Extremal results in random and pseudorandom structures

Pedro Campos Araújo

IMPA

Doutorado

Rio de Janeiro, 19/03/2021

This thesis is dedicated to
the crazy sequence of events that
somehow made me get to this point.
And also to everyone who helped me on the way.

Acknowledgements

The order in which I present my acknowledgments does not imply in any way an order of importance. Each of the contributions and inspirations I got from the people in my life are immeasurable and any sort of classification would be unfair. All the work you see in this theses was only possible because of

Luiz Paulo for waking me up so many days with a tapioca and forcing me to go to IMPA by bike, for suggesting me to take a Combinatorics course, for listening to all my doubts with patience and for sharing so many great moments with me during my time in Rio

Giulia for inspiring me to develop more aspects of my life rather than work

Letícia Mattos and Zoraida for, each in their own way, believing in and supporting my art aspirations

Roberto Viveiros for sharing our entire trajectory through Masters and PhD

Matías for being the right person at the right time and place

Simon for so much affection in my time in Hamburg

David for being in my life while it lasted

My family... For the gift of life, for my raising and all the love that made everything possible in the first place

Guilherme for being ahead of me, setting an example and many paths to be followed by mathematicians

Rodrigo and Martina for living with me and creating the best environment I could ever have wished for

All the people from IMPA... such a great community

To Rob for being the pressure relieve that I needed, for being a professor I aspired to be, for teaching me how to write well and for the infinitely many corrections in this work

To Mathias and all people in Hamburg for the wonderful work experience and for being part of our sitcom

And finally I thank myself for getting through this process. It was not easy...

Abstract

In this thesis we study extremal properties of random and pseudorandom structures. In $G(N,p)$ we focus on the class of bounded degree trees, proving an approximate random analogue of the Erdős-Sós conjecture and apply it to extend to this setting a theorem of Chvátal on Ramsey goodness of trees. In the pseudorandom setting, we focus on 3-uniform hypergraphs and provide asymptotically optimal conditions for Hamiltonicity.

In chapter 2 we prove an approximate version of the Erdős-Sós conjecture is true for $G(n,p)$. We show that for every $D \geqslant 2$ and $\delta, \varrho \in (0,1)$, there exists $C > 0$ such that if $p \geqslant C/N$, then $G = G(N,p)$ with high probability has the following property. Every subgraph $G' \subseteq G$ with $e(G') \geqslant (\varrho + \delta)\, e(G)$ contains every tree with ρN vertices and maximum degree D.

In chapter 3, we study Ramsey Goodness of bounded degrees trees in random graphs. For a graph G, we write $G \to \big(K_{r+1}, \mathcal{T}(n,D)\big)$ if every blue-red colouring of the edges of G contains either a blue copy of K_{r+1}, or a red copy of each tree with n edges and maximum degree at most D. We combine results from the previous chapter with a stability argument and the study of tree containment in expander graphs to prove the following. For each $r, D \geqslant 2$ there exist constants $C, C' > 0$ such that if $p \geqslant Cn^{-2/(r+2)}$ and $N \geqslant rn + C'/p$, then with high probability $G(N,p) \to \big(K_{r+1}, \mathcal{T}(n,D)\big)$.

In chapter 4, we study sufficient conditions for the existence of Hamilton cycles in hypergraphs. We consider 3-uniform hypergraphs $H = (V,E)$ such that for any set of vertices X and for any collection P of pairs of vertices, the number of hyperedges composed by a pair belonging to P and one vertex from X is at least $(1/4 + o(1))|X||P| - o(|V|^3)$ and H has minimum vertex degree at least $\Omega(|V|^2)$. We show that hypergraphs with these properties contain a tight Hamilton cycle. A probabilistic construction shows that the constant $1/4$ is optimal in this context.

Contents

Chapter 1

Introduction

The study of random and pseudorandom objects in Combinatorics has its origin in Ramsey Theory, which is commonly described as the study of finding structure among chaos. Even though there are earlier results with this flavor, for example the theorems of Schur [70] and van der Waerden [77], the area got its name from Ramsey's Theorem, which states that for every k there exists an integer n such that every blue-red coloring of $E(K_n)$ contains a monochromatic copy of K_k. We call the minimum such n the Ramsey number $R(k)$.

A few years later, Erdős and Székeres [29] rediscovered this theorem and proved that $R(k) \leqslant 4^k$. Their proof relies on a simple idea: if a red K_{k-1} is contained in the red neighborhood of any vertex, then we have a red K_k (and analogously for the color blue). Applying induction on the red and blue neighborhoods of a vertex yields the bound. It is worthwhile mentioning that every subsequent improvement on this upper bound was obtained by building up on this same argument.

An exponential lower bound was not proved until twelve years later, due to Erdős [24], but the wait was worthwhile: the idea behind the construction established the Probabilistic Method as a powerful tool for solving problems. By coloring each edge independently at random with probability $1/2$, the expected number of monochromatic copies of K_k is smaller than 1 if the graph has $\sqrt{2}^k$ vertices. Therefore, there must exist a coloring with no monochromatic copies of K_k, which shows that $R(k) \geqslant \sqrt{2}^k$.

The proof of this lower bound is considered to be the first appearance of the concept of the random graph $G(n, p)$, a model which is heavily studied in probabilistic combinatorics. In [26] and [25], Erdős and Rényi studied several properties of the random graph, such as subgraph containment, connectivity and largest component, and they wrote that "The study of evolution of random graphs leads to rather surprising results", referring to the existence of threshold functions for many graph properties.

Building up on the argument of Erdős and Székeres for the upper bound on $R(k)$, Thomason [74] proved in 1988 that

$$R(k) \leqslant k^{-1/2+o(1)}\binom{2k}{k}.$$

Around the same time, Thomason built the foundation of the study of pseudorandomness in Combinatorics. He defined in [73] the so-called *jumbled* graphs, in an attempt to encapsulate the randomness of $G(n, p)$ into a deterministic property. A graph G is said to be (p, β)-jumbled if for every set $U \subset V(G)$ we have that

$$\left| e(U) - p\binom{|U|}{2} \right| \leqslant \beta|U|.$$

Thomason studied several properties of jumbled graphs, such as diameter, connectivity, Hamiltonicity and subgraph containment. Another remarkable advance on the subject is due to Chung, Graham and Wilson [15] who showed that several notions of pseudorandomness are in some sense equivalent. In particular, they showed that if a graph has roughly the same number of copies C_4 as a random graph with the same density, then this property extends to any 'small' graph.

It is believed that extremal colourings of the Ramsey problem are pseudorandom. In 2009, Conlon [18] found a way to extend Thomason's approach by using pseudorandomness in a much more involved way. In particular, he obtained the first super-polynomial improvement on the bound of Erdos and Szekeres. He proved that

$$R(k) \leqslant k^{-C\log k/\log\log k}\binom{2k}{k},$$

for some $C > 0$. In a recent breakthrough, Sah [67] proved a theorem which is the state of the art for this problem. By improving Conlon's pseudorandomness argument to its limit, Sah proved that

$$R(k) \leqslant k^{-C\log k}\binom{2k}{k}.$$

for some $C > 0$. The best construction for the lower bound on $R(k)$ still is the one given by Erdős and this problem remains as one of the most important open problems in Combinatorics.

In this thesis we work with random and pseudorandom structures as we aim to extend to this setting classical results from extremal Combinatorics. For graphs G, H we write $\mathrm{ex}(G, H)$ for the maximum number of edges in an H-free subgraph of G. The celebrated Erdős-Stone Theorem [28] states that

$$\mathrm{ex}(K_n, H) = \left(1 - \frac{1}{\chi(H)-1} + o(1)\right)\binom{n}{2}, \tag{1.1}$$

where $\chi(H)$ is the chromatic number of H. For $H = K_r$, Turán [76] had previously proved that in fact $\mathrm{ex}(K_n, H)$ can only be achieved by a balanced complete $(r-1)$-partite graph.

In 1986, Frankl and Rödl [30] proved that, for $p \geqslant n^{-1/2+\varepsilon}$, with high probability the largest triangle subgraph of $G(n,p)$ has $pn^2/8 + o(pn^2)$ edges. For $p = o(n^{-1/2})$, the number of edges in $G(n,p)$ is much larger than the number of triangles, thus with high probability one can obtain a triangle-free subgraph of $G(n,p)$ with $(1-o(1))p\binom{n}{2}$ edges by removing an edge from each triangle. This same argument extends to a general graph H and motivates the definition of 2-density of H, which is defined by $m_2(H) = \max\left\{\frac{e(H')-1}{v(H')-2} : H' \subseteq H \text{ with } v(H') \geqslant 3\right\}$ and note that $m_2(K_3) = 2$. As in the triangle case, if $p = o(n^{-1/m_2(H)})$, then with high probability there exists a H-free subgraph of $G(n,p)$ with $(1-o(1))p\binom{n}{2}$ edges.

A systematic study of this problem was initiated in the 1990s, Haxell, Kohayakawa and Łuczak [37, 38] and Kohayakawa, Łuczak and Rődl [44]. It was conjectured that the above observation represents the only obstacle for a sparse analogue of (1.1) and they proved that to be true for the case of even cycles, odd cycles and K_4, respectively. More precisely, if H is one of these graphs, there exists a constant $C > 0$ such that if $p \geqslant Cn^{-1/m_2(H)}$, then with high probability $G(n,p)$ satisfies

$$\mathrm{ex}(G(n,p), H) = \left(1 - \frac{1}{\chi(H)-1} + o(1)\right) p\binom{n}{2}. \tag{1.2}$$

The Turán problem for random graphs, as it was called, attracted the interest of several researchers who proved partial results and it remained open for a long time. It was observed in [44] that a sparse variant of the so-called Embedding Lemma for any graph H, together with the Sparse Regularity Lemma proved by Kohayakawa [42], would imply (1.2). As in the dense case, these tools could also be used to prove a sparse version of Erdős-Simonovits Stability Theorem, that is, for $p \gg n^{-1/m_2(H)}$ almost surely every H-free subgraph of $G(n,p)$ with almost $(1 - 1/(\chi(H)))p(\binom{n}{2})$ edges must be very close to $(\chi(H)-1)$-partite. We briefly sketch below how such argument would work for the Turán problem.

Let H be a graph and $p \gg n^{-1/m_2(H)}$. We consider a typical outcome of $G = G(n,p)$ and a subgraph $G' \subset G$ with $(1 - 1/(\chi(H)-1) + \delta)p\binom{n}{2}$ edges. For some $\varepsilon > 0$, we apply the Sparse Regularity Lemma to G', we get a (ε, p)-regular partition (for definitions see Subsection 2.2) of $V(G') = V_0 \cup V_1 \cup \cdots \cup V_k$ with $k = O_\varepsilon(1)$. We define a reduced graph R with $[k]$ as the vertex set and as edges we consider pairs (i,j) such that (V_i, V_j) is an (ε, p)-regular graph with p-density at least δ', for some $\delta' > 0$ small compared to δ. One can show that R has density at least $(1 - 1/(\chi(H)-1) + \delta/2)$ and hence it contains a copy of H, by the Erdős-Stone Theorem. The proof ends if one shows that a copy of H in R ensures a copy of H in G'. In fact, a stronger (and more precise) statement of such an Embedding Lemma was conjectured in [44] and it remained open for 18 years as the KŁR conjecture.

In the 1980s, Frankl and Rödl [31] and Łuczak, Ruciński, and Voigt [52] initiated the study of Ramsey properties of $G(n,p)$. Let us write $G \to H$ to denote that every blue-red colouring of the edges of G contains either a monochromatic copy of H. An important early breakthrough by Rödl and Ruciński [62, 63] established the following threshold result for fixed non-acyclic graphs H:

$$\lim_{n\to\infty} \mathbb{P}\big(G(n,p) \to H\big) = \begin{cases} 1 & \text{if } p \gg n^{-1/m_2(H)}, \\ 0 & \text{if } p \ll n^{-1/m_2(H)}. \end{cases}$$

Sparse analogues are also studied in other discrete structures rather than graphs. For a set $X \subset [n]$ let us write $X \to_\varepsilon k$ for the statement that every subset $Y \subset X$ with $|X| \geqslant \varepsilon|Y|$ contains an arithmetic progression of length k. In 1970, Szemerédi [72] proved that for every integer $k \geqslant 3$ and $\varepsilon > 0$ there exists n_0 such that $[n] \to_\varepsilon k$ for every $n \geqslant n_0$ and the case $k = 3$ was proved earlier by Roth [65]. Ruzsa and Szemerédi [66] found a new proof of Roth's Theorem using the Triangle Remmoval Lemma. Inspired by this idea, Kohayakawa, Łuczak and Rödl [43] proved a sparse Triangle Remmoval Lemma, from which they drew an analogue of Roth's Theorem for the set $[n]_p \subset [n]$, in which each element is present with probability p, independently from each other. More precisely, they proved that for if $p \gg n^{-1/2}$, then for every $\varepsilon > 0$ with high probability $[n]_p \to_\varepsilon 3$. For a similar reason of the Turán problem for random graphs, $n^{-1/2}$ is best possible, since the expected number of 3-APs is of order $o(pn)$ for $p \ll n^{-1/2}$.

Extending classical results to sparse random settings draw the attention of many researchers and culminated in breakthroughs of Conlon and Gowers [20] and of Schacht [69], who developed a general approach for these kind of problems. In particular, their method was used to prove the aforementioned sparse analogues of extremal results, that is, the theorems of Erdős and Stone, of Erdős and Simonovits, of Ramsey and of Szemerédi. More recently, Balogh, Morris and Samotij [6] and Saxton and Thomason [68] introduced the so-called Container Method as a powerful tool to solve these problems. They translate several problems in Extremal Combinatorics can be translated into the study of independent sets in hypergraphs and developed a very general way to group them in a small number of sets of an 'appropriate' size. In particular, as a fairly straightforward consequence of their main theorems they were able to prove the results of Conlon and Gowers and of Schacht and more remarkably the KŁR Conjecture.

Far less is known of pseudorandom graphs. We consider a stronger notion of pseudorandomness than the jumbled graphs of Thomasson. We say that a graph G is (p,β)-*bijumbled* if for every disjoint pair of sets $X, Y \subset V(G)$ we have

$$\big|e(X,Y) - p|X||Y|\big| \leqslant \beta\sqrt{|X||Y|}. \tag{1.3}$$

It is not hard to show that with high probability $G(n,p)$ is $(p,\Theta(\sqrt{pn}))$-jumbled, while a result of Erdős and Spencer [27] states that this is best possible, that is for every (p,β)-bijumbled graph we have $\beta = \Omega(\sqrt{pn})$. A fair question is if $G(n,p)$ and $(p,\sqrt{pn})$-bijumbled graphs share the same threshold for any given graph property. To answer that question, we look at the property of triangle containment. By a First Moment argument, one can prove that with high probability $G(n,p)$ is triangle-free for $p \ll 1/n$ while a Second Moment argument shows that triangles are likely to appear for $p \gg 1/n$. Since the pseudorandom setting is deterministic, such arguments do not have place and triangles appear as long as (1.3) guarantees the existence of and edge in the neighborhood of a vertex with average degree. For $(p,\sqrt{pn})$-bijumbled graphs, this correspond to p being $\Omega(n^{-1/3})$ and a construction of Alon [4] shows that this is best possible up to a constant factor.

In the same way as in random graphs, extremal results to the pseudorandom setting can be proved by the sparse regularity lemma and an appropriate embedding lemma. By building on that idea, Kohayakawa, Rödl, Schacht and Skokan proved triangle removal lemma for bijumbled graphs. More precisely, they proved that fore every $\delta > 0$ there exists $\varepsilon, \gamma > 0$ such that every $(p,\gamma p^3 n)$-bijumbled n-vertex graph with $p \geqslant n^{-1/2}$ satisfies the following. Every graph with at most εpn^3 triangles can be made triangle-free by the removal of at most δpn^2 edges. As in the dense and random cases, this also implies Roth's Theorem for pseudorandom sets of $[n]$. They also conjectured that both the embedding and triangle removal lemmas are true for $(p,\gamma p^2 n)$-bijumbled graphs.

An important question in the area of pseudorandom graphs is whether any regular subgraph of a (p,β)-bijumbled graph contains a copy of a graph H whenever the values of p and β guarantees a copy of H in any bijumbled graph with same parameters. This question remains open even for triangles. In fact, optimal values of p and β that ensure a copy of a graph H in (p,β)-bijumbled graphs is not known. An important advance in answering this question is the breakthrough of Conlon, Fox and Zhao [19] who proved an embedding lemma for any graph H for (p,β)-bijumbled graphs, for certain values of p and β. However, they do not know whether their results are optimal, even for the case of H being a clique. For many graphs H, their results were improved later by Allen, Bötcher, Skokan and Stein [3].

It is important to remark that even without an embedding lemma, it is possible to prove extremal results in the pseudorandom setting. Sudakov, Szabó and Vu [71] proved that $\mathrm{ex}(G, K_{r+1}) = (1 - 1/r + o(1))e(G)$ for any $(p,\gamma p^r n)$-bijumbled graph G, which is known to be optimal for triangles. More recently, Berger, Lee and Schacht [7] proved optimal results for the Turán problem for odd cycles in (n,d,λ)-graphs, a different notion of pseudorandomness.

In the next three subsections we talk about our contributions to the research in the context of random and pseudorandom graphs.

