Resilience and anti-Ramsey properties of sparse random graphs and dense hypergraphs

Dissertation

zur Erlangung des Doktorgrades an der
Fakultät für Mathematik, Informatik und Naturwissenschaften

Fachbereich Mathematik
der Universität Hamburg

vorgelegt von
Jakob Schnitzer

Hamburg, 2017

Gutachter

- Prof. Dr. Mathias Schacht
- Prof. Dr. Tibor Szabó

Datum der Disputation: 22. März 2018

Contents

Acknowledgements

First and foremost, I would like to thank my supervisor, Mathias Schacht, who has shared his deep understanding of combinatorics with me, has always lent me an open ear, and taught me how to write in a concise and clear style. It has truly been a pleasure and a privilege to work with him.

I would like to thank all my co-authors for their collaboration, which was always productive and fun. I would also like to thank the discrete mathematics group at the Universidade de São Paulo for their repeated hospitality and the DAAD for their generosity.

The discrete mathematics group at the Universität Hamburg has provided a friendly environment and many fruitful discussions on mathematics and beyond – thank you.

Finally, I thank my family and friends for their love and support.

1 Introduction

1.1 Overview

This thesis focuses on extremal and probabilistic combinatorics and Ramsey theory. We start with an overview of these areas and their ties to each other, highlighting were the obtained results fit in. The overview is followed by more in-depth introductions to the individual results. We will consider simple undirected graphs and hypergraphs $G = (V, E)$, where V and E are the vertex set and edge set respectively and usually $n = |V|$, the size of a considered graph, is large. We assume that the reader is familiar with basic notions of graph theory and standard notation like, e.g., the minimum and maximum degree $\delta(G)$ and $\Delta(G)$ or the chromatic number $\chi(G)$. The reader is referred to one of the standard text books ([6, 9, 15]) for an introduction to graph theory.

Turán's theorem and resilience

A large part of extremal graph theory traces back to the landmark result of Turán [59]: any graph on n vertices not containing K_r, the clique or complete graph on r vertices, as a subgraph has at most $\left(1 - \frac{1}{r-1}\right)\frac{n^2}{2}$ edges. In other words, if the edge density of a graph is at least $1 - \frac{1}{r-1}$, one is guaranteed to find a clique of size r as a subgraph. Erdős and Stone [22] generalised Turán's theorem from cliques to complete partite graphs and later Erdős and Simonovits [21] generalised it to all graphs: any graph on n vertices not containing a given graph F as a subgraph contains at most $\left(1 - \frac{1}{\chi(F)-1}\right)\frac{n^2}{2} + o(n^2)$ edges. Again, an edge density of $1 - \frac{1}{r-1} + o(1)$ implies the existence of a chosen r-colourable graph as a subgraph. This result might also be viewed as a "resilience" of the complete graph; one may delete a sizeable fraction of edges from it and one may yet find cliques or really any chosen graph as a subgraph.

Dirac's theorem and local resilience

A slightly different question is which conditions force the existence of a spanning subgraph, that is, one covering all vertices. Clearly, a bound on the edge density may not be sufficient as a graph can have $\binom{n}{2} - (n-1)$ edges and contain an isolated vertex, which would make it impossible to find any spanning connected subgraphs. So studying minimum vertex degree conditions instead of density conditions is one natural way to obtain results on the existence of spanning subgraphs. One of the first results of this type is Dirac's theorem on Hamiltonian cycles [16]: any graph G on n vertices with a minimum degree of at least $n/2$ contains a Hamiltonian cycle, that is, a cycle containing all vertices. Corradi and Hajnal [14] proved that a graph on n vertices with a minimum degree of at least $2n/3$ contains a triangle factor, that is, a set of disjoint triangles covering all but at most two of the vertices. A result of Hajnal and Szemerédi [25] generalises this and gives optimal degree conditions guaranteeing the existence of a K_k-factor, that is, a set of disjoint cliques K_k covering all but at most $k-1$ vertices of the graph: a minimum vertex degree of $\frac{k-1}{k}n$ implies the existence of a K_k-factor. These types of results represent a "local resilience" of the complete graph as one may delete some fraction of edges at each vertex and yet find some spanning substructures.

In what be can considered a combination of spanning cycles and small cliques, Pósa (see [19]) conjectured that any graph with minimum degree at least $2n/3$ contains the square of a Hamiltonian cycle, where a square (a $(k-1)$-st power) is obtained from a cycle by connecting vertices at distance at most two (at most $k-1$). This was generalised by Seymour [57] who conjectured that a minimum degree $\frac{k-1}{k}n$ suffices for the existence of the $(k-1)$-st power of a Hamiltonian cycle. Komlós, Sárközy, and Szemerédi [39] resolved this conjecture. The main tools in their proof are the regularity lemma, which was first introduced in the proof of Szemerédi's theorem on arithmetic progressions, and the blowup lemma, which was a new tool by these authors. These two tools enabled proofs of a range of results on the existence of spanning subgraphs, such as trees, H-factors, and planar graphs (see, e.g., [40, 41, 43, 45]). Bollobás and Komlós [37] generalised the conjecture of Pósa and Seymour and conjectured that a similar minimum degree allows the embedding of subgraphs of chromatic number k that have bounded maximum degree and small bandwidth, where small bandwidth means that the vertices of the graph can be linearly ordered so that no edge connects vertices that are far apart in the linear order. This so-called bandwidth conjecture can be seen as a common generalisation of some of the results on spanning subgraphs that were proved using the blowup lemma. The bandwidth conjecture was proved by Böttcher, Schacht, and Taraz [11], mostly solving the question of the local resilience of the complete graph with respect to embedding spanning subgraphs.

Resilience and local resilience in hypergraphs

Both resilience and local resilience results can have generalisations to hypergraphs. Turán already considered the generalisation of his theorem to hypergraphs and conjectured that for $K_4^{(3)}$, the complete 3-uniform hypergraph on four vertices, the required edge density is 5/9. This conjecture is still open and in contrast to the graph case where most Turán-type questions are solved, very little is known. However, there has been more process on minimum degree conditions in hypergraphs forcing the existence of spanning subhypergraphs.

Rödl, Ruciński, and Szemerédi [53, 54] were able to extend Dirac's theorem to hypergraphs using (and introducing) the so-called absorbing method. For 3-uniform hypergraphs they proved that a minimum pair-degree of at least $n/2+o(n)$ implies the existence of a Hamiltonian tight cycle, that is, a cycle where any two consecutive edges intersect in two vertices. For 3-uniform hypergraphs one may consider vertex- and pair-degree conditions as well as tight and loose cycles, where in loose cycles consecutive edges intersect in single vertices. Kühn and Osthus [44] proved that a minimum pair-degree of at least $n/4 + o(n)$ implies the existence of a loose Hamiltonian cycle. Buß, Han, and Schacht [12] proved that a minimum vertex-degree of at least $\frac{7}{16}\binom{n}{2} + o(n^2)$ implies the existence of a loose Hamiltonian cycle. The remaining case was recently solved by Reiher, Rödl, Ruciński, Schacht, and Szemerédi [49], who proved that a minimum vertex degree of $\frac{5}{9}\binom{n}{2} + o(n^2)$ implies the existence of a tight Hamiltonian cycle.

As one can see from the described results, there are various degree conditions and types of cycles to consider in hypergraphs and the absorbing method has helped prove several Dirac-type results in hypergraphs. Whereas there are often many extremal examples for Turán problems in hypergraphs, for Dirac-type questions the picture is usually somewhat easier which has allowed optimal bounds to be proved. We contribute an optimal minimum $(k-2)$-degree condition for loose Hamiltonian cycles in k-uniform hypergraphs (see Section 1.2 in the introduction and Chapter 2).

The random graph and transference

In probabilistic combinatorics, the random graph $G(n,p)$ and other random models have over time evolved from a mere tool used in proofs to a subject of study itself. One of the first uses of $G(n,p)$, the random graph on n vertices where each edge is included with probability p independently of all other edges, and the probabilistic method in general was Erdős's proof [17] of an exponential lower bound for symmetric Ramsey numbers. Another early application was the proof by Erdős [18] of the existence of graphs both containing no short cycles and having high chromatic number.

A more recent development in probabilistic combinatorics has been the "transference" of properties from complete graphs to random graphs: understanding how much the random graph behaves like a complete graph or more formally, finding the threshold probability $p = p(n)$ that is required for $G(n,p)$ to have a property that the complete graph $K_n = G(n,1)$ has. Rödl and Ruciński [50, 51] obtained the threshold for the Ramsey property, one of the first transference results. The Turán problem in sparse random graphs was recently solved by Conlon and Gowers [13] and independently by Schacht [56].

Finding sparse transferences of spanning local resilience results however is a more open area of research. Lee and Sudakov [46] obtained the sparse analogue of Dirac's theorem. The main tools to prove more general local resilience statements for sparse random graphs are the sparse version of the regularity lemma by Kohayakawa and Rödl [32, 36] and the so-called blowup lemma for sparse graphs by Allen, Böttcher, Hàn, Kohayakawa, and Person [3]. Extending a result of Allen, Böttcher, Ehrenmüller, and Taraz [2], who proved a sparse analogue of the bandwidth theorem for almost spanning subgraphs, we contribute a sparse analogue of the bandwidth theorem for spanning subgraphs (see Section 1.3 in the introduction and Chapter 3).

Ramsey theory

Recall the theorem of Ramsey [48]: for any n, k and r there exists R such that for every r-colouring of $[R]^{(k)}$, the k-subsets of the first R integers, there exists $N \subset [R]$ of size n such that $N^{(k)}$ is monochromatic. If one does not want to restrict the number of colours that can be used, one cannot hope to have a monochromatic clique but in its absence there is still some structure to be found, as evidenced by the canonical Ramsey theorem of Erdős and Rado [20]: for every n there exists R such that for every colouring of the edges of the complete graph K_R there exists a subset N of $\{1, \dots R\}$ of size n satisfying the following. All edges on N are of the same colour or their colours are uniquely and distinctly determined either by their left vertices, by their right vertices or by both vertices. If one only considers proper colourings, i.e., colourings where no edges of the same colour meet in a vertex, only the last alternative can occur which implies the rainbow Ramsey theorem: for every n there exists R such that for every proper colouring of K_R there exists a rainbow K_n, i.e., one where no two edges are of the same colour.

As mentioned above, Rödl and Ruciński proved the threshold of $G(n, p)$ for the Ramsey property. Apart from a few exceptions, for a given graph H this threshold is roughly the value $p(n)$ for which the expected number of copies of H in $G(n, p)$ exceeds the expected number of edges in $G(n, p)$. There is no known result for the threshold for the canonical Ramsey property, one obstacle in this direction being that the proof of Erdős and Rado for the graph version of the canonical Ramsey theorem requires Ramsey's theorem for 4-uniform hypergraphs and no direct proof is known. The rainbow Ramsey theorem however can also be proved without resorting to higher uniformities and Kohayakawa, Konstadinidis, and Mota [33] proved an upper bound on the threshold, which is of the same order of magnitude as the threshold for the Ramsey property. Nenadov, Person, Škorić and Steger [47] proved the matching lower bound for cliques on at least 19 vertices and cycles on at least 7 vertices. We contribute a matching lower bound for cliques on at least five vertices and a new upper and matching lower bound for K_4 (see Section 1.4 in the introduction and Chapter 4).

1.2 Hamiltonian cycles

A frequent question in graph theory is which abstract graph properties force the existence of a Hamiltonian cycle, that is, a cycle containing all vertices of the graph. Recall Dirac's theorem on Hamiltonian cycles in graphs.

Theorem 1 (Dirac [16])**.** *Any graph G with minimum degree $\delta(G) \geqslant \frac{|G|}{2}$ contains a Hamiltonian cycle.*

For simple graphs, this result has been generalised to give optimal conditions for the existence of a Hamiltonian cycle based on the degree sequence of a graph. Obtaining minimum degree conditions forcing the existence of Hamiltonian cycles for hypergraphs is a more recent area of study. In the following, we will work with the following notion of cycle in a hypergraph. We say that a k-uniform hypergraph $\mathcal{C}$ is an *ℓ-cycle* if there exists a cyclic ordering of its vertices such that every edge of $\mathcal{C}$ is composed of k consecutive vertices, two (vertex-wise) consecutive edges share exactly ℓ vertices, and every vertex is contained in an edge. Moreover, if the ordering is not cyclic, then $\mathcal{C}$ is an *ℓ-path* and we say that the first and last ℓ vertices are the ends of the path. Just like in the graph case, we say that an ℓ-cycle in a hypergraph is Hamiltonian if it contains all vertices of the hypergraph.

Given a k-uniform hypergraph $\mathcal{H} = (V, E)$ and a set $S \in V^{(s)}$ of s vertices, we denote by $d(S)$ the number of edges in E containing S and we denote by $N(S)$ the $(k-s)$-element sets $T \in V^{(k-s)}$ such that $T \cup S \in E$, so $d(S) = |N(S)|$. The *minimum s-degree* of $\mathcal{H}$ is denoted by $\delta_s(\mathcal{H})$ and it is defined as the minimum of $d(S)$ over all sets $S \in V^{(s)}$.

In a k-uniform hypergraphs there are $k-1$ different types of ℓ-cycles to consider, namely for all values of ℓ between 1 and $k-1$. Similarly, there are various minimum degrees δ_s to consider, for all values of s between 1 and $k-1$, since the edge density δ_0 is not sufficient to find a spanning structure. Different types of ℓ-cycles might require different minimum degree conditions, so for any k there are $(k-1)^2$ possible generalisations of Dirac's theorem to k-uniform hypergraphs. One would expect that ℓ-cycles are somewhat "harder" to find for higher values of ℓ, i.e., they require a higher minimum degree, since for example a 2-cycle in a 3-uniform hypergraph contains a 1-cycle if the number of vertices is even. Similarly, having a minimum degree condition on larger sets of vertices is "stronger": for example a minimum 2-degree implies a minimum vertex degree.

If one obtains a minimum degree condition for the existence of a Hamiltonian cycle, ideally it should be *optimal* in the following sense: There is an extremal example, that is, a hypergraph that does not contain a Hamiltonian cycle, with minimum degree just below the obtained bound. We say that a δ_s-degree condition for k-uniform hypergraphs is *asymptotically optimal* if there is an example not containing a Hamiltonian cycle with just $o(n^{k-s})$ fewer edges. To see that the minimum degree condition in Dirac's theorem is optimal, one may consider a complete bipartite graph where one side has just below $n/2$ vertices. Clearly a cycle in a bipartite graph needs to contain the same number of vertices from both sides of the bipartion, so an unbalanced bipartite graph contains no Hamiltonian cycle.

The problem of finding minimum degree conditions that ensure the existence of Hamiltonian cycles in hypergraphs has been extensively studied over the last years (see, e.g., the surveys [52, 60]). Katona and Kierstead [30] started the study of this problem, posing a conjecture that was later confirmed by Rödl, Ruciński, and Szemerédi, who proved the following result.

Theorem 2 (Rödl, Ruciński, Szemerédi [53, 54])**.** *For every $k \geqslant 3$, if $\mathcal{H}$ is a k-uniform n-vertex hypergraph with $\delta_{k-1}(\mathcal{H}) \geqslant (1/2 + o(1))n$, then $\mathcal{H}$ contains a Hamiltonian $(k-1)$-cycle.*

This result is an asymptotically optimal generalisation of Dirac's theorem to hypergraphs. Kühn and Osthus [44] proved the asymptotically optimal condition $\delta_2(\mathcal{H}) \geqslant (1/4 + o(1))n$ for 1-cycles in 3-uniform hypergraphs $\mathcal{H}$. Hàn and Schacht [26] (see also [31]) generalised this result to a $(k-1)$-degree condition for arbitrary k and ℓ-cycles with $1 \leqslant \ell < k/2$. In [42], Kühn, Mycroft, and Osthus generalised this result to $1 \leqslant \ell < k$, settling the problem of the existence of Hamiltonian ℓ-cycles in k-uniform hypergraphs with large minimum $(k-1)$-degree.

The first conditions for minimum $(k-2)$-degree were obtained by Buß, Hàn, and Schacht who obtained the following result.

Theorem 3 (Buß, Hàn, Schacht [12])**.** *For every $\gamma > 0$ there exists an n_0 such that every 3-uniform hypergraph $\mathcal{H} = (V, E)$ on $|V| = n \geqslant n_0$ vertices with $n \in 2\mathbb{N}$ and*

$$\delta_1(\mathcal{H}) \geqslant \left(\frac{7}{16} + \gamma\right)\binom{n}{2}$$

contains a Hamiltonian 1-cycle. □

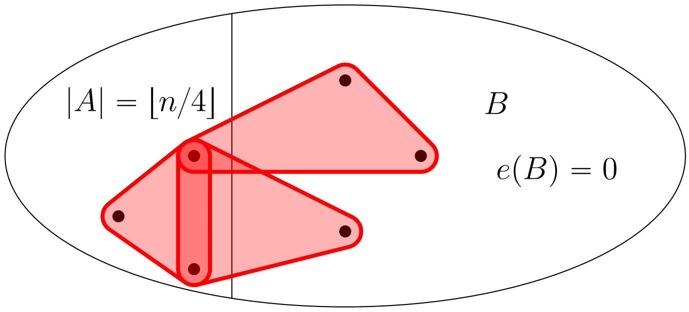


Figure 1.1: The 3-uniform extremal hypergraph $\mathcal{X}_{3,1}(n)$
(one edge per size of intersection with A drawn)

Recently, Reiher, Rödl, Ruciński, Schacht, and Szemerédi [49] showed the asymptotically optimal minimum 1-degree condition $\delta_1(\mathcal{H}) \geqslant (5/9 + o(1))\binom{n}{2}$ for the existence of 2-cycles in 3-uniform hypergraphs. In Theorem 4 below we have a generalisation of Theorem 3 to higher uniformities.

Theorem 4 (Bastos, Mota, Schacht, S., Schulenburg [4])**.** *For all integers $k \geqslant 3$ and $1 \leqslant \ell < k/2$ and every $\gamma > 0$ there exists an n_0 such that every k-uniform hypergraph $\mathcal{H} = (V, E)$ on $|V| = n \geqslant n_0$ vertices with $n \in (k-\ell)\mathbb{N}$ and*

$$\delta_{k-2}(\mathcal{H}) \geqslant \left(\frac{4(k-\ell)-1}{4(k-\ell)^2} + \gamma\right)\binom{n}{2}$$

contains a Hamiltonian ℓ-cycle. □

All the theorems seen so far for hypergraphs were only asymptotically optimal. The example showing the optimality of Dirac's theorem can be extended to hypergraphs by choosing a vertex subset of a small enough size and taking all edges incident to it: The construction of the example varies slightly depending on whether n, the size of the example hypergraph, is an odd or an even multiple of $k-\ell$. We first consider the case that n is an odd multiple of $(k-\ell)$ here, the (minimally) different constuction for even multiples is given below. Let $\mathcal{X}_{k,\ell}(n) = (V, E)$ be a k-uniform hypergraph on n vertices such that an edge belongs to E if and only if it contains at least one vertex from $A \subset V$, where $|A| = \left\lfloor \frac{n}{2(k-\ell)} \right\rfloor$ (see Figure 1.1). It is easy to see that $\mathcal{X}_{k,\ell}(n)$ contains no Hamiltonian ℓ-cycle for $\ell < k/2$, as it would have to contain $\frac{n}{k-\ell}$ edges and each vertex in A is contained in at most two of them.

Let us now consider the case that n is an even multiple of $k-\ell$. Similarly, let $\mathcal{X}_{k,\ell}(n) = (V, E)$ be a k-uniform hypergraph on n vertices that contains all edges incident to $A \subset V$, where $|A| = \frac{n}{2(k-\ell)} - 1$. Additionally, fix some $\ell + 1$ vertices of $B = V \setminus A$ and let $\mathcal{X}_{k,\ell}(n)$ contain all edges on B that contain all of these vertices, that is, an $(\ell+1)$-star. Again, of the $\frac{n}{k-\ell}$ edges that a Hamiltonian ℓ-cycle would have to contain, at most $\frac{n}{k-\ell} - 2$ can be incident to A. So two edges would have to be completely contained in B and be disjoint or intersect in exactly ℓ vertices, which is impossible since the induced subhypergraph on B only contains an $(\ell+1)$-star. Note that for the minimum $(k-2)$-degree the $(\ell+1)$-star on B is only relevant if $\ell = 1$, in which case this star increases the minimum $(k-2)$-degree by one.

In [28], Han and Zhao proved the optimal version of Theorem 3. We extend this to k-uniform hypergraphs in the following theorem, i.e., we prove the optimal version of Theorem 4.

Theorem 5. *For all integers $k \geqslant 4$ and $1 \leqslant \ell < k/2$ there exists n_0 such that every k-uniform hypergraph $\mathcal{H} = (V, E)$ on $|V| = n \geqslant n_0$ vertices with $n \in (k-\ell)\mathbb{N}$ and*

$$\delta_{k-2}(\mathcal{H}) > \delta_{k-2}(\mathcal{X}_{k,\ell}(n))$$

contains a Hamiltonian ℓ-cycle. In particular, if

$$\delta_{k-2}(\mathcal{H}) \geqslant \frac{4(k-\ell)-1}{4(k-\ell)^2}\binom{n}{2},$$

then $\mathcal{H}$ contains a Hamiltonian ℓ-cycle.

The proof of Theorem 5 is the topic of Chapter 2.

1.3 Spanning subgraphs in sparse graphs

A main area of study in extremal graph theory has been the transference of extremal results in dense graphs to sparse graphs. Another way to think about Dirac's theorem is the following: In K_n, the complete graph on n vertices, one may delete about half the edges, $\lfloor n/2 \rfloor$ to be precise, at each vertex and will obtain a graph containing a Hamiltonian cycle. We define the *local resilience* of a graph G with respect to a monotone increasing graph property $\mathcal{P}$ as the minimum number of edges m such that one may obtain a graph not having the property $\mathcal{P}$ by deleting at most m edges at each vertex of G. So, again rephrasing Dirac's theorem, the complete graph $K_n = G(n,1)$ has local resilience $\lfloor n/2 \rfloor$ with respect to Hamiltonicity. Sudakov and Vu [58] initiated the study of resilience in random (and pseudo-random) graphs and Lee and Sudakov resolved their following question concerning the local resilience of $G(n,p)$ with respect to Hamiltonicity. We say that a property $\mathcal{P}$ holds *asymptotically almost surely* (a.a.s.) for $G(n,p)$ if the probability that $G(n,p) \in \mathcal{P}$ tends to 1 as n tends to infinity.

Theorem 6 (Lee, Sudakov [46]). *For every positive ε, there exists a constant $C = C(\varepsilon)$ such that for $p \geqslant \frac{C \log n}{n}$ asymptotically almost surely every subgraph of $G(n,p)$ with minimum degree at least $(1/2+\varepsilon)np$ is Hamiltonian.*

Note that this result is best possible (up to the constant C), as for $p \leqslant \frac{\log n}{n}$ the random graph $G(n,p)$ will almost surely contain vertices of degree at most one. In addition to Hamiltonian cycles, resilience results for $G(n,p)$ have been obtained for a wide range of graphs such as powers of Hamiltonian cycles, trees, or F-factors for any fixed F. What all these graph have in common is bounded degree and small *bandwidth*; the bandwidth of a graph H is the minimum b such that there is an injective labelling of the vertex set of H by integers with $|i-j| < b$ for every edge $\{i,j\}$ in H. It was conjectured by Bollobás and Komlós that these resilience results for the complete graph could be extended to guarantee existence of subgraphs of small bandwidth, bounded maximum degree and chromatic number. This conjecture was resolved by Böttcher, Schacht, and Taraz who proved the following, which is also known as the bandwidth theorem.

Theorem 7 (Böttcher, Schacht, Taraz [11]). *For every $\gamma > 0$, $\Delta \geqslant 2$, and $k \geqslant 1$, there exist $\beta > 0$ and $n_0 \geqslant 1$ such that for every $n \geqslant n_0$ the following holds. If G is a graph on n vertices with minimum degree $\delta(G) \geqslant \left(\frac{k-1}{k} + \gamma\right) n$ and if H is a k-colourable graph on n vertices with maximum degree $\Delta(H) \leqslant \Delta$ and bandwidth at most βn, then G contains a copy of H.*

Similar to the transference obtained for Dirac's theorem, one would like to generalise it by transferring the bandwidth theorem to random graphs. This was done by Allen, Böttcher, Ehrenmüller, and Taraz who proved the following statement.

Theorem 8 (Allen, Böttcher, Ehrenmüller, Taraz [2]). *For each $\gamma > 0$, $\Delta \geqslant 2$, and $k \geqslant 1$, there exist constants $\beta^* > 0$ and $C^* > 0$ such that the following* *holds asymptotically almost surely for* $\Gamma = G(n,p)$ *if* $p \geqslant C^*\left(\frac{\log n}{n}\right)^{1/\Delta}$. *Let* G *be a spanning subgraph of* Γ *with* $\delta(G) \geqslant \left(\frac{k-1}{k}+\gamma\right)pn$, *and let* H *be a* k*-colourable graph on* n *vertices with* $\Delta(H) \leqslant \Delta$, *bandwidth at most* $\beta^* n$, *and with at least* $C^* p^{-2}$ *vertices which are not contained in any triangles of* H. *Then* G *contains a copy of* H.

Note that this is not a straightforward transference of the bandwidth theorem as it contains the additional restriction that some vertices are not contained in any triangles, or, alternatively one only finds an almost spanning embedding of low-bandwidth k-colourable graphs of maximum degree Δ. The additional restriction is necessary however, which can be seen as follows if $p \ll n$. By deleting all edges in the neighbourhood of a vertex $v \in G(n,p)$, which ensures that v is contained in no triangle, one would only remove $O(np^2) \ll np$ edges at each vertex. Indeed one can, only removing a small fraction of the edges at each vertex, ensure that $\Omega(p^{-2})$ vertices have independent sets as neighbourhoods. Similarly, one could make neighbourhoods of $\Omega(p^{-2})$ vertices bipartite, again by only removing a tiny fraction of edges at each vertex, which would prevent the existence of a K_4-factor. So keeping neighbourhoods non-independent is not enough to drop the requirement that some vertices are not contained in any triangles, but the following statement along those lines holds.

Theorem 9. *For each* $\gamma > 0$, $\Delta \geqslant 2$, $k \geqslant 2$ *and* $0 \leqslant s \leqslant k-1$, *there exist constants* $\beta^* > 0$ *and* $C^* > 0$ *such that the following holds asymptotically almost surely for* $\Gamma = G(n,p)$ *if* $p \geqslant C^*\left(\frac{\log n}{n}\right)^{1/\Delta}$. *Let* G *be a spanning subgraph of* Γ *with* $\delta(G) \geqslant \left(\frac{k-1}{k}+\gamma\right)pn$, *such that for each* $v \in V(G)$ *there are at least* $\gamma p^{\binom{s}{2}}(pn)^s$ *copies of* K_s *in* $N_G(v)$. *Let* H *be a graph on* n *vertices with* $\Delta(H) \leqslant \Delta$, *bandwidth at most* $\beta^* n$ *and suppose that there is a proper* k*-colouring of* $V(H)$ *and at least* $C^* p^{-2}$ *vertices in* $V(H)$ *whose neighbourhood contains only* s *colours. Then* G *contains a copy of* H.

The proof of Theorem 9 is the topic of Chapter 3.

1.4 Anti-Ramsey

Recall Ramsey's theorem.

Theorem 10 (Ramsey [48])**.** *For any positive integers n, k, and r there exists R such that the following holds. For any r-colouring of the k-subsets of $[R]$, the first R integers, there exists a subset $N \subset [R]$ of size n such that all k-sets on N are of the same colour.*

This result initiated Ramsey theory, the study of obtaining large homogeneous substructures in graphs and other combinatorial structures. For this topic, we will only consider 2-graphs. Let r be a positive integer and let G and H be graphs. We denote by $G \to (H)_r$ the property that any colouring of the edges of G with at most r colours contains a monochromatic H in G. Ramsey's theorem states that for every n and r there exists R such that $K_R \to (K_n)_r$. If one were to drop the requirement that only a bounded number of colours is used, it is clearly futile to hope that one might find a large monochromatic subgraph. One of the oldest generalisations of Ramsey's theorem is the so-called canonical Ramsey theorem, which we state here for the graph case.

Theorem 11 (Erdős, Rado [20])**.** *For any n there exists R such that the following holds. Suppose that the edges of the complete graph K_R are arbitrarily coloured. Then there exists a subset N of $V(K_R)$ of size n such that one of the four conditions holds for all $a < b, c < d \in N$.*

(i) *All edges on N are of the same colour.*

(ii) *$\{a, b\}$ and $\{c, d\}$ are of the same colour if and only if $a = c$.*

(iii) *$\{a, b\}$ and $\{c, d\}$ are of the same colour if and only if $b = d$.*

(iv) *$\{a, b\}$ and $\{c, d\}$ are of the same colour if and only if $a = c$ and $b = d$.*

If we require that the colouring be *proper*, i.e., the colours of edges containing any fixed vertex are all distinct, only item (iv) is possible. Note that item (iv) means that the colouring is *rainbow* on N, that is, all pairs in N are given distinct colours. Given graphs H and G, we are interested in the following 'anti-Ramsey' notion, denoted by $G \xrightarrow[\mathrm{p}]{\mathrm{rb}} H$: for every proper edge-colouring of G, there exists a rainbow H in G, i.e., a copy of H with no two edges of the same colour. The term 'anti-Ramsey' can be understood as follows: rather then trying to obtain a colouring which is constant on a large subgraph, we want to obtain a colouring which is injective on a large subgraph.

