\Lambda_k]|\mathbf{Q}^\ell)$. Thus $d(\mathcal{G}_k[\Lambda_k]|\mathbf{Q}) \leq$

$\max_{1\le\ell\le T} d(\mathcal{G}_k[\Lambda_k]|\mathbf{Q}^\ell) = \max_{1\le\ell\le T} d(\mathcal{B}^\ell_k|\mathbf{Q}^\ell) < d_{\Lambda_k} - \eta_{j+1}$. This is a contradiction, and so $(**)$ holds.

Note that $(**)$ implies that for all but at most $\binom{t}{k}\eta_{j+1}\prod_{\Lambda_j} a_{\Lambda_j}$ hypergraphs $\mathcal{C}_j$ the hypergraph $\mathcal{G}_k^{\mathcal{C}_j}$ is $(d_{\Lambda_k}, \eta_{j+1}, r_{j+1})$-regular with respect to $\mathcal{G}_{k-1}^{\mathcal{C}_j}$ – we call these $\mathcal{C}_j$ *nice*. Since $\eta_{j+1} \ll 1/t$, this completes the proof of $(*)$.

We are now ready to finish the proof of the induction step of the claim. The induction

$$|K_t^{(k)}|_{\mathcal{G}} \;\; \geq \;\; \sum_{\text{nice } \mathcal{C}_j} |K_t^{(k)}|_{\mathcal{G}^{\mathcal{C}_j}} \geq (1-\varepsilon_{j+1}) \sum_{\text{nice } \mathcal{C}_j} n^t \left(\prod_{i=2}^{j-1} d_i^{\binom{t}{i}}\right) d'^{\binom{t}{j}}_j \prod_{i=j+1}^{k} \prod_{\Lambda_i} d_{\Lambda_i}.$$

The summation is over all possible choices of nice $\mathcal{C}_j$. So the number of summands is at least $(1-\sqrt{\eta_{j+1}})\prod_{\Lambda_j} a_{\Lambda_j}$ and for each Λ_j we have $a_{\Lambda_j} d'_j = d_{\Lambda_j}$. Since $\eta_{j+1}, \varepsilon_{j+1} \ll \varepsilon_j$, the claim follows and hence the lower bound in Lemma 4.14 as well. □

Instead of making use of the parameter r in the proof of $(*)$ we could have also used the fact that the partitions guaranteed by the slicing lemma are obtained by considering random partitions.

It is straightforward to obtain a corresponding upper bound from the lower bound in Lemma 4.14.

Lemma 4.15 (Counting lemma for complete complexes – upper bound) *Under the conditions of Lemma 4.14,*

$$|K_t^{(k)}|_{\mathcal{G}} = (1 \pm \varepsilon) n^t \prod_{i=2}^{k} \prod_{\Lambda_i} d_{\Lambda_i}.$$

Proof. Clearly, all we have to prove is the upper bound. The proof is based on an argument that was used in [58] and later in [17] to derive a similar upper bound in the case of 3-complexes. Let $[t]^k$ denote the set of all k-subsets of $[t] = \{1, \ldots, t\}$. Given $S \subseteq [t]^k$, we let $\mathcal{G}^S$ denote the (k,t)-complex obtained from $\mathcal{G}$ as follows: for each $\{i_1, \ldots, i_k\} \in S$ we replace the set $E_k(\mathcal{G}[\Lambda_k])$ of all k-edges of $\mathcal{G}$ induced on $\Lambda_k := \{V_{i_1}, \ldots, V_{i_k}\}$ by $\mathcal{K}_k(\mathcal{G}_{k-1}[\Lambda_k]) \setminus E_k(\mathcal{G}[\Lambda_k])$. Thus the density of $\mathcal{G}_k^S[\Lambda_k]$ with respect to $\mathcal{G}_{k-1}^S[\Lambda_k]$ is now $1 - d_{\Lambda_k}$. Moreover,

$$|K_t^{(k-1)}|_{\mathcal{G}_{k-1}} = \sum_{S \subseteq [t]^k} |K_t^{(k)}|_{\mathcal{G}^S}.$$