1.1 Sparse Erdős-Sós Conjecture for bounded degree trees

Resilience is a measure of how much one has to perturb a graph in order to destroy a given property of it (see e.g. [11] for a discussion of resilience in the random graph) and it is a convenient way of phrasing extremal problems in general settings. For example, in the context of bounded degree trees, a classical result of Komlós, Sárközy and Szemerédi [47] says that every n-vertex graph G with $\delta(G) \geqslant (1/2+o(1))n$ contains all trees in $\mathcal{T}(n, D)$, for n large enough. In other words, one can say that if an adversary deletes a $(1/2-o(1))$-proportion of the edges incident at each vertex of K_n, the resulting graph still contains all trees in $\mathcal{T}(n, D)$. Balogh, Csaba and Samotij [5] proved that the same happens a.a.s. in the random graph for the class of almost spanning trees with bounded degree, provided that $p \gg 1/n$. That is, any subgraph of $G(n, p)$ obtained by deleting at most a $(1/2 - o(1))$-proportion of the edges incident to each vertex of $G(n, p)$ contains all trees in $\mathcal{T}((1 - o(1))n, D)$ with high probability.

In [5], the authors developed tools for embedding trees in "well-behaved" sparse bipartite graphs. We combine these tools with the approach of Besomi, Pavez-Signé and Stein [9] to the Erdős–Sós Conjecture[1], for bounded degree trees and dense host graphs, to obtain the following "global" resilience result.

Theorem 1.1.1. *For every $D \geqslant 2$ and $\delta, \varrho \in (0, 1)$, there exists $C > 0$ such that if $p \geqslant C/N$, then $G = G(N, p)$, with high probability, has the following property. Every subgraph $G' \subseteq G$ with $e(G') \geqslant (\varrho + \delta)\, e(G)$ is $\mathcal{T}(\varrho N, D)$-universal.*

Theorem 1.1.1 will follow by a stronger result, in which $G(N, p)$ can be replaced by a pseudorandom graph. More precisely, we ask that the number of edges between any disjoint pair of sets is roughly what one would expect in $G(N, p)$.

In terms of resilience, Theorem 1.1.1 says that if $pN \gg 1$, then a.a.s. one can delete a $(1 - \varrho - o(1))$-proportion of the edges of $G(N, p)$ so that the resulting graph still contains all trees in $\mathcal{T}(\varrho N, D)$. This result can be viewed as an approximate random analogue of the Erdős–Sós conjecture for bounded degree trees of linear size. We point out that Theorem 1.1.1 is sharp in the following senses: the value of p is best possible, up to a constant factor, since the largest connected component of $G(N, p)$ is sublinear when

[1]The Erdős–Sós Conjecture [23] from 1964 says that, given $k \in \mathbb{N}$, every graph G with average degree greater than k contains all trees with $k + 1$ edges. In particular, it says that if $k = \varrho n$ for some $n \in \mathbb{N}$, then every graph G with $e(G) > \varrho n^2/2 \approx \varrho\binom{n}{2}$ edges contains each tree with $\varrho n + 1$ edges.

$p \ll 1/N$. Moreover, for an integer $r \geqslant 2$ and $\varrho = 1/r$, the constant ϱ cannot be improved. Indeed, one can partition the vertex set in $r+1$ parts, one with at most r vertices and the others in a balanced way and thus with fewer than N/r vertices. If the edges between parts are deleted, then a.a.s. we get a subgraph $G' \subseteq G(N,p)$ which has $(1/r - o(1))e(G(N,p))$ edges but every connected component of G' has less than N/r vertices.

1.2 Ramsey goodness of trees in random graphs

Ramsey properties of random graphs involving sparse graphs have attracted significant attention also in recent years. To give just two examples, Letzter [51] proved that if $\varepsilon > 0$ and $pn \to \infty$, then

$$G\big((3/2+\varepsilon)n, p\big) \to P_n$$

with high probability (the constant $3/2$ is best possible), and Kohayakawa, Mota and Schacht [45] proved that $\big(\frac{\log N}{N}\big)^{1/2}$ is the threshold for the event that for any two-colouring of the edges of $G(N,p)$, there exist two monochromatic trees that partition the vertex set.

In this paper we will be interested in the problem of extending to the setting of sparse random graphs a theorem of Chvátal [16] from 1977, which states that if $r \in \mathbb{N}$, and T is a tree with n edges, then

$$K_N \to (K_{r+1}, T) \qquad \Leftrightarrow \qquad N \geqslant rn+1. \tag{1.4}$$

The necessity of the lower bound on N is easy to see, and (as was first observed by Burr [12]) holds in significantly greater generality. To be precise, if H is a connected graph, F is a graph with $\sigma(F) \leqslant |H|$, where $\sigma(F)$ is the minimum size of a colour class in a proper $\chi(F)$-colouring of F, and $N < \big(\chi(F)-1\big)\big(|H|-1\big) + \sigma(F)$, then

$$K_N \not\to (F,H).$$

Indeed, to see this it suffices to consider $\chi(F)-1$ disjoint red cliques of size $|H|-1$, and one additional disjoint red clique of size $\sigma(F)-1$. A (connected) graph H is said to be *Ramsey F-good* (or just *F-good*) if $K_N \to (F,H)$ whenever $N \geqslant \big(\chi(F)-1\big)\big(|H|-1\big) + \sigma(F)$. The systematic study of Ramsey goodness was initiated by Burr and Erdős [13] in 1983.

As far as we are aware, the problem of Ramsey goodness in random graphs was first studied only very recently, by Moreira [54], who considered the case in which F is a clique and H is a path. The main results of [54] identified two different thresholds for the event that $G(N,p) \to (K_{r+1}, P_n)$, for different values of N. More precisely, it was proved there that if $p \gg n^{-2/(r+2)}$ and $t \gg 1/p$, then

$$G\big(rn+t,p\big) \to \big(K_{r+1}, P_n\big),$$

while if $p \gg n^{-2/(r+1)}$ and $t = \Omega(n)$, then

$$G\big(rn+t,p\big) \to \big(K_{r+1}, P_n\big)$$

in both cases with high probability as $n \to \infty$. These results are sharp in the sense that whp $G(rn+t,p) \not\to (K_{r+1}, P_n)$ in three different settings. First, if $p \in (0,1)$ and $t \ll 1/p$, then one can partition $V(G(N,p)) = V_0 \cup V_1 \cup \cdots \cup V_r$ such that $|V_0| = t$ and $e(V_0, V_r) = 0$. This is possible since, with high probability, sets of size $o(1/p)$ have $o(n)$ external neighbours in $G(N,p)$. Then we can colour the edges in red if and only if they have both endpoints inside parts without creating a blue K_{r+1} or any red component with more than n vertices. Second, for $n^{-2/(r+1)} \ll p \ll n^{-2/(r+2)}$, one can show that there are values of $t \gg 1/p$ such that $G\big(rn+t,p\big) \nrightarrow \big(K_{r+1}, P_n\big)$. Finally, if $p \ll n^{-2/(r+1)}$ and $t = O(n)$, then, with high probability, $G(N,p)$ has $o(n)$ copies of K_{r+1}, whose edges can be all coloured in red without creating any red copy of P_n, see [54] for the details.

Our main theorems generalise the results of [54] from paths to arbitrary bounded degree trees. Let us denote by $\mathcal{T}(n,D)$ the class of all trees with n edges and maximum degree at most D. Let us write $G \to (K_{r+1}, \mathcal{T}(n,D))$ to denote that $G \to (K_{r+1}, T)$ for every $T \in \mathcal{T}(n,D)$.

Theorem 1.2.1. *For each $r, D \geqslant 2$, there exist $C, C' > 0$ such that the following holds. If*

$$p \geqslant C' N^{-2/(r+2)} \qquad \text{and} \qquad N \geqslant rn + C/p,$$

then $G(N,p) \to \big(K_{r+1}, \mathcal{T}(n,D)\big)$ with high probability as $n \to \infty$.

As mentioned above, it follows from the results of [54] that the bound on N is sharp up to the value of C, and the bound on p is sharp up to a the value of C'. For smaller values of p we obtain the following bound.

Theorem 1.2.2. *For every $r, D \geqslant 2$ and $\varepsilon > 0$ there exists $C' > 0$ such that the following holds. If*

$$p \geqslant C' N^{-2/(r+1)} \qquad \text{and} \qquad N \geqslant rn + \varepsilon n,$$

then $G(N,p) \to \big(K_{r+1}, \mathcal{T}(n,D)\big)$ with high probability as $n \to \infty$.

We will prove Theorem 1.2.2 by iteratively applying a theorem due to Haxell [36] to find either red copies of every tree in $\mathcal{T}(n, D)$, or $r+1$ large disjoint sets with only blue edges between them. The result will then follow by a straightforward application of the Janson inequalities. The proof of Theorem 1.2.1 is significantly more challenging, and is based on a stability argument. One of the key steps is to prove that the random graph not only contains all large bounded degree trees, but is also resilient with respect to this property.

1.3 Tight Hamiltonian cycles in pseudorandom hypergraphs

Dirac's theorem states that any graph on $n \geqslant 3$ vertices and minimum degree at least $n/2$ contains a Hamilton cycle. This is best possible in terms of minimum degree, since a graph composed by two disjoint cliques of sizes $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$ is not even connected. We investigate what kind of properties ensure the existence of Hamilton cycles in 3-uniform hypergraphs.

Since we restrict our attention to 3-uniform hypergraphs, if not mentioned otherwise, by a hypergraph we will mean a 3-uniform hypergraph. We denote an edge $\{u, v, w\} \in E(H)$ by uvw. An ordered set of distinct vertices $(v_1, v_2, \ldots, v_\ell)$ forms a *tight path* of length $\ell - 2$ if every three consecutive vertices form an edge. The pairs (v_1, v_2) and $(v_{\ell-1}, v_\ell)$ are the *starting pair* and the *ending pair* of the path, and we frequently call such a tight path a (v_1, v_2)-$(v_{\ell-1}, v_\ell)$-path. For simplicity we denote a tight path by listing its vertices. A tight path $v_1 v_2 \ldots v_\ell$ together with the edges $v_{\ell-1} v_\ell v_1$ and $v_\ell v_1 v_2$ forms a *tight cycle* of length ℓ. A tight cycle which covers all vertices of the hypergraph will be called *tight Hamilton cycle*. Similarly, a *loose Hamilton cycle* in an n-vertex hypergraph (with $n \geqslant 6$ even) is a cyclicly ordered collection of $n/2$ edges in such a way that two edges intersect if and only if they are consecutive and, consequently, they intersect in exactly one vertex.

There are more than one notion of degrees in hypergraphs. Given a hypergraph H and $v \in V(H)$, we define the *neighbourhood* and the *degree* of v by

$$N_H(v) = \{e \smallsetminus \{v\} \colon v \in e \in E(H)\} \quad \text{ and } \quad d_H(u) = |N(u)|,$$

respectively. Similarly, for $u, v \in V(H)$, we also define their *neighbourhood* and their *codegree* by

$$N_H(u, v) = \{w \in V(H) \colon \{u, v, w\} \in E(H)\} \quad \text{ and } \quad d_H(u, v) = |N(u, v)|.$$

Let $\delta_1(H)$ be the *minimum degree* and $\delta_2(H)$ the *minimum codegree* of H.

A possible extension of Dirac's theorem for hypergraphs was proposed in [41]. The optimal minimum degree and codegree conditions were obtained for *loose* Hamilton cycles [14, 49] and for *tight* Hamilton cycles [59, 64]. As the extremal examples for Dirac's theorem for graphs, the constructions that show optimality for those results have a very rigid structure. In the graph case, for instance, the extremal constructions contain large pairs of sets of vertices with no edges between them.

Motivated by this, we say an n-vertex graph G is (ρ, d)-dense if for every pair of vertex sets, X and Y, the number of edges between them is at least $d|X||Y| - \rho n^2$. Using a result from Chvátal and Erdős [17], it is not hard to prove that for every $\alpha, d > 0$ there is an $\rho > 0$ for which every sufficiently large (ρ, d)-dense n-vertex graph with minimum degree at least αn contains a Hamilton cycle. Note that the minimum degree condition can not be dropped, as this notion of (ρ, d)-density does not prevent the graph from having isolated vertices.

There are several ways to extend the notion of (ρ, d)-density to 3-uniform hypergraphs. Here we consider the following three notions that we symbolise by $\therefore$, $\dot{-}$, and $\curlywedge$ (see also [2, 58, 60, 61]).

Definition 1.3.1. *Let ρ, $d \in (0, 1]$ and let H be a 3-uniform hypergraph on n vertices.*

We say that H is $(\rho, d, \therefore)$-dense *if for every three sets of vertices X, Y, Z we have*

$$e(X, Y, Z) = |\{(x, y, z) \in X \times Y \times Z \colon \{x, y, z\} \in E(H)\}| \geqslant d|X||Y||Z| - \rho n^3.$$

We say that H is $(\rho, d, \dot{-})$-dense *if for every set of vertices X and every collection of pairs of vertices $P \subseteq V \times V$ we have*

$$e(X, P) = |\{(x, (y, z)) \in X \times P \colon \{x, y, z\} \in E(H)\}| \geqslant d|X||P| - \rho n^3.$$

We say that H is $(\rho, \varepsilon, \curlywedge)$-dense *if for every two collections of pairs of vertices $P, Q \subseteq V \times V$ we have*

$$e(P, Q) = |\{((x, y), (y, z)) \in P \times Q \colon \{x, y, z\} \in E(H)\}| \geqslant d|K_{\curlywedge}(Q, P)| - \rho n^3,$$

where $K_{\curlywedge}(Q, P) = \{((x, y), (y, z)) \in P \times Q\}$.

Observe that $\therefore$ is the weakest notion and $\curlywedge$ is the strongest (see [60] for details). Our main result concerns $\dot{-}$-dense hypergraphs. We consider this notion as a *localised codegree condition* since it implies that for every linear sized set X most pairs of vertices will have the same proportion of neighbours in X as in the whole hypergraph.

We are interested in (asymptotically) optimal assumptions for $\dot{-}$-dense hypergraphs to ensure Hamilton cycles. This line of research can be traced back to the work of

Lenz, Mubayi and Mycroft [50], who proved that for arbitrarily small $d, \alpha > 0$ there is an $\rho > 0$ such that every sufficiently large $(\rho, d, \therefore)$-dense n-vertex hypergraph with minimum degree αn^2 contains a loose Hamilton cycle (in fact they proved this result for r-uniform hypergraphs for $r \geqslant 2$). As this density condition is the weakest one, this theorem implies the same result for the stronger notions $\underline{\therefore}$ and $\wedge$.

Aigner-Horev and Levy [2] proved the same conclusion for tight cycles, but considering minimum codegree conditions instead of vertex degrees and assuming the strongest density notion $\wedge$. More precisely, they proved that for every $d, \alpha > 0$ there is a $\rho > 0$ such that every sufficiently large $(\rho, d, \wedge)$-dense hypergraph with minimum codegree αn contains a tight Hamilton cycle. It turns out that for the $\underline{\therefore}$-density an analogous result is not possible due the following counterexample.

Example 1.3.2. Let G be a random graph $G_{n-2,1/2}$ and define a 3-uniform hypergraph on the same set of vertices for which a triple of vertices is a hyperedge, if it forms a triangle in G or in $\overline{G}$. Observe that every tight cycle in H can only use edges, all of which induce triangles in G or they induce only triangles en $\overline{G}$. Finally, add two new vertices x, y in such a way that $N_H(x) = E(G)$ and $N_H(y) = E(\overline{G})$. Then x is covered only by cycles induced by triangles in G and y is covered only by cycles induced by triangles in $\overline{G}$. Hence H contains no tight Hamilton cycle. Obviously, adding all the edges containing the pair $\{x, y\}$, the hypergraph H only yields a tight Hamilton path, but not a tight Hamilton cycle. One can show for every that $\rho > 0$ with high probability H is $(\rho, 1/4, \underline{\therefore})$-dense and it has minimum degree $(1/4 - \rho)\binom{n}{2}$ and even minimum codegree $(1/4 - \rho)n$.

Our main result asserts that the previous example is essentially best possible.

Theorem 1.3.3. *For every $\varepsilon > 0$ there exist $\rho > 0$ and n_0 such that every $(\rho, 1/4+\varepsilon, \underline{\therefore})$-dense 3-uniform hypergraph H on $n \geqslant n_0$ vertices with $\delta_1(H) \geqslant \varepsilon\binom{n}{2}$ contains a tight Hamilton cycle.*

We also strengthen a result of Aigner-Horev and Levy [2] by showing that their codegree assumption for tight Hamilton cycles in $\wedge$-dense hypergraphs can be relaxed to a minimum vertex degree assumption.

Theorem 1.3.4. *For every d, $\alpha > 0$ there exist $\rho > 0$ and n_0 such that every $(\rho, d, \wedge)$-dense 3-uniform hypergraph H on $n \geqslant n_0$ vertices with $\delta_1(H) \geqslant \alpha\binom{n}{2}$ contains a tight Hamilton cycle.*

Theorem 1.3.4 was conjectured in [2] and was obtained independently in [33]. The main purpose result of this section is Theorem 1.3.3. The proof of Theorem 1.3.4 is based on similar ideas and we discuss the details in Section 4.6.

Chapter 2

Sparse Erdős-Sós Conjecture for bounded degree trees

This chapter is devoted to prove the global resilience of trees of linear size and bounded maximum degree in $G(N, p)$. Actually, we will prove the following stronger result.