For fixed G, the graph properties $G \to (H)_r$ and $G \xrightarrow[\text{p}]{\text{rb}} H$ are *monotone*, i.e., if they hold in a subgraph of $G' \subseteq G$ then they also hold in G. Since they hold in large enough complete graphs one could hope to obtain the *threshold* for this property. We say that $p_{\mathcal{P}} = p_{\mathcal{P}}(n)$ is the threshold for a graph property $\mathcal{P}$ if the following holds.

$$\mathbb{P}(G(n,p) \in \mathcal{P}) \to \begin{cases} 1, & \text{if } p \gg p_{\mathcal{P}}(n), \\ 0, & \text{if } p \ll p_{\mathcal{P}}(n), \end{cases}$$

Rödl and Ruciński determined the threshold for the property $G(n,p) \to (H)_r$ for all graphs H. The *maximum 2-density* $m^{(2)}(H)$ of a graph H on at least three vertices is denoted by

$$m^{(2)}(G) = \max\left\{\frac{|E(J)|-1}{|V(J)|-2} : J \subset H,\ |V(J)| \geqslant 3\right\}.$$

Theorem 12 (Rödl, Ruciński [50, 51])**.** *Let H be a graph that is not a forest of stars or if $r = 2$, paths of length 3. Then, the threshold $p_H = p_H(n)$ for the property $G(n,p) \to (H)_r$ is given by $p_H(n) = n^{-1/m^{(2)}(H)}$.*

Since the property $G(n,p) \xrightarrow[\text{p}]{\text{rb}} H$ is monotone for every fixed graph H, we know that it admits a threshold function $p_H^{\text{rb}} = p_H^{\text{rb}}(n)$ [8]. The study of anti-Ramsey properties of random graphs was initiated by Rödl and Tuza, who proved in [55] that for every ℓ there exists a fairly small p, such that $G(n,p) \xrightarrow[\text{p}]{\text{rb}} C_\ell$ almost surely. The following result gives an upper bound for the threshold p_H^{rb} for any fixed graph H.

Theorem 13 (Kohayakawa, Konstadinidis, Mota [33])**.** *Let H be a fixed graph. Then there exists a constant $C > 0$ such that for $p = p(n) \geqslant Cn^{-1/m^{(2)}(H)}$ we have $G(n,p) \xrightarrow[p]{rb} H$ almost surely.*

In particular, Theorem 13 implies $p_H^{\text{rb}} \leqslant n^{-1/m^{(2)}(H)}$. For the Ramsey property $G(n,p) \to (H)_r$, the graphs for which the threshold is not given by the 2-density are forests of stars or if $r = 2$, paths of length 3. For example for a star with k edges, the threshold is determined by the appearance of a star with $r(k-1)+1$ edges. For the rainbow Ramsey property, the triangle is the only example for which the threshold is easily seen to be below the function given by the 2-density as any properly coloured triangle is rainbow. In [34] it was proved that there are infinitely many graphs H for which the threshold is asymptotically smaller than $n^{-1/m^{(2)}(H)}$. These graphs consist of the "amalgamation" of a triangle and a sparse graph like a cycle, i.e., a cycle where two adjacent vertices have been connected to a new additional vertex. Recently, it was proved that for sufficiently large cycles and complete graphs the lower bound on the threshold matches the upper bound given in Theorem 13.

Theorem 14 (Nenadov, Person, Škorić, Steger [47])**.** *Let H be a cycle on at least 7 vertices or a complete graph on at least 19 vertices. Then $p_H^{rb} = n^{-1/m^{(2)}(H)}$.*

The authors of the above result remark that it could hold for all cycles and complete graphs of size at least 4. We prove that Theorem 14 can be extended to complete graphs of size at least 5, but not for K_4. In fact, we can show that if H is a connected graph on 4 vertices, then p_H^{rb} is asymptotically smaller than $n^{-1/m^{(2)}(H)}$.

Theorem 15. *For $k \geqslant 5$, $p_{K_k}^{rb} = n^{-1/m^{(2)}(K_k)}$. Furthermore, $p_{K_4}^{rb} = n^{-7/15}$.*

The proof of Theorem 15 is the topic of Chapter 4.

2 Hamiltonian cycles in hypergraphs

In this chapter, we will prove the following result.

Theorem 16. *For all integers $k \geqslant 4$ and $1 \leqslant \ell < k/2$ there exists n_0 such that every k-uniform hypergraph $\mathcal{H} = (V, E)$ on $|V| = n \geqslant n_0$ vertices with $n \in (k-\ell)\mathbb{N}$ and*

$$\delta_{k-2}(\mathcal{H}) > \delta_{k-2}(\mathcal{X}_{k,\ell}(n)) \tag{2.1}$$

contains a Hamiltonian ℓ-cycle. In particular, if

$$\delta_{k-2}(\mathcal{H}) \geqslant \frac{4(k-\ell)-1}{4(k-\ell)^2}\binom{n}{2},$$

then $\mathcal{H}$ contains a Hamiltonian ℓ-cycle.

The following notion of extremality is motivated by the extremal example hypergraph $\mathcal{X}_{k,\ell}(n)$. A k-uniform hypergraph $\mathcal{H} = (V, E)$ is called (ℓ, ξ)-*extremal* if there exists a partition $V = A \cup B$ such that

$$|A| = \left\lceil \frac{n}{2(k-\ell)} - 1 \right\rceil, \quad |B| = \left\lfloor \frac{2(k-\ell)-1}{2(k-\ell)} n + 1 \right\rfloor,$$

and $e(B) = |E \cap B^{(k)}| \leqslant \xi\binom{n}{k}$. We say that $A \cup B$ is an (ℓ, ξ)-*extremal partition* of V. Theorem 16 follows easily from the next two results, the so-called *extremal case* (see Theorem 18 below) and the *non-extremal case* (see Theorem 17).

Theorem 17 (Non-extremal case)**.** *For any $0 < \xi < 1$ and all integers $k \geqslant 4$ and $1 \leqslant \ell < k/2$, there exists $\gamma > 0$ such that the following holds for sufficiently large n. Suppose $\mathcal{H}$ is a k-uniform hypergraph on n vertices with $n \in (k-\ell)\mathbb{N}$ such that $\mathcal{H}$ is not (ℓ, ξ)-extremal and*

$$\delta_{k-2}(\mathcal{H}) \geqslant \left(\frac{4(k-\ell)-1}{4(k-\ell)^2} - \gamma\right)\binom{n}{2}.$$

Then $\mathcal{H}$ contains a Hamiltonian ℓ-cycle. □

The non-extremal case was the main result of [4].

Theorem 18 (Extremal case)**.** *For any integers $k \geqslant 3$ and $1 \leqslant \ell < k/2$, there exists $\xi > 0$ such that the following holds for sufficiently large n. Suppose $\mathcal{H}$ is a k-uniform hypergraph on n vertices with $n \in (k-\ell)\mathbb{N}$ such that $\mathcal{H}$ is (ℓ, ξ)-extremal and*

$$\delta_{k-2}(\mathcal{H}) > \delta_{k-2}(\mathcal{X}_{k,\ell}).$$

Then $\mathcal{H}$ contains a Hamiltonian ℓ-cycle.

In Section 2.1 we give an overview of the proof of Theorem 18 and state Lemma 19, the main result required for the proof. In Section 2.2 we first prove some auxiliary lemmas and then we prove Lemma 19.

2.1 Overview

Let $\mathcal{H} = (V, E)$ be a k-uniform hypergraph and let $X, Y \subset V$ be disjoint subsets. Given a vertex set $L \subset V$ we denote by $d(L, X^{(i)}Y^{(j)})$ the number of edges of the form $L \cup I \cup J$, where $I \in X^{(i)}$, $J \in Y^{(j)}$, and $|L| + i + j = k$. We allow for $Y^{(j)}$ to be omitted when j is zero and write $d(v, X^{(i)}Y^{(j)})$ for $d(\{v\}, X^{(i)}Y^{(j)})$.

The proof of Theorem 18 follows ideas from [27], where a corresponding result with a $(k-1)$-degree condition is proved. Let $\mathcal{H} = (V, E)$ be an extremal hypergraph satisfying (2.1). We first construct an ℓ-path $\mathcal{Q}$ in $\mathcal{H}$ (see Lemma 19 below) with ends L_0 and L_1 such that there is a partition $A_* \cup B_*$ of $(V \smallsetminus \mathcal{Q}) \cup L_0 \cup L_1$ composed only of "typical" vertices (see (ii) and (iii) below). The set $A_* \cup B_*$ is suitable for an application of Lemma 20 below, which ensures the existence of an ℓ-path $\mathcal{Q}'$ on $A_* \cup B_*$ with L_0 and L_1 as ends. Note that the existence of a Hamiltonian ℓ-cycle in $\mathcal{H}$ is guaranteed by $\mathcal{Q}$ and $\mathcal{Q}'$. So, in order to prove Theorem 18, we only need to prove the following lemma.

Lemma 19 (Main lemma)**.** *For any $\varrho > 0$ and all integers $k \geqslant 3$ and $1 \leqslant \ell < k/2$, there exists a positive ξ such that the following holds for sufficiently large $n \in (k-\ell)\mathbb{N}$. Suppose that $\mathcal{H} = (V, E)$ is an (ℓ, ξ)-extremal k-uniform hypergraph on n vertices and*

$$\delta_{k-2}(\mathcal{H}) > \delta_{k-2}(\mathcal{X}_{k,\ell}(n)).$$

Then there exists a non-empty ℓ-path $\mathcal{Q}$ in $\mathcal{H}$ with ends L_0 and L_1 and a partition $A_* \cup B_* = (V \smallsetminus \mathcal{Q}) \cup L_0 \cup L_1$ where $L_0, L_1 \subset B_*$ such that the following hold:

(i) $|B_*| = (2k-2\ell-1)|A_*| + \ell$,

(ii) $d(v, B_*^{(k-1)}) \geqslant (1-\varrho)\binom{|B_*|}{k-1}$ *for any vertex* $v \in A_*$,

(iii) $d(v, A_*^{(1)}B_*^{(k-2)}) \geqslant (1-\varrho)|A_*|\binom{|B_*|}{k-2}$ *for any vertex* $v \in B_*$,

(iv) $d(L_0, A_*^{(1)}B_*^{(k-\ell-1)}), d(L_1, A_*^{(1)}B_*^{(k-\ell-1)}) \geqslant (1-\varrho)|A_*|\binom{|B_*|}{k-\ell-1}$.

The next result, which we will use to conclude the proof of Theorem 18, was obtained by Han and Zhao (see [27, Lemma 3.10]).

Lemma 20. *For any integers $k \geqslant 3$ and $1 \leqslant \ell < k/2$ there exists $\varrho > 0$ such that the following holds. If $\mathcal{H}$ is a sufficiently large k-uniform hypergraph with a partition $V(\mathcal{H}) = A_* \cup B_*$ and there exist two disjoint ℓ-sets $L_0, L_1 \subset B_*$ such that* (i)–(iv) *hold, then $\mathcal{H}$ contains a Hamiltonian ℓ-path $\mathcal{Q}'$ with L_0 and L_1 as ends.* □

2.2 Proof of the main lemma

We will start this section by describing the setup for the proof, which will be fixed for the rest of the chapter. Then we will prove some auxiliary lemmas and finally prove Lemma 19. Let $\varrho > 0$ and integers $k \geqslant 3$ and $1 \leqslant \ell < k/2$ be given. Fix constants

$$\frac{1}{k}, \frac{1}{\ell}, \varrho \gg \delta \gg \varepsilon \gg \varepsilon' \gg \vartheta \gg \xi,$$

where "$\gg x$" denotes that x is chosen sufficiently small with respect to all constants to its left. Let $n \in (k-\ell)\mathbb{N}$ be sufficiently large and let $\mathcal{H}$ be an (ℓ, ξ)-extremal k-uniform hypergraph on n vertices that satisfies the $(k-2)$-degree condition

$$\delta_{k-2}(\mathcal{H}) > \delta_{k-2}(\mathcal{X}_{k,\ell}(n)).$$

Let $A \cup B = V(\mathcal{H})$ be a minimal extremal partition of $V(\mathcal{H})$, i.e. a partition satisfying

$$a = |A| = \left\lceil \frac{n}{2(k-\ell)} \right\rceil - 1, \quad b = |B| = n - a, \quad \text{and} \quad e(B) \leqslant \xi\binom{n}{k}, \tag{2.2}$$

which minimises $e(B)$. Recall that the extremal example $\mathcal{X}_{k,\ell}(n)$ implies

$$\delta_{k-2}(\mathcal{H}) > \binom{a}{2} + a(b-k+2). \tag{2.3}$$

Since $e(B) \leqslant \xi\binom{n}{k}$, we expect most vertices $v \in B$ to have low degree $d(v, B^{(k-1)})$ into B. Also, most $v \in A$ must have high degree $d(v, B^{(k-1)})$ into B such that the degree condition for $(k-2)$-sets in B can be satisfied. Thus, we define the sets A_ε and B_ε to consist of vertices of high respectively low degree into B by

$$\begin{aligned} A_\varepsilon &= \left\{ v \in V \colon d(v, B^{(k-1)}) \geqslant (1-\varepsilon)\binom{|B|}{k-1} \right\}, \\ B_\varepsilon &= \left\{ v \in V \colon d(v, B^{(k-1)}) \leqslant \varepsilon\binom{|B|}{k-1} \right\}, \end{aligned}$$

and set $V_\varepsilon = V \smallsetminus (A_\varepsilon \cup B_\varepsilon)$. We will write $a_\varepsilon = |A_\varepsilon|$, $b_\varepsilon = |B_\varepsilon|$, and $v_\varepsilon = |V_\varepsilon|$. It follows from these definitions that

$$\text{if } A \cap B_\varepsilon \neq \varnothing, \quad \text{then} \quad B \subset B_\varepsilon, \quad \text{while otherwise} \quad A \subset A_\varepsilon. \tag{2.4}$$

For the first inclusion, consider a vertex $v \in A \cap B_\varepsilon$ and a vertex $w \in B \smallsetminus B_\varepsilon$. Exchanging v and w would create a minimal partition with fewer edges in $e(B)$, a contradiction to the minimality of the extremal partition. The other inclusion is similarly implied by the minimality.

Actually, as we shall show below, the sets A_ε and B_ε are not too different from A and B respectively:

$$|A \smallsetminus A_\varepsilon|, |B \smallsetminus B_\varepsilon|, |A_\varepsilon \smallsetminus A|, |B_\varepsilon \smallsetminus B| \leqslant \vartheta b \quad \text{and} \quad |V_\varepsilon| \leqslant 2\vartheta b. \tag{2.5}$$

Note that by the minimum $(k-2)$-degree

$$\binom{a}{2}\binom{b}{k-2} + a\binom{b}{k-1}(k-1) < \binom{b}{k-2}\delta_{k-2}(\mathcal{H}) \leqslant \sum_{S \in B^{(k-2)}} d(S).$$

Every vertex $v \in |A \smallsetminus A_\varepsilon|$ satisfies $d(v, B^{(k-1)}) < (1-\varepsilon)\binom{b}{k-1}$, so we have

$$\begin{aligned} \sum_{S \in B^{(k-2)}} d(S) \leqslant & \binom{a}{2}\binom{b}{k-2} + a\binom{b}{k-1}(k-1) \\ & + e(B)\binom{k}{2} - |A \smallsetminus A_\varepsilon|\varepsilon\binom{b}{k-1}(k-1). \end{aligned}$$

Consequently $|A \smallsetminus A_\varepsilon| \leqslant \vartheta b$, as $e(B) < \xi\binom{n}{k}$ and $\xi \ll \vartheta, \varepsilon$.

Moreover, $|B \smallsetminus B_\varepsilon| \leqslant \vartheta b$ holds as a high number of vertices in $B \smallsetminus B_\varepsilon$ would contradict $e(B) < \xi\binom{b}{k}$. The other three inequalities (2.5) follow from the already shown ones, for example for $|A_\varepsilon \smallsetminus A| < \vartheta b$ observe that

$$A_\varepsilon \smallsetminus A = A_\varepsilon \cap B \subset B \smallsetminus B_\varepsilon.$$

Although the vertices in B_ε were defined by their low degree into B, they also have low degree into the set B_ε itself; for any $v \in B_\varepsilon$ we get

$$\begin{aligned} d(v, B_\varepsilon^{(k-1)}) &\leqslant d(v, B^{(k-1)}) + |B_\varepsilon \smallsetminus B| \binom{|B_\varepsilon| - 1}{k-2} \\ &\leqslant \varepsilon \binom{b}{k-1} + \vartheta b |B_\varepsilon|^{k-1} \\ &< 2\varepsilon \binom{|B_\varepsilon|}{k-1}. \end{aligned}$$

Since we are interested in ℓ-paths, the degree of ℓ-tuples in B_ε will be of interest, which motivates the following definition. An ℓ-set $L \subset B_\varepsilon$ is called *ε-typical* if

$$d(L, B^{(k-\ell)}) \leqslant \varepsilon \binom{|B|}{k-\ell}.$$

If L is not ε-typical, then it is called *ε-atypical*. Indeed, most ℓ-sets in B_ε are ε-typical; denote by x the number of ε-atypical sets in B_ε. We have

$$\frac{x \cdot \varepsilon \binom{|B|}{k-\ell}}{\binom{k}{\ell}} \leqslant e(B \cup B_\varepsilon) \leqslant \xi \binom{n}{k} + \vartheta |B|^k, \quad \text{implying} \quad x \leqslant \varepsilon' \binom{|B_\varepsilon|}{\ell}. \tag{2.6}$$

Lemma 21. *The following holds for any $B_\varepsilon^{(m)}$-set M if $m \leqslant k-2$.*

$$d(M, A_\varepsilon^{(1)} B_\varepsilon^{(k-m-1)}) + \frac{k-m}{2} d(M, B_\varepsilon^{(k-m)}) \geqslant (1-\delta)\, |A_\varepsilon| \binom{|B_\varepsilon| - m}{k-m-1}.$$

In particular, the following holds for any ε-typical $B^{(\ell)}$-set L.

$$d(L, A_\varepsilon^{(1)} B_\varepsilon^{(k-\ell-1)}) \geqslant (1-2\delta) |A_\varepsilon| \binom{|B_\varepsilon| - \ell}{k-\ell-1}.$$

In the proof of the main lemma we will connect two ε-typical sets only using vertices that are unused so far. Even more, we want to connect two ε-typical sets using exactly one vertex from A. The following corollary of Lemma 21 allows us to do this.

Corollary 22. *Let L and L' be two disjoint ε-typical sets in B_ε and $U \subset V$ with $|U| \leqslant \varepsilon n$. Then the following holds.*

(a) *There exists an ℓ-path disjoint from U of size two with ends L and L' that contains exactly one vertex from A_ε.*

(b) *There exist $a \in A_\varepsilon \smallsetminus U$ and a set $(k-\ell-1)$-set $C \subset B_\varepsilon \smallsetminus U$ such that $L \cup a \cup C$ is an edge in $\mathcal{H}$ and every ℓ-subset of C is ε-typical.*

Proof of Corollary 22. For (a), the second part of Lemma 21 for L and L' implies that they both extend to an edge with at least $(1-2\delta)|A_\varepsilon|\binom{|B_\varepsilon|-\ell}{k-\ell-1}$ sets in $A_\varepsilon^{(1)}B_\varepsilon^{(k-\ell-1)}$. Only few of those intersect U and by an averaging argument we obtain two sets $C, C' \in A_\varepsilon^{(1)}B_\varepsilon^{(k-\ell-1)}$ such that $|C \cap C'| = \ell$ and $L \cup C$ as well as $L' \cup C'$ are edges in $\mathcal{H}$, which yields the required ℓ-path. In view of (2.6), (b) is a trivial consequence of the second part of Lemma 21. $\square$

Proof of Lemma 21. Let $m \leqslant k-2$ and let $M \in B_\varepsilon^{(m)}$ be an m-set. We will make use of the following sum over all $(k-2)$-sets $D \subset B_\varepsilon$ that contain M.

$$\begin{aligned}\sum_{\substack{M\subset D\subset B_\varepsilon\\ |D|=k-2}} d(D) = \sum_{\substack{M\subset D\subset B_\varepsilon\\ |D|=k-2}} \Big(& d(D, A_\varepsilon^{(1)}B_\varepsilon^{(1)}) + d(D, (A_\varepsilon \cup V_\varepsilon)^{(2)}) \\ & + d(D, B_\varepsilon^{(2)}) + d(D, V_\varepsilon^{(1)}B_\varepsilon^{(1)})\Big)\end{aligned} \tag{2.7}$$

Note that we can relate the sums $\sum d(D, A_\varepsilon^{(1)}B_\varepsilon^{(1)})$ and $\sum d(D, B_\varepsilon^{(2)})$ in (2.7) to the terms in question as follows.

$$\begin{aligned} d(M, A_\varepsilon^{(1)}B_\varepsilon^{(k-m-1)}) &= \frac{1}{k-m-1}\sum_{\substack{M\subset D\subset B_\varepsilon\\ |D|=k-2}} d(D, A_\varepsilon^{(1)}B_\varepsilon^{(1)}),\\ d(M, B_\varepsilon^{(k-m)}) &= \frac{1}{\binom{k-m}{2}}\sum_{\substack{M\subset D\subset B_\varepsilon\\ |D|=k-2}} d(D, B_\varepsilon^{(2)}).\end{aligned} \tag{2.8}$$

We will bound some of the terms on the right-hand side of (2.7). It directly follows from (2.5) that $d(D, (A_\varepsilon \cup V_\varepsilon)^{(2)}) \leqslant \binom{a+3\vartheta b}{2}$; moreover, $d(D, V_\varepsilon^{(1)}B_\varepsilon^{(1)}) \leqslant 2\vartheta b b_\varepsilon$. Using the minimum $(k-2)$-degree condition (2.3) we obtain

$$\sum_{\substack{M\subset D\subset B_\varepsilon\\ |D|=k-2}} d(D) > \binom{b_\varepsilon - m}{k-m-2}\left(\binom{a}{2} + a(b-k+2)\right).$$

Combining these estimates with (2.7) and (2.8) yields

$$\begin{aligned} & d(M, A_\varepsilon^{(1)}B_\varepsilon^{(k-m-1)}) + \frac{k-m}{2}d(M, B_\varepsilon^{(k-m)}) \\ &\geqslant \frac{1}{k-m-1}\binom{b_\varepsilon - m}{k-m-2}\left(\binom{a}{2} + a(b-k+2) - \binom{a+3\vartheta b}{2} - 2\vartheta b b_\varepsilon\right) \\ &\geqslant (1-\delta)\, a_\varepsilon \binom{b_\varepsilon - m}{k-m-1}. \end{aligned}$$

For the second part of the lemma, note that the definition of ε-typicality and $\varepsilon \ll \delta$ imply that $\frac{k-\ell}{2}d(L, B_\varepsilon^{(k-\ell)})$ is smaller than $\delta a_\varepsilon \binom{b_\varepsilon - \ell}{k-\ell-1}$ for any ε-typical ℓ-set L, which concludes the proof. $\square$

For Lemma 19, we want to construct an ℓ-path $\mathcal{Q}$, such that $V_\varepsilon \subset V(\mathcal{Q})$ and the remaining sets $A_\varepsilon \smallsetminus \mathcal{Q}$ and $B_\varepsilon \smallsetminus \mathcal{Q}$ have the right relative proportion of vertices, i.e., their sizes are in a ratio of one to $(2k - 2\ell - 1)$. If $|A \cap B_\varepsilon| > 0$, then $B \subset B_\varepsilon$ (see (2.4)) and so $\mathcal{Q}$ should cover V_ε and contain the right number of vertices from B_ε. For this, we have to find suitable edges inside B_ε, which the following lemma ensures.

Lemma 23. *Suppose that $q = |A \cap B_\varepsilon| > 0$. Then there exist $2q + 2$ disjoint paths of size three, each of which contains exactly one vertex from A_ε and has two ε-typical sets as its ends.*

Proof. We say that an $(\ell - 1)$-set $M \subset B_\varepsilon$ is *good* if it is a subset of at least $(1 - \sqrt{\varepsilon'})b_\varepsilon$ ε-typical sets, otherwise we say that the set is *bad.* We will first show that there are $2q + 2$ edges in B_ε, each containing one ε-typical and one *good* $(\ell - 1)$-set. Then we will connect pairs of these edges to ℓ-paths of size three.

Suppose that $q = |A \cap B_\varepsilon| > 0$. So $B \subset B_\varepsilon$ by (2.4) and consequently $|B_\varepsilon| = |B| + q$ and $q \leqslant \vartheta|B|$. It is not hard to see from (2.6) that at most a $\sqrt{\varepsilon'}$ fraction of the $(\ell - 1)$-sets in $B_\varepsilon^{(l-1)}$ are bad. Hence, at least

$$\left(1 - \binom{k-2}{\ell}\varepsilon' - \binom{k-2}{\ell-1}\sqrt{\varepsilon'}\right)\binom{b}{k-2}$$

$(k - 2)$-sets in B_ε contain no ε-atypical or bad subset. Let $\mathcal{B} \subset B_\varepsilon^{(k)}$ be the set of edges inside B_ε that contain such a $(k - 2)$-set. For all $M \in B_\varepsilon^{(k-2)}$, by the minimum degree condition, we have $d(M, B_\varepsilon^{(2)}) \geqslant q(b - k + 2) + \binom{q}{2}$ and, with the above, we have

$$\begin{aligned}|\mathcal{B}| &\geqslant \left(1 - \binom{k-2}{\ell}\varepsilon' - \binom{k-2}{\ell-1}\sqrt{\varepsilon'}\right)\binom{b}{k-2}\frac{q(b-k+2)}{\binom{k}{2}}\\ &= \left(1 - \binom{k-2}{\ell}\varepsilon' - \binom{k-2}{\ell-1}\sqrt{\varepsilon'}\right)\binom{b}{k-1}\frac{2q}{k} \geqslant \frac{q}{k}\binom{b}{k-1}.\end{aligned}$$

On the other hand, for any $v \in B_\varepsilon$ we have $d(v, B_\varepsilon^{(k-1)}) < 2\varepsilon\binom{b_\varepsilon}{k-1}$ which implies that any edge in $\mathcal{B}$ intersects at most $2k\varepsilon\binom{b_\varepsilon}{k-1}$ other edges in $\mathcal{B}$. So, in view of $\varepsilon \ll \frac{1}{k}$ we may pick a set $\mathcal{B}'$ of $2q + 2$ disjoint edges in $\mathcal{B}$.

We will connect each of the edges in $\mathcal{B}'$ to an ε-typical set. Assume we have picked the first $i - 1$ desired ℓ-paths, say $\mathcal{P}_1, \ldots, \mathcal{P}_{i-1}$, and denote by U the set of vertices contained in one of the paths or one of the edges in $\mathcal{B}'$. For the rest of this proof, when we pick vertices and edges, they shall always be disjoint from U

and everything chosen before. Let e be an edge in $\mathcal{B}'$ we have not considered yet and pick an arbitrary ε-typical set $L' \subset B_\varepsilon \setminus U$.

We will first handle the cases that $2\ell + 1 < k$ or that $\ell = 1$, $k = 3$. In the first case, a $(k-2)$-set that contains no ε-atypical set already contains two disjoint ε-typical sets. In the second case, an ℓ-set $\{v\}$ is ε-typical for any vertex v in B_ε by the definition of ε-typicality. Hence in both cases e contains two disjoint ε-typical sets, say L_0 and L_1. We can use Corollary 22 (a), as $|U| \leqslant 6kq$, to connect L_1 to L' and obtain an ℓ-path $\mathcal{P}_i$ of size three that contains one vertex in A_ε and has ε-typical ends L_0 and L'.

So now assume that $2\ell + 1 = k$ and $k > 3$, in particular $k - 2 = 2\ell - 1$ and we may split the $(k-2)$-set considered in the definition of $\mathcal{B}$ into an ε-typical ℓ-set L and a good $(\ell - 1)$-set G. Moreover, let $w \in e \setminus (L \cup G)$ be one of the remaining two vertices and set $N = G \cup w$.

First assume that $d(N, A_\varepsilon^{(1)} B_\varepsilon^{(\ell)}) \geqslant \frac{\delta}{3} a_\varepsilon \binom{b_\varepsilon}{\ell}$. As $\vartheta \ll \delta$, at most $\frac{\delta}{3} a_\varepsilon \binom{b}{\ell}$ sets in $A_\varepsilon^{(1)} B_\varepsilon^{(\ell)}$ intersect U. So it follows from Lemma 21 that there exist $A_\varepsilon^{(1)} B_\varepsilon^{(\ell)}$-sets C, C' such that $N \cup C$ and $L' \cup C'$ are edges, $|C \cap C'| = \ell$ and $|C \cap C' \cap A_\varepsilon| = 1$.

Now assume that $d(N, A_\varepsilon^{(1)} B_\varepsilon^{(\ell)}) < \frac{\delta}{3} a_\varepsilon \binom{b_\varepsilon}{\ell}$. As the good set G forms an ε-typical set with most vertices in B_ε, there exists $v \in B_\varepsilon \setminus U$ such that

$$d(N \cup \{v\}, A_\varepsilon^{(1)} B_\varepsilon^{(\ell-1)}) < \delta a_\varepsilon \binom{b_\varepsilon}{\ell - 1}$$

and $G \cup \{v\}$ is an ε-typical set. Lemma 21 implies that

$$\begin{aligned} d(N \cup \{v\}, B_\varepsilon^{(\ell)}) &\geqslant \frac{2}{\ell} \left((1-\delta) a_\varepsilon \binom{b_\varepsilon - (\ell + 1)}{\ell - 1} - \delta a_\varepsilon \binom{b_\varepsilon}{\ell - 1} \right) \\ &\geqslant \frac{2}{\ell} \left(\frac{1}{2} - 2\delta \right) a_\varepsilon \binom{b_\varepsilon}{\ell - 1} \\ &\geqslant \delta \binom{b_\varepsilon}{\ell}. \end{aligned}$$

So there exists an ε-typical ℓ-set $L^* \subset (B_\varepsilon \setminus U)$ such that $N \cup L^* \cup \{v\}$ is an edge in $\mathcal{H}$. Use Lemma 22 (a) to connect L^* to L' and obtain an ℓ-path $\mathcal{P}_i$ of size three that contains one vertex in A_ε and has ε-typical ends $G \cup \{v\}$ and L'. $\square$

If the hypergraph we consider is very close to the extremal example then Lemma 23 does not apply and we will need the following lemma.