Observe that $|K_t^{(k)}|_{\mathcal{G}} = |K_t^{(k)}|_{\mathcal{G}^\emptyset}$ and hence

$$|K_t^{(k)}|_{\mathcal{G}} = |K_t^{(k-1)}|_{\mathcal{G}_{k-1}} - \sum_{S \subseteq [t]^k, S \neq \emptyset} |K_t^{(k)}|_{\mathcal{G}^S}.$$

Thus, to obtain an upper bound on $|K_t^{(k)}|_{\mathcal{G}}$ all we have to do now is to obtain an upper bound on $|K_t^{(k-1)}|_{\mathcal{G}_{k-1}}$ and a lower bound on $|K_t^{(k)}|_{\mathcal{G}^S}$, for every non-empty S.

But the former follows from the dense counting lemma (Lemma 4.6) and the latter follows from Lemma 4.14 above. (This is why in Lemma 4.14 we need to allow more general densities than just $1/a$, for $a \in \mathbb{N}$.)

We first fix a constant ε' such that $\delta_k \ll \varepsilon' \ll \varepsilon, d_j, 1/t$. Note that if ε' replaces ε, this hierarchy is more restrictive than that which is required by Lemma 4.6, and so we can apply Lemma 4.6 with ε' playing the role of ε.

One technical difficulty in these calculations is that we wish to apply Lemma 4.14 to a new hypergraph $\mathcal{G}^S$. However, the k-tuples of $\mathcal{G}^S$ corresponding to hyperedges in S now have density roughly $1 - d_e$, which may not be greater than d_k. Indeed, if $d_e(S)$ denotes the density of the k-tuple corresponding to the hyperedge e in $\mathcal{G}^S$ (so $d_e(S) = d_e$ if $e \notin S$ and $d_e(S) = 1 - d_e$ otherwise), there may not be any $d_k' = d_k'(S) \gg \delta_k$ such that $d_e(S) \geq d_k'$ for every e.

To overcome this difficulty, we first consider the set E_k' of all those hyperedges $e \in E_k(\mathcal{H})$ for which $1 - d_e \leq \varepsilon'$. We would like to be able to apply Lemma 4.14 to obtain the lower bound of

$$(1-\varepsilon')d_{k-1}(\mathcal{H})n^t \prod_{e\in E_k(\mathcal{H})} d_e(S)$$

for $|\mathcal{H}|_{\mathcal{G}^S}$ for any $\mathcal{D} \neq \phi$, where $d_{k-1}(\mathcal{H}) := \prod_{i=2}^{k-1} d_i^{e_i(\mathcal{H})}$. However as indicated above, this is not possible. Instead, we use the trivial lower bound $|\mathcal{H}|_{\mathcal{G}^S} \geq 0$, and we observe that if $S \cap E'_k \neq \phi$, then

$$(1-\varepsilon')d_{k-1}(\mathcal{H})n^t \prod_{e\in E_k(\mathcal{H})} d_e(S) \leq \varepsilon' d_{k-1}(\mathcal{H})n^t$$

so

$$|\mathcal{H}|_{\mathcal{G}^S} \geq 0 \geq (1-\varepsilon')d_{k-1}(\mathcal{H})n^t \prod_{e\in E_k(\mathcal{H})} d_e(S) - \varepsilon' d_{k-1}(\mathcal{H})n^t.$$

since $d_e(S) \leq \varepsilon'$ for some $e \in E_j(\mathcal{H})$. So by adding on an error term of $2^{\binom{t}{k}}\varepsilon'\left(\prod_{i=2}^{k-1} d_i^{e_i(\mathcal{H})}\right)n^t$, we may assume that for each S, $\mathcal{G}^S$ is $(d'_k, d_{k-1}, \ldots, d_2, \delta_k, \delta, r)$-regular, where $d'_k = \min\{d_k, \varepsilon'\}$. (The term $2^{\binom{t}{k}}$ is the number of possible sets S, a trivial upper bound on the number of S's satisfying $S \cap E'_k \neq \phi$.) We will now make this assumption and compensate for this by adding the above error term at the end of the argument.

Note that replacing ε, d_k with ε', d'_k, we still have constants satisfying the hierarchy of Lemma 4.6.