Theorem 2.0.1. *Let $\delta, \varrho \in (0, 1)$ and $D \geqslant 2$. There are positive constants n_0, η_0 and C_0 such that for all $\eta \leqslant \eta_0$ and $n \geqslant n_0$ the following holds. Let G be a (η, p)-uniform graph on n vertices and let $p \in [0, 1]$ with $pn \geqslant C_0$. Then every subgraph $G' \subseteq G$ with $e(G') \geqslant (\varrho + \delta)\, e(G)$ is $\mathcal{T}(\varrho n, D)$-universal.*

It turns out that Theorem 1.1.1 easily follows from Theorem 2.0.1. Indeed, given $\delta, \varrho \in (0, 1)$ and $D \geqslant 2$, by Lemma 3.3.3 we know that $G(N, p)$ is, with high probability, (η_0, p)-uniform for $p \geqslant C/N$ and therefore, by Theorem 2.0.1, any subgraph $G' \subseteq G(N, p)$ with $e(G') \geqslant (\varrho + \delta)e(G(N, p))$ is $\mathcal{T}(\varrho N, D)$-universal.

2.1 Overview

Let G be an (η, p)-uniform graph and let $G' \subseteq G$ be a subgraph of G such that $e(G') \geqslant (\varrho + \delta)e(G)$. Since we obtained G' by removing edges from G, it is clear that G' is (η, p)-upper uniform, and therefore, by the regularity lemma (Theorem 2.2.2), we know that $V(G')$ admits an (ε, p)-regular partition. We will work on the reduced graph R of G' in order to find a good structure into which any given bounded degree tree can be embedded. Let k be the number of vertices of R. As usual in the arguments envolving regularity, we show that R inherits some property of G', in this case the edge density. More precisely, we show that the average degree of R is at least $(\varrho + \delta/3)k$ and thus we can find a subgraph $R' \subseteq R$ such that $d(R') \geqslant (\varrho + \delta/3)k$ and that $\delta(R) \geqslant (\varrho + \delta/3)k/2$. Let $X \in V(R)$ be a vertex of degree at least the average. Note that $N(X)$ is larger than the size of the tree (scaled by the size of the clusters), and so our plan will be to use the neighbourhood of X to embed every tree in $\mathcal{T}(\varrho n, D)$. To do that we partition $N(X)$ into a maximal matching $\mathcal{M}$ and an independent set $\mathcal{Y}$. If we denote by $\mathcal{H}$ the bipartite graph induced by $\mathcal{Y}$ and $\mathcal{Z} = N(\mathcal{Y}) \setminus (X \cup N(X))$, then by the minimum degree of R we can prove that $\mathcal{Y}$ has large minimum degree in $\mathcal{H}$, as long as $\mathcal{M}$ is not larger than $(\varrho + \delta/16)k$.

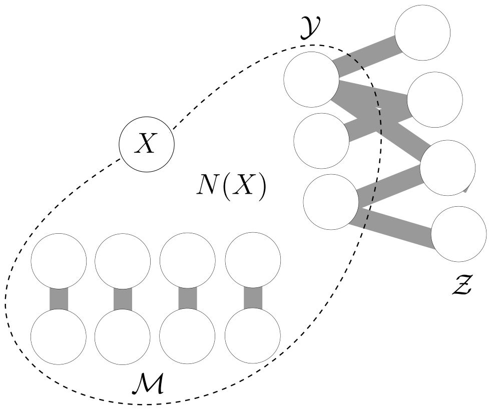

Figure 2.1: Structure in the reduced graph

Given a tree $T \in \mathcal{T}(\varrho n, D)$, our goal is to embed T using the structure that we have found in the neighbour of X. To do so, we first need to cut the tree into very small subtrees and then locate every such subtree into some edge of the reduced graph. If $\mathcal{M}$ is large enough, then we will locate each subtree into an edge of the matching, using both clusters of the edge in a balanced way. Otherwise, we will first locate subtrees into edges from $\mathcal{H}$, until a large proportion of $\mathcal{Y} \cup \mathcal{Z}$ is used. The leftover subtrees can be located into $\mathcal{M}$, always using both clusters from each edge in a balanced way. In any case, once we have located the subtrees, we will use an embedding technique due to Balogh, Csaba and Samotij [5], in order to embed each of this subtrees into the (ε, p)-regular pair that was assigned to this subtree. The role of X here is to connect the embedding, meaning that X will be used in order to go from one edge to another in $\mathcal{M} \cup \mathcal{H}$.

2.2 Sparse regularity

The proof of Theorem 1.1.1 relies on a sparse version of the Szemerédi Regularity lemma. In order to state this result we need some basic definitions. Let G be a graph and let $p \in (0, 1)$. Given two disjoint sets $A, B \subseteq V(G)$, we define the p-density of the pair (A, B) by

$$d_p(A, B) = \frac{e(A, B)}{p|A||B|}.$$

Given $\varepsilon > 0$, we say that (A, B) is (ε, p)-regular if for all $A' \subseteq A$ and $B' \subseteq B$, with $|A'| \geqslant \varepsilon|A|$ and $|B'| \geqslant \varepsilon|B|$, we have

$$|d_p(A', B') - d_p(A, B)| \leqslant \varepsilon.$$

Now we state some standard results regarding properties of regular pairs (we refer to the survey [34] for the proofs).

Lemma 2.2.1. *Given $\alpha > \varepsilon > 0$, let G be a graph and let $A, B \subseteq V(G)$ be disjoint sets such that (A, B) is (ε, p)-regular with $d_p(A, B) = d > 0$. Then the following are true.*

1. *For any $A' \subseteq A$ with $|A'| \geqslant \alpha|A|$ and $B' \subseteq B$ with $|B'| \geqslant \alpha|B|$, the pair (A', B') is $(\varepsilon/\alpha, p)$-regular with p-density at least $d - \varepsilon$.*

2. *There are at most $\varepsilon|A|$ vertices in A with less then $(d - \varepsilon)p|B|$ neighbours in B.*

A partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_k$ is said to be (ε, p)-regular if

1. $|V_0| \leqslant \varepsilon|V(G)|$,

2. $|V_i| = |V_j|$ for all $i, j \in [k]$, and

3. all but at most εk^2 pairs (V_i, V_j) are (ε, p)-regular.

We may now state a sparse version of Szemerédi's regularity lemma, due to Kohayakawa and Rödl [42, 46] .

Theorem 2.2.2. *Given $\varepsilon > 0$ and $k_0 \in \mathbb{N}$, there are $\eta > 0$ and $K_0 \geqslant k_0$ such that the following holds. Let G be an η-upper-uniform graph on $n \geqslant k_0$ vertices and let $p \in (0, 1)$, then G admits an (ε, p)-regular partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_k$ with $k_0 \leqslant k \leqslant K_0$.*

Let G be a graph that admits an (ε, p)-regular partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_k$. Let $d \in (0, 1)$. The (ε, p, d)-reduced graph R, with respect to this (ε, p)-regular partition of G, is the graph with vertex set $V(R) = \{V_i : i \in [k]\}$, called clusters, such that V_iV_j is an edge if and only if (V_i, V_j) is an (ε, p)-regular pair with $d_p(V_i, V_j) \geqslant d$. Next proposition establishes that the edge density of R is roughly the same as in G. Since its proof is fairly standard in the applications of the Regularity Lemma, we omit it.

Proposition 2.2.3. *Let $\varepsilon, \eta, p, d \in (0, 1)$ and let $k \in \mathbb{N}$ such that $k \geqslant 1/\varepsilon$. Let G be an (η, p)-upper uniform graph on n vertices that admits an (ε, p)-regular partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_k$, and let R be the (ε, p, d)-reduced graph of G with respect to this partition. Then*

$$e(R) \geqslant \frac{e(G)}{(1+\eta)p}\left(\frac{k}{n}\right)^2 - \frac{6\varepsilon + d}{1+\eta}k^2.$$

2.3 Cutting up a tree

Now we show how to cut a given tree T into a constant number of tiny rooted subtrees, such that the root of each of this subtrees is at even distance from the root of T. The following lemma, proved by Balogh, Csaba and Samotij [5], gives a partition of the tree into a constant number of subtrees such that each subtree has few vertices and is adjacent to a bounded number of others subtrees.

Lemma 2.3.1. *Let $D \geqslant 2$ and let (T, r) be a rooted tree with maximum degree at most D. If $\beta \geqslant 1/|V(T)|$, then there exists a family of $t \leqslant 4/\beta$ disjoint rooted subtrees $(T_i, r_i)_{i\in[t]}$ such that $V(T) = V(T_1) \cup \cdots \cup V(T_t)$ and for each $i \in [t]$ we have*

1. *$|V(T_i)| \leqslant D^2\beta|V(T)|$,*
2. *T_i is connected (by an edge) to at most D^3 others subtrees, and*
3. *T_i is rooted at r_i and all the children of r_i belong to T_i.*

Given a tree T, let $(T_i, r_i)_{i\in[t]}$ be the family given by Lemma 2.3.1. We may define an auxiliary graph T_Π, called *cluster tree*, with vertex set $V(T_\Pi) = [t]$ and edge set

$$E(T_\Pi) = \{ij \mid \ T_i \text{ and } T_j \text{ are adjacent in } T\}.$$

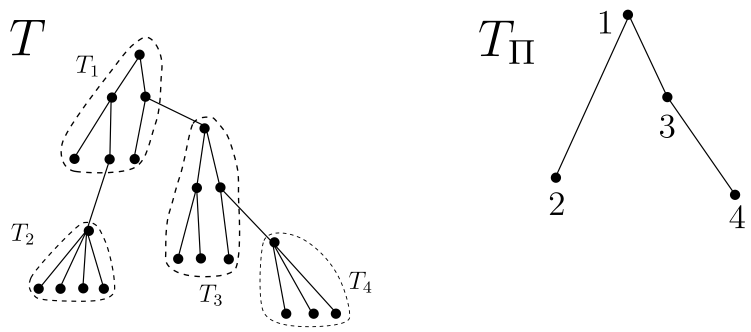

Figure 2.2: Cluster tree

Now we need to refine the partition given by Lemma 2.3.1 in order to impose that the root of each subtree is at even distance from the root of T.

Proposition 2.3.2. *Let $D \geqslant 2$ and let (T, r) be a rooted tree with maximum degree at most D. If $\beta \geqslant 1/|V(T)|$, then there exists a family of $t \leqslant 4D/\beta$ disjoint rooted subtrees $(T_i, r_i)_{i\in[t]}$ such that $V(T) = V(T_1) \cup \cdots \cup V(T_t)$ and for each $i \in [t]$ we have*

1. $|V(T_i)| \leqslant D^4\beta|V(T)|$*,*

2. T_i *is rooted at* r_i *and the distance from* r_i *to* r *is even,*

3. all the children of r_i *belong to* T_i*, and*

4. the corresponding cluster tree has maximum degree at most D^4*.*

Proof. Starting with the partition given by Lemma 2.3.1, we will refine this partition as we run a breadth first search on (T, r). Suppose that in this search we have reached a vertex v, which is the root of a subtree in the current partition, such that v and all roots before v are at even distance from each other in the current partition.

If there is a root u of some subtree in the current partition, which is at odd distance from v and such that the subtree pending from v is adjacent to u, then we may update the partition by splitting the tree pending from u (each neighbour of u is now the root of a subtree) and adding u to the subtree pending from v. Note that after this splitting, the root of each tree that is adjacent to the tree pending from v is at even distance from all the previous roots.

At the end of this process, each subtree of the original partition is split into at most D parts and hence we end up with at most $4D/\beta$ rooted subtrees. For the same reason, the maximum degree of the cluster tree cannot go higher than D^4. Moreover, the size of each subtree grows by at most D^3 when the roots are added, so at the end of the process each subtree has size at most $D^2\beta|V(T)| + D^3 \leqslant D^4\beta|V(T)|$. □

2.4 Structure in the reduced graph

In this section, we will follow a strategy inspired in the approach of Besomi, Stein and Pavez-Signé [9] to the Erdős–Sós conjecture for bounded degree trees and dense host graphs. We will prove that if H is an (η, p)-upper-uniform graph with $2e(H) \geqslant (\varrho + \delta/2)pn^2$, then H has an (ε, p, d)-reduced graph R with a useful substructure. That is, R contains a cluster X of large degree such that its neighbourhood can be partitioned as $N(X) = V(\mathcal{M}) \cup \mathcal{Y}$, where $\mathcal{M}$ is a matching and $\mathcal{Y}$ is an independent set. Moreover, if $\mathcal{H}$ denotes the bipartite graph induced by $\mathcal{Y}$ and $\mathcal{Z} = N(\mathcal{Y}) \setminus (X \cup N(X))$, then every cluster in $\mathcal{Y}$ has large degree in $\mathcal{H}$.

We need the following lemma (see [8] for a proof).

Lemma 2.4.1. *Given a graph* F*, there exists an independent set* I*, a matching* M *and a family of triangles* Γ*, such that* $V(F) = I \cup V(M) \cup V(\Gamma)$*. Moreover, we may write* $V(M) = M_1 \cup M_2$*, where each edge* $e \in M$ *is of the form* $e = v_1v_2$ *with* $v_i \in M_i$ *for* $i \in \{1, 2\}$*, so that* $N(I) \subseteq M_1$*.*

Proposition 2.4.2. *Let $\varepsilon, \delta, \varrho \in (0,1)$ and let $d = \delta/100$. There exist $n_0, K_0 \in \mathbb{N}$ and $n_0 > 0$ such that for all $0 < \eta \leqslant \eta_0$, $p \in (0,1)$ and $n \geqslant n_0$, the following holds. Let H be an (η, p)-upper uniform graph on n vertices such that $2e(H) \geqslant (\varrho + \delta/2)pn^2$. Then H admits an (ε, p)-regular partition $V(H) = V_0 \cup V_1 \cup \cdots \cup V_k$, with $1/\varepsilon \leqslant k \leqslant K_0$, such that its (ε, p, d)-reduced graph R contains a cluster X, a matching $\mathcal{M}$ and a bipartite subgraph $\mathcal{H}$, with vertex set $V(\mathcal{H}) = \mathcal{Y} \cup \mathcal{Z}$, satisfying the following properties:*

(a) $N(X) = V(\mathcal{M}) \cup \mathcal{Y}$ *and* $V(\mathcal{M}) \cap \mathcal{Y} = \emptyset$*;*

(b) $|V(\mathcal{M})| + |\mathcal{Y}| \geqslant (\varrho + \delta/3)\,k$*; and*

(c) for all $Y \in \mathcal{Y}$ *we have*

$$|N_{\mathcal{H}}(Y)| \geqslant \left(\varrho + \frac{\delta}{4}\right)\frac{k}{2} - \frac{|V(\mathcal{M})|}{2}.$$

Proof. Given $\varepsilon' = \min\{\varepsilon/5, \delta/1000\}$ and $k_0 = 1/\varepsilon'$, let η_0, n_0' and K_0' be the outputs of the regularity lemma (Theorem 2.2.2) with parameters ε' and k_0. Setting $n_0 = n_0'$ and $\eta_0 = \min\{\eta_0', \delta/1000\}$, let H be an (η, p)-upper uniform graph on $n \geqslant n_0$ vertices and $0 < \eta \leqslant \eta_0$. Then H admits an (ε', p)-regular partition $V(H) = V_0' \cup V_1' \cup \cdots \cup V_\ell'$, with $1/\varepsilon' \leqslant \ell \leqslant K_0$, and let us denote by R' the $(\varepsilon', p, 2d)$-reduced graph of H with respect to this regular partition. By Proposition 2.2.3 and the bound on $e(H)$ we have

$$e(R') \geqslant (1+\eta)^{-1}\left(\varrho + \frac{\delta}{2}\right)\frac{\ell^2}{2} - (1+\eta)^{-1}(6\varepsilon' + 2d)\ell^2 \geqslant \left(\varrho + \frac{\delta}{3}\right)\frac{\ell^2}{2}. \tag{2.1}$$

Note that (2.1) implies that the average degree of R' is at least $(\varrho + \delta/3)\ell$. Thus, by successively removing vertices of low degree, we may find a subgraph $R_0 \subseteq R'$ such that

$$d(R_0) \geqslant \left(\varrho + \frac{\delta}{3}\right)\ell \quad \text{and} \quad \delta(R_0) \geqslant \left(\varrho + \frac{\delta}{3}\right)\frac{\ell}{2}.$$

In particular, this implies that there exists a cluster $X' \in V(R_0)$ with degree at least $(\varrho + \delta/3)\ell$ in R_0. Applying Lemma 2.4.1 to $N_{R_0}(X')$, we find an independent set I, a matching $\mathcal{M}'$ and a collection of triangles Γ that partition $N_{R_0}(X') = I \cup V(\mathcal{M}') \cup V(\Gamma)$, and moreover, by writing $V(\mathcal{M}') = M_1 \cup M_2$ we have that $N_{R_0}(I) \subseteq M_1$. Note that the minimum degree on R_0 implies that for all $Y \in I$ we have

$$|N_{R_0}(Y) \setminus (X' \cup N_{R_0}(X))| \geqslant \left(\varrho + \frac{\delta}{3}\right)\frac{\ell}{2} - 1 - \frac{|V(\mathcal{M})|}{2} \geqslant \left(\varrho + \frac{\delta}{4}\right)\frac{\ell}{2} - \frac{|V(\mathcal{M})|}{2}. \tag{2.2}$$

If there are no triangles in this decomposition, then we would finish the proof by setting $\mathcal{M} = \mathcal{M}'$ and $\mathcal{H}$ as the bipartite graph induced by I and $N_{R'}(I) \setminus (X \cup N_{R'}(X))$. If is not the case, for each $i \in [\ell]$ we may arbitrarily partition $V_i = V_{i,0} \cup V_{i,1} \cup V_{i,2}$ so that

$|V_{0,i}| \leqslant 1$ and $|V_{i,1}| = |V_{i,2}|$. Notting that $|V_{i,1}| = |V_{i,2}| \geqslant |V_i|/3$ for every $i \in [\ell]$, because of Lemma 2.2.1, for each $V_iV_j \in E(R')$ and $a, b \in \{1, 2\}$ the pair $(V_{i,a}, V_{j,b})$ is (ε, p)-regular with density at least d. Moreover, by setting $V_0 = V_0' \cup V_{1,0} \cup \cdots \cup V_{\ell,0}$ we conclude that $V(H) = V_0 \cup V_{1,2} \cup V_{2,2} \cup \cdots \cup V_{\ell,1} \cup V_{\ell,2}$ is an (ε, p)-regular partition with $2\ell + 1$ parts. Let R be the (ε, p, d)-reduced graph of H with respect to this partition, and let $k = 2\ell$ be the number of vertices of R (note that R is a blow-up of R'). We set X as one of the clusters coming from X', and $\mathcal{Y}$ as the set of all the $V_{i,a}$ such that $V_i' \in I$ and $a \in \{1, 2\}$. Now note that each triangle in Γ can be decomposed as three disjoint edges in R. Then we set

$$\mathcal{M} = \bigcup_{V_iV_j \in \mathcal{M}'} \{V_{i,1}V_{j,1}, V_{i,2}V_{j,2}\} \cup \bigcup_{V_aV_bV_c \in \Gamma} \{V_{a,1}V_{b,1}, V_{b,2}V_{c,1}, V_{c,2}V_{a,2}\}$$

and $\mathcal{Z} = N_R(\mathcal{Y}) \setminus (X \cup N_R(X))$. Letting $\mathcal{H}$ as the bipartite graph induced by $\mathcal{Y}$ and $\mathcal{Z}$, is clear that X, $\mathcal{M}$ and $\mathcal{H}$ satisfy (a) and (b), (c) follows from (2.2).