Lemma 24. *Suppose that $B = B_\varepsilon$. If n is an odd multiple of $k - \ell$ then there exists a single edge on B_ε containing two ε-typical ℓ-sets. If n is an even multiple of*

$k-\ell$ then there either exist two disjoint edges on B_ε each containing two ε-typical ℓ-sets or an ℓ-path of size two with ε-typical ends.

Proof. For the proof of this lemma all vertices and edges we consider will always be completely contained in B_ε. First assume that there exists an ε-atypical ℓ-set L. Recall that this means that $d(L, B^{(k-\ell)}) > \varepsilon\binom{|B|}{k-\ell}$ so in view of (2.6) and $\varepsilon' \ll \varepsilon$ we can find two disjoint $(k-\ell)$-sets extending it to an edge, each containing an ε-typical set, which would prove the lemma.

So we may assume that all ℓ-sets in $B_\varepsilon^{(\ell)}$ are ε-typical. We infer from the minimum degree condition that B_ε contains a single edge, which proves the lemma in the case that n is an odd multiple of $k-\ell$ and for the rest of the proof we assume that n is an even multiple of $k-\ell$.

Assume for a moment that $\ell = 1$. Recall that in this case any $(k-2)$-set in B in the extremal hypegraph $\mathcal{X}_{k,\ell}(n)$ is contained in one edge. Consequently, the minimum degree condition implies that any $(k-2)$-set in B_ε extends to at least two edges on B_ε. Fix some edge e in B_ε; any other edge on B_ε has to intersect e in at least two vertices or the lemma would hold. Consider any pair of disjoint $(k-2)$-sets K and M in $B_\varepsilon \setminus e$ to see that of the four edges they extend to, there is a pair which is either disjoint or intersect in one vertex, proving the lemma for the case $\ell = 1$.

Now assume that $\ell > 1$. In this case the minimum degree condition implies that any $(k-2)$-set in B_ε extends to at least one edge on B_ε. Again, fix some edge e in B_ε; any other edge on B_ε has to intersect e in at least one vertex or the lemma would hold. Applying the minimum degree condition to all $(k-2)$-sets disjoint from e implies that one vertex $v \in e$ is contained in at least $\frac{1}{2k^2}\binom{|B_\varepsilon|}{k-2}$ edges. We now consider the $(k-1)$-uniform link hypergraph of v on B_ε. Since any two edges intersecting in $\ell-1$ vertices would finish the proof of the lemma, we may assume that there are no such pair of edges. However, a result of Frankl and Füredi [23, Theorem 2.2] guarantees that this $(k-1)$-uniform hypergraph without an intersection of size $\ell-1$ contains at most $\binom{|B_\varepsilon|}{k-\ell-1}$ edges, a contradiction. □

The following lemma will allow us to handle the vertices in V_ε.

Lemma 25. *Let $U \subset B_\varepsilon$ with $|U| \leqslant 4k\vartheta$. There exists a family $\mathcal{P}_1, \dots, \mathcal{P}_{v_\varepsilon}$ of disjoint ℓ-paths of size two, each of which is disjoint from U such that for all*

$i \in [v_\varepsilon]$

$$|V(\mathcal{P}_i) \cap V_\varepsilon| = 1 \quad \textit{and} \quad |V(\mathcal{P}_i) \cap B_\varepsilon| = 2k - \ell - 1,$$

and both ends of $\mathcal{P}_i$ are ε-typical sets.

Proof. Let $V_\varepsilon = \{x_1, \ldots, x_{v_\varepsilon}\}$. We will iteratively pick the paths. Assume we have already chosen ℓ-paths $\mathcal{P}_1, \ldots, \mathcal{P}_{i-1}$ containing the vertices $v_1, \ldots, v_{i-1}$ and satisfying the lemma. Let U' be the set of all vertices in U or in one of those ℓ-paths. From $v_i \notin B_\varepsilon$ we get

$$d(v_i, B_\varepsilon^{(k-1)}) \geqslant d(v_i, B) - |B \smallsetminus B_\varepsilon| \cdot \binom{|B|}{k-2} \geqslant \frac{\varepsilon}{2}\binom{b_\varepsilon}{k-1}.$$

From (2.6) we get that at most $k^\ell \varepsilon' \binom{b_\varepsilon}{k-1}$ sets in $B_\varepsilon^{(k-1)}$ contain at least one ε-atypical ℓ-set. Also, less than $\frac{\varepsilon}{8}\binom{b_\varepsilon}{k-1}$ sets in $B_\varepsilon^{(k-1)}$ contain one of the vertices of U'. In total, at least $\frac{\varepsilon}{4}\binom{b_\varepsilon}{k-1}$ of the $B_\varepsilon^{(k-1)}$-sets form an edge with v_i. So we may pick two edges e and f in $V_\varepsilon^{(1)} B_\varepsilon^{(k-1)}$ that contain the vertex v_i and intersect in ℓ vertices. In particular, these edges form an ℓ-path of size two as required by the lemma. □

We can now proceed with the proof of Lemma 19. Recall that we want to prove the existence of an ℓ-path $\mathcal{Q}$ in $\mathcal{H}$ with ends L_0 and L_1 and a partition

$$A_* \cup B_* = (V \smallsetminus \mathcal{Q}) \cup L_0 \cup L_1$$

satisfying properties (i)–(iv) of Lemma 19. Set $q = |A \cap B_\varepsilon|$. We will split the construction of the ℓ-path $\mathcal{Q}$ into two cases, depending on whether $q = 0$ or not.

First, suppose that $q > 0$. In the following, we denote by U the set of vertices of all edges and ℓ-paths chosen so far. Note that we will always have $|U| \leqslant 20k\vartheta n$ and hence we will be in position to apply Corollary 22. We use Lemma 23 to obtain paths $\mathcal{Q}_1, \ldots, \mathcal{Q}_{2q+2}$ and then we apply Lemma 25 to obtain ℓ-paths $\mathcal{P}_1, \ldots, \mathcal{P}_{v_\varepsilon}$. Every path $\mathcal{Q}_i$, for $i \in [2q+2]$, contains $3k - 2\ell - 1$ vertices from B_ε and one from A_ε, while every $\mathcal{P}_j$, for $j \in [v_\varepsilon]$, contains $2k - \ell - 1$ from B_ε and one from V_ε.

As the ends of all these paths are ε-typical, we apply Corollary 22 (a) repeatedly to connect them to one ℓ-path $\mathcal{P}$. In each of the $v_\varepsilon + 2q + 1$ steps of connecting two ℓ-paths, we used one vertex from A_ε and $2k - 3\ell - 1$ vertices from B_ε. Overall, we have that

$$|V(\mathcal{P}) \cap A_\varepsilon| = v_\varepsilon + 4q + 3,$$

as well as

$$|V(\mathcal{P}) \cap B_\varepsilon| = (4k - 4\ell - 2)v_\varepsilon + (5k - 5\ell - 2)(2q + 2) - (2k - 3\ell - 1).$$

Furthermore $|V(\mathcal{P})| \leqslant 10k\vartheta b$.

Using the identities $a_\varepsilon + b_\varepsilon + v_\varepsilon = n$ and $a_\varepsilon + q + v_\varepsilon = a$, we will now establish property (i) of Lemma 19. Set $s(\mathcal{P}) = (2k-2\ell-1)|A_\varepsilon \smallsetminus V(\mathcal{P})| - |B_\varepsilon \smallsetminus V(\mathcal{P})| - 2\ell$, so

$$\begin{aligned}
s(\mathcal{P}) &= (2k - 2\ell - 1)|A_\varepsilon \smallsetminus V(\mathcal{P})| - |B_\varepsilon \smallsetminus V(\mathcal{P})| - 2\ell \\
&= (2k - 2\ell - 1)(a_\varepsilon - (v_\varepsilon + 4q + 3)) - b_\varepsilon \\
&\quad + (4k - 4\ell - 2)v_\varepsilon + (5k - 5\ell - 2)(2q + 2) - (2k - 3\ell - 1) - 2\ell \\
&= (2k - 2\ell - 1)a_\varepsilon - b_\varepsilon + (2k - 2\ell - 1)v_\varepsilon + 2(k - \ell)q + 2k - 3\ell \\
&= 2(k - \ell)(a_\varepsilon + v_\varepsilon + q + 1) - n - \ell \\
&= 2(k - \ell)(a + 1) - n - \ell.
\end{aligned}$$

If $n/(k - \ell)$ is even, $s(\mathcal{P}) = -\ell$ (see (2.2)) and we set $\mathcal{Q} = \mathcal{P}$. Otherwise $s(\mathcal{P}) = k - 2\ell$ and we use Corollary 22 (b) to append one edge to $\mathcal{P}$ to obtain $\mathcal{Q}$. It is easy to see that one application of Corollary 22 (b) decreases $s(\mathcal{P})$ by $k - \ell$. Setting $A_* = A_\varepsilon \smallsetminus V(\mathcal{Q})$ and $B_* = (B_\varepsilon \smallsetminus V(\mathcal{Q})) \cup L_0 \cup L_1$ we get from $s(\mathcal{Q}) = -\ell$ that A_* and B_* satisfy (i).

Now, suppose that $q = 0$. Apply Lemma 25 to obtain ℓ-paths $\mathcal{P}_1, \dots, \mathcal{P}_{v_\varepsilon}$. If $B = B_\varepsilon$, apply Lemma 24 to obtain one or two more ℓ-paths contained in B_ε. We apply Corollary 22 (a) repeatedly to connect them to one ℓ-path $\mathcal{P}$.

Since $q = 0$, we have that $B_\varepsilon \subset B$ and $a_\varepsilon + v_\varepsilon = |V \smallsetminus B_\varepsilon| = a + |B \smallsetminus B_\varepsilon|$. We can assume without loss of generality that $V_\varepsilon \neq \varnothing$, otherwise just take $V_\varepsilon = \{v\}$ for an arbitrary $v \in V(\mathcal{H})$. If $B = B_\varepsilon$ let x be $2(k - \ell)$ or $k - \ell$ depending on whether n is an odd or even multiple of $k - \ell$; otherwise let $x = 0$. With similar calculations as before and the same definition of $s(\mathcal{P})$ we get that

$$s(\mathcal{P}) = 2(k - \ell)a + x + 2(k - \ell)|B \smallsetminus B_\varepsilon| - n - \ell \equiv -\ell \mod (k - \ell).$$

Extend the ℓ-path $\mathcal{P}$ to an ℓ-path $\mathcal{Q}$ by adding $\frac{s(\mathcal{P})+\ell}{k-l}$ edges using Corollary 22 (b). Thus $s(\mathcal{Q}) = -\ell$, and we get (i) as in the previous case.

In both cases, we will now use the properties of the constructed ℓ-path $\mathcal{Q}$ to show (ii)-(iv). We will use that $v(\mathcal{Q}) \leqslant 20k\vartheta b$, which follows from the construction.

Since $A_* \subset A_\varepsilon$, for all $v \in A_*$ we have $d(v, B^{(k-1)}) \geqslant (1-\varepsilon)B^{(k-1)}$. Thus

$$d(v, B_*^{(k-1)}) \geqslant d(v, B^{(k-1)}) - |B_* \setminus B| \binom{|B_*|-1}{k-2} \geqslant (1-2\varepsilon)\binom{|B_*|}{k-1},$$

which shows (ii).

For (iii), Lemma 21 yields for all vertices $v \in B_* \subset B_\varepsilon$ that

$$d(v, A_\varepsilon^{(1)} B_\varepsilon^{(k-2)}) + \frac{k-1}{2} d(v, B_\varepsilon^{(k-1)}) \geqslant (1-\delta)\,|A_\varepsilon| \binom{|B_\varepsilon|-1}{k-2}.$$

The second term on the left can be bounded from above by $2k\varepsilon\binom{b_\varepsilon}{k-1}$. So, as $\delta, \varepsilon \ll \varrho$ and $a_\varepsilon - |A_*| \ll \varrho|A_*|$ as well as $b_\varepsilon - |B_*| \ll \varrho|B_*|$, we can conclude (iii).

By Lemma 21, we know that

$$d(L_0, A_\varepsilon^{(1)} B_\varepsilon^{(k-1)}), d(L_1, A_\varepsilon^{(1)} B_\varepsilon^{(k-1)}) \geqslant (1-\delta) a_\varepsilon \binom{b_\varepsilon - \ell}{k-\ell-1}.$$

As $\delta \ll \varrho$ and $a_\varepsilon - |A_*| \ll \varrho|A_*|$ as well as $b_\varepsilon - |B_*| \ll \varrho|B_*|$, we can conclude (iv), concluding the proof of Lemma 19.

3 Spanning subgraphs in sparse random graphs

In this chapter, we will prove the following result.

Theorem 26. *For each $\gamma > 0$, $\Delta \geqslant 2$, $k \geqslant 2$ and $0 \leqslant s \leqslant k-1$, there exist constants $\beta^* > 0$ and $C^* > 0$ such that the following holds asymptotically almost surely for $\Gamma = G(n,p)$ if $p \geqslant C^*\big(\frac{\log n}{n}\big)^{1/\Delta}$. Let G be a spanning subgraph of Γ with $\delta(G) \geqslant \big(\frac{k-1}{k} + \gamma\big)pn$, such that for each $v \in V(G)$ there are at least $\gamma p^{\binom{s}{2}}(pn)^s$ copies of K_s in $N_G(v)$. Let H be a graph on n vertices with $\Delta(H) \leqslant \Delta$, bandwidth at most $\beta^* n$ and suppose that there is a proper k-colouring of $V(H)$ and at least $C^* p^{-2}$ vertices in $V(H)$ whose neighbourhood contains only s colours. Then G contains a copy of H.*

We prove Theorem 26 by making use of the sparse regularity lemma of Kohayakawa and Rödl [32, 36], the sparse blowup lemma of [3], and several lemmas from [2]. In Section 3.1 we give the definitions and results necessary to state and use the sparse regularity lemma and the sparse blowup lemma, and conclude with a few probabilistic lemmas. In Section 3.2 we give a somewhat more general statement (Theorem 43) than Theorem 26, which allows for graphs H which are not quite k-colourable, and outline briefly how to prove it using various lemmas. The basic strategy, and most of the lemmas, are taken from [2]. The exception is Lemma 46, which replaces the 'common neighbourhood lemma' of [2]. Proving this lemma is the main work of this chapter, and we do it in Section 3.3. We give the proof of Theorem 43 in Section 3.4; we should stress that this proof is a fairly minor modification of the corresponding proof in [2] which we include here mainly for completeness' sake. Finally, we give some remarks on the optimality of the results in Section 3.5.

3.1 Preliminaries

Throughout the chapter log denotes the natural logarithm. We assume that the order n of all graphs tends to infinity and therefore is sufficiently large whenever necessary. Let $G = (V, E)$ be graph. For disjoint vertex sets $A, B \subseteq V$ we denote the number of edges between A and B by $e(A, B)$. For a vertex $v \in V(G)$ we write $N_G(v)$ for the neighbourhood of v in G and $N_G(v, A) := N_G(v) \cap A$ for the neighbourhood of v restricted to A. Finally, let $\deg_G(v) := |N_G(v)|$ be the degree of v in G. For the sake of readability, we do not make any effort to optimise the constants in our theorems and proofs.

Now we introduce some definitions and results of the regularity method as well as related tools that are essential in our proofs. In particular, we state a minimum degree version and a refining version of the sparse regularity lemma (Lemma 30 and Lemma 31) and the sparse blowup lemma (Lemma 34). These lemmas use the concept of regular pairs. Let $G = (V, E)$ be a graph, $\varepsilon, d > 0$, and $p \in (0, 1]$. Moreover, let $X, Y \subseteq V$ be two disjoint nonempty sets. The *p-density* of the pair (X, Y) is defined as

$$d_{G,p}(X, Y) := \frac{e_G(X, Y)}{p|X||Y|}.$$

The pair (X, Y) is called *$(\varepsilon, p)_G$-regular* if $|d_{G,p}(X', Y') - d_{G,p}(X, Y)| \leqslant \varepsilon$ for all $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| \geqslant \varepsilon|X|$ and $|Y'| \geqslant \varepsilon|Y|$. Whereas this definition of regular pairs is used for instance in Lemma 31, we will mainly use the following definition of (super-)(lower-)regular pairs, the density of which only has to be bounded from below.

Definition 27 ((ε, d, p)-(super-)(lower-)regular pairs)**.** *The pair (X, Y) is called* $(\varepsilon, d, p)_G$-lower-regular *if for every $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| \geqslant \varepsilon|X|$ and $|Y'| \geqslant \varepsilon|Y|$ we have $d_{G,p}(X', Y') \geqslant d - \varepsilon$.*

It is called $(\varepsilon, d, p)_G$-regular *if there exists $d' \geqslant d$ such that for every $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| \geqslant \varepsilon|X|$ and $|Y'| \geqslant \varepsilon|Y|$ we have $d_{G,p}(X', Y') = d' \pm \varepsilon$.*

If (X, Y) is either $(\varepsilon, d, p)_G$-lower-regular or $(\varepsilon, d, p)_G$-regular, and in addition we have

$$|N_G(x, Y)| \geqslant (d - \varepsilon) \max\big(p|Y|, \deg_\Gamma(x, Y)/2\big) \quad \textit{and}$$
$$|N_G(y, X)| \geqslant (d - \varepsilon) \max\big(p|X|, \deg_\Gamma(y, X)/2\big)$$

*for every $x \in X$ and $y \in Y$, then the pair (X, Y) is called $(\varepsilon, d, p)_G$-*super-regular*. When we use super-regularity it will be clear from the context whether (X, Y) is lower-regular or regular.*

Note that a regular pair is by definition lower-regular, though the converse does not hold. Although the definition of super-regularity of G, which we need to use the results of [3], contains a reference to Γ, at each place in this chapter where we use super-regularity, we will see that the first term in the maximum is larger than the second. When it is clear from the context, we may omit the subscript G in $(\varepsilon, d, p)_G$-(super-)regular which is used to indicate with respect to which graph a pair is (super-)regular. A direct consequence of the definition of (ε, d, p)-lower-regular pairs is the following proposition about the sizes of neighbourhoods in lower-regular pairs.

Proposition 28. *Let (X, Y) be (ε, d, p)-lower-regular. Then there are less than $\varepsilon|X|$ vertices $x \in X$ with $|N(x, Y)| < (d - \varepsilon)p|Y|$.* □

The following proposition is another immediate consequence of Definition 27. It states that an (ε, d, p)-regular pair is still regular if only a linear fraction of its vertices is removed.

Proposition 29. *Let (X, Y) be (ε, d, p)-regular and suppose $X' \subseteq X$ and $Y' \subseteq Y$ satisfy $|X'| \geqslant \mu|X|$ and $|Y'| \geqslant \nu|Y|$ with some $\mu, \nu > 0$. Then (X', Y') is $(\frac{\varepsilon}{\min\{\mu,\nu\}}, d, p)$-regular.* □

In order to state the sparse regularity lemma, we need some more definitions. A partition $\mathcal{V} = \{V_i\}_{i \in \{0,\dots,r\}}$ of the vertex set of G is called an $(\varepsilon, p)_G$-*regular partition* of $V(G)$ if $|V_0| \leqslant \varepsilon|V(G)|$ and $(V_i, V_{i'})$ forms an $(\varepsilon, 0, p)_G$-regular pair for all but at most $\varepsilon\binom{r}{2}$ pairs $\{i, i'\} \in \binom{[r]}{2}$. It is called an *equipartition* if $|V_i| = |V_{i'}|$ for every $i, i' \in [r]$. The partition $\mathcal{V}$ is called (ε, d, p)-*(lower-)regular* on a graph R with vertex set $[r]$ if $(V_i, V_{i'})$ is $(\varepsilon, d, p)_G$-(lower-)regular for every $\{i, i'\} \in E(R)$. The graph R is referred to as the $(\varepsilon, d, p)_G$-*reduced graph* of $\mathcal{V}$, the partition classes V_i with $i \in [r]$ as *clusters*, and V_0 as the *exceptional set*. We also say that $\mathcal{V}$ is $(\varepsilon, d, p)_G$-*super-regular* on a graph R' with vertex set $[r]$ if $(V_i, V_{i'})$ is $(\varepsilon, d, p)_G$-super-regular for every $\{i, i'\} \in E(R')$. Again, when we talk about reduced graphs or super-regularity, whether we are using lower-regularity or regularity will be clear from the context. We will however always specify whether a partition is regular or only lower-regular on R.

Analogously to Szemeredi's regularity lemma for dense graphs, the sparse regularity lemma, proved by Kohayakawa and Rödl [32, 36], asserts the existence of an (ε, p)-regular partition of constant size of any sparse graph. We state a minimum degree version of this lemma, whose proof (following [10]) can be found in the appendix of [2].

Lemma 30 (Minimum degree version of the sparse regularity lemma)**.** *For each $\varepsilon > 0$, each $\alpha \in [0, 1]$, and $r_0 \geqslant 1$ there exists $r_1 \geqslant 1$ with the following property. For any $d \in [0, 1]$, any $p > 0$, and any n-vertex graph G with minimum degree αpn such that for any disjoint $X, Y \subset V(G)$ with $|X|, |Y| \geqslant \frac{\varepsilon n}{r_1}$ we have $e(X, Y) \leqslant \big(1 + \frac{1}{1000}\varepsilon^2\big)p|X||Y|$, there is an $(\varepsilon, p)_G$-regular equipartition of $V(G)$ with $(\varepsilon, d, p)_G$-reduced graph R satisfying $\delta(R) \geqslant (\alpha - d - \varepsilon)|V(R)|$ and $r_0 \leqslant |V(R)| \leqslant r_1$.* $\square$

We will need the following version of the sparse regularity lemma (see, e.g., [2, Lemma 29] for a proof), allowing for a partition equitably refining an initial partition with parts of very different sizes. Given a partition $V(G) = V_1 \cup \ldots \cup V_s$, we say a partition $\{V_{i,j}\}_{i\in[s],j\in[t]}$ is an equitable (ε, p)-regular refinement of $\{V_i\}_{i\in[s]}$ if $|V_{i,j}| = |V_{i,j'}| \pm 1$ for each $i \in [s]$ and $j, j' \in [t]$, and there are at most $\varepsilon s^2 t^2$ pairs $(V_{i,j}, V_{i',j'})$ which are not $(\varepsilon, 0, p)$-regular.

Lemma 31. *For each $\varepsilon > 0$ and $s \in \mathbb{N}$ there exists $t_1 \geqslant 1$ such that the following holds. Given any graph G, suppose $V_1 \cup \ldots \cup V_s$ is a partition of $V(G)$. Suppose that $e(V_i) \leqslant 3p|V_i|^2$ for each $i \in [s]$, and $e(V_i, V_{i'}) \leqslant 2p|V_i||V_{i'}|$ for each $i \neq i' \in [s]$. Then there exist sets $V_{i,0} \subset V_i$ for each $i \in [s]$ with $|V_{i,0}| < \varepsilon|V_i|$, and an equitable (ε, p)-regular refinement $\{V_{i,j}\}_{i\in[s],j\in[t]}$ of $\{V_i \setminus V_{i,0}\}_{i\in[s]}$ for some $t \leqslant t_1$.* $\square$

A key ingredient in the proof of our main theorem is the so-called sparse blowup lemma developed by Allen, Böttcher, Hàn, Kohayakawa, and Person. Given a subgraph $G \subseteq \Gamma = G(n, p)$ with $p \gg (\log n/n)^{1/\Delta}$ and an n-vertex graph H with maximum degree at most Δ with vertex partitions $\mathcal{V}$ and $\mathcal{W}$, respectively, the sparse blowup lemma guarantees under certain conditions a spanning embedding of H in G which respects the given partitions. In order to state this lemma we need to introduce some definitions.

Definition 32 ((ϑ, R')-buffer). *Let R' be a graph on r vertices and let H be a graph with vertex partition $\mathcal{W} = \{W_i\}_{i\in[r]}$. We say that the family $\widetilde{\mathcal{W}} = \{\widetilde{W}_i\}_{i\in[r]}$ of subsets $\widetilde{W}_i \subseteq W_i$ is an* (ϑ, R')-buffer *for H if*

(i) $|\widetilde{W}_i| \geqslant \vartheta|W_i|$ *for all $i \in [r]$, and*

(ii) *for each $i \in [r]$ and each $x \in \widetilde{W}_i$, the first and second neighbourhood of x go along R', i.e., for each $\{x, y\}, \{y, z\} \in E(H)$ with $y \in W_j$ and $z \in W_k$ we have $\{i, j\} \in E(R')$ and $\{j, k\} \in E(R')$.*

Let G and H be graphs on n vertices with partitions $\mathcal{V} = \{V_i\}_{i\in[r]}$ of $V(G)$ and $\mathcal{W} = \{W_i\}_{i\in[r]}$ of $V(H)$. We say that $\mathcal{V}$ and $\mathcal{W}$ are *size-compatible* if $|V_i| = |W_i|$ for all $i \in [r]$. If there exists an integer $m \geqslant 1$ such that $m \leqslant |V_i| \leqslant \kappa m$ for every $i \in [r]$, then we say that $(G, \mathcal{V})$ is κ-balanced. Given a graph R on r vertices, we call $(G, \mathcal{V})$ an *R-partition* if for every edge $\{x, y\} \in E(G)$ with $x \in V_i$ and $y \in V_{i'}$ we have $\{i, i'\} \in E(R)$.

Definition 33 (Restriction pair). *Let $\varepsilon, d > 0$, $p \in [0, 1]$, and let R be a graph on r vertices. Furthermore, let G be a (not necessarily spanning) subgraph of $\Gamma = G(n, p)$ and let H be a graph given with vertex partitions $\mathcal{V} = \{V_i\}_{i\in[r]}$ and $\mathcal{W} = \{W_i\}_{i\in[r]}$, respectively, such that $(G, \mathcal{V})$ and $(H, \mathcal{W})$ are size-compatible R-partitions. Let $\mathcal{I} = \{I_x\}_{x\in V(H)}$ be a collection of subsets of $V(G)$, called* image restrictions, *and $\mathcal{J} = \{J_x\}_{x\in V(H)}$ be a collection of subsets of $V(\Gamma) \smallsetminus V(G)$, called* restricting vertices. *For each $i \in [r]$ we define $R_i \subseteq W_i$ to be the set of all vertices $x \in W_i$ for which $I_x \neq V_i$. We say that $\mathcal{I}$ and $\mathcal{J}$ are a $(\varrho, \zeta, \Delta, \Delta_J)$*-restriction pair *if the following properties hold for each $i \in [r]$ and $x \in W_i$.*

($RP1$) *We have $|R_i| \leqslant \varrho|W_i|$.*

($RP2$) *If $x \in R_i$, then $I_x \subseteq \bigcap_{u\in J_x} N_\Gamma(u, V_i)$ is of size at least $\zeta(dp)^{|J_x|}|V_i|$.*

($RP3$) *If $x \in R_i$, then $|J_x| + \deg_H(x) \leqslant \Delta$ and if $x \in W_i \smallsetminus R_i$, then $J_x = \varnothing$.*

($RP4$) *Each vertex in $V(G)$ appears in at most Δ_J of the sets of $\mathcal{J}$.*

($RP5$) *We have $\left|\bigcap_{u\in J_x} N_\Gamma(u, V_i)\right| = (p \pm \varepsilon p)^{|J_x|}|V_i|$.*

($RP6$) *If $x \in R_i$, for each $xy \in E(H)$ with $y \in W_j$,*

$$\text{the pair} \quad \Big(V_i \cap \bigcap_{u\in J_x} N_\Gamma(u), V_j \cap \bigcap_{v\in J_y} N_\Gamma(v)\Big) \quad \text{is } (\varepsilon, d, p)_G\text{-lower-regular.}$$

Suppose $\mathcal{V}$ is an $(\varepsilon, d, p)_G$-regular partition of $V(G)$ with reduced graph R. We say $(G, \mathcal{V})$ has *one-sided inheritance* with respect to R if for every $\{i, j\}, \{j, k\} \in E(R)$ and every $v \in V_i$ the pair $\big(N_\Gamma(v, V_j), V_k\big)$ is $(\varepsilon, d, p)_G$-regular, and $V_i \in \mathcal{V}$ has *two-sided inheritance* with respect to $V_j, V_k \in \mathcal{V}$ if for every $v \in V_i$ the pair $\big(N_\Gamma(v, V_j), N_\Gamma(v, V_k)\big)$ is $(\varepsilon, d, p)_G$-regular.

Now we can finally state the sparse blowup lemma.

Lemma 34 ([3, Lemma 1.21])**.** *For each Δ, $\Delta_{R'}$, Δ_J, $\vartheta, \zeta, d > 0$, $\kappa > 1$ there exist $\varepsilon_{\text{BL}}, \varrho > 0$ such that for all r_1 there is a C_{BL} such that for $p \geqslant C_{\text{BL}}(\log n/n)^{1/\Delta}$ the random graph $\Gamma = G_{n,p}$ asymptotically almost surely satisfies the following.*

Let R be a graph on $r \leqslant r_1$ vertices and let $R' \subseteq R$ be a spanning subgraph with $\Delta(R') \leqslant \Delta_{R'}$. Let H and $G \subseteq \Gamma$ be graphs given with κ-balanced, size-compatible vertex partitions $\mathcal{W} = \{W_i\}_{i\in[r]}$ and $\mathcal{V} = \{V_i\}_{i\in[r]}$ with parts of size at least $m \geqslant n/(\kappa r_1)$. Let $\mathcal{I} = \{I_x\}_{x\in V(H)}$ be a family of image restrictions, and $\mathcal{J} = \{J_x\}_{x\in V(H)}$ be a family of restricting vertices. Suppose that

- *(BUL1) $\Delta(H) \leqslant \Delta$, for every edge $\{x, y\} \in E(H)$ with $x \in W_i$ and $y \in W_j$ we have $\{i, j\} \in E(R)$ and $\widetilde{\mathcal{W}} = \{\widetilde{W}_i\}_{i\in[r]}$ is an (ϑ, R')-buffer for H,*
- *(BUL2) $\mathcal{V}$ is $(\varepsilon_{\text{BL}}, d, p)_G$-lower-regular on R, $(\varepsilon_{\text{BL}}, d, p)_G$-super-regular on R', has one-sided inheritance on R', and two-sided inheritance on R' for $\widetilde{\mathcal{W}}$,*
- *(BUL3) $\mathcal{I}$ and $\mathcal{J}$ form a $(\varrho, \zeta, \Delta, \Delta_J)$-restriction pair.*

Then there is an embedding $\varphi\colon V(H) \to V(G)$ such that $\varphi(x) \in I_x$ for each $x \in H$. □

Observe that in the blowup lemma for dense graphs, proved by Komlós, Sárközy, and Szemerédi [38], one does not need to explicitly ask for one- and two-sided inheritance properties since they are always fulfilled by dense regular partitions. This is, however, not true in general in the sparse setting. The following two lemmas will be very useful whenever we need to redistribute vertex partitions in order to achieve some regularity inheritance properties.