Recall that

$$|\mathcal{H}|_{\mathcal{G}} = |\mathcal{H}^{(k-1)}|_{\mathcal{G}} - \sum_{S\subseteq E_k(\mathcal{H}), S\neq\phi} |\mathcal{H}|_{\mathcal{G}^S}.$$

By the dense counting lemma for $(k-1)$-complexes,

$$|\mathcal{H}^{(k-1)}|_{\mathcal{G}} \leq (1+\varepsilon')\left(\prod_{i=2}^{k-1} d_i^{e_i(\mathcal{H})}\right)n^t.$$

Furthermore, by Lemma 4.14, for each $S \neq \phi$,

$$|\mathcal{H}|_{\mathcal{G}^S} \geq (1-\varepsilon')\left(\prod_{i=2}^{k-1} d_i^{e_i(\mathcal{H})}\right) n^t \prod_{e\in E_k(\mathcal{H})} d_e(S).$$

Now

$$\sum_{S\subseteq E_k(\mathcal{H}), S\neq\phi} \prod_{e\in E_k(\mathcal{H})} d_e(S) = 1 - \prod_{e\in E_k(\mathcal{H})} d_e$$

and so

$$\begin{aligned}
|\mathcal{H}|_{\mathcal{G}} &= |\mathcal{H}^{(k-1)}|_{\mathcal{G}} - \sum_{S\subseteq E_k(\mathcal{H}), S\neq\phi} |\mathcal{H}|_{\mathcal{G}^S} \\
&\leq d_{k-1}(\mathcal{H}) n^t \left(1+\varepsilon' - (1-\varepsilon')(1 - \prod_{e\in E_k(\mathcal{H})} d_e)\right) \\
&\leq d_{k-1}(\mathcal{H}) n^t \left(2\varepsilon' + \prod_{e\in E_k(\mathcal{H})} d_e\right)
\end{aligned}$$

Hence we obtain

$$|\mathcal{H}|_{\mathcal{G}} \leq d_{k-1}(\mathcal{H}) n^t (1+\sqrt{\varepsilon'}) \prod_{e\in E_k(\mathcal{H})} d_e.$$

Now it only remains to add back in the error term which we obtained from our assumption that each $\mathcal{G}^S$ was $(d'_k, d_{k-1}, \ldots, d_2, \delta_j, \delta, r)$-regular. Observe that $2^{\binom{t}{k}}\varepsilon' \leq \sqrt{\varepsilon'}\prod_{e\in E_k(\mathcal{H})} d_e$. Thus the upper bound on the number of copies of $\mathcal{H}$ in $\mathcal{G}$ becomes

$$|\mathcal{H}|_{\mathcal{G}} \leq (1+2\sqrt{\varepsilon'}) d_{k-1}(\mathcal{H}) n^t \prod_{e\in E_k(\mathcal{H})} d_e \leq (1+\varepsilon) d_{k-1}(\mathcal{H}) n^t \prod_{e\in E_k(\mathcal{H})} d_e$$

as required. □

Lemma 4.4 now follows from Lemma 4.14 in exactly the same way as Lemma 4.6 followed from Lemma 4.12.

4.8 Proof of the Extension Lemmas 4.5 and 4.7

We now use Lemma 4.4 to derive Lemma 4.5 (Lemma 4.7 can be derived in the same way from Lemma 4.6). The proof idea is similar to that of [65, Cor. 14], [32, Lemma 6.6] and [17, Lemma 5]. Pick a copy H of $\mathcal{H}$ in $\mathcal{G}$ uniformly at random, and define $X := |H \to \mathcal{H}'|$. Then X is a random variable. We have $\mathbb{E}(X) = \frac{1}{|\mathcal{H}|_{\mathcal{G}}} \sum_{H \in \mathcal{G}} |H \to \mathcal{H}'| = |\mathcal{H}'|_{\mathcal{G}} / |\mathcal{H}|_{\mathcal{G}}$. (Here the sum $\sum_{H \in \mathcal{G}}$ is over all copies of $\mathcal{H}$ in $\mathcal{G}$.) We pick some constant ε satisfying $\delta_k \ll \varepsilon \ll \beta$. By applying the upper bound of the counting lemma (Lemma 4.4) to $\mathcal{H}$ and the lower bound to $\mathcal{H}'$ we obtain a lower bound for $\mathbb{E}(X)$. Similarly we obtain an upper bound. In this way we can easily deduce that

$$\mathbb{E}(X) = (1 \pm \sqrt{\varepsilon}) \overline{|\mathcal{H} \to \mathcal{H}'|}. \tag{4.23}$$