□

2.5 Proof of Theorem 1.1.1

In this section we put everything togheter in order to prove Theorem 2.0.1. As we mentioned in the sketch of the proof, the idea is to use the structure given by Proposition 2.4.2, that is, the cluster X, the matching $\mathcal{M}$ and the bipartite graph $\mathcal{H}$. To do so, we first need to cut the tree into a family $(T_i, r_i)_{i \in [t]}$ of tiny subtrees such that the root of all the subtrees are in the same color class (see Proposition 2.3.2). The main idea of the proof is to first assign each T_i to some edge of $\mathcal{M} \cup \mathcal{H}$. After this, we may remove some bad vertices from each cluster that is used, and thus each subtree T_i can be assigned to a pair $(Y_{i,1}, Y_{i,2})$ which induces a bipartite expander graph and that connects well with a large subset of X (see Claim 2.5.3). Finally, by using an embedding tool due to Balogh, Csaba and Samotij [5], we can embed each subtree into the pair that was assigned to that tree.

The following lemma, proved in [5], gives sufficient expansion conditions for a bipartite graph to contain all trees of a given size. This is the bipartite version of Theorem 3.2.1, and is useful because it is sensitive to the unbalance of the tree in question.

Lemma 2.5.1. *Let $D \geqslant 2$ and let H be a bipartite graph with colour classes V_1 and V_2, where $|V_1| \leqslant |V_2|$. Suppose that H is a bipartite $(m, D+1)$-expander with $0 < m < |V_1|/(2D+1)$. Then H contains all trees with maximum degree at most D and colour classes of sizes at most $|V_1| - (2D+1)m$ and $|V_2| - (2D+1)m$ respectively. Furthermore, any such tree can be embeddeded even if we require that a particular vertex of the tree is mapped to a particular vertex of H, as long as this mapping respect the colour classes.*

Although is not true that (ε, p)-regular pairs are bipartite expanders (since they can have isolated vertices), any large subgraphs of an (ε, p)-regular pairs contains an almost spanning subgraph which is a bipartite expander. The proof of the following result is similar as the proof of Proposition 3.4.2 and it was proved in [5].

Lemma 2.5.2. *Let (A, B) be an (ε, p)-regular pair such that $d_p(A, B) > \varepsilon$. Suppose that $|A| = |B| = m$ and let $A' \subseteq A$ and $B' \subseteq B$ be sets of size at least $(4D + 6)\varepsilon m$. Then there are subsets $A'' \subseteq A'$ and $B'' \subseteq B'$ such that*

(a) $|A' \setminus A''| \leqslant \varepsilon m$ and $|B' \setminus B''| \leqslant \varepsilon m$, and

(b) the subgraph induced by (A'', B'') is a bipartite $(\varepsilon m, 2D + 2)$-expander.

Now we are ready to prove Theorem 2.0.1.

Proof of Theorem 2.0.1. Let n_0', K_0 and η_0 be the outputs of Proposition 2.4.2 with inputs δ, ϱ and $\varepsilon = \delta^4/(2^{28}D^6)$. We set

$$\beta = \frac{\delta^2}{2^{12}kD^4} \quad \text{and} \quad C_0 = \frac{2^{17}10^2D^5K_0^2}{\delta^3}, \tag{2.3}$$

and let $n_0 = \max\{n_0', \beta^{-1}\}$ and $n \geqslant n_0$. Given $p \geqslant C_0/n$ and $0 < \eta \leqslant \eta_0$, let G be an (η, p)-uniform graph on n vertices and let $G' \subseteq G$ be a subgraph with

$$2e(G') \geqslant (\varrho + \delta)2e(G) \geqslant (1 - \eta)(\varrho + \delta)pn^2 \geqslant \left(\varrho + \frac{\delta}{2}\right)pn^2.$$

Since G' is (η, p)-upper uniform, by Proposition 2.4.2 we may find an (ε, p)-regular partition $V(G') = V_0 \cup V_1 \cup \cdots \cup V_k$, with $1/\varepsilon \leqslant k \leqslant K_0$, such that the $(\varepsilon, p, \delta/100)$-reduced graph R, with respect to this partition, contains a cluster X, a matching $\mathcal{M}$ and a bipartite subgraph $\mathcal{H}$, with vertex set $V(\mathcal{H}) = \mathcal{Y} \cup \mathcal{Z}$, satisfying the conclusions of Proposition 2.4.2.

Let $T \in \mathcal{T}(\varrho n, D)$ be given. We consider the bipartition of T that assigns colour 1 to the smaller partition class of T and colour 2 to the larger one, and then we choose an arbitrary vertex r in colour 1 as the root of T. We apply Proposition 2.3.2 to (T, r), with parameter β, obtaining a family $(T_i, r_i)_{i\in[t]}$ of $t \leqslant 4D/\beta$ rooted trees, each of size at most $D^4\beta\varrho n$. Furthermore, each root r_i is at even distance from r and therefore every root has colour 1. For $i \in [t]$, let us write $T_{i,j}$ for the set of vertices of T_i having colour $j \in \{1, 2\}$.

Let m denote the size of the clusters and observe that $m \geqslant (1 - \varepsilon)n/k$. The heart of the proof is the following claim.

Claim 2.5.3. *For each $i \in [t]$, there are sets $(Y_{i,1}, Y_{i,2})$ and $W_i \subseteq X$ such that the following holds.*

(1) *For* $\ell \neq i$ *and* $j, j' \in \{1,2\}$, $Y_{i,j} \cap Y_{\ell,j'} = \emptyset$.

(2) *For* $j \in \{1,2\}$, $|Y_{i,j}| \geqslant |T_{i,j}| + 13D\varepsilon m$.

(3) $G'[Y_{i,1}, Y_{i,2}]$ *is a bipartite* $(\varepsilon m, 2D+2)$*-expander.*

(4) *Every vertex of* $Y_{i,2}$ *has at least* $\delta pm/(200)$ *neighbours in* W_i.

(5) *If* T_ℓ *is a child of* T_i *in the cluster tree, then every vertex of* W_i *has at least* $D+1$ *neighbours in* $Y_{\ell,2}$.

Before proving Claim 2.5.3, let us show how to use it in order to finish the proof of Theorem 2.0.1. Assume that we have ordered $[t]$ so that if T_i is below T_ℓ, with respect to the root of T, then $i \leqslant \ell$. Starting with the subtree containing r, we will embed $(T_i)_{i\in[t]}$ following this ordering. Let us denote by φ the partial embedding of T. For every embedded subtree (T_i, r_i) we will ensure that

(a) $\varphi(r_i) \in W_s$ for some $s \leqslant i$, and

(b) $\varphi(T_{i,j} \setminus \{r_i\}) \subseteq Y_{i,j}$ for $j \in \{1,2\}$.

Suppose we are about to embed a subtree T_ℓ which is a child of some subtree T_i that was already embedded satisfying (a) and (b). Let $v_i \in V(T_i)$ be the parent of r_ℓ and note that v_i is embedded into some vertex $\varphi(v_i) \in Y_{i,2}$ (since v_i is adjacent to r_ℓ and every root has colour 1).

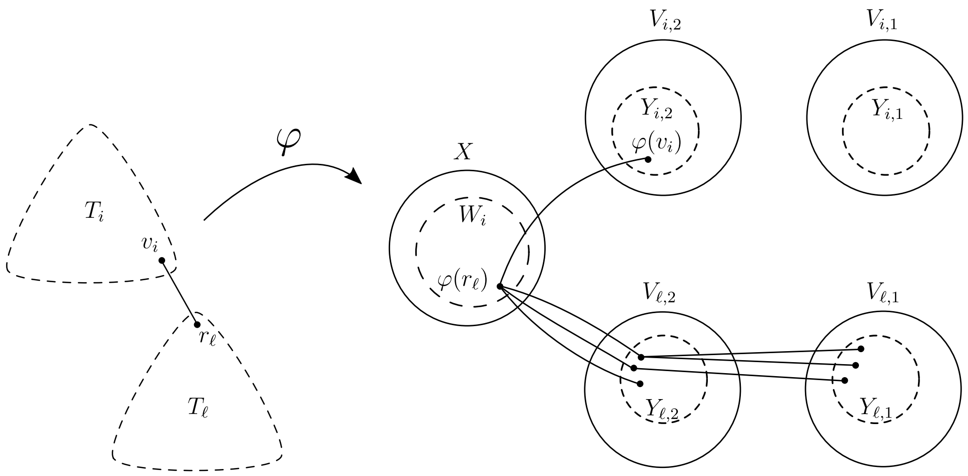

Figure 2.3: Embedding of T_ℓ

Then, because of Claim 2.5.3 (4)

$$|N_{G'}(\varphi(v_i)) \cap W_i| \geqslant \frac{\delta}{200}pm \geqslant (1-\varepsilon)\frac{\delta C_0}{200k} \geqslant \frac{8D}{\beta} \geqslant 2t$$

and therefore at least one neighbour of $\varphi(v_i)$ has not been used during the embedding. We choose any unused vertex $w_\ell \in W_i \cap N_{G'}(\varphi(v_i))$ and set $\varphi(r_\ell) = w_\ell$ (when we embed T_1, we choose any vertex vetex $w_1 \in W_1$ as the image of $r_1 = r$). By Claim 2.5.3 (3) we know that $G'[Y_{i,1}, Y_{i,2}]$ is a bipartite $(\varepsilon m, 2D+2)$-expander, we will prove now that

$$G'[Y_{\ell,1} \cup \{w_\ell\}, Y_{\ell,2}] \text{ is a bipartite } (\varepsilon m + 1, D+1)\text{-expander.}$$

Indeed, since $G'[Y_{i,1}, Y_{i,2}]$ is a bipartite $(\varepsilon m, 2D+2)$-expander is easy to see that the expansion conditions hold for every set $Y \subseteq Y_{\ell,1} \cup Y_{\ell,2}$. Let $Y' \subseteq Y_{\ell,1}$ non-empty and let us consider $Y = Y' \cup \{w_\ell\}$. If $|Y'| \leqslant \varepsilon m$ then we have

$$|N_{G'}(Y) \cap Y_{\ell,2}| \geqslant (2D+2)|Y'| \geqslant (D+1)|X|,$$

where the first inequality follows because $G'[Y_{\ell,1}, Y_{\ell,2}]$ is bipartite $(\varepsilon m, 2D+2)$-expander. Similarly, if $|Y'| \geqslant \varepsilon m$ then we have

$$|N_{G'}(Y) \cap Y_{\ell,2}| \geqslant |N_{G'}(Y') \cap Y_{\ell,2}| \geqslant |Y_{\ell,2}| - (\varepsilon m + 1).$$

Finally, if $Y = \{w_\ell\}$ then by Claim 2.5.3 (5) we know that $|N_{G'}(w_\ell) \cap Y_{\ell,2}| \geqslant D+1$, and therefore $G'[Y_{\ell,1} \cup \{w_\ell\}, Y_{\ell,2}]$ is a bipartite $(\varepsilon m+1, D+1)$-expander.

To complete the embedding of T_ℓ, note that because of Claim 2.5.3 (2) we have

$$|Y_{\ell,j}| - (2D+1)(\varepsilon m+1) \geqslant |T_{\ell,j}| + 13D\varepsilon m - 6D\varepsilon m \geqslant |T_{\ell,j}|$$

for $j \in \{1,2\}$. Thus, using Lemma 2.5.1 we may extend φ to T_ℓ, embedding T_ℓ into $(Y_{\ell,1} \cup \{w_\ell\}, Y_{\ell,2})$ so that $\varphi(T_{\ell,j} \setminus \{r_\ell\}) \subseteq Y_{\ell,j}$ for $j \in \{1,2\}$ and w_ℓ is fixed as the image of r_ℓ (we remark that Claim 2.5.3 (1) allows us to ensure that at every step of the embedding we are using unused vertices). □

Proof of Claim 2.5.3. Let σ be a permutation on $[t]$ such that for all $1 \leqslant i < j \leqslant t$ we have

$$|T_{\sigma(i),2}| - |T_{\sigma(i),1}| \geqslant |T_{\sigma(j),2}| - |T_{\sigma(j),1}|.$$

Recall that we chose colour 2 for the larger partition class of $V(T)$. Therefore, for every $\ell \in [t]$ we have

$$\sum_{i=1}^{\ell} (|T_{\sigma(i),2}| - |T_{\sigma(i),1}|) \geqslant 0. \tag{2.4}$$

The proof of Claim 2.5.3 will be done in two stages. In the first stage, for each $i \in [t]$ the subtree T_i will be assigned to a pair of sets $(X_{i,1}, X_{i,2})$, contained in some edge from $\mathcal{M} \cup E(\mathcal{H})$, such that $|X_{i,j}| = |T_{i,j}| + 16D\varepsilon m$ for $j \in \{1,2\}$. In the second stage, we will remove some vertices from each set in order to find the sets $W_i \subseteq X$ and $Y_{i,j} \subseteq X_{i,j}$ satisfying the properties $(1) - (6)$ from Claim 2.5.3.

Stage 1 (Assignation): In this stage we will prove that for each $i \in [t]$, there exist an edge $V_{i,1}V_{i,2} \in \mathcal{M} \cup E(\mathcal{H})$ and sets $X_{i,j} \subseteq V_{i,j}$, for $j \in \{1,2\}$, such that

(A) $X_{i,j} \cap X_{\ell,j'} = \emptyset$ if $\{i,j\} \neq \{\ell, j'\}$;

(B) $|X_{i,j}| = |T_{i,j}| + 16D\varepsilon m$; and

(C) if $(V_{i,1}, V_{i,2}) \in E(\mathcal{H})$ then $V_{i,2} \in \mathcal{Y}$.

The assignment will be done in two steps following the order given by σ. At step 1 we assign trees to edges from $\mathcal{H}$ until we use a large proportion of $\mathcal{Y} \cup \mathcal{Z}$, and at step 2 we will use edges from $\mathcal{M}$ ensuring that the clusters from each edge of $\mathcal{M}$ are used in a balanced way.

Step 1: We will assume that $|\mathcal{M}| \leqslant (\varrho + \delta/16)k$, as otherwise we just skip this step. Let us set $Q = (\varrho + \delta/4)k - |V(\mathcal{M})|$ and note that we have

$$|\mathcal{Y}| \geqslant Q \geqslant \frac{\delta}{16}k \quad \text{and} \quad d_{\mathcal{H}}(Y) \geqslant Q/2 \text{ for all } Y \in \mathcal{Y}.$$

We will choose sets in $\mathcal{Y} \cup \mathcal{Z}$ until we have assigned at least $(1 - \delta/16)Qm$ vertices to $\mathcal{Y} \cup \mathcal{Z}$. Following the order of σ, assume that we have made the assignation up to some $0 \leqslant \ell \leqslant t-1$ and we are about to assign the tree $T_{\sigma(\ell+1)}$. Suppose that there are $Y \in \mathcal{Y}$ such that

$$\sum_{X_{\sigma(i),2} \subseteq Y} |X_{\sigma(i),2}| \leqslant m - (D^4\beta n + 16D\varepsilon m), \tag{2.5}$$

and $Z \in N_{\mathcal{H}}(Y)$ with

$$\sum_{X_{\sigma(i),1} \subseteq Z} |X_{\sigma(i),1}| \leqslant m - (D^4\beta n + 16D\varepsilon m). \tag{2.6}$$

Since $|T_{\sigma(\ell+1)}| \leqslant D^4\beta\varrho n$, we can select sets $X_{\sigma(\ell+1),1} \subseteq Z$ and $X_{\sigma(\ell+1),2} \subseteq Y$, disjoints from the previously chosen sets, such that $|X_{\sigma(\ell+1),j}| = |T_{\sigma(\ell+1),j}| + 16D\varepsilon m$ for $j \in \{1,2\}$.