Lemma 35 (One-sided lower-regularity inheritance, [3])**.** *For each $\varepsilon_{\text{OSRIL}}, \alpha_{\text{OSRIL}} > 0$ there exist $\varepsilon_0 > 0$ and $C > 0$ such that for any $0 < \varepsilon < \varepsilon_0$ and $0 < p < 1$ asymptotically almost surely $\Gamma = G(n, p)$ has the following property. For any disjoint sets X and Y in $V(\Gamma)$ with $|X| \geqslant C \max\big(p^{-2}, p^{-1}\log n\big)$ and $|Y| \geqslant$*

$Cp^{-1}\log n$, *and any subgraph* G *of* $\Gamma[X,Y]$ *which is* $(\varepsilon, \alpha_{\text{OSRIL}}, p)_G$*-lower-regular, there are at most* $Cp^{-1}\log(en/|X|)$ *vertices* $z \in V(\Gamma)$ *such that* $(X \cap N_\Gamma(z), Y)$ *is not* $(\varepsilon_{\text{OSRIL}}, \alpha_{\text{OSRIL}}, p)_G$*-lower-regular.* $\square$

Lemma 36 (Two-sided lower-regularity inheritance, [3]). *For each* $\varepsilon_{\text{TSRIL}}, \alpha_{\text{TSRIL}} > 0$ *there exist* $\varepsilon_0 > 0$ *and* $C > 0$ *such that for any* $0 < \varepsilon < \varepsilon_0$ *and* $0 < p < 1$*, asymptotically almost surely* $\Gamma = G_{n,p}$ *has the following property. For any disjoint sets* X *and* Y *in* $V(\Gamma)$ *with* $|X|, |Y| \geqslant C\max\{p^{-2}, p^{-1}\log n\}$*, and any subgraph* G *of* $\Gamma[X,Y]$ *which is* $(\varepsilon, \alpha_{\text{TSRIL}}, p)_G$*-lower-regular, there are at most* $C\max\{p^{-2}, p^{-1}\log(en/|X|)\}$ *vertices* $z \in V(\Gamma)$ *such that* $\big(X \cap N_\Gamma(z), Y \cap N_\Gamma(z)\big)$ *is not* $(\varepsilon_{\text{TSRIL}}, \alpha_{\text{TSRIL}}, p)_G$*-lower-regular.* $\square$

Finally, we need a statement about random subpairs of regular pairs.

Corollary 37 ([24, Corollary 3.8]). *For any* d, β, $\varepsilon' > 0$ *there exist* $\varepsilon_0 > 0$ *and* C *such that for any* $0 < \varepsilon < \varepsilon_0$ *and* $0 < p < 1$*, if* (X,Y) *is an* (ε, d, p)*-lower-regular pair in a graph* G*, then the number of pairs* $X' \subseteq X$ *and* $Y' \subseteq Y$ *with* $|X'| = w_1 \geqslant C/p$ *and* $|Y'| = w_2 \geqslant C/p$ *such that* (X', Y') *is an* (ε', d, p)*-lower-regular pair in* G *is at least* $(1 - \beta^{\min(w_1,w_2)})\binom{|X|}{w_1}\binom{|Y|}{w_2}$. $\square$

We close this section with two of Chernoff's bounds for random variables that follow a binomial (Theorem 39) and a hypergeometric distribution (Theorem 40), respectively, and the following useful observation. Roughly speaking, it states that a.a.s. nearly all vertices in $G(n,p)$ have approximately the expected number of neighbours within large enough subsets.

Proposition 38 ([2]). *For each* $\varepsilon > 0$ *there exists a constant* $C > 0$ *such that for every* $0 < p < 1$ *asymptotically almost surely* $\Gamma = G(n,p)$ *has the following properties. For any disjoint* $X, Y \subset V(\Gamma)$ *with* $|X| \geqslant Cp^{-1}\log n$ *and* $|Y| \geqslant Cp^{-1}\log(en/|X|)$*, we have* $e(X,Y) = (1 \pm \varepsilon)p|X||Y|$ *and* $e(X) \leqslant 2p|X|^2$*. Furthermore, for every* $X \subseteq V(\Gamma)$ *with* $|X| \geqslant Cp^{-1}\log n$*, the number of vertices* $v \in V(\Gamma)$ *with* $\big||N_\Gamma(v,X)| - p|X|\big| > \varepsilon p|X|$ *is at most* $Cp^{-1}\log(en/|X|)$. $\square$

Note that in most of this chapter we will use the upper bound $\log(en/|X|) \leqslant \log n$ when applying this proposition, and Lemmas 35 and 36, valid since (in all applications) we have $|X| \geqslant e$. The full strength of these three results is only needed in the proof of the Lemma for G (Lemma 44).

The proof of Proposition 38 in [2] uses the following version of Chernoff's inequalities (see, e.g., [29, Chapter 2] for a proof).

Theorem 39 (Chernoff's inequality, [29])**.** *Let X be a random variable which is the sum of independent Bernoulli random variables. Then we have for $\varepsilon \leqslant 3/2$*

$$\mathbb{P}\big[|X - \mathbb{E}[X]| > \varepsilon\mathbb{E}[X]\big] < 2e^{-\varepsilon^2\mathbb{E}[X]/3}\,.$$

Furthermore, if $t \geqslant 6\mathbb{E}[X]$ then we have

$$\mathbb{P}\big[X \geqslant \mathbb{E}[X] + t\big] \leqslant e^{-t}\,. \qquad \square$$

Finally, let N, m, and s be positive integers and let S and $S' \subseteq S$ be two sets with $|S| = N$ and $|S'| = m$. The *hypergeometric distribution* is the distribution of the random variable X that is defined by drawing s elements of S without replacement and counting how many of them belong to S'. It can be shown that Theorem 39 still holds in the case of hypergeometric distributions (see, e.g., [29, Chapter 2] for a proof) with $\mathbb{E}[X] := ms/N$.

Theorem 40 (Hypergeometric inequality, [29])**.** *Let X be a random variable that follows the hypergeometric distribution with parameters N, m, and s. Then for any $\varepsilon > 0$ and $t \geqslant \varepsilon ms/N$ we have*

$$\mathbb{P}\big[|X - ms/N| > t\big] < 2e^{-\varepsilon^2 t/3}\,. \qquad \square$$

We require the following technical lemma, which is a consequence of the hypergeometric inequality stated in Theorem 40.

Lemma 41. *For each $\varepsilon_0^+, d^+ > 0$ there exists $\varepsilon^+ > 0$ such that for each $\varepsilon, d > 0$ there exists $\varepsilon^- > 0$ such that the following holds. For each $\eta > 0$ and Δ there exists C such that the following holds for each $p > 0$. Let $W \subset [n]$, let $t \leqslant 100n^\Delta$, and let $T_1, \dots, T_t$ be subsets of W. For each $m \leqslant |W|$ there is a set $S \subset W$ of size m such that*

$$|T_i \cap S| = \tfrac{m}{|W|}|T_i| \pm \big(\eta|T_i| + C\log n\big) \text{ for every } i \in [t]\,.$$

Suppose furthermore that for each $i \in [t]$ there is a pair (X_i, Y_i) which is either (ε^+, d^+, p)-lower-regular, or (ε^-, d, p)-lower-regular, in a graph G on W. If $m|X_i|/|W|, m|Y_i|/|W| \geqslant 2Cp^{-1}\log n$ for each $i \in [t]$, then we may choose S such that in addition the pair $\big(X_i \cap S, Y_i \cap S\big)$ is $\big(\varepsilon_0^+, d^+, p\big)$-lower-regular, or (ε, d, p)-lower-regular (respectively) in G for each $i \in [t]$.

Proof. Given ε_0^+, d^+, let ε^+ be returned by Corollary 37 for input d^+, $\beta = \frac{1}{2}$ and ε_0^+. Given ε, d, let ε^- be returned by Corollary 37 for input d, $\beta = \frac{1}{2}$ and ε. Let $C \geqslant 30\eta^{-2}\Delta$ be large enough for these applications of Corollary 37.

Observe that for each i, the size of $T_i \cap S$ is hypergeometrically distributed. By Theorem 40, for each i we have

$$\mathbb{P}\big[|T_i \cap S| \neq \tfrac{m}{|W|}|T_i| \pm \big(\eta|T_i| + C\log n\big)\big] < 2e^{-\eta^2 C\log n/3} < \frac{2}{n^{1+\Delta}}\,,$$

so taking the union bound over all $i \in [t]$ we conclude that the probability of failure is at most $2t/n^{1+\Delta} \leqslant 200/n \to 0$ as $n \to \infty$, as desired.

To obtain the 'furthermore' statement, observe that the same application of Theorem 40 implies that we have $|X_i \cap S|, |Y_i \cap S| \geqslant Cp^{-1}\log n$ for each $i \in [t]$ with probability tending to one as $n \to \infty$. Conditioning on the size of $|X_i \cap S|$, the set $X_i \cap S$ is a uniformly distributed subset of X_i of size $|X_i \cap S|$, and the same applies to $Y_i \cap S$. Now Corollary 37 says that, conditioning on $|X_i \cap S|, |Y_i \cap S| \geqslant Cp^{-1}\log n$, the probability that $\big(X_i \cap S, Y_i \cap S\big)$ fails to have the desired lower-regularity in G is at most $2^{-Cp^{-1}\log n}$, and taking a union bound over the choices of i the result follows. □

3.2 Main lemmas

As in [2], we deduce Theorem 26 from a slightly more technical statement, see Theorem 43 below. As there, this result is more general (if harder to parse) in that it allows for an extra colour, *zero*, in the colouring of H, provided that this colour does not appear too often.

Definition 42 (Zero-free colouring)**.** *Let H be a $(k+1)$-colourable graph on n vertices and let $\mathcal{L}$ be a labelling of its vertex set of bandwidth at most βn. A proper $(k+1)$-colouring $\sigma : V(H) \to \{0, \ldots, k\}$ of its vertex set is said to be (z, β)-*zero-free *with respect to $\mathcal{L}$ if any z consecutive blocks contain at most one block with colour zero, where a block is defined as a set of the form $\{(t-1)4k\beta n + 1, \ldots, t4k\beta n\}$ with some $t \in [1/(4k\beta)]$.*

We can now state the following technical statement, from which one can easily deduce Theorem 26.

Theorem 43. *For each $\gamma > 0$, $\Delta \geqslant 2$, $k \geqslant 2$ and $1 \leqslant s \leqslant k-1$, there exist constants $\beta > 0$, $z > 0$, and $C > 0$ such that the following holds asymptotically almost surely for $\Gamma = G(n,p)$ if $p \geqslant C\big(\frac{\log n}{n}\big)^{1/\Delta}$. Let G be a spanning subgraph of Γ with $\delta(G) \geqslant \big(\frac{k-1}{k} + \gamma\big)pn$ such that for each $v \in V(G)$ there are at least $\gamma p^{\binom{s}{2}}(pn)^s$ copies of K_s in $N_G(v)$ and let H be a graph on n vertices with $\Delta(H) \leqslant \Delta$ that has a labelling $\mathcal{L}$ of its vertex set of bandwidth at most βn, a $(k+1)$-colouring that is (z,β)-zero-free with respect to $\mathcal{L}$ and where the first $\sqrt{\beta}n$ vertices in $\mathcal{L}$ are not given colour zero and the first βn vertices in $\mathcal{L}$ include Cp^{-2} vertices whose neighbourhood contains only s colours. Then G contains a copy of H.*

The proof of this theorem is quite similar to the corresponding [2, Theorem 23]. Eventually, we will apply Lemma 34 to embed H into G, and we need to obtain the necessary conditions for this lemma. As in [2], this is as such not possible; whatever regular partition of G we take, there may be some exceptional vertices which are 'badly behaved' with respect to this partition. Our first main lemma, the following Lemma for G, states that there is a partition with only few such vertices.

Lemma 44 (Lemma for G, [2, Lemma 24])**.** *For each $\gamma > 0$ and integers $k \geqslant 2$ and $r_0 \geqslant 1$ there exists $d > 0$ such that for every $\varepsilon \in \big(0, \frac{1}{2k}\big)$ there exist $r_1 \geqslant 1$ and $C^* > 0$ such that the following holds a.a.s. for $\Gamma = G(n,p)$ if $p \geqslant C^*\,(\log n/n)^{1/2}$. Let $G = (V,E)$ be a spanning subgraph of Γ with $\delta(G) \geqslant \big(\frac{k-1}{k} + \gamma\big)\,pn$. Then there exists an integer r with $r_0 \leqslant kr \leqslant r_1$, a subset $V_0 \subseteq V$ with $|V_0| \leqslant C^*p^{-2}$, a k-equitable vertex partition $\mathcal{V} = \{V_{i,j}\}_{i\in[r],j\in[k]}$ of $V(G) \setminus V_0$, and a graph R^k_r on the vertex set $[r] \times [k]$ with $K^k_r \subseteq B^k_r \subseteq R^k_r$, with $\delta(R^k_r) \geqslant \big(\frac{k-1}{k} + \frac{\gamma}{2}\big)\,kr$, and such that the following is true.*

- $(G1)$ $\frac{n}{4kr} \leqslant |V_{i,j}| \leqslant \frac{4n}{kr}$ *for every $i \in [r]$ and $j \in [k]$,*
- $(G2)$ *$\mathcal{V}$ is $(\varepsilon,d,p)_G$-lower-regular on R^k_r and $(\varepsilon,d,p)_G$-super-regular on K^k_r,*
- $(G3)$ *both $\big(N_\Gamma(v,V_{i,j}),V_{i',j'}\big)$ and $\big(N_\Gamma(v',V_{i,j}),N_\Gamma(v,V_{i',j'})\big)$ are $(\varepsilon,d,p)_G$-lower-regular pairs for every $\{(i,j),(i',j')\} \in E(R^k_r)$ and $v \in V \setminus V_0$,*
- $(G4)$ $|N_\Gamma(v,V_{i,j})| = (1 \pm \varepsilon)p|V_{i,j}|$ *for every $i \in [r]$, $j \in [k]$ and every $v \in V \setminus V_0$.*

$\square$

Following the proof strategy in [2], the next step is to find a partition of H which more or less matches that of G. In other words, we colour $V(H)$ with the colours

(i, j) which are the indices of the partition $\mathcal{V}$, such that about $|V_{i,j}|$ vertices get colour (i, j) and all edges of H are given colours corresponding to edges of R_r^k.

Lemma 45 (Lemma for H, [2, Lemma 25])**.** *Given $D, k, r \geqslant 1$ and $\xi, \beta > 0$ the following holds if $\xi \leqslant 1/(kr)$ and $\beta \leqslant 10^{-10}\xi^2/(Dk^4r)$. Let H be a D-degenerate graph on n vertices, let $\mathcal{L}$ be a labelling of its vertex set of bandwidth at most βn and let $\sigma : V(H) \to \{0, \dots k\}$ be a proper $(k+1)$-colouring that is $(10/\xi, \beta)$-zero-free with respect to $\mathcal{L}$, where the colour zero does not appear in the first $\sqrt{\beta}n$ vertices of $\mathcal{L}$. Furthermore, let R_r^k be a graph on vertex set $[r] \times [k]$ with $K_r^k \subseteq B_r^k \subseteq R_r^k$ such that for every $i \in [r]$ there exists a vertex $z_i \in \big([r] \setminus \{i\}\big) \times [k]$ with $\big\{z_i, (i, j)\big\} \in E(R_r^k)$ for every $j \in [k]$. Then, given a k-equitable integer partition $\{m_{i,j}\}_{i\in[r],j\in[k]}$ of n with $n/(10kr) \leqslant m_{i,j} \leqslant 10n/(kr)$ for every $i \in [r]$ and $j \in [k]$, there exists a mapping $f : V(H) \to [r] \times [k]$ and a set of special vertices $X \subseteq V(H)$ such that we have for every $i \in [r]$ and $j \in [k]$*

$(H1)$ $m_{i,j} - \xi n \leqslant |f^{-1}(i, j)| \leqslant m_{i,j} + \xi n$,

$(H2)$ $|X| \leqslant \xi n$,

$(H3)$ $\{f(x), f(y)\} \in E(R_r^k)$ for every $\{x, y\} \in E(H)$,

$(H4)$ $y, z \in \cup_{j' \in [k]} f^{-1}(i, j')$ for every $x \in f^{-1}(i, j) \setminus X$ and $xy, yz \in E(H)$, and

$(H5)$ $f(x) = \big(1, \sigma(x)\big)$ for every x in the first $\sqrt{\beta}n$ vertices of $\mathcal{L}$. □

During the pre-embedding, we embed a vertex x of H onto a vertex v of V_0, and we also embed all vertices at distance at most s from x. This creates restrictions on the vertices of G to which we can embed the vertices at distance $s + 1$, and for the application of the sparse blowup lemma (Lemma 34) we need certain conditions to be satisfied. The next lemma states that we can find vertices in G, to which we can embed the vertices at distance at most s from x in H, satisfying these conditions. This is the main difference in the proof in comparison to [2] and the place where we need that the neighbourhood of every vertex in G has a certain density of K_s's.

Lemma 46 (Partial embedding lemma)**.** *For $\Delta, k \geqslant 2$, $2 \leqslant s \leqslant k - 1$, and $\gamma, d > 0$ with $d \leqslant \frac{\gamma}{32}$ there exists $\zeta > 0$ such that for every $\varepsilon' > 0$ there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$, all $\mu > 0$ and $r \geqslant 1$, there exists a constant*

$C^* > 0$ such that the random graph $\Gamma = G(n,p)$ a.a.s. has the following property if $p \geqslant C^*\big(\frac{\log n}{n}\big)^{1/\Delta}$.

Suppose that G' is a subgraph of Γ with $|V(G')| = (1\pm\varepsilon)\mu n$, with $\delta(G') \geqslant \big(\frac{k-1}{k} + \gamma\big)p|V(G')|$, and such that for a vertex $v \in V(G')$ there are at least $\gamma p^{\binom{s+1}{2}}(\mu n)^s$ copies of K_s in $N_{G'}(v)$ and $|N_{G'}(T)| \leqslant 2\mu n p^t$ for any set $T \subset V(G')$ of size $t \leqslant \Delta$.

Suppose that G is a spanning subgraph of Γ with $\delta(G) \geqslant \big(\frac{k-1}{k} + \gamma\big)pn$, and we have an (ε,p)-lower-regular partition $V(G) = V_0 \mathbin{\dot\cup} V_1 \mathbin{\dot\cup} \ldots \mathbin{\dot\cup} V_r$ with (ε,d,p)-reduced graph R, such that $|V_i \cap V(G')| = (1\pm\varepsilon)\mu|V_i|$ for each i, and such that $V_0 \cap V(G'), \ldots, V_r \cap V(G')$ is also an (ε,p)-lower-regular partition of G' with (ε,d,p)-reduced graph R. Suppose that $\frac{n}{4r} \leqslant |V_i| \leqslant \frac{4n}{r}$ for all $i \in [r]$.

Suppose that H' is a graph with $\Delta(H') \leqslant \Delta$, with a root vertex x, and no vertex at distance greater than $s+1$ from x. Let ϱ be a proper k-colouring of $V(H')$ in which $N(x)$ receives at most s colours, and let T be the set of vertices in H' at distance exactly $s+1$ from x.

Then there exists $\varphi : V(H') \setminus T \to V(G')$ which is a partial embedding of H' into G', and a subset $\{V'_1, \ldots, V'_k\} \subset \{V_1, \ldots, V_r\}$ with the following properties (where we let $\Pi(u) = \varphi\big(N_{H'}(u) \cap \mathrm{Dom}(\varphi)\big)$ for each $u \in T$).

- $(P1)$ *$\varphi(x) = v$,*
- $(P2)$ *$V'_1, \ldots, V'_k$ form a clique in R,*
- $(P3)$ *for all $u \in T$ we have $\big|N_\Gamma\big(\Pi(u)\big) \cap V'_{\varrho(u)}\big| = (1\pm\varepsilon')p^{|\Pi(u)|}|V'_{\varrho(u)}|$,*
- $(P4)$ *for all $u \in T$ we have $\big|N_G\big(\Pi(u)\big) \cap V'_{\varrho(u)}\big| \geqslant \zeta p^{|\Pi(u)|}|V'_{\varrho(u)}|$,*
- $(P5)$ *for all $u \in T$ and $j \in [k]$ with $j \neq \varrho(u)$ the pair $\big(N_\Gamma(\Pi(u), V'_{\varrho(u)}), V'_j\big)$ is $(\varepsilon', d, p)_G$-lower-regular, and*
- $(P6)$ *for all $uu' \in H'$ with $u, u' \in T$ the pair $\big(N_\Gamma(\Pi(u), V'_{\varrho(u)}), N_\Gamma(\Pi(u'), V'_{\varrho(u')})\big)$ is $(\varepsilon', d, p)_G$-lower-regular.*

Returning to the proof strategy of [2], the sizes of the clusters $V_{i,j}$ from Lemma 44 do not quite match the sizes of the sets $X_{i,j}$ from Lemma 45. Also, Lemma 46 embeds some vertices, creating a little further imbalance, and we need to slightly alter the mapping f from Lemma 45 to accommodate these pre-embedded vertices. The next lemma allows us to change the sizes of the clusters $V_{i,j}$ slightly to match the partition of H, without destroying the properties of the partition of G and of the pre-embedded vertices we worked to achieve.

Lemma 47 (Balancing lemma, [2, Lemma 27])**.** *For all integers $k \geqslant 1$, $r_1, \Delta \geqslant 1$, and reals $\gamma, d > 0$ and $0 < \varepsilon < \min\{d, 1/(2k)\}$ there exist $\xi > 0$ and $C^* > 0$ such that the following is true for every $p \geqslant C^* (\log n/n)^{1/2}$ and every $10\gamma^{-1} \leqslant r \leqslant r_1$ provided that n is large enough. Let Γ be a graph on the vertex set $[n]$ and let $G = (V, E) \subseteq \Gamma$ be a (not necessarily spanning) subgraph with vertex partition $\mathcal{V} = \{V_{i,j}\}_{i\in[r],j\in[k]}$ that satisfies $n/(8kr) \leqslant |V_{i,j}| \leqslant 4n/(kr)$ for each $i \in [r]$, $j \in [k]$. Let $\{n_{i,j}\}_{i\in[r],j\in[k]}$ be an integer partition of $\sum_{i\in[r],j\in[k]} |V_{i,j}|$. Let R^k_r be a graph on the vertex set $[r] \times [k]$ with minimum degree $\delta(R^k_r) \geqslant \big((k-1)/k + \gamma/2\big)kr$ such that $K^k_r \subseteq B^k_r \subseteq R^k_r$. Suppose that the partition $\mathcal{V}$ satisfies the following properties for each $i \in [r]$, each $j \neq j' \in [k]$, and each $v \in V$.*

($B1$) *We have $n_{i,j} - \xi n \leqslant |V_{i,j}| \leqslant n_{i,j} + \xi n$,*

($B2$) *$\mathcal{V}$ is $\big(\frac{\varepsilon}{4}, d, p\big)_G$-lower-regular on R^k_r and $\big(\frac{\varepsilon}{4}, d, p\big)_G$-super-regular on K^k_r,*

($B3$) *both $\big(N_\Gamma(v, V_{i,j}), V_{i,j'}\big)$ and $\big(N_\Gamma(v, V_{i,j}), N_\Gamma(v, V_{i,j'})\big)$ are $\big(\frac{\varepsilon}{4}, d, p\big)_G$-lower-regular pairs, and*

($B4$) *we have $|N_\Gamma(v, V_{i,j})| = \big(1 \pm \frac{\varepsilon}{4}\big)p|V_{i,j}|$.*

Then, there exists a partition $\mathcal{V}' = \{V'_{i,j}\}_{i\in[r],j\in[k]}$ of V such that the following properties hold for each $i \in [r]$, each $j \neq j' \in [k]$, and each $v \in V$.

($B1$') *We have $|V'_{i,j}| = n_{i,j}$,*

($B2$') *We have $|V_{i,j} \triangle V'_{i,j}| \leqslant 10^{-10}\varepsilon^4 k^{-2} r_1^{-2} n$,*

($B3$') *$\mathcal{V}'$ is $(\varepsilon, d, p)_G$-lower-regular on R^k_r and $(\varepsilon, d, p)_G$-super-regular on K^k_r,*

($B4$') *both $\big(N_\Gamma(v, V'_{i,j}), V'_{i,j'}\big)$ and $\big(N_\Gamma(v, V'_{i,j}), N_\Gamma(v, V'_{i,j'})\big)$ are $(\varepsilon, d, p)_G$-lower-regular pairs, and*

($B5$') *For each $1 \leqslant s \leqslant \Delta$ and vertices $v_1, \ldots, v_s \in [n]$ we have*

$$\big|N_\Gamma(v_1, \ldots, v_s; V_{i,j}) \triangle N_\Gamma(v_1, \ldots, v_s; V'_{i,j})\big| \leqslant 10^{-10}\varepsilon^4 k^{-2} r_1^{-2} \deg_\Gamma(v_1, \ldots, v_s) + C^* \log n\,.$$

Furthermore, if for any two disjoint vertex sets $A, A' \subset V(\Gamma)$ with $|A|, |A'| \geqslant \frac{1}{50000kr_1}\varepsilon^2 \xi pn$ we have $e_\Gamma(A, A') \leqslant \big(1 + \frac{1}{100}\varepsilon^2\xi\big)p|A||A'|$, and if 'lower-regular' is replaced with 'regular' in ($B2$), and ($B3$), then we can replace 'lower-regular' with 'regular' in ($B3$') and ($B4$'). □

After applying Lemma 47 it remains only to check that the conditions of Lemma 34 are met to complete the embedding of H.

3.3 Proof of the partial embedding lemma

To prove Lemma 46, we proceed as follows. After choosing the constants required for the lemma and its proof, we select a bounded number of clusters of the regular partition of G so that we can refine the partition on $W = N_{G'}(v)$ and those chosen clusters with Lemma 31. The chosen clusters should behave like the whole partition in some ways which we can get by picking them at random. The lower bound on the number of K_s in W will then allow us to find a clique of size s in the refined reduced graph on W. This would already allow us to embed the s-coloured $N_{H'}(x)$ to W. However, it would yield image restrictions that are too small for the remaining vertices of H' and so we need to embed all vertices up to distance s. The minimum degree of G' allows us to find clusters extending the K_s to a K_{k+1} in the refined reduced graph and iteratively we find cliques in the refined reduced graph, swapping out the clusters in W one by one with subclusters of the regular partition of G'. We can then embed the vertices of H' up to distance s iteratively, following a certain order defined by the colouring, maintaining properties that allow us to keep the embedding going and ensuring that what we end up with valid image restricting sets for the vertices at distance $s + 1$ from x.

Proof of Lemma 46. First we fix all constants that we need throughout the proof. As before, "$x \ll$" ("$x \ll$") denotes that x is chosen sufficiently small (big) with respect to all constants to its right. Let $\Delta, k \geqslant 2$ and $\gamma, d > 0$ be given.

Let $d' = \min(d, 10^{-k}\gamma)$ and choose a small $\xi \ll \gamma, \frac{1}{k}$ and an integer $\ell \gg \frac{1}{\xi}, \frac{1}{\gamma}, k$. Let $\nu^{**}_{\Delta} = \frac{1}{100\Delta^{2k}}$ and for every $i \in (\Delta - 1, \ldots, 1, 0)$, let $\nu^{**}_i \leqslant \nu^{**}_{i+1}$ be returned by Lemma 35 with input $\varepsilon_{\text{OSRIL}} = \nu^{**}_{i+1}$ and $\alpha_{\text{OSRIL}} = d'$. Next, let $\nu^*_{\Delta-1,\Delta-1} = \nu^*_{i,\Delta} = \nu^*_{\Delta,i} = 1$ for $i \in [\Delta]$. For each $(i, j) \in \{0, \ldots, \Delta - 1\}^2 \setminus \{(\Delta - 1, \Delta - 1)\}$ in reverse lexicographic order, we choose $\nu^*_{i,j} \leqslant \nu^*_{i+1,j}, \nu^*_{i,j+1}, \nu^*_{i+1,j+1}$ not larger than the ε_0 returned by Lemma 35 for both input $\nu^*_{i+1,j}$ and d', and for input $\nu^*_{i,j+1}$ and d', and not larger than the ε_0 returned by Lemma 36 for input $\nu^*_{i+1,j+1}$ and d'. Choose $\nu_0 \ll \nu^{**}_0, \nu^*_{0,0}, \gamma, d', \frac{1}{k}$. Now, Lemma 31 with input ν_0^2/ℓ^2 and ℓ returns t_1.