Now consider $\mathbb{E}(X^2)$. We aim to show that its value is approximately $\overline{|\mathcal{H} \to \mathcal{H}'|}^2$, and so X has a low variance. Using Chebyshev's inequality, this will then imply that X is concentrated around its mean. In other words, only a few copies of $\mathcal{H}$ do not extend to the correct number of copies of $\mathcal{H}'$ in $\mathcal{G}$.

Observe that $\mathbb{E}(X^2) = \frac{1}{|\mathcal{H}|_{\mathcal{G}}} \sum_{H \in \mathcal{G}} |H \to \mathcal{H}'|^2$. We view $|H \to \mathcal{H}'|^2$ as the number of pairs H_1', H_2' of copies of $\mathcal{H}'$ which extend H. Here the pairs are allowed to overlap, but we first obtain a rough estimate by insisting that they intersect precisely in H. So let $\mathcal{H}^*$ be the (k, ℓ)-complex obtained from two disjoint copies of $\mathcal{H}'$ by identifying them on $\mathcal{H}$. Thus any copy of $\mathcal{H}^*$ in $\mathcal{G}$ extending H corresponds to a pair H_1', H_2'. However, we will later need to take account of those pairs H_1', H_2' which do not arise from a copy of $\mathcal{H}^*$. These pairs are exactly those whose intersection is strictly larger than H.

By applying the counting lemma (Lemma 4.4) to $\mathcal{H}^*$ and to $\mathcal{H}$, as before we obtain

$$\frac{1}{|\mathcal{H}|_{\mathcal{G}}}\sum_{H\in\mathcal{G}}|H\to\mathcal{H}^*| = (1\pm\sqrt{\varepsilon})\overline{|\mathcal{H}\to\mathcal{H}^*|} = (1\pm\sqrt{\varepsilon})\overline{|\mathcal{H}\to\mathcal{H}'|}^2.$$

On the other hand, the number of pairs H'_1, H'_2 which do not arise from a copy of $\mathcal{H}^*$ is at most $(t'-t)^2 n^{2(t'-t)-1} < \varepsilon((\prod_{i=2}^{k} d_i^{e_i(\mathcal{H}')-e_i(\mathcal{H})})n^{t'-t})^2 = \varepsilon\overline{|\mathcal{H}\to\mathcal{H}'|}^2$. Thus

$$\frac{1}{|\mathcal{H}|_{\mathcal{G}}}\sum_{H\in\mathcal{G}}|H\to\mathcal{H}'|^2 = (1\pm 2\sqrt{\varepsilon})\overline{|\mathcal{H}\to\mathcal{H}'|}^2. \tag{4.24}$$

Putting (4.23) and (4.24) together, we obtain

$$var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 < 5\sqrt{\varepsilon}\overline{|\mathcal{H}\to\mathcal{H}'|}^2.$$

Now recall Chebyshev's inequality: $\mathbb{P}(|X-\mathbb{E}(X)|\geq t)\leq var(X)/t^2$. We apply this inequality with $t := \beta\overline{|\mathcal{H}\to\mathcal{H}'|}$. This implies that the probability that a randomly chosen copy of $\mathcal{H}$ in $\mathcal{G}$ does not satisfy the conclusion of the extension lemma is at most $var(X)/\beta^2\overline{|\mathcal{H}\to\mathcal{H}'|}^2 < 5\sqrt{\varepsilon}/\beta^2 < \beta$, and so at most $\beta|\mathcal{H}|_{\mathcal{G}}$ copies of $\mathcal{H}$ do not satisfy the conclusion, as required.

LIST OF REFERENCES

[1] S. Abbasi, *The solution of the El-Zahar problem*, Ph.D. Thesis, Rutgers University 1998.

[2] S. Abbasi, How tight is the Bollobás-Komlós conjecture?, *Graphs Combin.* **16** (2000), 129-137.