So, if there is no $Y \in \mathcal{Y}$ satisfying (2.5), then we have

$$\begin{aligned}\sum_{i=1}^{\ell} |T_{\sigma(i)}| \geqslant \sum_{i=1}^{\ell} |T_{\sigma(i),2}| = \sum_{i=1}^{\ell} \big(|X_{\sigma(i),2}| - 16D\varepsilon m\big) \\ \geqslant |\mathcal{Y}|m - t \cdot 16D\varepsilon m - k \cdot (D^4\beta n + 16D\varepsilon m) \\ \geqslant |\mathcal{Y}|m - \frac{\delta^2}{16^2}km \\ \geqslant \Big(1 - \frac{\delta}{16}\Big)Qm.\end{aligned}$$

This means that we have already used enough vertices from $\mathcal{Y} \cup \mathcal{Z}$. On the other hand, if every Y satisfying (2.5) has no neighbours satisfying (2.6), we may use (2.4) to deduce

$$\begin{aligned}\sum_{i=1}^{\ell} |T_\sigma(i)| \geqslant 2\sum_{i=1}^{\ell} |T_{\sigma(i),1}| = 2\sum_{i=1}^{\ell} \big(|X_{\sigma(i),1}| - 16D\varepsilon m\big) \\ \geqslant 2d_{\mathcal{H}}(Y)m - t \cdot 32D\varepsilon m - k \cdot 2(D^4\beta n + 16D\varepsilon m) \\ \geqslant Qm - \frac{\delta^2}{16^2}km \\ \geqslant \Big(1 - \frac{\delta}{16}\Big)Qm.\end{aligned}$$

This means that if at step $\ell + 1 \in [t]$ we could not find a pair (Y, Z) satisfying (2.5) and (2.6), then we have used vertices at least $(1 - \delta/16)Qm$ vertices from $\mathcal{Y} \cup \mathcal{Z}$ at step ℓ.

Step 2: Let $0 \leqslant \ell_0 \leqslant t$ be such that $T_{\sigma(1)}, \dots, T_{\sigma(\ell_0)}$ have been assigned to $\mathcal{Y} \cup \mathcal{Z}$, satisfying (A),(B) and (C), and

$$\Big(1 - \frac{\delta}{16}\Big)Qm \leqslant \sum_{i=1}^{\ell_0} |T_{\sigma(i)}| \leqslant \Big(1 - \frac{\delta}{16}\Big)Qm + D^4\beta\varrho n. \tag{2.7}$$

Assume that $\ell_0 < t$, otherwise we are done. For $\ell_0 + 1 \leqslant i \leqslant t$ we will assign each $T_{\sigma(i)}$ to some edge $AB \in \mathcal{M}$. At each step we will ensure that for every edge $AB \in \mathcal{M}$ we have

$$\left| \sum_{X_{\sigma(i),j} \subseteq A} |X_{\sigma(i),j}| - \sum_{X_{\sigma(i),j} \subseteq B} |X_{\sigma(i),j}| \right| \leqslant D^4\beta\varrho n. \tag{2.8}$$

Suppose we are about to assign a subtree $T_{\sigma(\ell)}$, for some $\ell \geqslant \ell_0 + 1$, and that (2.8) holds at step $i = \ell - 1$ (note that (2.8) holds trivially at step ℓ_0). Suppose that there is an edge $AB \in \mathcal{M}$ such that

$$\max\Big\{ \sum_{X_{\sigma(i),j} \subseteq A} |X_{\sigma(i),j}|, \sum_{X_{\sigma(i),j} \subseteq B} |X_{\sigma(i),j}| \Big\} \leqslant m - (D^4\beta\varrho n + 16D\varepsilon m). \tag{2.9}$$

Assuming that $\sum_{X_{\sigma(i),j} \subseteq A} |X_{\sigma(i),j}| \leqslant \sum_{X_{\sigma(i'),j'} \subseteq B} |X_{\sigma(i'),j'}|$, we let $j^\star = \operatorname{argmax}_{j \in \{1,2\}} |T_{\sigma(\ell),j}|$ and then we may take sets

- $X_{\sigma(\ell),j^\star} \subseteq A$ with $|X_{\sigma(\ell),j^\star}| = |T_{\sigma(\ell),j^\star}| + 16D\varepsilon m$, and
- $X_{\sigma(\ell),3-j^\star} \subseteq B$ with $|X_{\sigma(\ell),3-j^\star}| = |T_{\sigma(\ell),3-j^\star}| + 16D\varepsilon m$.

disjoints from the previously chosen sets. Note that we have assigned the larger colour class of $T_{\sigma(\ell)}$ to the less occupied cluster in $\{A, B\}$. Furthermore, since (2.8) holds at step $\ell-1$ and as $|T_{\sigma(\ell)}| \leqslant D^4\beta\varrho n$, the assignment of $T_{\sigma(\ell)}$ implies that (2.8) holds at step ℓ. So suppose that (2.9) does not hold at step $\ell-1$ for any $AB \in \mathcal{M}$. Then we have

$$\begin{aligned}\sum_{i=\ell_0+1}^{\ell-1} |T_{\sigma(i)}| &\geqslant |V(\mathcal{M})|m - t\cdot 32D\varepsilon m - k\cdot(3D^4\beta\varrho n + 32D\varepsilon m)\\ &\geqslant |V(\mathcal{M})|m - \frac{\delta}{16}km\end{aligned}$$

that together with (2.7) yields

$$\begin{aligned}\sum_{i=1}^{\ell-1} |T_{\sigma(i)}| &\geqslant \Big(1-\frac{\delta}{16}\Big)Qm + |V(\mathcal{M})|m - \frac{\delta}{16}km\\ &\geqslant \Big(1-\frac{\delta}{16}\Big)\Big(\varrho+\frac{\delta}{4}\Big)km - \frac{\delta}{16}km\\ &\geqslant \Big(\varrho+\frac{\delta}{8}\Big)km\\ &\geqslant \Big(\varrho+\frac{\delta}{16}\Big)n,\end{aligned}$$

which is impossible since $|T| = \varrho n$. This implies that we can make the assignation for each $\ell \in [t]$.

Stage 2 (Cleaning): Assume that the cluster tree is ordered according to a BFS starting from the subtree which the root of T. Starting with a leaf of the cluster tree, suppose that we have found the sets $Y_{i,j}$ satisfying properties (1) – (6) for all subtrees T_i below T_ℓ in the order of the cluster tree. Let us define

$$W_\ell := \{v \in X : d(v, Y_{i,2}) \geqslant D+1 \text{ for all } i \text{ such that } T_i \text{ is a child of } T_\ell\},$$

we want to prove that W_ℓ has a reasonable size. Given a child T_i of T_ℓ in the cluster tree, we have that

$$|Y_{i,2}| \geqslant |T_{i,j}| + 13D\varepsilon m \geqslant (D+1)\varepsilon m$$

and therefore, since $(X, V_{i,2})$ is (ε, p)-regular, by Lemma 2.2.1 there are at most $(D+1)\varepsilon m$ vertices in X with less than $D+1$ neighbours in $Y_{i,2}$. Since the auxiliary tree has maximum degree D^4, then W_ℓ has at least

$$|X| - (D+1)D^4\varepsilon|X| \geqslant \frac{m}{2}$$

vertices. Now, since $(X, V_{\ell,2})$ is (ε, p)-regular, then by Lemma 2.2.1 the pair $(W_\ell, V_{\ell,2})$ is $(2\varepsilon, p)$-regular with p-density at least $\delta/(100) - \varepsilon$. By Lemma 2.2.1 there are at most $2\varepsilon m$ vertices of $V_{\ell,2}$ with less than

$$\left(\frac{\delta}{100} - 3\varepsilon\right) p|W_\ell| \geqslant \frac{\delta}{200} pm$$

neighbours in W_ℓ. We remove each such vertex from $X_{\ell,2}$ thus obtaining a set $X'_{\ell,2}$ such that every vertex in $X'_{\ell,2}$ has at least $\delta pm/200$ neighbours in W_ℓ. Now, we need to find an expander subgraph of $(X_{\ell,1}, X'_{\ell,2})$. Since $(V_{\ell,1}, V_{\ell,2})$ is (ε, p)-regular with $d_p(V_{\ell,1}, V_{\ell,2}) \geqslant \delta/100$ and

$$|X_{\ell,1}|, |X'_{\ell,2}| \geqslant 16D\varepsilon m - 2\varepsilon m \geqslant (4D + 6)\varepsilon m,$$

we may use Lemma 2.5.2 to obtain a pair $(Y_{\ell,1}, Y_{\ell,2})$, with $Y_{\ell,1} \subseteq X_{\ell,1}$ and $Y_{\ell,2} \subseteq X'_{\ell,2}$, such that $G'[Y_{\ell,1}, Y_{\ell,2}]$ is bipartite $(\varepsilon m, 2D + 2)$-expander and satisfies $|Y_{\ell,j}| \geqslant |X_{\ell,j}| - 3\varepsilon m \geqslant |T_{\ell,j}| + 13D\varepsilon m$ for $j \in \{1, 2\}$. $\square$

Chapter 3

Ramsey goodness of trees in random graphs

Throughout this chapter, we will always assume that Theorem 1.1.1 holds.

3.1 Overview

For $N \geqslant rn + \Omega(1/p)$ and $p = \Omega(N^{-2/(r+2)})$, our aim is to prove that with high probability $G(N,p) \to \big(K_{r+1}, \mathcal{T}(n,D)\big)$. To do so, we rely on a stability argument, in which we show that any colouring of a *typical* outcome of the edges $G = G(N,p)$ is either sufficiently close to a extremal colouring or it contains a blue K_{r+1} or the red graph is $\mathcal{T}(n,D)$-universal.

Let G_B and G_R be the blue and red graphs, respectively, in a coloring of $E(G)$ and let us assume that $K_{r+1} \nsubseteq G_B$ and that G_R is not $T \in \mathcal{T}(n,D)$-universal. By Theorem 1.1.1, we have that $e(G_R) \leqslant (1/r + o(1))e(G)$ and consequently that

$$e(G_B) \geqslant \left(1 - \frac{1}{r} - o(1)\right) e(G). \tag{3.1}$$

In words, in this scenario the blue graph roughly has at least the number of edges of the intersection of a Turán graph with G.

By a result of Conlon and Gowers [20] and of Schacht [69], with high probability, every K_{r+1} subgraph of G with as many edges as in (3.1) is almost r-partite. More precisely, there exists a partition $V(G) = V_1' \cup \cdots \cup V_r'$ with $o(pN^2)$ blue edges within the parts.

We first define V_0 as the set of vertices with $\Omega(pN)$ blue neighbours inside the part to which it belongs and B the vertices with $o(pN)$ neighbours in any of the parts. One can show that $|V_0| = o(N)$, by the number of blue edges inside parts, and that $|B| = O(1/p)$ by the properties of G. By setting $V_i = V_i' \setminus (V_0 \cup B)$, we get a new partition $V(G) = (B \cup V_0) \cup V_1 \cup \cdots \cup V_r$, with $|B \cup V_0| = o(N)$ and such that for each $i \in [r]$ every vertex $v \in V_i$ has $\Omega(pN)$ red neighbours and only few blue neighbours, both in V_i.

We show that the red graphs induced in each part is an expander graph, which roughly means that every set has a large red external neighbourhood. In particular, these red graphs satisfy the hypothesis of a theorem of Haxell [36], which imply that for every $T \in \mathcal{T}\big((1-o(1))|V_i|, D)$ and for every $v \in T$ and $u \in V_i$ there exists an embedding of T in $G_R[V_i]$ that maps v to u. This already implies that $|V_i| = (1+o(1))n$ for every $i \in [r]$ and also that there are no red edges between different parts. The second property is true because of the flexibility given by Haxell's Theorem to choose the starting vertices to embed the trees. Indeed, suppose that there exists a red edge between different parts, say V_1 and V_2. We can split the tree in two subtrees connected by an edge and then we may embed the tree by mapping one part of the tree into V_1 and the other part into V_2 and complete the embedding with this red edge.

This part of the argument is captured by Proposition 3.5.2 and by its statement it is possible to conclude Theorem 1.2.1 if $N \geqslant rn + o(n)$, since at least one of the V_i' would have more than $(1+o(1))n$ vertices. However, in order to prove Theorem 1.2.1 we need to push further this stability argument, which is the second part of the proof. Our aim at this point was to prove that all vertices from $V_0 \setminus B$ can be relocated to some V_i, with $i \in [r]$ so that the expansions propertys of the red graphs remains an expander, as in the previous step, to guarantee that we have no blue edges inside parts in this new partition.

Let $v \in V_0 \setminus B$. Since all edges between parts are blue, then if v has $\Omega(pN)$ blue neighbours in each of the r parts, then we get a blue copy of K_{r+1} in a typical G, by Janson's inequality. On the other hand, if v has $\Omega(pN)$ red neighbours in more than one part, then, in a similar way as mentioned above, Haxell's Theorem would yield all trees of $\mathcal{T}(n, D)$ in red. Therefore, for every vertex $v \in V_0 \setminus B$ there is exactly one $i \in [r]$ in which v has $\Omega(pN)$ red neighbours and $o(pN)$ blue neighbors in V_i

These conditions on the blue and red degree are enough to guarantee that $G_R[V_i \cup \{v\}]$ is an expander and therefore there are no blue edges between parts, as before. Repeating this process, we can relocate all vertices of V_0 except the ones in B, which is consistent with the construction shown in Section [REF], as $|B| = O(1/p)$. At this point, we are left to show that if the largest part, say V_1, had $n + \Omega(1/p)$ vertices, then $G_R[V_1]$ would be $\mathcal{T}(n, D)$-universal. This is the final aspect of the proof of Theorem 1.2.1 and it is covered in the next section.

As usual in problems of tree embeddings in expander graphs, with treat differently the cases of trees with few or many leaves. For trees with less than $n/(\log^3 n)$ leaves, we use a result of Montgomery [53] that in fact yields embeddings of this class of spanning trees in expander graphs. In the case of trees with many leaves, previous results do not fit in our context. In particular, in the proof of Haxell's Theorem the vertices, except the leaves, of the tree are embedded inductively, followed by an application of Hall's Theorem to embed the leaves. However, in our context this strategy reaches the following barrier. There might be disjoint sets $X, Y \subset V_1$ of sizes $\omega(1/p)$ and $n/(\log^3 n)$, respectively, with no edges of $G(N,p)$ between them. To see why this is an impediment, let $T \in \mathcal{T}(n,D)$ be a tree with at least $n/(\log^3 n)$ leaves and let $T' \subseteq T$ be the tree obtained by removing the leaves from T. If we carelessly embed T' in $G_R[V_1]$, we could have the bad luck that this embedding maps the neighbors of the leaves to Y, while at the same time X is contained in the set of unused vertices. In this situation, we can extend it to an embedding of T if and only if one can guarantee a Hall-type condition in the bipartite graph induced by the image of the neighbors of the leaves in V_1 and the set of unused vertices. However, since X has no edges to Y and since we have only $O(1/p)$ "extra" vertices, there is no way to guarantee the extention of this embedding.

We deal with this problem beforehand in the proof of Theorem 3.2.4 in Section 3. The basic idea is to choose a random set $R \subseteq V_1$ of size roughly $n/(\log^3 n)$. We prove that there exists a realisation of R such that for every set $X \subseteq V_1$ of size $\Omega(1/p)$ and for every set $Y \subseteq R$ of size $n/(\log^3 n)$ there is at least one edge between them. With an intermediate result of Haxell's Theorem, we inductively embed T' in H, requiring that the neighbors of the leaves are mapped to R. Finally, we are able to show that the aforementioned bipartite graph satisfies the hypothesis of Hall's Theorem.

3.2 Trees in expanders

For a graph H and a subset $X \subseteq V(H)$, we denote by $\Gamma(X) = \bigcup_{x \in X} N(x)$ the set of neighbours of X and write $N(X) = \Gamma(X) \setminus X$ for the external neighbourhood of X. In this section, we study the family of graphs called *expanders* in which subsets of vertices have a large external neighbourhood. The notion of expander graphs has a plentiful number of applications in combinatorics and it is particularly useful for embedding trees. Indeed, Friedman and Pippenger [32] proved that given integers m and D, if a graph H satisfies

$$|\Gamma(X)| \geqslant (D+1)|X| \text{ for all } X \subseteq V(H) \text{ with } 1 \leqslant |X| \leqslant 2m,$$

then H contains all trees with m vertices and maximum degree D. A limitation of this result is that it only works for trees of size at most $|V(H)|/(2D+2)$. In a successful attempt to overcome this issue, Haxell [36] considered a different notion of expansion in order to prove the following result.

Theorem 3.2.1. *Let $D, m, t \in \mathbb{N}$ and let H be a graph with the following properties:*

(*i*) $|N(X)| \geqslant D|X| + 1$, *for all $X \subseteq V(H)$ with $1 \leqslant |X| \leqslant m$.*

(*ii*) $|N(X)| \geqslant t + D|X| + 1$, *for all $X \subseteq V(H)$ with $m + 1 \leqslant |X| \leqslant 2m$.*

Then H contains a copy of every tree T with t vertices and maximum degree at most D. Furthermore, given $v \in V(H)$ and $u \in V(T)$, there exists an embedding of T mapping u to v.