Set $\zeta = \left(\frac{d'}{4}\right)^{\Delta}/2t_1$. Given ε', let $\varepsilon^{**}_{\Delta} = \varepsilon'$ and for every $i \in (\Delta-1, \ldots, 1, 0)$, let $\varepsilon^{**}_i \leqslant \varepsilon^{**}_{i+1}$ be returned by Lemma 35 with input $\varepsilon_{\text{OSRIL}} = \varepsilon^{**}_{i+1}$ and $\alpha_{\text{OSRIL}} = d$. Next, let $\varepsilon^{*}_{\Delta-1,\Delta-1} = \varepsilon'$ and $\varepsilon^{*}_{i,\Delta} = \varepsilon^{*}_{\Delta,i} = 1$ for $i \in [\Delta]$. For each $(i,j) \in \{0,\ldots,\Delta-1\}^2 \setminus \{(\Delta-1,\Delta-1)\}$ in reverse lexicographic order, we choose $\varepsilon^{*}_{i,j} \leqslant \varepsilon^{*}_{i+1,j}, \varepsilon^{*}_{i,j+1}, \varepsilon^{*}_{i+1,j+1}$ not larger than the ε_0 returned by Lemma 35 for both input $\varepsilon^{*}_{i+1,j}$ and d, and for input $\varepsilon^{*}_{i,j+1}$ and d, and not larger than the ε_0 returned by Lemma 36 for input $\varepsilon^{*}_{i+1,j+1}$ and d.

We choose $\varepsilon_0 \leqslant \varepsilon^{**}_0, \varepsilon^{*}_{0,0}, \frac{\nu_0}{t_1}$ small enough such that $(1+\varepsilon_0)^{\Delta} \leqslant 1+\varepsilon'$ and $(1-\varepsilon_0)^{\Delta} \geqslant 1-\varepsilon'$. Given $r \geqslant 1$, ε with $0 < \varepsilon \leqslant \varepsilon_0$, and $\mu > 0$, let C be a large enough constant for all of the above calls to Lemmas 35 and 36, and for Proposition 38 with input ν_0 and ε_0. Finally, we choose

$$C^* \gg k, t_1, r, \frac{1}{\mu}, \frac{1}{\varepsilon}, \Delta, C.$$

Let $\Gamma = G(n,p)$ with $p \geqslant C^*(\log n/n)^{1/\Delta}$. Then Γ satisfies a.a.s. the properties stated in Lemma 35, Lemma 36, Proposition 38 and Lemma 31 with the parameters specified above. We assume from now on that Γ satisfies these good events and has these properties. Let G', $v \in V(G')$, G, $\{V_i\}_{i\in\{0,\ldots,r\}}$, H', and $x \in V(H')$ be as in the statement of the lemma.

To be able to apply Lemma 31 we need to choose a suitable subset of the clusters $\{V_i\}$ of bounded size. As the clusters $\{V_i\}$ might be of different sizes and we will want to have a minimum degree condition on the reduced graph, we will consider a weighted version of this degree that takes the cluster sizes into account.

Claim 48. *There exists $L \subset [r]$ of size ℓ satisfying the following. The (ε_0, d, p)-regular graph R^* on to the sets $\{V_i\}$ indexed by L, satisfies the following weighted minimum degree condition.*

$$\forall i \in L : \sum_{j \in N_R(i) \cap L} \frac{|V_j|}{|V^*|} \geqslant \left(\frac{k-1}{k} + \frac{\gamma}{5}\right),$$

where $V^* = \bigcup_{i\in L} V_i$. Additionally, we have that

$$W := \left\{ w \in N_{G'}(v) : |N_{G'}(w) \cap V^*| \geqslant \left(\frac{k-1}{k} + \frac{\gamma}{5}\right) p |V^* \cap V(G')| \right\}$$

has size at least $(1-\xi)|N_{G'}(v)|$ and there are at least $\frac{1}{2}\gamma p^{\binom{s+1}{2}}(\mu n)^s$ copies of K_s in W.

Proof. We choose a subset $L \subset [r]$ of size ℓ at random. First, we will transfer the minimum degree of G to the reduced graph and show that with high probability the minimum degree is preserved on the chosen clusters. Recall that G satisfies a minimum degree of $\delta(G) \geqslant (\frac{k-1}{k} + \gamma)pn$ and that we have the following bounds on the sizes of the clusters.

$$\frac{4n}{r} \geqslant |V_i| \geqslant \frac{n}{4r} \geqslant Cp^{-1} \log n \tag{3.1}$$

Without loss of generality, we may assume that no V_i forms an irregular pair with more than $\sqrt{\varepsilon}$ of the clusters, otherwise, add it to V_0, which over all clusters increases the size of V_0 by at most $4\sqrt{\varepsilon}n$. Fix $i \in [r]$. Proposition 38 applied to the edges between V_i and V_0 implies that

$$e(V_i, V_0) \leqslant 2p(\varepsilon + 4\sqrt{\varepsilon})n|V_i| \quad \text{and} \quad e(V_i) \leqslant 2p|V_i|^2 \leqslant 2p\frac{16}{r}n|V_i|$$

Also, we can bound the number of edges from V_i to other clusters that are in pairs which are not dense or (ε, p)-lower-regular as follows.

$$e(V_i, \bigcup_{j \in R \setminus N_R(i)} V_j) \leqslant dpn|V_i| + 2p \cdot 4\sqrt{\varepsilon}n|V_i|.$$

Putting the above together, we obtain that

$$e(V_i, \bigcup_{j \in N_R(i)} V_j) \geqslant \left(\frac{k-1}{k} + \gamma - 2\varepsilon - 16\sqrt{\varepsilon} - d - \frac{32}{r}\right) pn|V_i|$$

As, again by Proposition 38, the number of edges between any V_i and V_j is at most $(1+\varepsilon_0)|V_i||V_j|$, we get that

$$\begin{aligned}\sum_{j \in N_R(i)} \frac{|V_j|r}{|V(G)|} &\geqslant \left(\frac{k-1}{k} + \gamma - 2\varepsilon - 16\sqrt{\varepsilon} - d - \frac{32}{r}\right)(1+\varepsilon_0)^{-1} r \\ &\geqslant \left(\frac{k-1}{k} + \frac{\gamma}{2}\right) r.\end{aligned}$$

By the size conditions on the clusters the relative sizes $w_j := \frac{|V_j|r}{|V(G)|}$ take values in $(\frac{1}{4}, 4)$. We now consider

$$w'_j = \xi \left\lfloor w_j/\xi \right\rfloor,$$

the discretisation of w_j into steps of size ξ. Of these discretised weights, we will ignore those that occur fewer than $\xi^2 r$ times. We lose at most a factor of 4ξ due to the discretisation as all weights are at least $\frac{1}{4}$. Also weights in $(\frac{1}{4}, 4)$ occuring fewer than $\xi^2 r$ times contribute at most $16\xi r$ to the sum, so we get the following lower bound.

$$\sum_{j\in N_R(i)} w'_j \geqslant (1-4\xi)\left(\frac{k-1}{k}+\frac{\gamma}{2}\right)r - 16\xi r \geqslant \left(\frac{k-1}{k}+\frac{\gamma}{3}\right)r.$$

We can now apply the hypergeometric inequality (Theorem 40) to all possible rounded weight values separately. For any $j \in [r]$ the probability that j is in L is ℓ/r and so for a given density in $(\frac{1}{4}, 4)$, which occurs, say, ϑr times, the probability that this density is chosen fewer than $(1-\xi)\vartheta\ell$ times is at most $2e^{-\xi^2\cdot\xi\vartheta\ell/3} \leqslant 2e^{-\xi^5\ell/3}$. This implies by the union bound that with probability at most $4\xi^{-1}2e^{-\xi^5\ell/3}$ we do not have have

$$\sum_{j\in N_R(i)\cap L} w_j \geqslant (1-\xi)\left(\frac{k-1}{k}+\frac{\gamma}{3}\right)\frac{\ell}{r}r \geqslant \left(\frac{k-1}{k}+\frac{\gamma}{4}\right)\ell. \tag{3.2}$$

So by the union bound the expected number of vertices in R^* that do not satisfy (3.2) is at most $\ell 8\xi^{-1}e^{-\xi^5\ell/3} < 1/10$. By Markov's inequality, the probability that there is any such vertex in R^* is thus at most $1/10$. By the same discretisation of w_j and application of the hypergeometric inequality to the discretised weights, we can also deduce that

$$|V^*| = \frac{|V(G)|}{r}\sum_{i\in L} w_i = (1\pm 100\xi)\frac{\ell}{r}\sum_{i\in[r]} w_i = (1\pm 100\xi)(1\pm\varepsilon)\frac{\ell|V(G)|}{r} \tag{3.3}$$

with probability at least $9/10$. Putting (3.2) and (3.3) together implies that with probability at least $8/10$ the first claimed statement holds.

For the claim, we also require that the minimum degree condition of the vertices in $N_{G'}(v)$ carries over to the chosen clusters for most vertices. Fix w in $N_{G'}$. For $j \in [r]$ we consider the following weighted p-density, which may take values in $(0, 5)$.

$$d_{w,j} = d_{G,p}(\{w\}, V_j \cap V(G'))\frac{|V_j \cap V(G')|r}{|V(G')|}.$$

Accounting for the exceptional set V_0 with Proposition 38, the minimum degree condition on G' of $(\frac{k-1}{k}+\gamma)p|V(G')|$ implies that these weighted p-densities satisfy

$$\sum_{j\in[r]} d_{w,j} \geqslant \left(\frac{k-1}{k}+\gamma-2(\varepsilon+4\sqrt{\varepsilon})\right)r \geqslant \left(\frac{k-1}{k}+\frac{\gamma}{2}\right)r.$$

Similarly to before, we consider $d'_{w,i} = \xi\lfloor d_{w,i}/\xi\rfloor$, the discretisation of $d_{w,i}$ into steps of size ξ. Of these discretised weighted densities, we ignore those that occur fewer than $\xi^2 r$ times and those that are smaller than $\sqrt{\xi}$. The small densities contribute at most $\sqrt{\xi}r$ to the sum and we lose a factor of at most $\sqrt{\xi}$ due to the discretisation for larger values. Also weights in $(\sqrt{\xi}, 5)$ occuring fewer than $\xi^2 r$ times contribute at most $25\xi r$ to the sum, so we get the following lower bound.

$$\sum_{i \in [r]} d'_{w,i} \geqslant (1 - \sqrt{\xi}) \left(\frac{k-1}{k} + \frac{\gamma}{2} - \sqrt{\xi} - 25\xi \right) r \geqslant \left(\frac{k-1}{k} + \frac{\gamma}{3} \right) r.$$

Applying the hypergeometric inequality to all density values separately as before, we get that for any $w \in N_{G'}(v)$ with probability at most $5\xi^{-1} 2e^{-\xi^5 \ell/3} \geqslant \xi/10$ we do not have

$$\sum_{i \in L} d'_{w,i} \geqslant (1 - \xi) \left(\frac{k-1}{k} + \frac{\gamma}{3} \right) \frac{\ell}{r} r \geqslant \left(\frac{k-1}{k} + \frac{\gamma}{4} \right) \frac{\ell}{r} r. \tag{3.4}$$

So the expected number of vertices in $N_{G'}(v)$ not satisfying (3.4) is at most $\xi |N_{G'}(v)|/10$. By Markov's inequality, with probability at least $9/10$ at most a fraction ξ of vertices in $N_{G'}(v)$ violate (3.4). In particular all vertices satisfying (3.4) have at least

$$(1 - 100\xi)(1 - \varepsilon) \left(\frac{k-1}{k} + \frac{\gamma}{4} \right) (1 - \varepsilon)\mu p |V^*| \geqslant \left(\frac{k-1}{k} + \frac{\gamma}{5} \right) p |V^* \cap V(G')|$$

neighbours in $V^* \cap V(G')$ if (3.3) holds. So indeed with probability at least $7/10$ the first two claimed statements hold, so assume we chose L such that they do.

For the claim it only remains to show the lower bound on the number of cliques in W. It follows, by inductively building up cliques, from the assumption in the lemma that any $t \leqslant \Delta$ vertices of G' have at most $2p^t \mu n$ common neighbours in G', that v and each $w \in N_{G'}(v)$ are contained in at most

$$\prod_{t=2}^{s} 2p^t \mu n = p^{\binom{s+1}{2} - 1} (2\mu n)^{s-1}$$

copies of K_{s+1}. The choice of L implies $|N_{G'}(v) \setminus W| \leqslant \xi 2\mu n p$ and so there are at least

$$\gamma p^{\binom{s+1}{2}} (\mu n)^s - \xi 2\mu n p p^{\binom{s+1}{2} - 1} (2\mu n)^{s-1} \geqslant \tfrac{1}{2} \gamma p^{\binom{s+1}{2}} (\mu n)^s$$

copies of K_s in W. □

Let $\{W_i\}_{i \in [\ell]}$ be an arbitrary equipartition of W into ℓ parts. We apply Lemma 31 to G' and $\{(V_i \cap V(G')) \setminus W\}_{i \in L} \cup \{W_i\}_{i \in [\ell]}$ with input parameter ν_0^2/s^2. This returns a partition refining each of these sets into $1 \leqslant t \leqslant t_1$ clusters together with small exceptional sets $\{W_{i,0}, V_{i,0}\}$. It follows directly from the definition of a regular refinenment that at most $\nu_0 t = \sqrt{\nu_0^2/\ell^2}\ell t$ of the clusters do not form a regular pair with more than $\nu_0 t$ of the clusters. Include the vertices of all those clusters in the exceptional sets, which now make up a fraction of at most $2\nu_0$ of the vertices.

We now want to obtain s clusters $W'_1, \dots, W'_s$ in $\{W'_{i,j}\}_{i\in[\ell],j\in[t]}$ that are pairwise (ν_0, d', p)-regular. Assume for a contradiction that no such clique exists in the reduced graph. So each clique in W must either contain an edge meeting an exceptional set, one which does not lie in a (ν_0, d', p)-regular pair or one that is contained completely in a $W'_{i,j}$. Note that we have for all $i \in [\ell]$ and $j \in [t]$ that

$$|W_{i,j}| \geqslant \frac{1}{2\ell t_1}|W| \geqslant \frac{\mu n p}{4\ell t_1} \geqslant Cp^{-1}\log n.$$

So we may apply Proposition 38 to bound the number of edges within and between clusters. Using the upper bound on common neighbourhoods in G' given in the lemma to bound the number of edges meeting the exceptional sets, we obtain that deleting at most

$$2\nu_0|W|2p^2\mu n + 2p(\nu_0 + d')|W|^2 + \ell 2p(|W|/\ell)^2 \leqslant (8\nu_0 + 8\nu_0 + 8d' + 2/\ell)p^3\mu^2 n^2$$

edges would remove all cliques from W. By the upper bound on common neighbourhoods in G' given in the lemma any of these edges is contained in at most

$$\prod_{t=3}^{s} 2p^t\mu n = p^{\binom{s+1}{2}-3}(2\mu n)^{s-2}$$

copies of K_{s+1} together with v. So there would be at most

$$(16\nu_0 + 8d' + 2/\ell)p^3\mu^2 n^2 p^{\binom{s+1}{2}-3}(2\mu n)^{s-2} < \tfrac{1}{2}\gamma p^{\binom{s+1}{2}}(\mu n)^s$$

copies of K_s in W, a contradiction. So let $W'_1, \dots, W'_s$ in $\{W'_{i,j}\}_{i\in[\ell],j\in[t]}$ be pairwise (ν_0, d', p)-regular.

Just like in the proof of Claim 48, the minimum degree condition on the vertices in W implies that each W'_i satisfies

$$\sum_{V_{i',j'}\colon (W'_i, V_{i',j'}) \text{ is } (\nu_0,d',p)-\text{regular}} \frac{|V_{i',j'}|}{|V^* \cap V(G')|} \geqslant \left(\frac{k-1}{k} + \frac{\gamma}{8}\right). \tag{3.5}$$

Also if $(V_i, V_{i'})$ is (ε_0, d, p)-regular then so is $(V_i \cap V(G'), V_{i'} \cap V(G'))$ as the reduced graphs are assumed to be identical in the lemma and by the choice of ε_0 and

$d' \leqslant d$ it follows that $(V_{i,j}, V_{i',j'})$ is (ν_0, d', p)-regular for all $j, j' \in [t]$. So a weighted minimum degree of $\left(\frac{k-1}{k} + \frac{\gamma}{8}\right)$ on the (ν_0, d', p)-reduced graph on the clusters $\{V_{i,j}\}$ is inherited from R^*, i.e., (3.5) holds for $V_{i,j}$ too.

Now we can choose the clusters into which we will embed the vertices of H'. Note that (3.5) allows one to find, for every clique of size at most k, a $(k+1)$-clique containing it. So we can choose

$$(i_{s+1}, j_{s+1}), \ldots, (i_{k+1}, j_{k+1}) \in L \times [t]$$

such that

$$W'_1, \ldots, W'_s, V_{i_{s+1}, j_{s+1}}, \ldots, V_{i_{k+1}, j_{k+1}}$$

are a clique in the (ν_0, d', p)-reduced graph and $i_{s+1}, \ldots, i_{k+1}$ are all distinct. Next, we choose pairs $(i_s, j_s), \ldots, (i_1, j_1)$ in that order sequentially such that for each $a \in \{s, \ldots, 1\}$ the clusters

$$W'_1, \ldots, W'_{a-1}, V_{i_a, j_a}, V_{i_{a+1}, j_{a+1}}, \ldots, V_{i_{k+1}, j_{k+1}}$$

form a K_{k+1} in the (ν_0, d', p)-reduced graph. Since the degree condition was inherited from R^* we may assume that $V_{i_1}, \ldots, V_{i_{k+1}}$ are pairwise (ε_0, d, p)-regular too.

Let H', ϱ be as in the statement of the lemma. We define a proper $(k+1)$-vertex colouring $\varrho' : V(H') \to [k+1]$ inductively as follows. Initially we set $\varrho'(w) = \varrho(w)$ for all w in H'. Let

$$U_{\varrho'} = \bigcup_{i=2}^{s-1} \left\{ w \in N^i(x) : \varrho'(w) \leqslant s - i + 1 \right\},$$

where $N^i(x)$ refers to the vertices at distance i from x. If $U_{\varrho'}$ contains a vertex w with no neighbour in ${\varrho'}^{-1}(i)$ for some $i \in \{s+1, \ldots, k+1\}$, we set $\varrho'(w) = i$. We repeat this step until $U_{\varrho'}$ contains no such vertices. With this recolouring procedure we ensure that every vertex in H' at distance $i \geqslant 2$ from x with colour at most $s - i + 1$ has at least two neighbours in the colour classes $s+1, \ldots, k+1$. Note that the colouring remains unchanged on $N(x)$ and the vertices at distance $s+1$ from x.

We define an order $<_{\varrho'}$ on H' given by the following enumeration of its vertices. First, we take an arbitrary enumeration of the vertices in $N^s(x) \cap {\varrho'}^{-1}(1)$, then for $i \in [s-1]$, we continue with the vertices in $N^{s-i}(x) \cap {\varrho'}^{-1}([i+1])$. The rest of the vertices in H' we then enumerate arbitrarily. With the colouring ϱ' defined as above, this gives us, for all u at distance at least two from x with $\varrho'(u) + d(x, u) \leqslant s + 1$:

$$|\mathrm{pred}_{<_{\varrho'}}(u) \cap N(u)| = |\{u' : u' <_{\varrho'} u, u' \in N(u)\}| \leqslant \Delta - 2. \tag{3.6}$$

Now we can assign the vertices of H' to clusters. For $u \in V(H')$, let

$$V_u = V_{i_{\varrho'(u)}} \quad \text{and} \quad C_u = \begin{cases} W'_{\varrho'(u)} & \text{if } \varrho'(u) + d(x,u) \leqslant s+1 \\ V_{i_{\varrho'(u)}, j_{\varrho'(u)}} & \text{otherwise.} \end{cases}$$

We now iteratively embed the vertices of H' in the order specified above respecting the assignments to clusters. More precisely, we claim the following. Here, as in the statement of the lemma, we set $\Pi(u) = \varphi\big(N_{H'}(u) \cap \mathrm{Dom}(\varphi)\big)$ and let T be the vertices in H' at distance exactly $s+1$ from v.

Claim 49. *For each integer* $0 \leqslant z \leqslant |H' \smallsetminus (T \cup \{x\})|$ *there exists an embedding* φ *of the first* z *vertices of* $H' \smallsetminus (T \cup \{x\})$ *(w.r.t. to the order* $<_{\varrho'}$*) into* G *such that the following holds. For every* $u, u' \in H' \smallsetminus (\mathrm{Dom}(\varphi) \cup \{x\})$*, where* $u' \in N_{H'}(u)$ *we have the following.*

($I1$) *For all* $u'' \in \mathrm{Dom}(\varphi)$ *we have* $\varphi(u'') \in C_{u''}$,

($I2$) $\big(N_\Gamma(\Pi(u), C_u), C_{u'}\big)$ *is* $(\nu^{**}_{|\Pi(u)|}, d', p)_G$*-lower-regular,*

($I3$) $|N_G(\Pi(u), C_u)| \geqslant \big(\frac{d'}{4}\big)^{|\Pi(u)|} p^{|\Pi(u)|} |C_u|$,

($I4$) $|N_\Gamma(\Pi(u), C_u)| = (1 \pm \nu_0)^{|\Pi(u)|} p^{|\Pi(u)|} |C_u|$,

($I5$) $\big(N_\Gamma(\Pi(u), C_u), N_\Gamma(\Pi(u'), C_{u'})\big)$ *is* $(\nu^{*}_{|\Pi(u)|, |\Pi(u')|}, d', p)_G$*-lower-regular.*

Also,

($L1$) $\big(N_\Gamma(\Pi(u), V_u), V_{u'}\big)$ *is* $(\varepsilon^{**}_{|\Pi(u)|}, d, p)_G$*-lower-regular,*

($L2$) $|N_\Gamma(\Pi(u), V_u)| = (1 \pm \varepsilon_0)^{|\Pi(u)|} p^{|\Pi(u)|} |V_u|$,

($L3$) $\big(N_\Gamma(\Pi(u), V_u), N_\Gamma(\Pi(u'), V_{u'})\big)$ *is* $(\varepsilon^{*}_{|\Pi(u)|, |\Pi(u')|}, d, p)_G$*-lower-regular.*

Proof. We prove the claim inductively and start with the empty embedding. If $\varphi = \varnothing$, then $\Pi(u) = \varnothing$ for all $u \in H'$, so ($I1$), ($I3$), ($I4$), and ($L2$) are trivial statements. By construction, for every edge uu' the clusters $(C_u, C_{u'})$ form an

$(\nu_0, d', p)_G$-regular pair so $(I2)$ and $(I5)$ hold. Similarly $(V_u, V_{u'})$ is $(\varepsilon_0, d, p)_G$-lower-regular, which implies $(L1)$ and $(L3)$.

Assume we have already embedded the first z vertices such that $(I1)$-$(I5)$ and $(L1)$-$(L3)$ hold and let w be the $(z+1)$th vertex. We will now prove there exists an embedding φ' extending φ with $\varphi'(w) \in C_w$ such that the statement of the claim holds for $z+1$. For this we will show that the number of choices $\varphi'(w)$ in C_w for which one of $(I2)$-$(I5)$ or $(L1)$-$(L3)$ does not hold is smaller than $|C_w|$.

By the construction we can use the following lower bounds on the sizes of W'_a, V'_{i_a,j_a}, and V_u.

$$|W'_a| \geqslant \frac{\mu np}{4\ell t_1} \geqslant \frac{\mu np}{5\ell r t_1}, \quad |V'_{i_a,j_a}| \geqslant \frac{\mu n}{5rt_1} \geqslant \frac{\mu n}{5\ell r t_1}, \quad \text{and} \quad |V_u| \geqslant \frac{\mu n}{5r}.$$

For $(I2)$, consider an edge uu' in $H' \smallsetminus (\mathrm{Dom}(\varphi) \cup \{w, x\})$. We only need to check $(I2)$ if $w \in N(u)$, as $\Pi(u)$ does not change otherwise. In particular, this implies that $|\Pi(u)| < \Delta - 1$ and if $C_u \in \{W'_i\}_{i\in[s]}$, then even $|\Pi(u)| < \Delta - 2$, by (3.6) if u is at distance two or more from x and otherwise by the fact that x is not in $\mathrm{Dom}(\varphi)$ for $u \in N(x)$. We want to apply Lemma 35 to $N_\Gamma(\Pi(u), C_u)$ and $C_{u'}$. By the inductive assumption $(I2)$ this pair is $(\nu^{**}_{|\Pi(u)|}, d', p)_G$-lower-regular and

$$\begin{aligned}|N_\Gamma(\Pi(u), C_u)| &\overset{(I4)}{\geqslant} (1-\nu_0)^{|\Pi(u)|} p^{|\Pi(u)|} |C_u| \geqslant (1-\nu_0)^{\Delta-2} p^{\Delta-2} \frac{\mu n}{5\ell r t_1} \\ &\geqslant C \max(p^{-2}, p^{-1}\log n).\end{aligned}$$

So we can apply Lemma 35, obtaining that for at most $Cp^{-1}\log n$ vertices v, the pair $(N_\Gamma(\Pi(u) \cap N(v)), C_{u'})$ is not $(\nu^{**}_{|\Pi(u)|+1}, d', p)_G$-lower-regular. Summing this over all possible $uu' \in H'$, at most $|H'|^2 Cp^{-1}\log n$ choices for $\varphi'(w)$ would violate $(I2)$. By the same argument with a similar calculation, at most $|H'|^2 Cp^{-1}\log n$ choices for $\varphi'(w)$ would violate $(L1)$.

For $(I3)$, consider $u \in H' \smallsetminus (\mathrm{Dom}(\varphi) \cup \{w, x\})$. As before, we only need to consider the case $w \in N(u)$ for $(I3)$. The inductive assumption $(I2)$ implies that $(N_\Gamma(\Pi(u), C_u), C_w)$ is $(\nu^{**}_{|\Pi(u)|}, d', p)$-lower-regular. Also,

$$\begin{aligned}|N_G(\Pi(u), C_u)| &\overset{(I3)}{\geqslant} \left(\tfrac{d'}{4}\right)^{|\Pi(u)|} p^{|\Pi(u)|} |C_u| \overset{(I4)}{\geqslant} \tfrac{1}{2}\left(\tfrac{d'}{4}\right)^{|\Pi(u)|} |N_\Gamma(\Pi(u), C_u)| \\ &\geqslant \nu^{**}_{|\Pi(u)|} |N_\Gamma(\Pi(u), C_u)|\end{aligned}$$

and so, using the lower-regularity of the pair, at most $\nu^{**}_{|\Pi(u)|}|C_w| \leqslant \nu^{**}_\Delta |C_w|$ choices for $\varphi'(w)$ violate the inequality in $(I3)$ for some u. So in total at most $|H'|\nu_\Delta^{**}|C_w|$ choices would violate $(I3)$.

For $(I4)$, consider $u \in H' \setminus (\mathrm{Dom}(\varphi) \cup \{w, x\})$. Once again, we only need to consider the case $w \in N(u)$ for $(I4)$, so $|\Pi(u)| < \Delta$ and if $C_u \in \{W_i'\}_{i\in[s]}$, then $|\Pi(u)| < \Delta - 1$. Similarly to above, we have that $|N_\Gamma(\Pi(u), C_u)| \geqslant Cp^{-1}\log n$. Therefore, by Proposition 38, there are at most $Cp^{-1}\log n$ vertices v such that

$$|N_\Gamma(\Pi(u) \cap N(v), C_u)| \neq (1 \pm \nu_0)p|N_\Gamma(\Pi(u), C_u)| \overset{(I4)}{=} (1 \pm \nu_0)^{|\Pi(u)|+|}p^{|\Pi(u)|+1}|C_u|$$

Summing this over all $u \in H'$, at most $|H'|Cp^{-1}\log n$ choices for $\varphi'(w)$ would violate $(I4)$. And again by a similar calculation at most $|H'|Cp^{-1}\log n$ choices for $\varphi'(w)$ would violate $(L2)$.

For $(I5)$, consider an edge uu' in $H' \setminus (\mathrm{Dom}(\varphi) \cup \{w, x\})$. Here, three cases are to be considered; $w \in N(u), w \in N(u')$, or both. First, assume that $w \in N(u)$ and $w \notin N(u')$. By the same arguments as before, we obtain that

$$|N_\Gamma(\Pi(u), C_u)| \geqslant C\max(p^{-2}, p^{-1}\log n) \quad \text{and} \quad |N_\Gamma(\Pi(u'), C_{u'})| \geqslant Cp^{-1}\log n.$$

So using the induction hypothesis and Lemma 35, for at most $Cp^{-1}\log n$ vertices, $(I5)$ would be violated. It follows from symmetry that only $Cp^{-1}\log n$ vertices violate $(I5)$ for u and u' if $w \notin N(u)$ and $w \in N(u')$. Now, assume that $w \in N(u)$ and $w \in N(u')$, which implies that $\Delta \geqslant 3$ as $\{u, u', w\}$ would yield an isolated triangle for $\Delta = 2$ which cannot exist in H', so

$$|N_\Gamma(\Pi(u), C_u)| \geqslant C\max(p^{-2}, p^{-1}\log n)$$
$$\text{and}$$
$$|N_\Gamma(\Pi(u'), C_{u'})| \geqslant C\max(p^{-2}, p^{-1}\log n).$$

We use the induction hypothesis of $(I5)$ and, hence, we can apply Lemma 36, obtaining that for at most $C\max(p^{-2}, p^{-1}\log n)$ choices, $(I5)$ would not hold for u and u'. In total, if $\Delta = 2$, then at most $2|H'|^2Cp^{-1}$ choices violate $(I5)$, and if $\Delta \geqslant 3$, at most $2|H'|^2C(p^{-1} + \max(p^{-2}, p^{-1}\log n))$ choices would. Again, the number of choices that violate $(L3)$ can be bounded by the same number.