[3] M. Ajtai, J. Komlós, E. Szemerédi, On a Conjecture of Loebl, *Proceedings of the 7th International Conference on Graph Theory, Combinatorics, and Algorithms*, Wiley, New York (1995), 1135–1146.

[4] N. Alon and R. Yuster, H-factors in dense graphs, *J. Combin. Theory B* **66** (1996), 269–282.

[5] O. Barr, R. Johansson, Another Note on the Loebl-Komlós-Sós conjecture, *Research reports* **22**, (1997), Umeå University, Sweden.

[6] C. Bazgan, H. Li, M. Woźniak, On the Loebl-Komlós-Sós conjecture, *Journal of Graph Theory,* **34** (2000) 269-276.

[7] B. Bollobás, S. Eldridge, Packing of graphs and applications to computational complexity, *J. Combin. Theory B* **25** (1978), 105-124.

[8] J. Böttcher, M. Schacht, A. Taraz, The bandwidth conjecture for 3-colourable graphs, *Journal of Combinatorial Theory B* **98** (2008): 752–777.

[9] J. Böttcher, M. Schacht, A. Taraz, Proof of the bandwidth conjecture of Bollobás and Komlós, *Math. Ann.*, 343 **1** (2009), 175–205.

[10] S.A. Burr and P. Erdős, On the magnitude of generalized Ramsey numbers for graphs. In *Infinite and Finite Sets, Colloquia Mathematica Societatis János Bolyai vol. 10* **1** (1975), 214–240.

[11] G. Chen and R. Schelp, Graphs with linearly bounded Ramsey numbers, *J. Combinatorial Theory B* **57** (1993), 138–149.

[12] V. Chvátal, V. Rödl, E. Szemerédi and W.T. Trotter, Jr., The Ramsey number of a graph with a bounded maximum degree, *J. Combinatorial Theory B* **34** (1983), 239–243.

[13] D. Conlon, J. Fox and B. Sudakov, Ramsey numbers of sparse hypergraphs, *Random Structures & Algorithms*, **35** (2009), 1-14.

[14] D. Conlon, J. Fox and B. Sudakov, Hypergraph Ramsey numbers, *Journal of the American Mathematical Society*, to appear.

[15] O. Cooley, Proof of the Loebl-Komlós-Sós Conjecture for large, dense graphs, *Discrete Mathematics*, to appear.

[16] O. Cooley, D. Kühn, D. Osthus, *Perfect packings with complete graphs minus an edge*, European J. Combinatorics 28 (2007), 2143–2155.

[17] O. Cooley, N. Fountoulakis, D. Kühn, D. Osthus, 3-uniform hypergraphs of bounded degree have linear Ramsey numbers, *J. Combinatorial Theory Series B* 98 (2008), 484–505.

[18] O. Cooley, N. Fountoulakis, D. Kühn, D. Osthus, Embeddings and Ramsey numbers of sparse k-uniform hypergraphs, *Combinatorica* **29** (2009), 263–297.

[19] K. Corrádi and A. Hajnal, On the maximal number of independent circuits in a graph, *Acta Math. Acad. Sci. Hungar.* **14** (1963), 423–439.

[20] B. Csaba, On the Bollobás-Eldridge conjecture for bipartite graphs, *Combinatorics, Probability & Computing* 16 **5** (2007) 661–691.

[21] D. de Caen, Extension of the theorem of Moon and Moser on complete subgraphs, *Ars Combin.* **16** (1983), 5-10.

[22] R. Diestel, *Graph Theory* (third edition) Springer-Verlag, (2005).

[23] E. Dobson, Constructing trees in graphs whose complement has no $K_{2,s}$, *Combinatorics, Probability and Computing* **11**, Issue 4, (2002), 343-347.

[24] P. Dorbec, S. Gravier, G. Sárközy, Monochromatic Hamiltonian t-tight Berge-cycles in Hypergraphs, *Journal of Graph Theory* 59, (2008), 34–44.