A different and convenient way of phrasing property ii of Theorem 3.2.1 is as follows. Let H be a graph such that every pair of disjoint sets $X, Y \subseteq V(H)$, with $|X| = m_1$ and $|Y| = m_2$, satisfies $e(X,Y) > 0$. Then for every $Z \subseteq V(H)$, with $m_1 \leqslant |Z| \leqslant 2m_1$, there are at most $m_2 - 1$ vertices in the non-neighbourhood of Z. By discounting the non-neighbours of Z and the vertices in Z, we get

$$|N(Z)| \geqslant |V(H)| - |Z| - m_2 + 1. \tag{3.2}$$

Therefore, when $|V(H)| - m_2 \geqslant t + 2(D+1)m_1$ we recover property ii. The main result of this section considers the case where m_1 and m_2 have different orders of magnitude, which leads us to the following definition.

Definition 3.2.2. *Let D, m_1, m_2 be integers. We say that a graph H is an (m_1, m_2, D)-expander if*

E1 $|N(X)| \geqslant D|X| + 1$ *for all $X \subseteq V(H)$ with $1 \leqslant |X| \leqslant m_1$, and*

E2 $e(X,Y) > 0$ *for all disjoint sets $X, Y \subseteq V(H)$ with $|X| = m_1$ and $|Y| = m_2$.*

Moreover, if only property 2 holds, then we say that H is a weak (m_1, m_2)-expander. We will often omit D when it is clear from context.

As is usual with tree embedding problems, we deal separately with trees having either too many or too few leaves. For trees with few leaves, we will use the following result of Montgomery [**?**, 53].

Theorem 3.2.3. *Let n be sufficiently large, let D be a positive integer, and set $d = D\log^4 n/20$. If H is a $(n/2d, n/2d, d)$-expander on n vertices, then H contains a copy of every tree on n vertices, maximum degree bounded by D, and at most n/d leaves.*

We remark that although Theorem 3.2.3 is not stated explicitly in [53] it follows directly from Montgomery's proof (see [53, Section 4.2]), where it is only used that $G(n,p)$ is an expander as in Theorem 3.2.3. The main result of this section deals with the case of (non-spanning) trees with many leaves.

Theorem 3.2.4. *Let m_1, m_2, n, D be positive integers such that $6m_1 \log n < m_2$ and $16Dm_2 \leqslant n$, and assume that n is sufficiently large. Let H be a graph on n vertices such that H is* *(i) a weak $(m_1, n/32D)$-expander, and*

(ii) a weak (m_2, m_2)-expander.

Then H contains every tree $T \in \mathcal{T}(n - m_1, D)$ with at least $24Dm_2$ leaves.

A first approach to Theorem 3.2.4 is to follow the proof of Haxell's embedding theorem (Theorem 3.2.1) to embed a tree with its leaves removed, and then use a Hall-type argument in order to embed the leaves. However, the hypotheses of Theorem 3.2.4 do not enable a straightforward modification of this proof for the following reason. Given a tree T, let $L \subseteq V(T)$ be the set of leaves of T and let $P = N(L)$ be their parents. Note that if $T \in \mathcal{T}(n - m_1, D)$ is a tree with $|L| = \Omega(m_2)$ leaves, then we also have $|P| = \Omega(m_2)$. Suppose that we have a partial embedding of $T - L$ which we want to extend to T. By the hypothesis of Theorem 3.2.4, it might be that the image of P has $m_2 - 1$ non-neighbours in the leftover vertices, in which case is impossible to extend the embedding of $T - L$ since $m_1 < m_2$.

We address this obstacle by finding a set $W \subseteq V(H)$ with $\Theta(m_2)$ vertices such that every subset $X \subseteq W$ with $|X| = m_2$ has less than m_1 non-neighbours in H. We then manage to find an embedding $\varphi : V(T - L) \to V(H)$ such that $\varphi(P) \subseteq W$, in which case we would have that

$$|N(X) \setminus \varphi(V(T - L))| \geqslant n - |T - L| - m_1 + 1 > |L|$$

for every $X \subseteq \varphi(P)$ with $|X| \geqslant m_2$. However, in order to use a Hall-type argument, we will also need to guarantee that small subsets of $\varphi(P)$ have enough neighbours in the set of unused vertices. This idea is captured by the following definition, which has previously appeared in the works of Friedman and Pippenger [32], Haxell [36], and Balogh, Csaba, and Samotij [5].

Definition 3.2.5. *Let m be a positive integer, let T be a tree with maximum degree at most D, and let H be a bipartite graph with parts V_1 and V_2. We say that an embedding $\varphi : V(T) \to V(H)$ is m-good in H if for every $i \in \{1, 2\}$ and $X \subseteq V_i$, with $1 \leqslant |X| \leqslant m$, we have*

$$|N_H(X) \setminus \varphi(V(T))| \geqslant \sum_{v \in \varphi^{-1}(X)} \big(D - d_T(v)\big) + D|X \setminus \varphi(V(T))|.$$

In the previous definition we considered H as being bipartite for technical reasons. More specifically, as we want to embed the set of parents of leaves into a set W, we have to alternate the embedding of T between W and $V(H) \setminus W$ and thus it is easier to consider H as being a bipartite graph. The next lemma gives sufficient conditions to extend good embeddings, and it was proved in [5] as the induction step[1] in the proof of a bipartite analogue of Theorem 3.2.1 (see Theorem 2.5.1).

Lemma 3.2.6. *Let m, n, D be positive integers, let T be a tree with maximum degree at most D, and let H be a bipartite graph with parts V_1 and V_2. Suppose that there exists an m-good embedding $\varphi : V(T) \to V(H)$, and that for $i \in \{1, 2\}$ and any subset $X \subseteq V_i$, with $m \leqslant |X| \leqslant 2m$, we have*

$$|N_H(X) \setminus \varphi(V(T))| \geqslant 2Dm + 2. \tag{3.3}$$

Then for every vertex $v \in T$, with $d_T(v) < D$, there exists an m-good embedding of the tree obtained by adding to T a leaf adjacent to v.

We will be able to use Lemma 3.2.6 in graphs satisfying the following notion of *bipartite expansion.*

Definition 3.2.7. *Let $D \geqslant 2$ and let H be a bipartite graph with parts V_1 and V_2 such that $|V_1| \leqslant |V_2|$. Let m be a positive integer with $m < |V_1|$. We say that H is a bipartite (m, D)-expander if the following two properties hold.*

(i) For $i \in \{1, 2\}$, every set $X \subseteq V_i$, with $1 \leqslant |X| \leqslant m$, satisfies $|N_H(X)| \geqslant D|X|$.

(ii) For every pair of sets $X_1 \subseteq V_1$ and $X_2 \subseteq V_2$, each of size at least m, we have $e(X_1, X_2) > 0$.

Note that property ii implies that for every subset $X \subseteq V_i$, with $|X| \geqslant m$, we have

$$|N(X)| \geqslant |V_{3-i}| - m + 1.$$

This will guarantee that (3.3) holds for the embedding of any tree with small enough bipartition classes. Now we can state one of the main results that we need for the proof of Theorem 3.2.4.

Lemma 3.2.8. *Let $m, D \in \mathbb{N}$ with $D \geqslant 2$, and let T be a tree with maximum degree at most D. Let $U_1 \cup U_2$ be any partition of one the bipartition classes of T and let U_3 be the other bipartition class. Let H be a graph on n vertices and let $V_1, V_2, V_3 \subseteq V(H)$ be disjoint sets such that $|V_i| \geqslant |U_i| + 3Dm$ for $i \in \{1, 2, 3\}$. If $H[V_1, V_3]$, $H[V_2, V_3]$ and $H[V_1 \cup V_2, V_3]$ are bipartite (m, D)-expanders, then there exists an m-good embedding $\varphi : V(T) \to V(H)$ such that $\varphi(U_i) \subseteq V_i$ for $i \in \{1, 2, 3\}$.*

[1]Under the hypothesis Theorem 7 from [5], the authors state that good embeddings can be extended as "Property 2" in page 6 from [5]. Moreover, the only place where they use the size of neighbours of sets with more than m vertices is in the proof of Claim 8. One can check that (3.3) is enough to get the same proof.

The strategy of the proof of Lemma 3.2.8 is to iteratively apply Lemma 3.2.6 in order to extend a partial embedding of the tree by adding a leaf at each step. Since we will alternate between vertices of V_1, V_2 and V_3, we will need to keep track that the embeddings are m-good in the graphs $H[V_1, V_3]$, $H[V_2, V_3]$ and $H[V_1 \cup V_2, V_3]$, respectively. This will guarantee that, at any stage of the embedding, small subsets of $V_1 \cup V_2$ have enough neighbours in the unused vertices of V_3, and that small subsets of V_3 have enough neighbours in the unused vertices of both V_1 and V_2.

In the context of Lemma 3.2.8, for a subtree $S \subseteq T$ we say that $\varphi : V(S) \to V(H)$ is m-great if

A1 $U_i \cap V(S)$ is mapped to V_i, for $i \in \{1, 2, 3\}$, and

A2 φ is m-good in both $H[V_1 \cup V_2, V_3]$ and $H[V_i, V_3]$, for $i \in \{1, 2\}$.

Proof of Lemma 3.2.8. We start by showing that there exists an m-great embedding of any single vertex subtree $S \subseteq T$.

Claim 3.2.9. *Let $S \subseteq T$ be a single vertex subtree. If $\varphi : V(S) \to V(H)$ is an embedding which satisfies property 1, then φ is m-great.*

Proof of Claim 3.2.9. We will only prove that φ is m-good in $H[V_1, V_3]$, as the other cases are completely analogous. Since $H[V_1, V_3]$ is a bipartite (m, D)-expander, then for $X \subseteq V_1$, with $m \leqslant |X| \leqslant 2m$, we have

$$|(N(X) \cap V_3) \setminus \varphi(V(S))| \geqslant |V_3| - |S| - m + 1,$$

which is larger than the required lower bound in the definition of m-goodness. Since the same bound holds if $X \subseteq V_3$, it follows that φ is m-good in $H[V_1, V_3]$. $\square$

Now that we have proved the base case, we will prove that any m-great embedding of a subtree $S \subset T$ can be extended by adding a leaf. Let $s \in V(S)$ and $v \in V(T - S)$ satisfy $sv \in E(T)$. Assume we have an m-great embedding $\varphi : V(S) \to V(H)$ and we want to add v. We deal separately with the cases when $v \in U_3$ or $v \in U_1 \cup U_2$.

Suppose that $v \in U_3$. Since $H[V_1 \cup V_2, V_3]$ is a (m, D)-expander, then for $X \subseteq V_1 \cup V_2$ (and analogously for $X \subseteq V_3$), with $m \leqslant |X| \leqslant 2m$, we have that

$$|(N(X) \cap V_3) \setminus \varphi(V(S))| \geqslant |V_3| - m + 1 - |U_3| \geqslant 3Dm - m + 1 \geqslant 2Dm + 2. \quad (3.4)$$

Thus, by Lemma 3.2.6, there exists an m-good embedding $\varphi' : V(S + sv) \to V(H[V_1 \cup V_2, V_3])$. We argue now that φ' is m-good in $H[V_i, V_3]$, for $i \in \{1, 2\}$. Indeed, given $X \subseteq V_i$ for some $i \in \{1, 2\}$, we already know that $|(N(X) \cap V_3) \setminus \varphi'(V(S))| \geqslant 2Dm + 2$

since φ' is m-good in $H[V_1 \cup V_2, V_3]$. For $X \subseteq V_3$ there is nothing to prove, since φ was m-great and we did not use any additional vertices from either V_1 or V_2.

The case when $v \in U_1$ (resp. $v \in U_2$) is analogous, but we apply Lemma 3.2.6 to φ in the bipartite graph $H[V_1, V_3]$ (resp. $H[V_2, V_3]$), together with the same calculation as in (3.4), to get an m-good embedding φ'. Note that $\varphi'(v) \in V_1$ (resp. $\varphi'(v) \in V_2$). This guarantees that φ' is m-good in $H[V_1, V_3]$ and $H[V_2, V_3]$. Moreover, for $H[V_1 \cup V_2, V_3]$ we only need to guarantee the neighbourhood expansion for $X \subseteq V_3$ with $m \leqslant |X| \leqslant 2m$. Note that since φ' is m-good in $H[V_i, V_3]$ for $i \in \{1, 2\}$ we have

$$|(N(X) \cap (V_1 \cup V_2)) \setminus \varphi'(V(S))| \geqslant |(N(X) \cap V_1) \setminus \varphi'(V(S))| \geqslant 2Dm + 2,$$

and thus φ' is m-good in $H[V_1 \cup V_2, V_3]$. □

The last ingredient that we need for Theorem 3.2.4 is a well-known generalisation of Hall's theorem.

Lemma 3.2.10. *Let G be a bipartite graph with parts $A = \{a_1, \dots, a_\ell\}$ and B. Let $(d_i)_{i\in[\ell]}$ be a sequence of non-negative integers, and let $(S_i)_{i\in[\ell]}$ be a collection of vertex-disjoint stars such that S_i has a central vertex s_i and d_i leaves for each $i \in [\ell]$. Then G contains an embedding of $(S_i)_{i\in[\ell]}$, with s_i copied to a_i for each $i \in [\ell]$, if and only if*

$$|N(X)| \geqslant \sum_{a_i \in X} d_i \text{ for all } X \subseteq A. \tag{3.5}$$

Proof of Theorem 3.2.4. Let L be a set of $12Dm_2$ leaves of T in the same bipartition class and let U_1 be the set of parents of L in T. Note that $12m_2 \leqslant |U_1| \leqslant 12Dm_2$. We choose, uniformly at random, a set $W \subseteq V$ with $r = |U_1| + 4Dm_2$ vertices, and note that $r \leqslant 16Dm_2 \leqslant n$. For each set $X \subseteq V(H)$ with m_1 vertices, let $Z_X = \{y \in W \setminus X : d(y, X) = 0\}$. Since H is a weak $(m_1, n/32D)$-expander, then

$$\mathbb{E}[|Z_X|] \leqslant \frac{r}{n} \cdot \frac{n}{32D} \leqslant \frac{m_2}{2}.$$

By standard tail bounds for the hypergeometric distribution (see Theorem 2.10 in [39]), we have

$$\mathbb{P}(|Z_X| \geqslant m_2) \leqslant \exp\left(-\frac{m_2}{6}\right).$$

Denoting by Z the number of sets $X \subseteq V(H)$ of size m_1 such that $|Z_X| \geqslant m_2$, we have

$$\mathbb{E}[Z] \leqslant n^{m_1} \exp(-m_2/6) < 1,$$

since $6m_1 \log n < m_2$. This implies that there is a realisation of W, denoted by W_1, such that every subset $X \subseteq V(H)$ of size m_1 has less than m_2 non-neighbours in W_1. Set $T' = T - L$ and let us denote one of the bipartition classes of T' by $U_1 \cup U_2$ and the other by U_3. We take two disjoint sets $W_2, W_3 \subseteq V(H) \setminus W_1$ such that $|W_i| = |U_i| + 4Dm_2$ for $i \in \{2, 3\}$, which is possible since in this case we have

$$|W_1| + |W_2| + |W_3| = |T| - |L| + 12Dm_2 \leqslant n.$$

Claim 3.2.11. *For each $i \in \{1, 2, 3\}$ there exists $V_i \subseteq W_i$, with $|W_i \setminus V_i| \leqslant 2m_2$, such that the graphs $H[V_1 \cup V_2, V_3]$, $H[V_1, V_3]$ and $H[V_2, V_3]$ are bipartite (m_2, D)-expanders.*

Proof of Claim 3.2.11. Since H is a weak (m_2, m_2)-expander, property 2 implies that the second property of the bipartite expansion is already satisfied for all the three bipartite graphs. We will find the sets V_i's iteratively. We initialise by setting $X_i = \emptyset$ and $V_i := W_i$ for $i \in \{1, 2, 3\}$.

- While there exists a set $X \subseteq V_3$ with $|X| \leqslant m_2$ and $|N(X) \cap V_i| < D|X|$ for some $i \in \{1, 2\}$, we set $X_i := X_i \cup X$ and $V_3 := V_3 \setminus X$, and
- while there exists a set $X \subseteq V_1 \cup V_2$ with $|X| \leqslant m_2$ and $|N(X) \cap V_3| < D|X|$, we set $X_3 := X_3 \cup X$ and $V_i := V_i \setminus X$ for $i \in \{1, 2\}$.

First, we show that at each step we have $|X_i| \leqslant m_2$ and $|N(X_i) \cap V_i| < D|X|$ for $i \in \{1, 2, 3\}$. Indeed, if this is satisfied at some step for X_1, X_2, X_3 and there exists $X \subseteq V_1 \cup V_2$ (or analogously for $X \subseteq V_3$) with $|N(X) \cap V_3| < D|X|$, then we have that

$$|N(X_3 \cup X) \cap V_3| \leqslant |N(X_3) \cap V_3| + |N(X) \cap V_3| < D|X_3| + D|X| = D|X_3 \cup X|.$$

On the other hand, if $|X_3 \cup X| \geqslant m_2$, then by property 2, $X_3 \cup X$ would have fewer than m_2 non-neighbours in V_3 and therefore we would have that

$$|N(X_3 \cup X) \cap V_3| \geqslant |V_3| - m_2 + 1 \geqslant 2Dm_2 + 1 \geqslant D|X \cup X_3| + 1,$$

which contradicts the choice of X. This finishes the proof since $|X_1 \cup X_2|, |X_3| \leqslant 2m_2$. $\square$

Let $V_i \subseteq W_i$ be the sets given by Claim 3.2.11 for $i \in \{1, 2, 3\}$ so that $H[V_1 \cup V_2, V_3]$, $H[V_1, V_3]$ and $H[V_2, V_3]$ are bipartite (m_2, D)-expanders. Observe that

$$|V_i| \geqslant |U_i| + 4Dm_2 - 2m_2 \geqslant |U_i| + 3Dm_2,$$

for $i \in \{1, 2, 3\}$ which, by Lemma 3.2.8, implies that we can find an m_2-good embedding $\varphi' : V(T') \to V(H)$ such that $\varphi'(U_i) \subseteq V_i$ for $i \in \{1, 2, 3\}$.