To wrap up the proof of the claim, we will now sum the number of all bad choices described above. This yields a total of

$$2|H'|^2Cp^{-1}\log n + |H'|\nu_\Delta^{**}|C_w| + 2|H'|^2Cp^{-1}\log n + 4|H'|^2Cp^{-1}\log n.$$

if $\Delta = 2$ and otherwise a total of

$$2|H'|^2 Cp^{-1} \log n + |H'|\nu_\Delta^{**}|C_w| + 2|H'|^2 Cp^{-1} \log n \\ + 4|H'|^2 C(p^{-1} + \max(p^{-2}, p^{-1} \log n)).$$

Note that we can bound the size of H' by $\Delta^{s+1} + 1 \leqslant \Delta^{2k}$, so the second term is at most $\frac{1}{100}|C_w|$. The other terms can be bounded similarly, which implies that the total number of bad choices is smaller than $C_w/2$. Hence, there is a suitable choice of $\varphi(w)$ in C_w, concluding the proof of the claim. $\square$

Now we can conclude the proof of Lemma 46. We complete the embedding by setting $\varphi(x) = v$, which is valid since we embedded all neighbours of x to W, so clearly $(P1)$ holds. By setting $V'_j = V_{i_j}$ for $j \in [k]$, we get $(P2)$. For every vertex u in T we have that $C_u = V_{i_{\varrho'(u)}, j_{\varrho'(u)}}$ and $|C_u| \geqslant |V_{i_{\varrho'(u)}}|/2t_1$. So by the choice of ζ, $(P4)$ follows from $(I3)$. The choice of constants ensures that the remaining statements in the lemma are a direct consequence of $(L1)$-$(L3)$. $\square$

3.4 Proof of the main theorem

The proof of Theorem 43 is broadly similar to the proof of [2, Theorem 23]. Again, basically the idea is that we apply the lemmas of Section 3.2 in order to first find a well-behaved partition of G and a corresponding partition of H. We then deal with the few badly-behaved vertices of G by sequentially pre-embedding onto them some vertices of H whose neighbourhoods contain at most s colours. Lemma 46 deals with this pre-embedding, and sets up for the vertices which are not pre-embedded but which have pre-embedded neighbours restriction sets in the sense of Definition 33. We then adjust the partition of H to fit this pre-embedding, and balance the partition of G to match. Finally, we see that the conditions of Lemma 34 are met, and that lemma completes the desired embedding of H in G.

As in [2], there are two slightly subtle points. The first is that for $\Delta = 2$ we can have $Cp^{-2} > pn$, so that we should be worried that we come to some badly-behaved vertex of G onto which we wish to pre-embed and discover that all its neighbours have already been used in pre-embedding. As in [2], this is easy to handle: at each step we choose the badly-behaved vertex with most neighbours already embedded to. It is easy to check that this ordering avoids the above problem. The second, more serious, problem is that we need restriction sets fulfilling the conditions of Definition 33. Although Lemma 46 gives us pre-embeddings satisfying these conditions, we might destroy the conditions when we pre-embed later vertices. The condition we could destroy is simply that we need each restriction set to be reasonably large; the danger is that we pre-embed many vertices to some restriction set. The solution to this is (as in [2]) to select a set S, whose size is linear in n but small, using Lemma 41 to avoid large intersections with any possible restriction set. When we apply Lemma 46 to cover a badly-behaved vertex v, we will pre-embed to v and to some vertices chosen from S, and not to any other vertex. The badly-behaved vertices are not (by construction) in any restriction set, while S has small intersection with all restriction sets, so that even removing all of S would not make the restriction sets too small.

The only point in the proof where we really need to do more than in [2] (apart from using Lemma 46 to pre-embed) is that we need to ensure the conditions of Lemma 46 are met. When we wish to cover a badly-behaved v, its neighbourhood within the set S must contain many copies of K_s. Further, some vertices of S will have been used in earlier pre-embeddings, and we need to ensure that these used vertices do not hit too many of the copies of K_s. For this, we apply the sparse regularity lemma, Lemma 31, to $G\big[N_G(v)\big]$ before choosing S. We will see that (since $N_G(v)$ contains many copies of K_s) we find a set of s clusters in $N_G(v)$ such that all the pairs are relatively dense and regular. When we use Lemma 41 to choose S, we also insist that S contains a significant fraction of each of these clusters. The order in which we cover badly-behaved vertices ensures that a (slightly smaller but still) significant fraction of each cluster is not used by the previous pre-embedding; and we find the desired many copies of K_s in $N_G(v) \cap S$ as a result.

Proof of Theorem 43. Given $\gamma > 0$, we set $d^+ = 2^{-s-5}\gamma$ and $\varepsilon^+_{s-2} = 16^{-s}(d^+)^{2s}/s$. For each $i = s-3, s-4, \dots, 0$ sequentially, let $0 < \varepsilon^+_i \leqslant \varepsilon^+_{i-1}$ be sufficiently small for Lemma 36 with input d^+ and ε^+_{i+1}. Let $\varepsilon^+ \leqslant \varepsilon^+_0$ be small enough for an application of Lemma 41 with input d^+ and ε^+_0. Let t^+_1 be returned by Lemma 31 for input ε^+ and $\lceil 1/d^+ \rceil$, and let $\alpha^+ = \frac{1}{4}d^+/t^+_1$. Let $\gamma^+ = 2^{-4s^2}(d^+)^{-2s^2}(t^+_1)^{-s}$. Note we have $\gamma^+ < \gamma$.

We now choose $d \leqslant \frac{\gamma^+}{32}$ not larger than the d given by Lemma 44 for input γ, k

and $r_0 := 10\gamma^{-1}$. We let α be the ζ returned by Lemma 46 for input Δ, k, s, γ^+ and d. We set $D = \Delta$ and let $\varepsilon_{\rm BL}$ be returned by Lemma 34 for input Δ, $\Delta_{R'} = 3k$, $\Delta_J = \Delta$, $\vartheta = \frac{1}{100D}$, $\zeta = \frac{1}{4}\alpha$, d and $\kappa = 64$. Next, putting $\varepsilon^* := \frac{1}{8}\varepsilon_{\rm BL}$ into Lemma 46 (with earlier parameters as above) returns $\varepsilon_0 > 0$. We set $\varepsilon = \min(\varepsilon_0, d, \varepsilon^*/4\Delta, 1/100k)$, and set $\varepsilon^- \leqslant \varepsilon$ small enough for Lemma 41 with input as above and d, ε. Now Lemma 44, for input ε^- and earlier constants as above, returns r_1. At last, Lemma 47, for input k, r_1, Δ, γ, d and 8ε, returns $\xi > 0$. Without loss of generality, we may assume $\xi < 10(10kr_1)$, and set $\beta = 10^{-12}\xi^2/(\Delta k^4 r_1^2)$. Let $\mu = \varepsilon^2/(100000kr_1)$. Next, suppose C^* is large enough to play the rôle of C in each of these lemmas, and also for Proposition 38 with input ε, for Lemma 36 with input d^+ and each of ε_i^+ for $i = 1, \dots, s-2$, and for Lemma 41 with input $\varepsilon\mu^2$, ε, $\min(d, d^+)$ and Δ.

We set $C = 10^{100}k^2 r_1^2 \varepsilon^{-2}\xi^{-1}\Delta^{1000k^3}\mu^{-\Delta}C^*$ and $z = 10/\xi$. Given $p \geqslant C\big(\frac{\log}{n}\big)$, a.a.s. $\Gamma = G(n,p)$ satisfies the good events of each of the lemmas and propositions listed above with each of the specified inputs.

In addition, for each set W of at most Δ vertices of $G(n,p)$, the size of the common neighbourhood $N_{G(n,p)}(W)$ is distributed as a binomial random variable with mean $p^{|W|}(n-|W|)$. By Theorem 39, the probability that the outcome is $(1\pm\varepsilon)p^{|W|}n$ is at least $1-n^{-(\Delta+1)}$ for sufficiently large n. By the union bound, we conclude that a.a.s. $G(n,p)$ satisfies

$$\text{for each } W \subset V\big(G(n,p)\big) \text{ with } |W| \leqslant \Delta \text{ we have } \big|N_{G(n,p)}(W)\big| = (1\pm\varepsilon)p^{|W|}n\,. \tag{3.7}$$

Suppose that $\Gamma = G(n,p)$ satisfies these good events. Let G be a spanning subgraph of Γ such that $\delta(G) \geqslant \big(\frac{k-1}{k} + \gamma\big)pn$ and such that for each $v \in V(G)$ the neighbourhood $N_G(v)$ contains at least $\delta p^{\binom{s}{2}}(pn)^s$ copies of K_s. Let H be a graph on n vertices with $\Delta(H) \leqslant \Delta$. Let σ be a proper colouring of $V(H)$ using colours $\{0, \dots, k\}$, and let $\mathcal{L}$ be a labelling of $V(H)$ with bandwidth at most βn with the following properties. The colouring σ is (z,β)-zero-free with respect to $\mathcal{L}$, the first $\sqrt{\beta}n$ vertices of $\mathcal{L}$ do not use the colour zero, and the first βn vertices of $\mathcal{L}$ contain Cp^{-2} vertices whose neighbourhood contains only s colours.

We now claim that for each $v \in V(G)$ we can find s large subsets of $N_G(v)$ all pairs of which are dense and regular in G. This forms a 'robust witness' that each vertex neighbourhood in G contains many copies of K_s.

Claim 50. *For each $v \in V(G)$, there exist sets $Q_{v,1}, \ldots, Q_{v,s} \subset N_G(v)$ each of size at least $\alpha^+ pn$ such that for each $i < j$ the pair $(Q_{v,i}, Q_{v,j})$ is (ε^+, d^+, p)-regular in G.*

Proof. We apply Lemma 31 with input ε^+ and $\lceil 1/d^+ \rceil$ to $G\big[N_G(v)\big]$, with an arbitrary equipartition into $\lceil 1/d^+ \rceil$ sets as an initial partition. Note that the conditions of Lemma 31 are satisfied because the good event of Proposition 38 holds. We obtain an (ε, p)-regular partition of $N_G(v)$ whose non-exceptional parts are of size between $\alpha^+ pn$ and $8\alpha^+ pn$, by choice of α^+ and since $\big|N_G(v)\big| > \frac{1}{2}pn$. If there exist s parts in this partition all pairs of which form (ε^+, d^+, p)-regular pairs, then these parts form the desired $Q_{v,1},\ldots,Q_{v,s}$. So we may assume for a contradiction that no such s parts exist. It follows that when we delete all edges within parts, meeting the exceptional sets, in irregular pairs, and in pairs of density less than d^+p, we remove all copies of K_s from $G\big[N_G(v)\big]$.

The total number of such edges is, since the good event of Proposition 38 holds, at most

$$\begin{aligned}(d^+)^{-1} \cdot 8p^3n^2(d^+)^2 + 2p(2\varepsilon^+ pn)(2pn) + 4\varepsilon^+ p^3n^2 + 4d^+p^3n^2 &\leqslant (12\varepsilon^+ + 12d^+)p^3n^2 \\ &\leqslant 2^{-s}\gamma p^3 n^2\,,\end{aligned}$$

where the final inequality is by choice of d^+ and ε^+. We now estimate simply how many copies of K_{s+1} a given edge e, together with v, can make in Γ. Since by (3.7) any ℓ-tuple of vertices of Γ has at most $2p^\ell n$ common neighbours, the number of copies of K_4 containing e and v is at most $2p^3n$, and inductively the number of copies of K_{s+1} containing e and v is at most

$$\prod_{\ell=3}^{s} 2p^\ell n = 2^{s-2}p^{\binom{s+1}{2}-3}n^{s-2}\,.$$

Putting these estimates together we see that the total number of copies of K_s in $G\big[N_G(v)\big]$ is at most $\frac{1}{2}\gamma p^{\binom{s+1}{2}}n^s$. This is the desired contradiction, completing the proof. □

We apply Lemma 44 to G, with input γ, k, r_0 and ε^-, to obtain an integer r with $10\gamma^{-1} \leqslant kr \leqslant r_1$, a set $V_0 \subset V(G)$ with $|V_0| \leqslant C^*p^{-2}$, a k-equitable partition $\mathcal{V} = \big\{V_{i,j}\big\}_{i\in[r],j\in[k]}$ of $V(G) \setminus V_0$, and a graph R^k_r on $[r]\times[k]$ with minimum degree $\delta(R^k_r) \geqslant \big(\frac{k-1}{k} + \frac{\gamma}{2}\big)kr$, such that $K^k_r \subset B^k_r \subset R^k_r$ and such that the following hold.

(G1a) $\frac{n}{4kr} \leqslant |V_{i,j}| \leqslant \frac{4n}{kr}$ for every $i \in [r]$ and $j \in [k]$,

(G2a) $\mathcal{V}$ is $(\varepsilon^-, d, p)_G$-lower-regular on R_r^k and $(\varepsilon^-, d, p)_G$-super-regular on K_r^k,

(G3a) both $\big(N_\Gamma(v, V_{i,j}), V_{i',j'}\big)$ and $\big(N_\Gamma(v, V_{i,j}), N_\Gamma(v, V_{i',j'})\big)$ are $(\varepsilon^-, d, p)_G$-lower-regular pairs for every $\{(i,j), (i',j')\} \in E(R_r^k)$ and $v \in V \setminus V_0$, and

(G4a) $|N_\Gamma(v, V_{i,j})| = (1 \pm \varepsilon)p|V_{i,j}|$ for every $i \in [r]$, $j \in [k]$ and every $v \in V \setminus V_0$.

Given $i \in [r]$, because $\delta(R_r^k) > (k-1)r$, there exists $v \in V(R_r^k)$ adjacent to each (i,j) with $j \in [k]$. This, together with our assumptions on H, allow us to apply Lemma 45 to H, with input D, k, r, $\frac{1}{10}\xi$ and β, and with $m_{i,j} := |V_{i,j}| + \frac{1}{kr}|V_0|$ for each $i \in [r]$ and $j \in [k]$, choosing the rounding such that the $m_{i,j}$ form a k-equitable integer partition of n. Since $\Delta(H) \leqslant \Delta$, in particular H is Δ-degenerate. Let $f\colon V(H) \to [r] \times [k]$ be the mapping returned by Lemma 45, let $W_{i,j} := f^{-1}(i,j)$, and let $X \subseteq V(H)$ be the set of special vertices returned by Lemma 45. For every $i \in [r]$ and $j \in [k]$ we have

(H1a) $m_{i,j} - \frac{1}{10}\xi n \leqslant |W_{i,j}| \leqslant m_{i,j} + \frac{1}{10}\xi n$,

(H2a) $|X| \leqslant \xi n$,

(H3a) $\{f(x), f(y)\} \in E(R_r^k)$ for every $\{x,y\} \in E(H)$,

(H4a) $y, z \in \bigcup_{j' \in [k]} f^{-1}(i,j')$ for every $x \in f^{-1}(i,j) \setminus X$ and $xy, yz \in E(H)$, and

(H5a) $f(x) = \big(1, \sigma(x)\big)$ for every x in the first $\sqrt{\beta}n$ vertices of $\mathcal{L}$.

We let F be the first βn vertices of $\mathcal{L}$. By definition of $\mathcal{L}$, in F there are at least Cp^{-2} vertices whose neighbourhood in H receives at most s colours from σ.

Next, we apply Lemma 41, with input $\varepsilon\mu^2$ and Δ, to choose a set $S \subset V(G)$ of size μn. We let the T_i of Lemma 41 be all sets which are common neighbourhoods in Γ of at most Δ vertices of Γ, together with the sets $V_{i,j}$ for $i \in [r]$ and $j \in [k]$, and the sets $Q_{v,i}$ for $v \in V(G)$ and $i \in [s]$. We let the regular pairs (X_i, Y_i) of Lemma 41 be the pairs $(Q_{v,i}, Q_{v,j})$ for $1 \leqslant i < j \leqslant s$ and $v \in V(G)$, and all regular pairs $(V_{i,j}, V_{i',j'}) \in R_r^k$.

The result of Lemma 41 is that for any $1 \leqslant \ell \leqslant \Delta$ and vertices $u_1, \dots, u_\ell$ of $V(G)$, we have

$$\begin{gathered}\Big|S \cap \bigcap_{1\leqslant i\leqslant \ell} N_\Gamma(u_i)\Big| = (1 \pm \varepsilon\mu)\mu\Big|\bigcap_{1\leqslant i\leqslant \ell} N_\Gamma(u_i)\Big| \pm \varepsilon\mu p^\ell n\,, \quad \text{and} \\ \big|S \cap V_{i,j}\big| = (1 \pm \tfrac{1}{2}\varepsilon)\mu|V_{i,j}| \quad \text{for each } i \in [r] \text{ and } j \in [k],\end{gathered} \tag{3.8}$$

where we use the fact $p \geqslant C\left(\frac{\log n}{n}\right)^{1/\Delta}$ and choice of C to deduce $C^* \log n < \varepsilon\mu p^\Delta n$. Furthermore, for each $v \in V(G)$ and $1 \leqslant i < j \leqslant s$ the pair $\big(Q_{v,i} \cap S, Q_{v,j} \cap S\big)$ is $\big(\varepsilon_0^+, d^+, p)$-regular in G, and for each $(V_{i,j}, V_{i',j'}) \in R_r^k$ the pair $\big(V_{i,j} \cap U, V_{i',j'} \cap U\big)$ is (ε, d, p)-regular in G.

Our next task is to create the pre-embedding that covers the vertices of V_0. We use the following algorithm, starting with φ_0 the empty partial embedding. Suppose this algorithm does not fail, terminating with $t = t^*$. Then the final φ_{t^*} is

Algorithm 1: Pre-embedding

Set $t := 0$;
while $V_0 \setminus \mathrm{Im}(\varphi_t) \neq \varnothing$ **do**
  Let $v_{t+1} \in V_0 \setminus \mathrm{Im}(\varphi_t)$ maximise $\big|N_G(v) \cap S \cap \mathrm{Im}(\varphi_t)\big|$ over $v \in V_0 \setminus \mathrm{Im}(\varphi_t)$;
  Choose $x_{t+1} \in F$ such that $\big|\sigma\big(N_H(x)\big)\big| \leqslant s$ and $\mathrm{dist}\big(x_{t+1}, \mathrm{Dom}(\varphi_t)\big) \geqslant 100k^2$;
  Set $H_{t+1} := H\left[\left\{y \in V(H) : \mathrm{dist}(x_{t+1}, y) \leqslant s+1\right\}\right]$;
  Set G'_{t+1} the maximum subgraph of $G\big[(S \cup \{v_{t+1}\}) \setminus \mathrm{Im}(\varphi_t)\big]$ with minimum degree $\left(\frac{k-1}{k} + \frac{\gamma}{4}\right)\mu pn$;
  Let φ be given by Lemma 46 with input G'_{t+1}, H'_{t+1} and colouring $\sigma|_{V(H')}$;
  Set $\varphi_{t+1} := \varphi_t \cup \varphi$;
  $t := t+1$;
end

an embedding of some vertices of H into $V(G)$ which covers V_0 and is contained in $V_0 \cup S$. We will see that, because φ_{t^*} is obtained by successively using Lemma 46 and by (3.8), the vertices of $H \setminus \mathrm{Dom}(\varphi_{t^*})$ which have neighbours in $\mathrm{Dom}(\varphi_{t^*})$ have image restriction sets matching the requirement of Definition 33. Before we justify this, we first claim that the algorithm does not fail, and the requirements of Lemma 46 are met at each iteration.

Claim 51. *Algorithm 1 does not fail, and the conditions of Lemma 46 are met at each iteration.*

Proof. Observe that in total we embed at most Δ^{s+2} vertices in each iteration, and the number of iterations is at most $|V_0| \leqslant C^* p^{-2}$, so that the total number of vertices we embed is at most $C^*\Delta^{s+2}p^{-2}$.

We begin by discussing the choice of v_{t+1}. Suppose that at some time t we pick a vertex $v = v_{t+1}$ such that $\big|N_G(v) \cap S \cap \mathrm{Im}(\varphi_t)\big| > \frac{1}{2}\alpha^+\mu pn$. For each $t - \frac{1}{4}\Delta^{-s-2}\mu\alpha^+pn \leqslant t' < t$, we have $\big|N_G(v) \cap S \cap \mathrm{Im}(\varphi_{t'})\big| > \frac{1}{4}\alpha^+\mu pn$, yet at each of these times v is not picked, so that the vertex picked at each time t' has at least $\frac{1}{4}\alpha^+\mu pn$ neighbours in $\mathrm{Im}(\varphi_t) \cap S$, and in particular in $\mathrm{Im}(\varphi_t)$, a set of size at most $C^*\Delta^{s+2}p^{-2}$. Let Z be a superset of $\mathrm{Im}(\varphi_t)$ of size at least $C^*p^{-1}\log n$. Now the good event of Proposition 38 states that in Γ at most $C^*p^{-1}\log n$ vertices of Γ have more than $2p|Z| < \frac{1}{4}\alpha^+\mu pn$ neighbours in Z. Since $\frac{1}{4}\Delta^{-s-2}\mu\alpha^+pn > C^*p^{-1}\log n$ by choice of p, this is a contradiction. We conclude that at each time t, the vertex v_{t+1} picked at time t satisfies $\big|N_G(v) \cap S \cap \mathrm{Im}(\varphi_t)\big| \leqslant \frac{1}{2}\alpha^+\mu pn$.

From this point on we consider a fixed time t, and write v rather than v_{t+1}, and φ for φ_t, and so on.

Since we cover at most $C^*\Delta^{s+2}p^{-2}$ vertices, so we have $|S \setminus \mathrm{Im}(\varphi)| = (1 \pm \frac{1}{2}\varepsilon)\mu n$. Now, to obtain the maximum subgraph of $G\big[(S \cup \{v\}) \setminus \mathrm{Im}(\varphi)\big]$ with minimum degree $\big(\frac{k-1}{k} + \frac{\gamma}{4}\big)\mu pn$, we successively remove vertices whose degree is too small until no further remain. We claim that less than $\frac{1}{8}\mu\alpha^+pn$ vertices are removed, and v is not one of the vertices removed. To see this, observe that every vertex has at least $\big(\frac{k-1}{k} + \frac{\gamma}{2}\big)\mu pn$ neighbours in S by (3.8). Suppose for a contradiction that there is a set Z of $\frac{1}{8}\mu\alpha^+pn$ vertices which are the first removed from S in this process. Then each vertex of Z has at least $\frac{1}{4}\gamma\mu pn$ neighbours in $Z \cup \mathrm{Im}(\varphi)$, which by choice of α^+ is a contradiction to the good event of Proposition 38.

We conclude $\big|(S \cup \{v\}) \setminus \mathrm{Im}(\varphi)\big| = (1 \pm \varepsilon)\mu n$. Since v has at least $\big(\frac{k-1}{k} + \frac{\gamma}{2}\big)\mu pn$ neighbours in S, of which at most $\frac{1}{2}\alpha^+\mu pn$ are in $\mathrm{Im}(\varphi)$ and at most $|Z|$ are in Z, v is not removed. Furthermore, for each $i \in [s]$ we have $\big|Q_{v,i} \cap V(G')\big| \geqslant \frac{1}{2}\big|Q_{v,i} \cap S\big|$. We now use this to count copies of K_s in $N_{G'}(v)$. We choose for $i = 1, \ldots, s$ sequentially vertices in $Q_{v,i} \cap V(G')$, at each step choosing a vertex w_i which is adjacent to the previous vertices, and which is such that $w_1, \ldots, w_i$ have at least $(d^+ - \varepsilon^+_{s-2})^i p^i |Q_{v,j}|$ common G-neighbours in each $Q_{v,j}$ for $j > i$, and have $(1 \pm \varepsilon)^i p^i |Q_{v,j}|$ common Γ-neighbours in each $Q_{v,j}$ for $j > i$, and the pair $\big(N_\Gamma(w_1, \ldots, w_i; Q_{v,j}), N_\Gamma(w_1, \ldots, w_i; Q_{v,j'})\big)$ is $(\varepsilon^+_i, d^+, p)$-lower-regular in G for each $i < j < j' \leqslant s$. Note that all these properties hold when $i = 0$ vertices have been chosen. Assuming these properties hold when we come to choose w_i, there are at least $2^{1-i}(d^+)^{i-1}p^{i-1}|Q_{v,i}|$ vertices of $Q_{v,i}$ which are adjacent to all previously chosen vertices. If $i = s$ then all of these are valid choices. If $i < s$, by Propositions 28 and 29, and because the good event of Proposition 38 holds, at most

$$s \cdot 4^i (d^+)^{1-i} \varepsilon^+_{s-2} p^{i-1} |Q_{v,i}| + s \cdot C^* p^{-1} \log n$$

vertices of $Q_{v,i}$ cause the numbers of G- or Γ-common neighbours in some $Q_{v,j}$ for $j > i$ to go wrong. Finally, if $i = s-1$ then there is no choice of $i < j < j' \leqslant s$ and so no failure of lower-regularity can occur, while if $i < s-1$ then by the good event of Lemma 36 the number of vertices which cause a failure of lower-regularity is at most $s^2 C^* p^{-2} \log n$. By choice of ε^+_{s-2} and p, in total at least $2^{-i}(d^+)^{i-1} p^{i-1} |Q_{v,i}|$ vertices of $Q_{v,i}$ are thus valid choices for w_i. Finally, by choice of γ^+ the total number of copies of K_s in $N_{G'}(v)$ is at least $2\gamma^+ p^{\binom{s}{2}} \big(p|S|\big)^s \geqslant \gamma^+ p^{\binom{s+1}{2}} (\mu n)^s$, as desired.

The remaining conditions of Lemma 46 are simpler to check. By (3.8) we have $\big|N_{G'}(W)\big| \leqslant \big|N_\Gamma(W) \cap S\big| \leqslant 2\mu n p^{|W|}$ for any $W \subset V(G')$ of size at most Δ. The graph G with the regular partition $(V_{i,j})_{i\in[r],j\in[k]}$, with reduced graph R^k_r, has the required minimum degree. By (3.8) the intersection of the part $V_{i,j}$ with S has size $(1 \pm \frac{1}{2}\varepsilon)\mu|V_{i,j}|$, so that $|V_{i,j} \cap V(G')| = (1 \pm \varepsilon)\mu|V_{i,j}|$ as required. Furthermore the regular pairs of R intersected with S are regular, and so by Proposition 29 the subpairs obtained by intersecting with $V(G')$ (which is, except for v, contained in S; and v is in V_0 hence not in any of these pairs) are also sufficiently regular. Finally, the graph H_{t+1} chosen at each time t satisfies the conditions of Lemma 46 by definition. Note that we can at each step choose x_{t+1} and hence H_{t+1} because there are at least Cp^{-2} vertices of F whose neighbourhood is coloured with at most s colours; even after embedding all of V_0, the domain of φ contains at most $C^*\Delta^{s+2}p^{-2}$ vertices, and hence at most $C^*\Delta^{s+100k^2+3}p^{-2} < Cp^{-2}$ vertices of H are too close to $\mathrm{Dom}(\varphi)$. □

Let $\varphi := \varphi_{t^*}$ and $H' = H \setminus \mathrm{Dom}(\varphi)$. We next define image restricting vertex sets and create an updated homomorphism $f^* : V(H') \to [r] \times [k]$. The former is easier. The vertices of $\mathrm{Dom}(\varphi)$ are partitioned according to the x_t chosen at each time in Algorithm 1, and because these vertices are chosen far apart in H, any vertex of $V(H')$ with a neighbour in $\mathrm{Dom}(\varphi)$ is at distance $s+1$ from some x_t. Its neighbours in H' are either also at distance $s+1$ in H from x_t and not adjacent to any vertices of $\mathrm{Dom}(\varphi)$ corresponding to other $x_{t'}$, or they are not adjacent to any vertex of $\mathrm{Dom}(\varphi)$ at all. Now items $(P3)$, $(P4)$, $(P5)$ and $(P6)$ immediately give valid image restriction sets for all the vertices $N_H\big(\mathrm{Dom}(\varphi)\big) \cap V(H')$, with the image restricting vertices for each such image restricted y being the vertices $J_y := \varphi\big(N_H(u) \cap \mathrm{Dom}(\varphi)\big)$.

We construct the updated homomorphism as follows. We will have $f^*(y) = f(y)$ for all vertices which are not within distance $s + \binom{k+1}{2}$ of $\mathrm{Dom}(\varphi)$ in H. Given a vertex x of H chosen at some time t in Algorithm 1, we set $f^*(y)$ for each y at distance between $s+1$ and $s + \binom{k+1}{2}$ from x in H as follows. We will generate a collection $Z_1, \dots, Z_{\binom{k+1}{2}}$ of copies of K_k in R_r^k, each labelled with the integers $1, \dots, k$. For each $i = 1, \dots, \binom{k+1}{2}$, if y is at distance $s + i$ from x in H, then we set $f^*(y)$ to be the label $\sigma(y)$ cluster of Z_i. The properties of the sequence $Z_1, \dots, Z_{\binom{k+1}{2}}$ we require are the following. First, Z_1 is the clique returned by the application of Lemma 46 at x with the labelling given by that lemma. Second, $Z_{\binom{k+1}{2}}$ is the clique $\big(V_{1,1}, \dots, V_{1,k}\big)$, labelled $1, \dots, k$ in that order. Third, for each $i = 2, \dots, \binom{k+1}{2}$, each cluster of Z_i is adjacent in R_r^k to each differently-labelled cluster of Z_{i-1}. Assuming such a sequence of cliques exists, the resulting f^* has the properties that each neighbour of $\mathrm{Dom}(\varphi)$ in H is assigned by f^* to the cluster of R_r^k in which Lemma 46 created an image restriction set, that each edge of H' is mapped by f^* to an edge of R_r^k, and that f and f^* disagree on at most $C^* p^{-2} \Delta^{s+\binom{k+1}{2}+3}$ vertices of H', all in the first $\sqrt{\beta} n$ vertices of $\mathcal{L}$. These will be the properties we need of f^*. Note that this definition is consistent, in that it does not attempt to set $f^*(y)$ to two different clusters for any y, because the vertices chosen at each step of Algorithm 1 are at pairwise distance at least $100k^2$. It remains only to show that the desired sequence of cliques always exists.