[25] H. Enomoto, A. Kaneko and Z. Tuza, P_3-factors and covering cycles in graphs of minimum degree $n/3$, *Combinatorics* (A. Hajnal, L. Lovász and V. Sós eds.), Colloq. Math. Soc. J. Bolyai **52**, Eger (Hungary) (1987), 213–220.

[26] P. Erdős, Z. Füredi, M. Loebl, V. T. Sós, Discrepancy of Trees, *Studia Sci. Math. Hungar.* **30** (1995), 47–57.

[27] P. Erdős and R. Rado, Combinatorial theorems on classifications of subsets of a given set, *Proc. London Mathematical Society* **3** (1952), 417–439.

[28] P. Frankl and V. Rödl, Extremal problems on set systems, *Random Structures & Algorithms* **20** (2002), 131–164.

[29] L. Gerencsér, A. Gyarfas, On Ramsey-type problems, *Ann. Univ. Sci. Budapest Eötvös Sect. Math.* **10** (1967), 167-170.

[30] S. Gerke, A. Steger, The sparse regularity lemma and its applications, *Surveys in Combinatorics 2005*, B. Webb ed., Cambridge University Press (2005), 227–258.

[31] W.T. Gowers, Hypergraph regularity and the multidimensional Szemerédi theorem, *Ann. of Math.* **166** (2007), 897–946.

[32] W.T. Gowers, Quasirandomness, counting and regularity for 3-uniform hypergraphs, *Combinatorics, Probability & Computing* **15** (2006), 143–184.

[33] R.L. Graham, B.L. Rothschild and J.H. Spencer, *Ramsey Theory*, John Wiley & Sons, 1980.

[34] R.L. Graham, V. Rödl and A. Ruciński, On graphs with linear Ramsey numbers, *J. Graph Theory* **35** (2000), 176-192.

[35] A. Gyárfás, J. Lehel, G.N. Sárközy and R. Schelp, Monochromatic Hamiltonian Berge-cycles in colored complete uniform hypergraphs, *J. Combinatorial Theory B* **98** (2008), 342–358.

[36] A. Hajnal and E. Szemerédi, Proof of a conjecture of Erdős, *Combinatorial Theory and its Applications* Vol. II (P. Erdős, A. Rényi and V. Sós eds.), Colloq. Math. Soc. J. Bolyai **4**, North-Holland, Amsterdam (1970), 601–623.

[37] P.E. Haxell, T. Łuczak, Y. Peng, V. Rödl, A. Ruciński, M. Simonovits and J. Skokan, The Ramsey number for hypergraph cycles I, *J. Combinatorial Theory A* **113** (2006), 67–83.

[38] P.E. Haxell, T. Łuczak, Y. Peng, V. Rödl, A. Ruciński, and J. Skokan, The Ramsey Number for 3-Uniform Tight Hypergraph Cycles, *Combinatorics, Probability, Computing* **18** (1-2) (2009), 165–203.

[39] P.E. Haxell, B. Nagle and V. Rödl, Integer and fractional packings in dense 3-uniform hypergraphs, *Random Structures & Algorithms* **22** (2003), 248–310.

[40] J. Hladky, D. Piguet, Loebl-Komlós-Sós Conjecture: dense case, *preprint*.

[41] Y. Ishigami, Linear Ramsey numbers for bounded-degree hypergraphs, *preprint*.

[42] K. Kawarabayashi, K_4^--factors in a graph, *J. Graph Theory* **39** (2002), 111–128.

[43] P. Keevash, A hypergraph blowup lemma, *preprint*.

[44] Y. Kohayakawa, V. Rödl and J. Skokan, Hypergraphs, quasi-randomness, and conditions for regularity, *J. Combinatorial Theory A* **97** (2002), 307–352.

[45] J. Komlós, The Blow-up lemma, *Comb. Probab. Comput.* **8** (1999), 161–176.

[46] J. Komlós, Tiling Turán theorems, *Combinatorica* **20** (2000), 203–218.

[47] J. Komlós, G. Sárkőzy and E. Szemerédi, The blow-up lemma, *Combinatorica* **17** (1997), 109–123.

[48] J. Komlós, G. N. Sárközy and E. Szemerédi, Proof of the Alon-Yuster conjecture, *Discrete Math.* **235** (2001), 255–269.