In order to finish the embedding of L, we will use Lemma 3.2.10 in the bipartite graph $H[\varphi'(U_1), V(H) \setminus \varphi'(V(T'))]$. Note that the condition of Lemma 3.2.10 is satisfied for every subset $X \subseteq \varphi'(U_1)$ with $|X| \leqslant m_2$, since by property i of the m_2-good embedding we have

$$N(S, V(H) \setminus \varphi'(V(T'))) \geqslant D|X| \geqslant \Delta(T)|X|.$$

Moreover, since $\varphi'(U_1) \subseteq W_1$ and by the choice of W_1, every subset $X \subseteq \varphi'(U_1)$, with $|X| \geqslant m_2$, has fewer than m_1 non-neighbours and therefore

$$|N(X) \cap V(H) \setminus \varphi'(V(T'))| \geqslant |V(H) \setminus \varphi'(V(T'))| - m_1 \geqslant |L|,$$

as $|T'| = |T| - |L| = n - m_1 - |L|$. Then Lemma 3.2.10 implies we can finish the embedding of L and thus finish the proof. □

3.3 Facts about random graphs

The following result is one of a series of random analogues of extremal results proved, independently, by Conlon and Gowers [20] and by Schacht [69].

Theorem 3.3.1. *For every $r \geqslant 2$ and $\varepsilon > 0$, there exist positive numbers C' and δ such that if $p \geqslant C'N^{-2/(r+2)}$ then a.a.s. the following holds. Every K_{r+1}-free subgraph G of $G(N, p)$ with*

$$e(G) \geqslant \left(1 - \frac{1}{r} - \delta\right) p \binom{N}{2}$$

can be made r-partite by removing at most $\varepsilon p N^2$ edges.

Definition 3.3.2. *Let $\eta, p \in (0, 1)$. We say that an n-vertex graph G is (η, p)-uniform, if all disjoint sets $A, B \subseteq V(G)$ with $|A|, |B| \geqslant \eta n$ satisfy*

$$(1 - \eta)p|A||B| \leqslant e_G(A, B) \leqslant (1 + \eta)p|A||B| \tag{3.6}$$

and

$$(1 - \eta)p\binom{|A|}{2} \leqslant e_G(A) \leqslant (1 + \eta)p\binom{|A|}{2}. \tag{3.7}$$

Furthermore, we say that G is (η, p)-upper-uniform if (possibly) only the upper bounds in (3.6) and (3.7) hold for all $A, B \subseteq V(G)$ as above.

The proof of the following lemma is a straightforward application of Chernoff's inequality and union bound and therefore we omit it.

Lemma 3.3.3. *For every $\eta > 0$ there exists $C > 0$ such that if $p \geqslant C/N$ then a.a.s. $G(N, p)$ is (η, p)-uniform.*

In particular, since any spanning subgraph of an (η, p)-uniform graph is (η, p)-upper-uniform, then, with high probability, every spanning subgraph of $G(N, p)$ is (η, p)-upper-uniform, as long as $p \geqslant C/N$.

Lemma 3.3.4. *For every $\gamma > 0$, $G = G(N,p)$ a.a.s satisfies the following properties.*

(*i*) *For every set $U \subseteq V$ with $|U| \geqslant \gamma N$, there are at most $64/\gamma p$ vertices in V with less than $\gamma pN/8$ neighbours in U.*

(*ii*) *For every $c > 0$, there exists $0 < c' < 1$ such that G is a weak $(c/p, c'N)$-expander. Moreover, $c' \to 0$ as $c \to \infty$.*

Proof. To prove (i), let us fix U with $|U| \geqslant \gamma N$ and let W be the set of vertices with less than $\gamma pN/8$ neighbours in U. Note that W is not necessarily disjoint from U, but $e(W \cap U) \leqslant \gamma pN|W \cap U|/8$ and Lemma 3.3.3 gives that either $|W \cap U| = o(N)$ or

$$e(W \cap U) \geqslant p\frac{|W \cap U|^2}{4}.$$

In both cases, we have that $|W \cap U| \leqslant \gamma N/2$ and $|U \setminus W| \geqslant \gamma N/2$, which let us bound the size of W, since

$$e(W, U \setminus W) < \frac{\gamma pN}{8}|W| \leqslant \frac{p}{4}|U \setminus W| \cdot |W|,$$

which will be shown to be unlikely if $|W| \geqslant 64/\gamma p$. Indeed, for a fixed pair of disjoint sets U', W', Chernoff's inequality states that

$$\mathbb{P}\left(e(U', W') \leqslant \frac{p|U'||W'|}{2}\right) \leqslant \exp\left(-\frac{1}{8}p|U'||W'|\right),$$

which implies, by the union bound over all pairs U', W' satisfying $|U'| \geqslant \gamma N/2$ and $|W'| \geqslant 64/\gamma p$, that the probability that one of them have the wrong number of edges is less than

$$2^N \cdot 2^N \cdot \exp\left(-\frac{p}{8} \cdot \frac{\gamma N}{2} \cdot \frac{64}{\gamma p}\right) \leqslant \exp\left(-2N\right).$$

Finally, the proof of (ii) follows from the fact that the set of non-neighbours of a fixed set U, with $|U| = c/p$, is dominated by a variable $X =^d \mathrm{Bin}(N, e^{-c})$, by the fact that $1 - p \leqslant e^{-p}$. Another application of Chernoff's inequality yields

$$\mathbb{P}\left(X \geqslant (2/e)^c N\right) \leqslant \exp\left(-\Omega(N)\right),$$

that finishes the proof, since the number of such sets is less then $\exp(c \log n/p)$. $\square$

Lemma 3.3.5. *Let $r \geqslant 1$ and let $G = G(N,p)$, with $p \gg N^{-2/(r+1)}$. Fix a disjoint collection $V_1, \ldots, V_{r+1} \subseteq V(G)$, with $|V_i| = m_i$. Then the probability that $V_1, \cdots, V_{r+1}$ spans a canonical copy of K_{r+1} is at least*

$$1 - \exp\left(-\Omega\left(p^{\binom{r+1}{2}} \prod_{i=1}^{r+1} m_i\right)\right)$$

In particular, there exists a constant $C > 0$ such that if an integer m satisfies

$$m^{r+1}p^{\binom{r+1}{2}} \geqslant C \log \binom{N}{m},$$

then with high probability there exists a canonical copy of K_{r+1} in every collection of $r+1$ disjoint m-sets.

Lemma 3.3.6. *For every $\gamma > 0$ there exists $C' > 0$ such that if $p \geqslant C'N^{-2/(r+2)}$, then $G = G(N,p)$ with high probability has the following property. For every $v \in V(G)$ and any r disjoint sets $W_1, \dots, W_r \subseteq N(v)$, with $|W_i| \geqslant \gamma pN$ for each $i \in [r]$, there exists a copy of K_{r+1} containing v and one vertex in each W_i, for $i \in [r]$.*

Proof. The proof follows by union bound and an application of Lemma 3.3.5. First, we may restrict ourselves to the case where each set W_i is of size $\Theta(pN)$, since with high probability every vertex $v \in V(G)$ has at least $2pN$ neighbours. Therefore, the probability that the event considered fails is at most

$$\sum_{w_1,\cdots,w_r} \prod_{i=1}^{r} \binom{N}{w_i} p^{w_i} \exp\left(- \prod_{i=1}^{r} p^{-\binom{r}{2}} \right) \leqslant$$
$$\sum_{w_1,\cdots,w_r} \prod_{i=1}^{r} \left(\frac{epN}{w_i} \right)^{w_i} \exp\left(-\Omega\left(p^{\binom{r+1}{2}} N^r \right) \right).$$

Since we are only considering sets of size $w_i = \Theta(pN)$, then last expression is at most

$$\exp\left(O(pN) - \Omega\left(p^{\binom{r+1}{2}} N^r \right) \right),$$

which goes to 0 as n tends to infinity, as long as C' is chosen appropriately. $\square$

3.4 Proof of Theorem 1.2.2

The proof of Theorem 1.2.2 follows by applying Proposition 3.4.2 $r+1$ times. For an appropriate choice of m_1 and m_2 there will be two possibilities. If the red graph is a weak (m_1, m_2)-expander, then we show that it is $\mathcal{T}(n, D)$-universal, using Theorem 3.2.1. Otherwise it will contain two disjoint sets of size m_1 and m_2 with all edges in between coloured in blue. We repeat this argument r times in the induced graph on the set with m_2 vertices. At the end, if the red graph is not $\mathcal{T}(n, D)$-universal, then we get $r+1$ disjoint sets, each of size m_1, with all the edges in between coloured in blue. This reasoning is made precise in the proof of the following lemma.

Lemma 3.4.1. *Let n, m, r, D be positive integers and let H be a graph on $N = rn + 10Drm$ vertices. Then one of the following holds:*

1. *H is $\mathcal{T}(n, D)$-universal.*

2. *There are disjoint sets $U_1, \ldots, U_{r+1} \subseteq V(H)$, each of size m, such that $e(V_i, V_j) = 0$ for $1 \leqslant i < j \leqslant r+1$.*

Before proving Lemma 3.4.1 we deal with the technical part of showing the weak expanders are almost expanders.

Proposition 3.4.2. *Let D, m_1, m_2 be integers and let $H = (V, E)$ be a graph with $|V| \geqslant m_2 + (2D+2)m_1$. If H is a weak (m_1, m_2)-expander, then there exists a set $V' \subseteq V$, with $|V \setminus V'| \leqslant m_1$, such that $H[V']$ is a (m_1, m_2)-expander.*

Proof. Take a maximal set $Z \subseteq V$ with $|Z| < m_1$ and $|N(Z)| \leqslant D|Y|$, and set $V' = V \setminus Z$. We will prove that for any $X \subseteq V'$ with $|N(X) \cap V'| \leqslant D|X|$ we have that $|X| > m_1$, which shows that $H[V']$ is a (m_1, m_2)-expander. For such X we have that

$$|N(Z \cup X)| < D|Z \cup X|,$$

since we are only counting external neighbours of $Z \cup X$. By the maximality of Z, we conclude that $|Z \cup X| \geqslant m_1$. Since H is a weak (m_1, m_2)-expander, then there are less than m_2 non-neighbours of $Z \cup X$ in V'. Therefore

$$\begin{aligned} D|X| &\geqslant |N(X) \cap V'| \\ &\geqslant |N(Z \cup X) \cap V'| - |N(Z) \cap V| \\ &> |V'| - m_2 - |X| - D|Z| \\ &\geqslant |V| - m_2 - (D+1)m_1 - |X| \geqslant (D+1)m_1 - |X|, \end{aligned}$$

which implies that $|X| > m_1$ and finishes the proof. $\square$

Now we move to the proof of Lemma 3.4.1.

Proof of Lemma 3.4.1. We assume that H is not $\mathcal{T}(n, D)$-universal and set $V_0 = V(H)$. We will prove that for $s \in [r]$ there exist disjoint sets U_s, V_s, with

$$|U_s| = m \quad \text{and} \quad |V_s| = (r-s)n + (r-s+1)5Dm,$$

such that $e(U_s, V_s) = 0$ and $U_s, V_s \subseteq V_{s-1}$. Indeed, if this is true, we set $U_{r+1} = V_r$ and get that $e(U_i, U_j) = 0$ for every $1 \leqslant i < j \leqslant r+1$, which is what we want to prove.

Suppose we have sets $V_0, U_1, V_1, \cdots, U_s, V_s$ as above for $s \in [r]$, or just V_0 for $s = 0$. Let $m_s = (r-s-1)n + (r-s)5Dm$. We show that if $H[V_s]$ were a weak (m, m_s)-expander, then it would be $\mathcal{T}(n, D)$-universal, which we assumed not to be true. To prove that, first we check that

$$|V_s| - m_2 \geqslant n + 5Dm.$$

In particular, $|V_s| \geqslant (D+2)m + m_s$, which is the requirement to apply Proposition 3.4.2. Therefore, there exists $V_s' \subseteq V_s$ such that $|V_s \setminus V_s'| \leqslant m$ and $H[V_s']$ is (m, m_s)-expander. As reasoned in (3.2), for sets $X \subseteq V_{r-1}$, with $m \leqslant |X| \leqslant 2m$, the (m, m_s)-expansion implies that

$$\begin{aligned} |N(X) \cap V_s'| &\geqslant |V_s'| - m_2 - |X| + 1 \\ &\geqslant |V_s| - m - m_2 - 2m + 1 \\ &\geqslant n + 5Dm - 3m + 1 \\ &\geqslant n + 2Dm + 1 \geqslant n + D|X| + 1. \end{aligned}$$

The above inequality and the first property of (m, m_s)-expansion imply that $H[V_i']$ is $\mathcal{T}(n, D)$-universal, by Theorem 3.2.1. □

Lemma 3.4.1 reduces the proof of Theorem 1.2.2 to finding the minimum value m such that with high probability every collection of $r+1$ disjoint m-sets span in $G(N, p)$ a copy of K_{r+1} with one vertex in each set, which we will call a *canonical copy.* Such value of m is determined by Lemma 3.3.5, but we repeat it here for convenience.

$$m^{r+1} p^{\binom{r+1}{2}} \geqslant C \log \binom{N}{m}, \tag{3.8}$$

Now we may state a stronger version of Theorem 1.2.2, with $t = O(m)$ and m satisfying (3.8). Note that when $t = \Omega(N)$, (3.8) is equivalent to say that $p \geqslant CN^{-2/(r+1)}$, for some $C > 0$.

Theorem 3.4.3. *For every $r, D \geqslant 2$ and for every $p = p(n)$ and m satisfying* (3.8)*, if*

$$N \geqslant rn + 10Drm,$$

then $G(N, p) \to \big(K_{r+1}, \mathcal{T}(n, D)\big)$ with high probability.

Proof. Let $G = G(N, p)$, where $N = rn + 10Drm$, and consider the even in which every collection of $r+1$ disjoint sets of size m span a canonical copy of K_{r+1}. By Lemma 3.3.5 and the hypothesis on m, this happens with high probability. Let G_R, $G_B \subseteq G$ be the red and blue graphs in a given edge colouring of G. By Lemma 3.4.1, if G_R is not $\mathcal{T}(n, D)$-universal, then there are disjoint sets $U_1, \dots, U_{r+1}$ of size m such that $e_R(U_i, U_j) = 0$ for all $1 \leqslant i < j \leqslant r+1$. In other words, all the edges in between these sets are coloured blue, which spans a blue copy of K_{r+1}, by the choice of m. □

3.5 Proof of Theorem 1.2.1

The proof of Theorem 1.2.1 follows from the following stability result.

Theorem 3.5.1. *For every $r, D \geqslant 2$ there exist $\delta, C, C' > 0$ such that if $N \geqslant (1-\delta)rn$ and $p \geqslant C'N^{-2/(r+2)}$, then $G = G(N,p)$ with high probability has the following property. For every blue-red colouring of $E(G)$, at least one of the following holds:*

a) G contains a blue copy of K_{r+1}.

b) G contains a red copy of every $T \in \mathcal{T}(n, D)$.

c) There exists a partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$, with $|V_0| \leqslant C/p$ and $|V_i| \leqslant n + C/p$ for each $i \in [r]$, and such that all edges of $G[V_i, V_j]$ are coloured in blue for each $1 \leqslant i < j \leqslant r$.

Note that Theorem 3.5.1 implies Theorem 1.2.1, as c cannot occur if $N > rn + (r+1)C/p$. As an intermediate step towards Theorem 3.5.1, we will provide a rough structure of a colouring of a typical outcome of $G(n,p)$ by combining Theorems 1.1.1 and 3.3.1.

Proposition 3.5.2. *For every $\alpha, \varepsilon > 0$ and integers $r, D \geqslant 2$, there exist $C', \delta > 0$ such that if $N \geqslant (1-\delta)rn$ and $p \geqslant C'N^{-2/(r+2)}$, then $G = G(N,p)$ has, with high probability, the following property. For every blue-red colouring of $E(G)$, at least one of the following holds:*

a) G contains a blue copy of K_{r+1}.

b) G contains a red copy of every $T \in \mathcal{T}(n, D)$.

c) There exists a partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$ such that $|V_0| \leqslant \alpha n$ and for each $i \in [r]$ we have $\big||V_i| - n\big| \leqslant \alpha n$ and $e_B(V_i) \leqslant \varepsilon p N^2$.

Proof. Without loss of generality, we may ask that ε is small enough for calculations. Let C' and δ' be the numerical outputs from Theorem 3.3.1 with inputs ε and r. Let $\delta = \alpha/(2r^2)$, $\varrho = 1/r + 2\delta$, $N \geqslant (1-\delta)rn$ and $p \geqslant C'N^{-2/(r+2)}$. Since $p \gg 1/N$, Theorem 1.1.1 implies that, with high probability, if $e(G_R) \geqslant (\varrho + \delta')e(G)$ then G_R contains all trees with maximum degree D and $\varrho N \geqslant n$ edges, and thus we may assume that

$$e(G_B) \geqslant \left(1 - \frac{1}{r} - \delta'\right) e(G).$$

Theorem 3.3.1 implies that, with high probability, all K_{r+1}-free subgraphs of G with this many edges are $\varepsilon p N^2$-close to being r-partite. Therefore, we may assume that there exists a partition $V(G) = W_1 \cup \cdots \cup W_r$ such that $e_B(W_i) \leqslant \varepsilon p N^2$ for each $i \in [r]$.