Claim 52. *For any k-cliques Z_1 and $Z_{\binom{k+1}{2}}$ in R_r^k a sequence $Z_1, \dots, Z_{\binom{k+1}{2}}$ with the above properties exists.*

Proof. By the minimum degree of R_r^k, any k-set in $V(R_r^k)$ has at least one common neighbour. We will use this fact at each step in the following algorithm. Set $t = 2$. We loop through $j = 1, \dots, k-1$ sequentially. For each value of j we perform the following operation.

For each $i = j+1, \dots, k$ sequentially, choose a cluster w_t of R_r^k which is adjacent to all the clusters of Z_{t-1} except possibly that labelled i, and which is also adjacent to the cluster of $Z_{\binom{k+1}{2}}$ labelled j. We let Z_t be the clique obtained from Z_{t-1} by replacing the label i cluster with w_t, which we label i; all other clusters keep their previous label. We increment t.

After performing the $i = k$ operation, we let Z_t be obtained from Z_{t-1} by replacing the label j cluster of Z_{t-1} with the label j cluster of $Z_{\binom{k+1}{2}}$, and increment t. We now proceed with the next round of the j-loop.

Observe that after the completion of each j-loop, the clusters of Z_{t-1} labelled $1, \ldots, j$ are the same as those of $Z_{\binom{k+1}{2}}$. In particular the given $Z_{\binom{k+1}{2}}$ has the required adjacencies in $Z_{\binom{k+1}{2}-1}$ (the final clique constructed in the $j = k-1$ loop), while the remaining required adjacencies hold by construction. □

At this point we complete the proof almost exactly as in [2]. What follows is taken from there, with only trivial changes, for completeness' sake.

For each $i \in [r]$ and $j \in [k]$, let $W'_{i,j}$ be the set of vertices $w \in V(H')$ with $f^*(w) \in V_{i,j}$, and let X' consist of X together with all vertices of H' at H-distance $100k^2$ or less from some x_t with $t \in [t^*]$. The total number of vertices $z \in V(H)$ at distance at most $100k^2$ from some x_t is at most $2\Delta^{200k^2}|V_0| < \frac{1}{100}\xi n$. Since $W_{i,j} \triangle W'_{i,j}$ contains only such vertices, we have

(H1b) $m_{i,j} - \frac{1}{5}\xi n \leqslant |W'_{i,j}| \leqslant m_{i,j} + \frac{1}{5}\xi n$,

(H2b) $|X'| \leqslant 2\xi n$,

(H3b) $\{f^*(x), f^*(y)\} \in E(R^k_r)$ for every $\{x, y\} \in E(H')$, and

(H4b) $y, z \in \bigcup_{j' \in [k]} W'_{i,j'}$ for every $x \in W'_{i,j} \setminus X'$ and $xy, yz \in E(H')$.

where (H2b), (H3b) and (H4b) hold by (H2a) and definition of X', by definition of f^*, and by (H4a) and choice of X' respectively.

Furthermore, we have

(G1a) $\frac{n}{4kr} \leqslant |V_{i,j}| \leqslant \frac{4n}{kr}$ for every $i \in [r]$ and $j \in [k]$,

(G2a) $\mathcal{V}$ is $(\varepsilon, d, p)_G$-lower-regular on R^k_r and $(\varepsilon, d, p)_G$-super-regular on K^k_r,

(G3a) both $\big(N_\Gamma(v, V_{i,j}), V_{i',j'}\big)$ and $\big(N_\Gamma(v, V_{i,j}), N_\Gamma(v, V_{i',j'})\big)$ are $(\varepsilon, d, p)_G$-lower-regular pairs for every $\{(i,j),(i',j')\} \in E(R^k_r)$ and $v \in V \setminus V_0$, and

(G4a) $|N_\Gamma(v, V_{i,j})| = (1 \pm \varepsilon)p|V_{i,j}|$ for every $i \in [r]$, $j \in [k]$ and every $v \in V \setminus V_0$.

(G5a) $\big|V_{f^*(x)} \cap \bigcap_{u \in J_x} N_G(u)\big| \geqslant \alpha p^{|J_x|}|V_{f^*(x)}|$ for each $x \in V(H')$,

(G6a) $\big|V_{f^*(x)} \cap \bigcap_{u \in J_x} N_\Gamma(u)\big| = (1 \pm \varepsilon^*)p^{|J_x|}|V_{f^*(x)}|$ for each $x \in V(H')$, and

($G7$a) $\big(V_{f^*(x)} \cap \bigcap_{u \in J_x} N_\Gamma(u), V_{f^*(y)} \cap \bigcap_{v \in J_y} N_\Gamma(v)\big)$ is $(\varepsilon^*, d, p)_G$-lower-regular for each $xy \in E(H')$.

($G8$a) $\big|\bigcap_{u \in J_x} N_\Gamma(u)\big| \leqslant (1 + \varepsilon^*)p^{|J_x|}n$ for each $x \in V(H')$,

Properties ($G1$a) to ($G4$a) are repeated for convenience (replacing ε^- with the larger ε). Properties ($G5$a), ($G6$a) and ($G8$a), are trivial when $J_x = \varnothing$, and are otherwise guaranteed by Lemma 46. Finally ($G7$a) follows from ($G2$a) when $J_x, J_y = \varnothing$, and otherwise is guaranteed by Lemma 46.

For each $i \in [r]$ and $j \in [k]$, let $V'_{i,j} = V_{i,j} \setminus \mathrm{Im}(\varphi_{t^*})$, and let $\mathcal{V}' = \{V'_{i,j}\}_{i \in [r], j \in [k]}$. Because $V_{i,j} \setminus V'_{i,j} \subset S$ for each $i \in [r]$ and $j \in [k]$, using (3.8) and Proposition 29, and our choice of μ, we obtain

($G1$b) $\frac{n}{6kr} \leqslant |V'_{i,j}| \leqslant \frac{6n}{kr}$ for every $i \in [r]$ and $j \in [k]$,

($G2$b) $\mathcal{V}'$ is $(2\varepsilon, d, p)_G$-lower-regular on R^k_r and $(2\varepsilon, d, p)_G$-super-regular on K^k_r,

($G3$b) both $\big(N_\Gamma(v, V'_{i,j}), V'_{i',j'}\big)$ and $\big(N_\Gamma(v, V'_{i,j}), N_\Gamma(v, V'_{i',j'})\big)$ are $(2\varepsilon, d, p)_G$-lower-regular pairs for every $\{(i, j), (i', j')\} \in E(R^k_r)$ and $v \in V \setminus V_0$, and

($G4$b) $|N_\Gamma(v, V'_{i,j})| = (1 \pm 2\varepsilon)p|V_{i,j}|$ for every $i \in [r]$, $j \in [k]$ and every $v \in V \setminus V_0$.

($G5$b) $\big|V'_{f^*(x)} \cap \bigcap_{u \in J_x} N_G(u)\big| \geqslant \frac{1}{2}\alpha p^{|J_x|}|V'_{f^*(x)}|$,

($G6$b) $\big|V'_{f^*(x)} \cap \bigcap_{u \in J_x} N_\Gamma(u)\big| = (1 \pm 2\varepsilon^*)p^{|J_x|}|V'_{f^*(x)}|$, and

($G7$b) $\big(V'_{f^*(x)} \cap \bigcap_{u \in J_x} N_\Gamma(u), V'_{f^*(y)} \cap \bigcap_{v \in J_y} N_\Gamma(v)\big)$ is $(2\varepsilon^*, d, p)_G$-lower-regular.

($G8$b) $\big|\bigcap_{u \in J_x} N_\Gamma(u)\big| \leqslant (1 + 2\varepsilon^*)p^{|J_x|}n$ for each $x \in V(H')$,

We are now almost finished. The only remaining problem is that we do not necessarily have $|W'_{i,j}| = |V'_{i,j}|$ for each $i \in [r]$ and $j \in [k]$. Since

$$|V'_{i,j}| = |V_{i,j}| \pm 2\Delta^{200k^2}|V_0| = m_{i,j} \pm 3\Delta^{200k^2}|V_0|,$$

by ($H1$b) we have $|V'_{i,j}| = |W'_{i,j}| \pm \xi n$. We can thus apply Lemma 47, with input k, r_1, Δ, γ, d, 8ε, and r. This gives us sets $V''_{i,j}$ with $|V''_{i,j}| = |W'_{i,j}|$ for each $i \in [r]$ and $j \in [k]$ by ($B1$'). Let $\mathcal{V}'' = \{V''_{i,j}\}_{i \in [r], j \in [k]}$. Lemma 47 guarantees us the following.

($G1$c) $\frac{n}{8kr} \leqslant |V''_{i,j}| \leqslant \frac{8n}{kr}$ for every $i \in [r]$ and $j \in [k]$,

($G2$c) $\mathcal{V}''$ is $(4\varepsilon^*, d, p)_G$-lower-regular on R^k_r and $(4\varepsilon^*, d, p)_G$-super-regular on K^k_r,

($G3$c) both $\big(N_\Gamma(v, V''_{i,j}), V''_{i',j'}\big)$ and $\big(N_\Gamma(v, V''_{i,j}), N_\Gamma(v, V''_{i',j'})\big)$ are $(4\varepsilon^*, d, p)_G$-lower-regular pairs for every $\{(i, j), (i', j')\} \in E(R^k_r)$ and $v \in V \setminus V_0$, and

($G4$c) we have $(1-4\varepsilon)p|V''_{i,j}| \leqslant |N_\Gamma(v, V''_{i,j})| \leqslant (1+4\varepsilon)p|V''_{i,j}|$ for every $i \in [r]$, $j \in [k]$ and every $v \in V \setminus V_0$.

($G5$c) $\big|V''_{f^*(x)} \cap \bigcap_{u \in J_x} N_G(u)\big| \geqslant \frac{1}{4}\alpha p^{|J_x|}|V''_{f^*(x)}|$,

($G6$c) $\big|V''_{f^*(x)} \cap \bigcap_{u \in J_x} N_\Gamma(u)\big| = (1 \pm 4\varepsilon^*)p^{|J_x|}|V'_{f^*(x)}|$, and

($G7$c) $\big(V''_{f^*(x)} \cap \bigcap_{u \in J_x} N_\Gamma(u), V''_{f^*(y)} \cap \bigcap_{v \in J_y} N_\Gamma(v)\big)$ is $(4\varepsilon^*, d, p)_G$-lower-regular.

Here ($G1$c) comes from ($G1$b) and ($B2$'), while ($G2$c) comes from ($B3$') and choice of ε. ($G3$c) is guaranteed by ($B4$'). Now, each of ($G4$c), ($G5$c) and ($G6$c) comes from the corresponding ($G4$b), ($G5$b) and ($G6$b) together with ($B5$'). Finally, ($G7$c) comes from ($G7$b) and ($G8$b) together with Proposition 29 and ($B5$').

For each $x \in V(H')$ with $J_x = \varnothing$, let $I_x = V''_{f^*(x)}$. For each $x \in V(H')$ with $J_x \neq \varnothing$, let $I_x = V''_{f^*(x)} \cap \bigcap_{u \in J_x} N_G(u)$. Now $\mathcal{W}'$ and $\mathcal{V}''$ are κ-balanced by ($G1$c), size-compatible by construction, partitions of respectively $V(H')$ and $V(G) \setminus \mathrm{Im}(\varphi_{t^*})$, with parts of size at least $n/(\kappa r_1)$ by ($G1$c). Letting $\widetilde{W}_{i,j} := W'_{i,j} \setminus X'$, by ($H2$b), choice of ξ, and ($H4$b), $\{\widetilde{W}_{i,j}\}_{i \in [r], j \in [k]}$ is a $\big(\vartheta, K_r^k\big)$-buffer for H'. Furthermore since f^* is a graph homomorphism from H' to R_r^k, we have ($BUL1$). By ($G2$c), ($G3$c) and ($G4$c) we have ($BUL2$), with $R = R_r^k$ and $R' = K_r^k$. Finally, the pair $(\mathcal{I}, \mathcal{J}) = \big(\{I_x\}_{x \in V(H')}, \{J_x\}_{x \in V(H')}\big)$ form a $\big(\varrho, \frac{1}{4}\alpha, \Delta, \Delta\big)$-restriction pair. To see this, observe that the total number of image restricted vertices in H' is at most $\Delta^2|V_0| < \varrho|V_{i,j}|$ for any $i \in [r]$ and $j \in [k]$, giving ($RP1$). Since for each $x \in V(H')$ we have $|J_x| + \deg_{H'}(x) = \deg_H(x) \leqslant \Delta$ we have ($RP3$), while ($RP2$) follows from ($G5$c), and ($RP5$) follows from ($G6$c). Finally, ($RP6$) follows from ($G7$c), and ($RP4$) follows since $\Delta(H) \leqslant \Delta$. Together this gives ($BUL3$). Thus, by Lemma 34 there exists an embedding φ of H' into $G \setminus \mathrm{Im}(\varphi_{t^*})$, such that $\varphi(x) \in I_x$ for each $x \in V(H')$. Finally, $\varphi \cup \varphi_{t^*}$ is an embedding of H in G, as desired. □

3.5 Remarks on the optimality

In Theorems 8 and 26, the requirement for C^*p^{-2} vertices in H whose neighbourhood contains few colours is optimal up to the value of C^*. However the value of C^* we obtain derives from (multiple applications of) the sparse regularity lemma and is hence very far from optimal. One can use the methods of this proof to obtain an improved (but still far from sharp) constant, and we expect that one can use the methods of this proof to determine an optimal C^* asymptotically, at least for special cases.

The way to obtain this improvement is the following. We work exactly as in the proof of Theorem 43, except that for each $v \in V(G)$ we identify the largest $1 \leqslant s \leqslant k-1$ for which there are many copies of K_s in $N_G(v)$, and obtain a robust witness for this property as in that proof. Now when we come to cover the vertices of the set V_0 returned by Lemma 44, we use vertices from zero-free regions of $\mathcal{L}$ which are not in the first few vertices of $\mathcal{L}$ whenever possible: in particular this is always possible when we are to cover a vertex which is in many copies of K_k. Our proof, with trivial modification, shows that this pre-embedding method succeeds. The result is that we can reduce C^* to a quantity on the order of Δ^{100k^2}; this number comes from our requirement to choose vertices in $\mathcal{L}$ which are widely separated in H for the pre-embedding onto the vertices of V_0 which are not in many copies of K_k.

When H contains many isolated vertices, this requirement disappears and we can further improve. We believe (but have not attempted to prove) that there is some C_k with the following property. Let Γ be a typical instance of $G(n,p)$, where $p \gg n^{-1/k}$. Suppose $G \subset \Gamma$ has minimum degree $\big(\frac{k-1}{k} + o(1)\big)pn$. Then any choice of G contains at most $\big(C_k + o(1)\big)p^{-2}$ vertices which are in $o\big(p^{\binom{k}{2}}n^{k-1}\big)$ copies of K_k; on the other hand there is a choice of G which has $\big(C_k - o(1)\big)p^{-2}$ vertices not in any copy of K_k.

Assuming the above statement to be true, it follows that C_k is the asymptotically optimal C^* whenever all vertices of H are either isolated or contained in a copy of K_k; for example when H consists of a $(k-1)$st power of a cycle together with some isolated vertices. Further generalisation to (for example) try to establish an optimal value of C^* in Theorem 8 would be possible; but it would also presumably depend on the graph structure of H. If the vertices of H which are not in triangles are far apart in H, then the generalisation is easy (and the answer is the same) but if they are not generally far apart it seems likely that one would have to use several such vertices to cover one badly-behaved vertex of G, and hence C^* would need to be larger than the above C_k.

4 Anti-Ramsey thresholds in sparse random graphs

In this chapter, we will prove the following result.

Theorem 53. *For $k \geqslant 5$, $p^{rb}_{K_k} = n^{-1/m^{(2)}(K_k)}$. Furthermore, $p^{rb}_{K_4} = n^{-7/15}$.*

We give the proof for complete graphs with at least 5 vertices in Section 4.1. The proof for K_4 is given in Section 4.2.

4.1 Complete graphs on at least five vertices

In this section we describe a strategy to prove lower bounds for p^{rb}_H when H is a complete graph with at least five vertices. The aim of this section is to prove the following result.

Lemma 54. *If H is a complete graph on at least five vertices, then $p^{rb}_H \geqslant n^{-m^{(2)}(H)}$.*

Beside the maximum 2-density of a graph, we will also need the *maximum density* $m(H)$ of a graph H, defined by

$$m(H) = \max\left\{\frac{|E(J)|}{|V(J)|} \colon J \subset H,\ |V(J)| \geqslant 1\right\}.$$

Theorem 14 is limited to complete graphs on at least 19 vertices only because of the following lemma [47, Lemma 24].

Lemma 55. *Let H be a complete graph on at least 19 vertices, then for any graph G with $m(G) < m^{(2)}(H)$ we do not have $G \xrightarrow[p]{rb} H$.*

We extend this for complete graphs by proving Lemma 56 below. Lemma 54 then follows by replacing Lemma 55 with Lemma 56 in the proof of [47, Theorem 7].

Lemma 56. *Let H be a complete graph on at least five vertices, then for any G with $m(G) < m^{(2)}(H)$ we do not have $G \xrightarrow[p]{rb} H$.*

In the remainder of this section we prove Lemma 56. In what follows we outline the ideas of our proof, analysing the structure of some subgraphs that will be important in our proof strategy (see Proposition 58 and Definition 59). We finish by proving an inductive result (Lemma 60) that directly implies Lemma 56.

From now on, let $k \geqslant 5$ and let G be a connected graph with $m(G) < m^{(2)}(K_k) = (k+1)/2$. Since we are interested in obtaining a colouring such that every K_k is non-rainbow, we may assume that all vertices and edges of G are contained in a K_k. We say that a subgraph of G is a K_k*-component* if any edge and vertex is contained a K_k and any pair of K_k's is K_k*-connected* in the following sense: two K_k's are K_k-connected if they are connected in the auxiliary graph that has K_k's in G as vertices and edge-intersecting K_k's as edges. Clearly, we may assume that G contains only a single K_k-component, as we might otherwise combine colourings of all its K_k-components to a colouring of G.

Let v be a vertex of minimum degree. A simple but important observation is that since the average degree in G is less than $k+1$, the vertex v has degree at most k. The following induced subgraphs of v and its neighbourhood in G play a special role in our proof:

- $K(v)$: on $\{v\} \cup N(v)$;
- $R(v)$: on $\{v\} \cup \{w \in N(v) : \text{every } K_k \text{ that contains } w \text{ also contains } v\}$;
- $S(v)$: on $V(K(v)) \setminus V(R(v))$.

We denote by G_v^* the induced graph on the vertices $V(G) \setminus V(R(v))$ and by G_v the graph obtained from G_v^* by removing all edges that are not contained in a K_k. In the inductive colouring strategy for Lemma 60, the induction step will be from G_v to G. The following simple fact provides useful information about the structure of G_v.

Fact 57. *Let $k \geqslant 5$ and let G be a graph on at least $k+1$ vertices with $m(G) < m^{(2)}(K_k) = (k+1)/2$ such that all vertices and edges of G are contained in a K_k. Let v be a vertex of minimum degree in G. Then the following hold.*

(i) *If $v(G_v) \leqslant k$ then G_v is a K_k;*

(ii) $|R(v)| \leqslant k-1$.

Proof. First suppose that $|R(v)| = k+1$. Thus, since $d(v) \leqslant k$, we know that v has exactly k neighbours (recall that v is a vertex of minimum degree). A clique K_k on $N(v)$ would contradict the definition of $R(v)$ so there is a non-edge in $R(v)$, say between u and w. Since w has degree at least k, there is an edge $\{w, z\}$ between w and a vertex z outside of $N(v)$. However, $\{w, z\}$ must also be contained in a K_k, so w cannot be in $R(v)$, a contradiction, so $|R(v)| <\leqslant k$.

For item (i), it is enough to show that any vertex that is contained in G_v is contained in a K_k. Since $|V(G)| \geqslant k+1$ and at most k vertices are removed, at least one vertex is left and if G_v contains at most k vertices, it is actually a K_k.

For item (ii), suppose now that $|R(v)| = k$. If $d(v) = k-1$, then all vertices in $R(v)$ have no neighbour outside of $\{v\} \cup R(v)$ and hence G is a K_k, a contradiction. If $d(v) = k$, then all vertices in $R(v)$ have no neighbour outside of $R(v)$ and since they have degree at least k, there is a clique K_k on $N(v)$, contradicting the assumption that $|R(v)| = k$. □

Note that since $d(v) \leqslant k$ and all vertices and edges are in a K_k, the subgraph $K(v)$ is either a K_k, a K_{k+1}^-, or a K_{k+1}. In the following proposition we categorise $K(v)$ according to its structure.

Proposition 58. *Let $k \geqslant 5$ and let G be a connected graph on at least $k+1$ vertices with $m(G) < (k+1)/2$ such that all vertices and edges of G are contained in a single K_k-component. Let v be a vertex of minimum degree in G. We say that $K(v)$ is either X_ℓ, Y_ℓ or U_1 according to the following, which are all possible configurations of $K(v)$.*

- *X_ℓ: $K(v) = K_k$, and $R(v) = K_\ell$, and $S(v) = K_{k-\ell}$ $(1 \leqslant \ell \leqslant k-2)$;*
- *Y_ℓ: $K(v) = K_{k+1}^-$, and $R(v) = K_\ell$, and $S(v) = K_{k-\ell+1}^-$ $(1 \leqslant \ell \leqslant k-2)$;*
- *U_1: $K(v) = K_{k+1}$, and $R(v) = K_1$, and $S(v) = K_k$.*

Proof. Clearly we have $R(v) = K_1$ when $K(v) = K_{k+1}$. So we only have to worry about the cases $K(v) = K_k$ and $K(v) = K_{k+1}^-$.

First, we will show that since G is a single K_k-component on at least $k+1$ vertices, we have $|R(v)| \leqslant k-2$. In fact, from 57(ii) we already know that $|R(v)| \leqslant k-1$. Suppose that $|R(v)| = k-1$. In this case, $K(v)$ is either a K_k or a K_{k+1}^-.

If $K(v) = K_k$, then $|S(v)| = 1$ and G is not a single K_k-component, since $K(v)$ is not K_k-connected with the other K_k's of G. If $K(v) = K_{k+1}^-$, then $|S(v)| = 2$. Moreover, $S(v)$ is an edge as otherwise G would not be a single K_k-component. But then, there is a missing edge xy with $x \in R(v)$ and $y \in S(v)$, which implies $d(x) < d(v)$, a contradiction. Therefore, we conclude that $|R(v)| \leqslant k - 2$.

It is left to show that if $K(v) = K_{k+1}^-$ and $R(v)$ has ℓ vertices (for any $1 \leqslant \ell \leqslant k - 2$), then $R(v) = K_\ell$, and $S(v) = K_{k-\ell+1}^-$. Suppose that $R(v)$ is not a K_ℓ. Then, since there is only one missing edge in $K(v)$, we have that $R(v) = K_\ell^-$, from where we conclude that there is a vertes in $R(v)$ with degree smaller than $d(v)$, a contradiction. Then, $R(v) = K_\ell$. Now, we just note that if $S(v)$ is not a $S(v) = K_{k-\ell+1}^-$, then the missing edge xy of $K(v)$ is such that $x \in R(v)$ and $y \in S(v)$, which implies $d(x) < d(v)$, a contradiction, which concludes the proof. $\square$

In Figure 4.1 we show all possible structures for $K(v)$ when $k = 5$. We will use the fact that $m(G) < (k + 1)/2$ to bound the number of occurrences of X_ℓ, Y_ℓ, and U_1 as $K(v)$ in the induction. Let $d^{(v)} = e(G_v^*) - e(G_v)$, i.e., the number of edges that are not contained in a K_k after removing $R(v)$. Recall that G_v is the graph obtained by removing $R(v)$ and all these edges.

Using the characterisation given in Proposition 58 the number of vertices removed is given in the subscripts and we can write the change in the number of edges from G_v to G as follows.

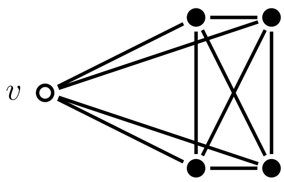
$$e(G) - e(G_v) - d^{(v)} = \begin{cases} k & \text{if } K(v) \text{ is } U_1, \\ \binom{\ell}{2} + \ell(k - \ell) & \text{if } K(v) \text{ is } X_\ell, \\ \binom{\ell}{2} + \ell(k - \ell + 1) & \text{if } K(v) \text{ is } Y_\ell. \end{cases} \tag{4.1}$$

We will use the following measure of how close G is to the upper bound $(k+1)/2$ on the density $m(G)$. Set

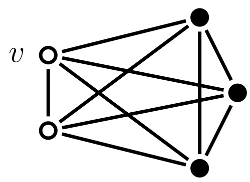
$$b(G) := 2e(G) - (k + 1)v(G) + 2k.$$

Clearly, from $e(G)/v(G) \leqslant m(G) < (k + 1)/2$ we have

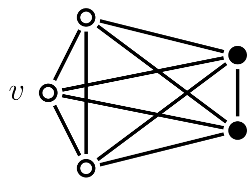
$$b(G) < 2k$$



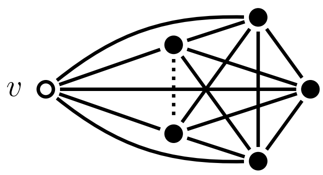
(a) X_1



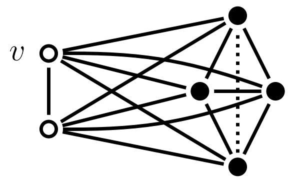
(b) X_2



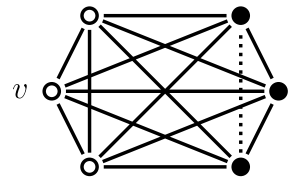
(c) X_3



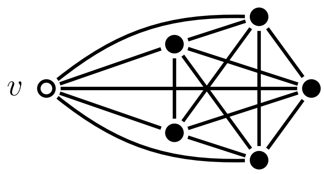
(d) Y_1



(e) Y_2



(f) Y_3



(g) U_1

Figure 4.1: Possible configurations of $K(v)$ for $k = 5$. Dotted lines represent non-edges, the vertices of $R(v)$ are white and the vertices of $S(v)$ are black.

and using (4.1) we get

$$b(G) - b(G_v) - 2d^{(v)} = \begin{cases} k-1 & \text{if } K(v) \text{ is } U_1, \\ (k-\ell-2)\ell & \text{if } K(v) \text{ is } X_\ell, \\ (k-\ell)\ell & \text{if } K(v) \text{ is } Y_\ell. \end{cases} \tag{4.2}$$

Note that there can be an arbitrary number of X_{k-2}'s in G (they contribute 0 to $b(G)$), but all other types of $K(v)$ are limited to a small number of occurrences.

We will describe an inductive colouring strategy, which will always lead to an edge-colouring of G with no rainbow K_k. To keep track of some additional properties of the colouring that will help us during the induction, we introduce five stages $P_0, \ldots, P_4$.

Definition 59 (Stages)**.** *Let $0 \leqslant j \leqslant 4$. We say that G is in stage P_j or $G \in P_j$ if there exists a partial proper colouring of G such that the following properties hold.*

(i) *Any K_k in G is non-rainbow;*

(ii) *If $G \in P_0$ then any K_3, K_4, and K_k contains at most 2 coloured edges, each colour is used exactly twice and any 4 vertices span at most 3 coloured edges; Also, any two K_k's intersect in at most one edge;*

(iii) *If $G \in P_j$ $(1 \leqslant j \leqslant 3)$ then any 4 vertices span at most $j+2$ coloured edges;*

Property (i) is the main property of the colouring we want to ensure. Properties (ii) and (iii) will allow us to keep the induction proof for Lemma 60 going. We will show that G is in some stage and that a certain amount of $b(G)$ is needed for G to not be in a smaller stage. Lemma 56 follows trivially from Definition 59(i) and Lemma 60 below.

Lemma 60. *Let $k \geqslant 5$ and let G be a connected graph on at least k vertices with $m(G) < (k+1)/2$ such that all vertices and edges of G are contained in a single K_k-component. There exists $0 \leqslant j \leqslant 4$ and a proper partial edge-colouring of G such that G is in stage P_j under this colouring. Furthermore,*

- $b(G) \geqslant 0$*;*
- *If $b(G) < 2$, then $G \in P_0$;*
- *If $b(G) < k-1$, then $G \in P_1$;*

- *If* $b(G) < k+1$*, then* $G \in P_2$*;*
- *If* $b(G) < 2k-2$*, then* $G \in P_3$*;*

Proof. We prove the lemma by induction on the size of the graph. If $|G| = k$ then G is a K_k by Fact 57 and we can colour two non-adjacent edges of G with the same colour and therefore G is in P_0. Also, $b(K_k) = 0$, so the lemma holds.

Now consider a graph G on at least $k+1$ vertices satisfying the assumptions of the lemma. Depending on $b(G)$, we have to show that G is in a certain stage. Let j be the maximal index in $\{0, \dots, 4\}$ such that the lemma holds if we prove $G \in P_j$. Let v be a vertex of minimum degree in G. Fact 57(i) implies that G_v has at least k vertices. We will first handle the case that G_v is a single K_k-component, so by the inductive hypothesis the lemma holds for G_v and G_v is in some stage $P_{j'}$.

If $d^{(v)} > 0$ and $j' = 0$, then $b(G_v^*) \geqslant b(G_v) + 2$ and it is easy to see that G_v^* is in stage P_1. In this case we set $j^* = 1$, otherwise we set $j^* = j'$. As no edge in $E(G_v^*) \setminus E(G_v)$ is contained in a K_k, the partial edge-colouring of G_v ensures that G_v^* is in P_{j^*} under the same partial edge-colouring.