[49] J. Komlós, G. N. Sárközy and E. Szemerédi, On the Pósa-Seymour conjecture, *Journal of Graph Theory* **29** (1998), 167-176.

[50] J. Komlós, G. N. Sárközy and E. Szemerédi, Proof of the Seymour Conjecture for large graphs, *Ann. Comb.* **2**, (1998), 43-60.

[51] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, *Bolyai Society Mathematical Studies 2, Combinatorics, Paul Erdős is Eighty (Vol. 2)* (D. Miklós, V. T. Sós and T. Szőnyi eds.), Budapest (1996), 295–352.

[52] A. Kostochka and V. Rödl, On Ramsey numbers of uniform hypergraphs with given maximum degree, *J. Combinatorial Theory A* **113** (2006), 1555–1564.

[53] D. Kühn and D. Osthus, The minimum degree threshold for perfect graph packings, *Combinatorica*, **29** (2009), 65–107.

[54] D. Kühn and D. Osthus, Critical chromatic number and the complexity of perfect packings in graphs, *17th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, (2006), 851–859.

[55] D. Kühn and D. Osthus, Loose Hamilton cycles in 3-uniform hypergraphs of large minimum degree, *J. Combinatorial Theory B* 96 (2006), 767–821.

[56] A. McLennan, The Erdős-Sós conjecture for trees of diameter four, *Journal of Graph Theory* **49** (2005), 291-301.

[57] B. Nagle, S. Olsen, V. Rödl and M. Schacht, On the Ramsey number of sparse 3-graphs, *Graphs and Combinatorics* **24** (2008), 205–228.

[58] B. Nagle and V. Rödl, Regularity properties for triple systems, *Random Structures & Algorithms* **23** (2003), 264–332.

[59] B. Nagle, V. Rödl and M. Schacht, Note on the 3-graph counting lemma, *Discrete Mathematics*, 308 **19** 4501–4517.

[60] B. Nagle, V. Rödl and M. Schacht, The counting lemma for k-uniform hypergraphs, *Random Structures & Algorithms* **28** (2006), 113–179.

[61] D. Piguet, M. J. Stein, An approximate version of the Loebl-Komlós-Sós conjecture, *preprint*.

[62] D. Piguet, M. J. Stein, The Loebl-Komlós-Sós conjecture for trees of diameter 5 and for certain caterpillars, *The Electronic Journal on Combinatorics* **15** (2008) # R106.

[63] J. Polcyn, V. Rödl, A. Ruciński and E. Szemerédi, Short paths in quasi-random triple systems with sparse underlying graphs, *J. Combinatorial Theory B* **96** (2006), 584–607.

[64] V. Rödl and M. Schacht, Regular partitions of hypergraphs: Regularity Lemma, *Combinatorics, Probability & Computing* **16** (2007), 833–885.

[65] V. Rödl and M. Schacht, Regular partitions of hypergraphs: Counting Lemmas, *Combinatorics, Probability & Computing* **16** (2007), 887–901.

[66] V. Rödl and J. Skokan, Regularity lemma for k-uniform hypergraphs, *Random Structures & Algorithms* **2**5 (2004), 1–42.

[67] J. F. Sacle, M. Woźniak, The Erdős-Sós Conjecture for Graphs Without C_4, *Journal of Combinatorial Theory, Series B* **70** (1997), 367–372.

[68] P. Seymour, *Problem section*, Combinatorics: Proceedings of the British Combinatorial Conference 1973 (T. P. Mcdonough, V. C. Mavron, eds.), Cambridge University Press (1974), 201-204.

[69] S. N. Soffer, The Komlós-Sós conjecture for graphs of girth 7, *Discrete Mathematics* **214**, no. 1 (2000), 279-283.

[70] L. Sun, On the Loebl-Komlós-Sós conjecture, *Australasian Journal of Combinatorics* **37** (2007) 271-276.

[71] E. Szemerédi, Regular partitions of graphs, *Colloques Internationaux C.N.R.S. No. 260 - Problèmes Combinatoires et Théorie des Graphes, Orsay* (1976), 399-401.

[72] Y. Zhao, Proof of the $(n/2 - n/2 - n/2)$ Conjecture for large n, *preprint*.