For most graphs $K(v)$, we will get an increase in $b(G)$ from $b(G_v^*)$ which tells us that G_v^* is in a lower stage. The graph X_{k-2} is the only configuration where j might be equal to j^*, because it contributes zero to $b(G)$, and in this case we will show that G is also in P_{j^*}. The information about the possible transitions between stages is encapsulated in (4.2). Note that the difference in the bound on $b(G)$ between two consecutive stages P_i and P_{i+1} in Lemma 60 is at most $k-3$ and between two stages P_i and P_{i+2} it is at most $k-1$. In other words, if $b(G) \geqslant b(G_v^*) + k - 3$, we have $j^* \leqslant j - 1$ and we advance at least one stage from G_v^* to G. Since $(k - \ell - 2)\ell \geqslant k - 3$ for $1 \leqslant \ell < k - 2$ we have that if $K(v)$ is an X_ℓ then $j^* \leqslant j - 1$. Since $(k - \ell)\ell \geqslant k - 1$ for $\ell \geqslant 2$ we advance two stages, i.e., $j^* \leqslant j - 2$, for all but X_ℓ. We will now give the desired partial edge-colouring that extends the edge-colouring of G_v^* to G such that $G \in P_j$.

If $K(v) = X_{k-2}$ then $K(v)$ intersects G_v^* in exactly one edge. We colour two disjoint edges, one contained in $R(v)$ and the other with one endpoint in $R(v)$, with a new colour. These two edges do not close a coloured triangle and no K_4 intersecting both $K(v)$ and G_v^* in more than one edge exists, and clearly all K_4's and K_k's in $K(v)$ contain at most two coloured edges. Also any four vertices containing one of the newly coloured edges can contain at most three coloured edges, so if G_v^* was in P_0 then Property 59(ii) still holds. By the last part of this argument, Property 59(iii) holds in G if it did in G_v^*, so G is in P_j.

If $K(v) = X_\ell$ for $2 \leqslant \ell < k-2$, then we extend the current colouring in the following way: if there is any coloured edge in $S(v)$, then we give this colour to one of the edges in $R(v)$, which contains an edge since $\ell \geqslant 2$. Otherwise we choose a new colour and colour two disjoint edges that both intersect $R(v)$ with this colour. In the first case it is trivial that G is in P_{j*+1} and in the second it is easy to see as the only set of four vertices where we added two edges contains at most three coloured edges.

If $K(v) = X_1$, recall that we advance one stage. If $G_v^* \in P_0$ and $S(v)$ already contains two coloured edges, they must be of the same colour and we are done. Since in P_0 any two K_k's in G_v^* intersect in at most one edge, any K_{k-1} must be contained in a K_k and hence $S(v)$, a K_{k-1}, contains at least one coloured edge, say e. As any colour is used at most twice, there is an edge incident to v that we can colour with the colour of the edge e. Now any four vertices containing this newly coloured edge can contain at most three coloured edges and G is in P_1. If $G_v^* \in P_1, P_2, P_3$ we choose an uncoloured edge in $S(v)$ and, disjoint from the first one, an edge that is incident to v. We colour these two edges with a new colour. On any four vertices not containing v we increase the number of coloured edges by at most one, and any four vertices containing v have at most four coloured edges. Therefore, $G \in P_{j*+1}$.

In the remaining cases, we always advance two stages so G_v^* is at most in stage P_2. Also note that unless G_v^* is in stage P_0, we are allowed to colour two or three disjoint edges: on any four vertices the number of coloured edges can increase by at most two, which is fine with Property 59(iii) as j increases by two. In case that G_v^* is in P_0 we will separately verify that Property 59(iii) holds for $j = 2$ in G, i.e., that any four vertices contain at most three edges.

If $K(v) = Y_\ell$ for $2 \leqslant \ell < k-2$, we have to deal with two K_k's. There exist two disjoint edges incident to $R(v)$ that are contained in both K_k's that we colour with a new colour.

If $K(v) = Y_1$, we have $S(v) = K_k^-$. We have to deal with two K_k's that are created by v and $S(v)$ and which intersect in a K_{k-2}. If G_v^* is in stage P_0, this K_{k-2} contains a triangle and hence an uncoloured edge. We colour this edge and the edge between v and the third vertex in the triangle with a new colour, which ensures that both K_k's are non-rainbow. Now any four vertices that contain v contain at most three coloured edges; also, any other four vertices we only added one coloured edge so the number of coloured edges might have increased from three to four, so G is in P_2. For $G_v^* \in P_1$, no matter how the coloured edges are distributed, only using Property 59(iii), we can always find three disjoint uncoloured edges such that each of the K_k's contains two of them. These three edges we colour with a new colour. Finally, if $G_v^* \in P_2$, it follows from Property 59(iii) that there are two edges not incident to v (but not necessarily disjoint) that hit both K_k's. For both K_k's we can choose an additional edge incident to v such that colouring the two pairs of edges makes both K_5's non-rainbow. This is the only property we have to ensure for P_4, which completes the case that $K(v) = Y_1$.

We now deal with the case that $K(v) = U_1$. Recall that in this case G_v^* is at most in stage P_2 and we advance two stages. It is easy to show that if any four vertices contain at most four coloured edges then five or more vertices contain two disjoint uncoloured edges. So Property 59(iii) implies that together with an edge incident to v we find three disjoint uncoloured edges that we colour with a new colour. Note that for $G_v^* \in P_0$, all four-sets of vertices where we added two edges either contain v or were K_4's in G_v^* already, so they contain at most four coloured edges in G. For $G_v^* \in P_1, P_2$ we observe as before that G is in P_{j^*+2}.

Finally, we have to handle the case that G_v contains more than one K_k-component, so removing $K(v)$ splits into edge-disjoint K_k-components $G_1, \ldots, G_m$. First note that this cannot happen if $K(v)$ is X_{k-2} or U_1, so we only have to deal with the other configurations. We apply the induction hypothesis to these K_k-components and without loss of generality, we may assume that the components are vertex-disjoint in $G_v \smallsetminus K(v)$: intersecting in vertices would yield a denser graph and since all K_k-components can use different colours, combining the partial colourings would still yield a proper colouring at the vertices in which they intersect.

If one of the components, say G_1, is in P_0, we will use the induction hypothesis in a slightly stronger version. By the colouring procedure outlined above for X_{k-2}, which is the only $K(v)$ that can occur if the graph is in P_0, note that we may pick any edge $e \in G_1$ and ensure that it is uncoloured. So for all components G_i that are in P_0 and intersect $S(v)$ in a single edge, we can ensure that this edge is uncoloured.

As any component G_i intersects $K(v)$ in at least one edge we get the following lower bound on $b(G)$.

$$b(G) \geqslant b(K(v)) + \sum_{i=1}^{m} (b(G_i) + r_i),$$

where $r_i \geqslant 0$ accounts for the edges and vertices in the intersection of G_i and $S(v)$ that we would otherwise double-count. It is easy to calculate that r_i is zero if $S(v) \cap G_i$ consists of a single edge and is at least $k-3$ otherwise, because starting with G_i and then adding $K(v)$ would not be X_{k-2}. Since we have

$$b(G) \geqslant \max_{i \in [m]} b(G_i),$$

the properties of the colouring are ensured in all K_k-components. In $K(v)$ we also cannot have any four-set of vertices with too many coloured edges, as any coloured edge in $K(v)$ belongs to a K_k-component G_i with $b(G_i)$ at least $k-3$ or with $r_i \geqslant k-3$, which also contributes at least $k-3$ to $b(G)$. If $K(v)$ is X_ℓ then any four vertices in $K(v)$ contain at most $j+1$ coloured edges. So the same colouring strategy as in the case with one K_k-component can be used to obtain a colouring with no rainbow K_k such that on any four vertices there are at most $j+2$ coloured edges and hence $G \in P_j$. If $K(v)$ is Y_ℓ then $b(K(v)) \geqslant k-3$ so we even get that any four vertices in $K(v)$ contain at most j coloured edges. Again, the same colouring strategy as in the one K_k-component case can be used to show that $G \in P_j$. $\square$

4.2 Complete graph on four vertices

In this section we analyse the anti-Ramsey threshold for K_4, showing that $p_{K_4}^{\mathrm{rb}} = n^{-7/15}$. For the upper bound on $p_{K_4}^{\mathrm{rb}}$, let J be the graph obtained from $K_{3,4}$ with partition classes $\{a, b, c\}$ and $\{w, x, y, z\}$ by adding the edges ab, ac and bc. It is easy to see that in any proper colouring of J there is a rainbow K_4. Therefore the upper bound

$$p_{K_4}^{\mathrm{rb}} \leqslant n^{-7/15} \tag{4.3}$$

follows from Theorem 61 below applied with $H = J$.

Theorem 61 (Bollobás [7]). *Let H be a fixed graph. Then, $p = n^{-1/m(H)}$ is the threshold for the property that G contains H.*

To show that $p_{K_4}^{\mathrm{rb}} \geqslant n^{-7/15}$ we follow a similar strategy as before, but we do not need the framework of [47], because we now have an even smaller upper bound $p \ll n^{-7/15} \ll n^{-2/(4+1)}$.

Let G be a K_4-component with $m(G) < \frac{15}{7}$. Observe that there always is a vertex v of degree 4 in G and that the assertion of Fact 57 still holds. The only options for $K(v)$ are X_1, X_2 and U_1. Theoretically Y_1 and Y_2 would also be possible, but Y_1 could only occur alone and Y_2 is already to dense. We define $b_{K_4}(G) := 7e(G) - 15v(G) + 18$ and note that $b_{K_4}(G) < 18$ and $b_{K_4}(K_4) = 0$. Then

$$b_{K_4}(G) - b_{K_4}(G_v) - 7d^{(v)} = \begin{cases} 6 & \text{if } K(v) \text{ is } X_1, \\ 5 & \text{if } K(v) \text{ is } X_2, \\ 13 & \text{if } K(v) \text{ is } U_1. \end{cases}$$

Thus we can bound the number of occurences of X_1, X_2 and U_1. Configuration X_1 is the only case where G_v could contain more than one K_4-component and there can be at most two different K_4-components, which both have one edge in common with $K(v)$. It is thus easy to see, that any K_4-component G with $m(G) < \frac{15}{7}$ contains at most 10 vertices.

Now consider $G(n,p)$ with $p \ll n^{-7/15}$. It follows from Markov's inequality and the union bound, that $G(n,p)$ asymptotically almost surely does not contain a subgraph G such that $m(G) \geqslant \frac{15}{7}$ and $|V(G)| \leqslant 12$. Therefore $G(n,p)$ asymptotically almost surely does not contain a K_4-component G with $m(G) \geqslant \frac{15}{7}$.

It remains to give the colouring of G depending on the sequence of $K(v)$'s. If $K(v)$ is U_1 then we are left with a single K_4 and it is easy to colour the whole K_5. Now we claim that if $b_{K_4}(G) < 6$ at most one edge is coloured in any K_3 and if $b_{K_4}(G) < 12$ at most two edges are coloured in any K_3. If $K(v)$ is X_2 we repeat the colour of the edge in $K(v) \cup G_v$ if that edge is coloured or otherwise we colour two new disjoint edges with a new colour, which both is fine with the above. Only the case that $K(v)$ is X_1 is left to check. If G_v consists of only one K_4-component, than we colour one edge on the triangle $K(v) \cup G_v$ and a new edge with the same colour. Since X_1 adds 6 to $b_{K_4}(G)$ this is fine with our condition. If G_v splits in more than one K_4-component it is enough to observe that either we can ensure that the intersecting edges are uncoloured or we already have $b_{K_4}(G_v) > 5$ and thus $b_{K_4}(G)$ will be at least 11.

Bibliography

[1] P. Allen, J. Böttcher, J. Ehrenmüller, J. Schnitzer, and A. Taraz, *A spanning bandwidth theorem in random graphs*, in preparation.

[2] P. Allen, J. Böttcher, J. Ehrenmüller, and A. Taraz, *The Bandwidth Theorem in sparse graphs*, available at arXiv:1612.00661.

[3] P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person, *Blow-up lemmas for sparse graphs*, available at arXiv:1612.00622.

[4] J. de O. Bastos, G. O. Mota, M. Schacht, J. Schnitzer, and F. Schulenburg, *Loose Hamiltonian cycles forced by large (k-2)-degree – approximate version*, SIAM J. Discrete Math. **31** (2017), no. 4, 2328–2347, DOI 10.1137/16M1065732.

[5] ———, *Loose Hamiltonian cycles forced by large (k-2)-degree – sharp version*, Contributions to Discrete Mathematics, to appear, available at arXiv:1705.03707.

[6] B. Bollobás, *Modern graph theory*, Graduate Texts in Mathematics, vol. 184, Springer-Verlag, New York, 1998. MR1633290

[7] ———, *Random graphs*, 2nd ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001. MR1864966

[8] B. Bollobás and A. Thomason, *Threshold functions*, Combinatorica **7** (1987), no. 1, 35–38, DOI 10.1007/BF02579198. MR905149

[9] J. A. Bondy and U. S. R. Murty, *Graph theory*, Graduate Texts in Mathematics, vol. 244, Springer, New York, 2008. MR2368647

[10] J. Böttcher, Y. Kohayakawa, and A. Taraz, *Almost spanning subgraphs of random graphs after adversarial edge removal*, Combin. Probab. Comput. **22** (2013), no. 5, 639–683, DOI 10.1017/S0963548313000199. MR3094477

[11] J. Böttcher, M. Schacht, and A. Taraz, *Proof of the bandwidth conjecture of Bollobás and Komlós*, Math. Ann. **343** (2009), no. 1, 175–205, DOI 10.1007/s00208-008-0268-6. MR2448444

[12] E. Buß, H. Hàn, and M. Schacht, *Minimum vertex degree conditions for loose Hamilton cycles in 3-uniform hypergraphs*, J. Combin. Theory Ser. B **103** (2013), no. 6, 658–678, DOI 10.1016/j.jctb.2013.07.004. MR3127586

[13] D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. of Math. (2) **184** (2016), no. 2, 367–454, DOI 10.4007/annals.2016.184.2.2. MR3548529

[14] K. Corradi and A. Hajnal, *On the maximal number of independent circuits in a graph*, Acta Math. Acad. Sci. Hungar. **14** (1963), 423–439, DOI 10.1007/BF01895727. MR0200185

[15] R. Diestel, *Graph theory*, 5th ed., Graduate Texts in Mathematics, vol. 173, Springer, Berlin, 2017. MR3644391

[16] G. A. Dirac, *Some theorems on abstract graphs*, Proc. London Math. Soc. (3) **2** (1952), 69–81, DOI 10.1112/plms/s3-2.1.69. MR0047308

[17] P. Erdős, *Some remarks on the theory of graphs*, Bull. Amer. Math. Soc. **53** (1947), 292–294, DOI 10.1090/S0002-9904-1947-08785-1. MR0019911

[18] ———, *Graph theory and probability*, Canad. J. Math. **11** (1959), 34–38, DOI 10.4153/CJM-1959-003-9. MR0102081

[19] ———, *Problem 9, Theory of Graphs and its Applications*, Proceedings of Symposium on Smolenice 1963, 1964, pp. 159.

[20] P. Erdős and R. Rado, *A combinatorial theorem*, J. London Math. Soc. **25** (1950), 249–255, DOI 10.1112/jlms/s1-25.4.249. MR0037886

[21] P. Erdős and M. Simonovits, *A limit theorem in graph theory*, Studia Sci. Math. Hungar **1** (1966), 51–57. MR0205876

[22] P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091. MR0018807

[23] P. Frankl and Z. Füredi, *Forbidding just one intersection*, J. Combin. Theory Ser. A **39** (1985), no. 2, 160–176, DOI 10.1016/0097-3165(85)90035-4. MR793269

[24] S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger, *Small subsets inherit sparse ε-regularity*, J. Combin. Theory Ser. B **97** (2007), no. 1, 34–56, DOI 10.1016/j.jctb.2006.03.004. MR2278123

[25] A. Hajnal and E. Szemerédi, *Proof of a conjecture of P. Erdős*, Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969), North-Holland, Amsterdam, 1970, pp. 601–623. MR0297607

[26] H. Hàn and M. Schacht, *Dirac-type results for loose Hamilton cycles in uniform hypergraphs*, J. Combin. Theory Ser. B **100** (2010), no. 3, 332–346, DOI 10.1016/j.jctb.2009.10.002. MR2595675

[27] J. Han and Y. Zhao, *Minimum codegree threshold for Hamilton ℓ-cycles in k-uniform hypergraphs*, J. Combin. Theory Ser. A **132** (2015), 194–223, DOI 10.1016/j.jcta.2015.01.004. MR3311344

[28] ———, *Minimum vertex degree threshold for loose Hamilton cycles in 3-uniform hypergraphs*, J. Combin. Theory Ser. B **114** (2015), 70–96, DOI 10.1016/j.jctb.2015.03.007. MR3354291

[29] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. MR1782847

[30] G. Y. Katona and H. A. Kierstead, *Hamiltonian chains in hypergraphs*, J. Graph Theory **30** (1999), no. 3, 205–212, DOI 10.1002/(SICI)1097-0118(199903)30:3<205::AID-JGT5>3.3.CO;2-F. MR1671170

[31] P. Keevash, D. Kühn, R. Mycroft, and D. Osthus, *Loose Hamilton cycles in hypergraphs*, Discrete Math. **311** (2011), no. 7, 544–559, DOI 10.1016/j.disc.2010.11.013. MR2765622

[32] Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230, DOI 10.1007/978-3-642-60539-0_16. MR1661982

[33] Y. Kohayakawa, P. B. Konstadinidis, and G. O. Mota, *On an anti-Ramsey threshold for random graphs*, European J. Combin. **40** (2014), 26–41, DOI 10.1016/j.ejc.2014.02.004. MR3191486

[34] ———, *On an anti-Ramsey threshold for sparse graphs with one triangle*, Journal of Graph Theory, to appear, DOI 10.1002/jgt.22150.

[35] Y. Kohayakawa, G. O. Mota, O. Parczyk, and J. Schnitzer, *Anti-Ramsey thresholds of complete graphs for sparse graphs*, in preparation.

[36] Y. Kohayakawa and V. Rödl, *Szemerédi's regularity lemma and quasi-randomness*, Recent advances in algorithms and combinatorics, CMS Books Math./Ouvrages Math. SMC, vol. 11, Springer, New York, 2003, pp. 289–351, DOI 10.1007/0-387-22444-0_9. MR1952989

[37] J. Komlós, *The blow-up lemma*, Combin. Probab. Comput. **8** (1999), no. 1-2, 161–176, DOI 10.1017/S0963548398003502. Recent trends in combinatorics (Mátraháza, 1995). MR1684627

[38] J. Komlós, G. N. Sárközy, and E. Szemerédi, *Blow-up lemma*, Combinatorica **17** (1997), no. 1, 109–123, DOI 10.1007/BF01196135. MR1466579

[39] ———, *Proof of the Seymour conjecture for large graphs*, Ann. Comb. **2** (1998), no. 1, 43–60, DOI 10.1007/BF01626028. MR1682919

[40] ———, *Proof of the Alon-Yuster conjecture*, Discrete Math. **235** (2001), no. 1-3, 255–269, DOI 10.1016/S0012-365X(00)00279-X. Combinatorics (Prague, 1998). MR1829855

[41] ———, *Spanning trees in dense graphs*, Combin. Probab. Comput. **10** (2001), no. 5, 397–416, DOI 10.1017/S0963548301004849. MR1869048

[42] D. Kühn, R. Mycroft, and D. Osthus, *Hamilton ℓ-cycles in uniform hypergraphs*, J. Combin. Theory Ser. A **117** (2010), no. 7, 910–927, DOI 10.1016/j.jcta.2010.02.010. MR2652102

[43] D. Kühn and D. Osthus, *Spanning triangulations in graphs*, J. Graph Theory **49** (2005), no. 3, 205–233, DOI 10.1002/jgt.20076. MR2145509

[44] ———, *Loose Hamilton cycles in 3-uniform hypergraphs of high minimum degree*, J. Combin. Theory Ser. B **96** (2006), no. 6, 767–821, DOI 10.1016/j.jctb.2006.02.004. MR2274077

[45] D. Kühn, D. Osthus, and A. Taraz, *Large planar subgraphs in dense graphs*, J. Combin. Theory Ser. B **95** (2005), no. 2, 263–282, DOI 10.1016/j.jctb.2005.04.004. MR2171366

[46] C. Lee and B. Sudakov, *Dirac's theorem for random graphs*, Random Structures Algorithms **41** (2012), no. 3, 293–305, DOI 10.1002/rsa.20419. MR2967175

[47] R. Nenadov, Y. Person, N. Škorić, and A. Steger, *An algorithmic framework for obtaining lower bounds for random Ramsey problems*, J. Combin. Theory Ser. B **124** (2017), 1–38, DOI 10.1016/j.jctb.2016.12.007. MR3623165

[48] F. P. Ramsey, *On a Problem of Formal Logic*, Proc. London Math. Soc. (2) **30** (1930), 264–286, DOI 10.1112/plms/s2-30.1.264. MR1576401

[49] C. Reiher, V. Rödl, A. Ruciński, M. Schacht, and E. Szemerédi, *Minimum vertex degree condition for tight Hamiltonian cycles in 3-uniform hypergraphs*, available at arXiv:1611.03118.

[50] V. Rödl and A. Ruciński, *Lower bounds on probability thresholds for Ramsey properties*, Combinatorics, Paul Erdős is eighty, Vol. 1, Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 1993, pp. 317–346. MR1249720

[51] ———, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8** (1995), no. 4, 917–942, DOI 10.2307/2152833. MR1276825

[52] ———, *Dirac-type questions for hypergraphs—a survey (or more problems for Endre to solve)*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 561–590, DOI 10.1007/978-3-642-14444-8_16. MR2815614

[53] V. Rödl, A. Ruciński, and E. Szemerédi, *A Dirac-type theorem for 3-uniform hypergraphs*, Combin. Probab. Comput. **15** (2006), no. 1-2, 229–251, DOI 10.1017/S0963548305007042. MR2195584

[54] ———, *An approximate Dirac-type theorem for k-uniform hypergraphs*, Combinatorica **28** (2008), no. 2, 229–260, DOI 10.1007/s00493-008-2295-z. MR2399020

[55] V. Rödl and Z. Tuza, *Rainbow subgraphs in properly edge-colored graphs*, Random Structures Algorithms **3** (1992), no. 2, 175–182, DOI 10.1002/rsa.3240030207. MR1151360

[56] M. Schacht, *Extremal results for random discrete structures*, Ann. of Math. (2) **184** (2016), no. 2, 333–365, DOI 10.4007/annals.2016.184.2.1. MR3548528

[57] P. Seymour, *Problem section*, Combinatorics: Proceedings of the British Combinatorial Conference 1973, 1974, pp. 201–202.

[58] B. Sudakov and V. H. Vu, *Local resilience of graphs*, Random Structures Algorithms **33** (2008), no. 4, 409–433, DOI 10.1002/rsa.20235. MR2462249

[59] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452 (Hungarian, with German summary). MR0018405

[60] Y. Zhao, *Recent advances on Dirac-type problems for hypergraphs*, Recent trends in combinatorics, IMA Vol. Math. Appl., vol. 159, Springer, 2016, pp. 145–165, DOI 10.1007/978-3-319-24298-9_6. MR3526407

Appendix

Summary

This thesis contains three theorems in graph theory and their proofs. The first result is a optimal $(k-2)$-degree condition for the existence of Hamiltonian cycles in hypergraphs. We describe a well-known extremal example $\mathcal{X}_{k,\ell}$ for $k \geqslant 3$ and $\ell < k/2$, a k-uniform hypergraph which contains no Hamiltonian ℓ-cycle, and prove that any sufficiently large hypergraph $\mathcal{H}$ with $\delta_{k-2}(\mathcal{H}) \geqslant \delta_{k-2}(\mathcal{X}_{k,\ell})$ contains a Hamiltonian ℓ-cycle.

The second result is a transference of the bandwidth theorem to sparse random graphs. For $p \gg \left(\frac{\log n}{n}\right)^{1/\Delta}$, we show that asymptotically almost surely for any subgraph G of $G(n,p)$ with a minimum degree of at least $\left(\frac{k-1}{k} + o(n)\right) pn$ and where each vertex neighbourhood contains at least $\Omega(p^{\binom{s}{2}}(pn)^s)$ copies of K_s the following holds: Let H be a graph on n vertices with $\Delta(H) \leqslant \Delta$, bandwidth at most $\beta^* n$ and suppose that there is a proper k-colouring of $V(H)$ and at least $\Omega(p^{-2})$ vertices in $V(H)$ whose neighbourhood contains only s colours. Then G contains H.

The third result gives the thresholds for $G(n,p)$ to have the rainbow Ramsey property for cliques. An upper bound on the threshold for general graphs was proved before and for cliques on at least 19 vertices the matching lower bound was also known. We prove a matching lower bound on the threshold for all cliques on at least 5 vertices and prove matching lower and upper bounds for K_4.

Zusammenfassung

Diese Dissertation enthält drei graphentheoretische Sätze und deren Beweis. Das erste Ergebnis ist eine optimale $(k-2)$-Gradbedingung für die Existenz von Hamiltonkreisen in Hypergraphen. Wir beschreiben ein bekanntes Beispiel $\mathcal{X}_{k,\ell}$ für $k \geqslant 3$ und $\ell < k/2$, einen k-uniformen Hypergraphen der keinen Hamilton-ℓ-Kreis enthält und beweisen, dass jeder hinreichend große Hypergraph $\mathcal{H}$ mit $\delta_{k-2}(\mathcal{H}) \geqslant \delta_{k-2}(\mathcal{X}_{k,\ell})$ einen Hamilton-ℓ-Kreis enthält.

Das zweite Ergebnis ist eine Übertragung des Bandbreitensatzes auf dünne Zufallsgraphen. Für $p \gg \left(\frac{\log n}{n}\right)^{1/\Delta}$ zeigen wir, dass asymptotisch fast sicher jeder Teilgraph G von $G(n,p)$ mit Minimalgrad mindestens $\left(\frac{k-1}{k} + o(n)\right) pn$ folgendes erfüllt, wenn jede Eckennachbarschaft in G mindestens $\Omega\left(p^{\binom{s}{2}}(pn)^s\right)$ Kopien von K_s enthält: Sei H ein Graph auf n Ecken mit $\Delta(H) \leqslant \Delta$ und Bandbreite höchstens $\beta^* n$ mit einer k-Färbung, so dass für $\Omega(p^{-2})$ Ecken in $V(H)$ die Nachbarschaft höchstens s Farben enthält. Dann enthält G eine Kopie von H.

Das dritte Ergebnis bestimmt den Schwellenwert für $G(n,p)$ für die Regenbogen-Ramsey-Eigenschaft für vollständige Graphen. Eine obere Schranke für alle Graphen war bereits bekannt und für vollständige Graphen auf mindestens 19 Ecken gab es auch eine passende untere Schranke. Wir zeigen eine passende untere Schranke für vollständige Graphen auf mindestens fünf Ecken und beweisen obere und untere Schranken für den K_4.

Publications related to this thesis

Articles

[A1] P. Allen, J. Böttcher, J. Ehrenmüller, J. Schnitzer, and A. Taraz, *A spanning bandwidth theorem in random graphs*, in preparation.

[A2] J. de O. Bastos, G. O. Mota, M. Schacht, J. Schnitzer, and F. Schulenburg, *Loose Hamiltonian cycles forced by large (k-2)-degree – approximate version*, SIAM J. Discrete Math. **31** (2017), no. 4, 2328–2347, DOI 10.1137/16M1065732.

[A3] ______, *Loose Hamiltonian cycles forced by large (k-2)-degree – sharp version*, Contributions to Discrete Mathematics, to appear, available at arXiv:1705.03707.

[A4] Y. Kohayakawa, G. O. Mota, O. Parczyk, and J. Schnitzer, *Anti-Ramsey thresholds of complete graphs for sparse graphs*, in preparation.

Extended Abstracts

[E1] J. de O. Bastos, G. O. Mota, M. Schacht, J. Schnitzer, and F. Schulenburg, *Loose Hamiltonian cycles forced by large (k-2)-degree – approximate version*, Electron. Notes Discrete Math. **54** (2016), 325–330, DOI 10.1016/j.endm.2016.09.056. Discrete Mathematics Days - JMDA16.

[E2] ______, *Loose Hamiltonian cycles forced by large (k-2)-degree – sharp version*, Electron. Notes Discrete Math. **61** (2017), 101–106, DOI 10.1016/j.endm.2017.06.026. The European Conference on Combinatorics, Graph Theory and Applications (EUROCOMB'17).

Declaration on my contributions

Chapter 2 is based on the paper *Loose Hamiltonian cycles forced by large (k-2)-degree – sharp version* [5], which is joint work with Josefran Bastos, Guilherme Mota, Mathias Schacht, and Fabian Schulenburg. We started working on the problem of finding minimum degree conditions for the existence of Hamiltonian cycles when all authors were in Hamburg in 2015. This work led to the approximate result, which we proved in [4] and which appeared in the PhD thesis of Fabian Schulenburg. In this paper we laid the groundwork for proving a sharp bound by proving a theorem that takes the extremal example into account. I drafted most of the proof for the sharp version and we jointly proofread the proof.

Chapter 3 is based on the paper *A spanning bandwidth theorem in random graphs* [1], which is joint work with Peter Allen, Julia Böttcher, Julia Ehrenmüller, and Anusch Taraz. After having worked on a similar problem with my supervisor, Mathias Schacht, I visited the first two co-authors in London in 2016, where we came up with the idea of how to modify the proof in [2] to provide a spanning result. I then drafted and proofread the proof of the *partial embedding lemma* (Lemma 46) with helpful input from Peter Allen and Julia Ehrenmüller and we jointly drafted the remainder of the paper.

Chapter 4 is based on the paper *Anti-Ramsey thresholds of complete graphs for sparse graphs* [35], which is joint work with Yoshiharu Kohayakawa, Olaf Parczyk, and Guilherme Mota. Work on this topic started on a DAAD-funded research visit of Olaf Parczyk and myself to São Paulo in January 2016. We came up with an inital proof strategy for this result during this visit. We subsequently jointly drafted and proofread the proof after the research visit.

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.