Since $p \gg 1/N$, we may also rule out the event in which G is not (η, p)-uniform for some $0 < \eta \ll \alpha$.

Claim 3.5.3. *In the events considered above, for each $i \in [r]$ the following holds. If $|W_i| \geqslant N/2r$, then there exists $V_i \subseteq W_i$, with $|W_i \setminus V_i| \leqslant \eta N$, such that $G_R[V_i]$ is a $(\eta N, \eta N, D)$-expander.*

Proof of Claim 3.5.3. We prove first that $G_R[W_i]$ is a weak $(\eta N, \eta N)$-expander. Since G is (η, p)-uniform, then for every pair of disjoint sets $X, Y \subseteq V(G)$, with $|X|, |Y| \geqslant \eta N$, we have

$$e_R(X, Y) = e(X, Y) - e_B(X, Y) \geqslant \frac{p}{2}|X||Y| - \varepsilon p N^2 > 0,$$

as long as $2\varepsilon < \eta^2$. Since $|W_i| \geqslant (D + 3)\eta N$, provided η is small enough, we may apply Proposition 3.4.2 to find a set $V_i \subseteq W_i$, with $|W_i \setminus V_i| \leqslant \eta N$, such that $G_R[V_i]$ is an $(\eta N, \eta N, D)$-expander. $\square$

For each $i \in [r]$ such that $|W_i| \geqslant N/2r$, by Claim 3.5.3 we know that $G_R[V_i]$ is an $(\eta N, \eta N, D)$ -expander and then for all $X \subseteq V_i$, with $\eta N \leqslant |X| \leqslant 2\eta N$, we have

$$|N_R(X) \cap V_i| \geqslant |V_i| - \eta N - |X| + 1 \geqslant (|V_i| - 3D\eta N) + D|X| + 1.$$

Suppose that V_1 is the largest of the V_i's and note that $|W_1| \geqslant |V_1| \geqslant N/r - \eta N \geqslant N/2r$. Therefore, if $G_R[V_1]$ is not $\mathcal{T}(n, D)$-universal, then Theorem 3.2.1 implies that $|V_i| \leqslant |V_1| \leqslant n + 3D\eta N$ for all $i \in [r]$. Set $V_0 = V(G) \setminus (V_1 \cup \cdots \cup V_r)$ and choose η small enough so that

$$|V_0| \leqslant \frac{\alpha n}{2r} \quad \text{and} \quad |V_i| \leqslant \left(1 + \frac{\alpha}{r}\right) n$$

for each $i \in [r]$. To finish the proof we only need to show that $|V_i| \geqslant (1 - \alpha)n$ for each $i \in [r]$. We suppose without loss of generality that $|V_r| < (1 - \alpha)n$. Then there exists $j \in [r - 1]$ such that

$$|V_j| \geqslant \frac{N - |V_r| - |V_0|}{r - 1} > \frac{1}{r - 1}\left((1 - \delta)rn - (1 - \alpha)n - \frac{\alpha n}{2r}\right) \geqslant \left(1 + \frac{\alpha}{r}\right) n,$$

which is a contradiction and thus $\big||V_i| - n\big| \leqslant \alpha n$ for all $i \in [r]$. $\square$

Now we push the stability even further. It is convenient to relate expansion properties of the red graphs on each part solely to the red and blue degrees inside that part. We prove that if a set induces a graph with high minimum red degree and roughly the expected codegree, then it satisfies property 1 of expansion.

Lemma 3.5.4. *For every $C, \gamma > 0$ there exists $\gamma' > 0$ such that the following holds for $p = \omega(\log N/N)$. Let G be an N-vertex graph such that for all $u, v \in V(G)$ we have $d(u) \geqslant \gamma p N$ and $|N(u) \cap N(v)| \leqslant 2p^2 N \log N$. Then for every $X \subseteq V(G)$, with $1 \leqslant |X| \leqslant C/p$, we have $|N(X)| \geqslant \gamma' p N|X|/\log N$.*

Proof. For $X \subseteq V(G)$ with $1 \leqslant |X| \leqslant C/p$, take a subset $X' \subseteq X$ with $1 \leqslant |X'| \leqslant \gamma/(4p\log N)$. By inclusion-exclusion, $|N(X)|$ is at least

$$\begin{aligned}\sum_{u\in X'} |N(u)| - \sum_{v\neq w} |N(v)\cap N(w)| - |X| &\geqslant \gamma pN|X'| - |X'|^2 \cdot (2p^2 N \log N) - |X| \\ &\geqslant \gamma pN|X'| - \frac{\gamma pN}{2}|X'| - |X| \\ &\geqslant \Omega\left(\frac{pN}{\log N}\right)|X|,\end{aligned}$$

where in the last inequality we used that $pN = \omega(\log N)$. □

Definition 3.5.5. *Let $\varepsilon > 0$ and let $r, D \geqslant 2$ be integers. For a blue-red coloured N-vertex graph G, we say that a partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$ is ε-good if for every $i \in [r]$*

a) $|V_i| \geqslant (1 - 1/2D)N/r$,

b) $d_R(v, V_i) \geqslant pN/32r$ for every $v \in V_i$, and

c) $d_B(v, V_i) \leqslant \varepsilon pN$ for every $v \in V_i$.

We will prove now that for any ε-good partition of $V(G(N,p))$ we have that $e_R(V_i, V_j) = 0$ for all $1 \leqslant i < j \leqslant r$. First, we prove that $G_R[V_i]$ is an expander for each $i \in [r]$. Thus, by Haxell's theorem (Theorem 3.2.1), we can embed any tree of size $(1 - o(1))n$ into any of the V_i's. Suppose there is a red edge between V_i and V_j. We may split any given tree $T \in \mathcal{T}(n, D)$ in two trees T_1 and T_2, connected by an edge and both having at most $(1 - 1/D)n$ vertices. Then, we can embed T_1 into V_i and T_2 into V_j, and complete the embedding of T using the red edge between V_i and V_j.

Using this fact we can prove that $G[V_i]$ has even stronger expansion properties. That is, for each $i \in [r]$ we may show that every pair of large disjoint subsets of V_i always have at least one red edge in between. Indeed, if for some $i \in [r]$ there exist a pair of disjoint sets $X, Y \subseteq V_i$ each of size $\Theta(N/\log^4 N)$ and no red edges in between, then, with high probability, X and Y and the remaining V_j's would span a canonical blue-copy of K_{r+1}. Combining this information with results of Section 3.2, we show that $G_R[V_i]$ is $\mathcal{T}(|V_i| - C/p, D)$-universal for every $i \in [r]$.

Proposition 3.5.6. *For integers $r, D \geqslant 2$ there exist $C, C', \delta, \varepsilon > 0$ such that if $N \geqslant (1-\delta)rn$ and $p \geqslant C'N^{-2/(r+2)}$, then $G = G(N,p)$ has, with high probability, the following property. For every blue-red colouring of $E(G)$ that admits an ε-good partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$, at least one of the following holds:*

a) G contains a blue copy of K_{r+1}.

b) G contains a red copy of every $T \in \mathcal{T}(n, D)$.

c) For every $1 \leqslant i < j \leqslant r$ we have $e_R(V_i, V_j) = 0$. Moreover, for each $i \in [r]$ the graph $G_R[V_i]$ is $\mathcal{T}(|V_i| - C/p, D)$-universal.

Proof. Assume that neither a nor b hold. For $\alpha = 1/32D$, we take C from Lemma 3.3.4 so that, with high probability, G is a weak $(C/p, \alpha N/4r)$-expander, and set $\varepsilon = \alpha/(6CD)$. Moreover, there exists a constant C' such that if $p \geqslant C' N^{-1/2}$, then, with high probability, every pair of vertices in G has at most $2p^2 N \log N$ common neighbours. Finally, because of the first property of the ε-good partition, we deduce that $N \leqslant 2r|V_i|$. Our first goal is to prove that each V_i satisfies the hypothesis of Theorem 3.2.1 in order to show that $G_R[V_i]$ is $\mathcal{T}((1 - 1/D)n, D)$-universal. For $i \in [r]$, we apply Lemma 3.5.4 to $G_R[V_i]$, with parameters $\gamma = 1/32r$ and C, so that for every $X \subseteq V_i$, with $1 \leqslant |X| \leqslant C/p$, we have

$$|N_R(X) \cap V_i| = \Omega\left(\frac{pN}{\log N}\right)|X| \geqslant D|X| + 1. \tag{3.9}$$

For $X \subseteq V_i$, with $C/p \leqslant |X| \leqslant 2C/p$, since G is a weak $(C/p, \alpha N/4r)$-expander we have

$$|N_R(X) \cap V_i| \geqslant |V_i| - \frac{\alpha N}{4r} - \varepsilon p N|X| - |X| \geqslant (1 - \alpha)\,|V_i| + D|X| + 1. \tag{3.10}$$

Since $\alpha \leqslant 1/D$, then $(1 - \alpha)|V_i| \geqslant (1 - 1/D)n$, and thus we may use Theorem 3.2.1 on each $G_R[V_i]$ in order to find trees of size $(1 - 1/D)n$ and maximum degree at most D.

Given a tree $T \in \mathcal{T}(n, D)$, there exists a cut edge $u_1u_2 \in E(T)$ which splits T into two trees T_1 and T_2, both with at least n/D vertices and, consequently, at most $(1 - 1/D)n$ vertices (see [9, Lemma 2.5]). Suppose that exists a red edge v_1v_2 between two different parts, say $v_1 \in V_1$ and $v_2 \in V_2$. By Theorem 3.2.1, we may find an embedding of T_i in $G_R[V_i]$ that maps u_i to v_i, for $i \in \{1, 2\}$, and thus, together with the red edge v_1v_2, yield an embedding of T. Therefore, there are no red edges between different parts. Now we move to prove the second part of c.

Set $d = D\log^4 n/20$. We will show now that $G_R[V_i]$ is an $(|V_i|/2d, |V_i|/2d, d)$-expander for each $i \in [r]$. Indeed, given $X \subseteq V_i$, with $1 \leqslant |X| \leqslant C/p$, by (3.9) we get $|N_R(X) \cap V_i| \geqslant d|X| + 1$. For $C/p \leqslant |X| \leqslant |V_i|/2d$, by (3.10) we have that

$$|N_R(X) \cap V_i| \geqslant (1 - \alpha)|V_i| - |X| \geqslant d|X| + 1,$$

as $\alpha < 1/2$. To show the second expansion property, suppose that there exists a pair of disjoint sets $X, Y \subseteq V_i$, with $|X| = |Y| = |V_i|/2d$, such that $e_R(X, Y) = 0$. By Lemma 3.3.5, with high probability there is a copy of K_{r+1} with one vertex in each of the sets X, Y and the V_j's with $j \neq i$ (we can apply Janson's inequality since $|V_i|/2d =$

$\Omega(N/\log^4 N))$. This is a contradiction and therefore $G_R[V_i]$ is an $(|V_i|/2d, |V_i|/2d, d)$-expander. Now, Theorem 3.2.3 implies that $G_R[V_i]$ contains all spanning trees with maximum degree bounded by D and at most $|V_i|/d$ leaves.

For trees with at least $|V_i|/d$ leaves, we know that $G_R[V_i]$ is a weak $(|V_i|/2d, |V_i|/2d)$-expander, and so we only need to show that it is also a weak $(C/p, |V_i|/32D)$-expander. But this is already guaranteed by (3.10) since $\alpha \leqslant 1/32D$. Now, Theorem 3.2.4 implies that $G_R[V_i]$ is $\mathcal{T}(|V_i| - C/p, D)$-universal. □

Now we are ready to prove Theorem 3.5.1.

Proof of Theorem 3.5.1. We apply Proposition 3.5.6, with parameters r and D, to get $\delta_1, \varepsilon, C, C_1'$, and let $\alpha \leqslant 1/6D$ be sufficiently small. Without loss of generality, we assume that $0 < \varepsilon \leqslant \alpha/r$ and apply Proposition 3.5.2, with parameters $\varepsilon^2/4$ and α, to get C_2' and δ_2. Let C_3' be given by Lemma 3.3.6 and set $C_4' = 10^5 r^2$. Finally, we set $\delta = \min\{\delta_1, \delta_2\}$ and $C' = \max\{C_1', C_2', C_3', C_4'\}$, and consider $N \geqslant (1-\delta)rn$ and $p \geqslant C' N^{-2/(r+2)}$.

By Proposition 3.5.2, with high probability, if $K_{r+1} \nsubseteq G_B$ and if G_R is not $\mathcal{T}(n, D)$-universal, then there exists a partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$ such that $|V_0| \leqslant \alpha n$, and for each $i \in [r]$ we have $\big||V_i| - n\big| \leqslant \alpha n$ and $e_B(V_i) \leqslant \varepsilon^2 p N^2/4$. We want to define a new partition by removing from each V_i a set of "bad" vertices. First, for $i \in [r]$ let B_i be the set of those vertices $v \in V_i$ having at least $\varepsilon p N$ blue neighbours in V_i and set $B = B_1 \cup \cdots \cup B_r$. Secondly, let B' be the set of those vertices $v \in V(G)$ such that $d(v, V_i \setminus B) \leqslant pN/16r$ for some $i \in [r]$.

Let $V(G) = W_0 \cup W_1 \cdots \cup W_r$ be the partition defined by $W_i = V_i \setminus (B \cup B')$ for $i \in [r]$ and $W_0 = V(G) \setminus (W_1 \cup \cdots \cup W_r)$. We will show that this partition is ε-good. Since $e_B(V_i) \leqslant \varepsilon^2 p N^2/4$, a double counting argument shows that $|B \cap V_i| \leqslant \varepsilon N/2$ and thus $|V_i \setminus B| \geqslant |V_i| - \varepsilon N/2 \geqslant (1-2\alpha)N/r$ as $\varepsilon \leqslant \alpha/r$. By Lemma 3.3.4, there are at most $128r/p$ vertices of G with less than $pN/16r$ neighbours in $V_i \setminus B$. Then we have

$$|W_i| \geqslant (1-2\alpha)\frac{N}{r} - \frac{128r^2}{p} \geqslant (1-3\alpha)\frac{N}{r} \geqslant \left(1 - \frac{1}{2D}\right)\frac{N}{r}.$$

By definition of W_i, each vertex $v \in W_i$ satisfies $d_B(v, W_i) \leqslant \varepsilon p N$. On the other hand, for $v \in W_i$ we have

$$d_R(u, W_i) \geqslant \frac{pN}{16r} - \varepsilon p N - \frac{128r^2}{p} \geqslant \frac{pN}{32r},$$

where we used that $\varepsilon \leqslant 1/20r$ and $pN \geqslant C_4/p$. To finish the proof, take an ε-good partition $V(G) = U_0 \cup U_1 \cup \cdots \cup U_r$ such that $W_i \subseteq U_i$ for $i \in [r]$ and that minimises $|U_0|$. We will prove that if $U_0 \nsubseteq B'$, then this partition would not be maximal. By contradiction, suppose there exists $u \in U_0 \setminus B'$. If $d_B(u, U_i) \geqslant \varepsilon p N$ for all $i \in [r]$, then by Lemma 3.3.6 we can find a blue copy of K_{r+1} containing u, which is not possible.

Then there must exist some $i \in [r]$ such that $d_R(u, U_i) \geqslant pN/32r$, in which case we update $U_i := U_i \cup \{u\}$. We claim that $V(G) = U_0 \cup U_1 \cup \cdots \cup U_r$ is still ε-good. Since the blue degree of each vertex in $U_i \setminus \{u\}$ grows in at most 1, it follows that the new partition is 2ε-good. This fact and Proposition 3.5.6 imply that $e_R(U_i, U_j) = 0$ for every $1 \leqslant i < j \leqslant r$. Finally, we may use Lemma 3.3.6 as before to show that the maximum blue degree inside each part is at most εpN, which makes this partition ε-good. This contradicts the maximality of the initial partition and thus $U_0 \subseteq B'$. In particular, we have $|U_0| \leqslant |B'| \leqslant 128r/p$. Note that if $|U_i| > (n + C/p)$ for some $i \in [r]$, then, by Proposition 3.5.6, $G_R[U_i]$ contains all trees with maximum degree at most D and $|U_i| - C/p \geqslant n$ edges, which is a contradiction. This finishes the proof. □

Chapter 4

Hamiltonian cycles in pseudorandom hypergraphs

4.1 Absorption Method

In [64], Rödl, Ruciński and Szemerédi introduced the Absorption Method, which turned out to be a very useful approach for embedding spanning cycles in hypergraphs. This method reduces the problem to finding an almost spanning cycle with a small special path in it, called the *absorbing path*. The absorbing path A can absorb any small set of vertices into a new bigger path, with the same ends as A, completing the almost spanning cycle into a Hamilton cycle.

The almost spanning cycle will be composed from smaller tight paths, which will be connected to longer paths. For that it would be useful if any given two pairs of vertices (x, y) and (w, z), being the ends of such smaller paths, can be connected by a short tight path. However, in view of the assumptions of Theorem 1.3.3, it is easy to see that not *any* pair of pairs can be connected in this way (in particular, there could be pairs with codegree zero). For that we introduce the following notion of *connectable pairs* and we will show that for those pairs there actually exist tight connecting paths between them (see Lemma 4.1.4 below).

Definition 4.1.1. *Let $H = (V, E)$ be a hypergraph. We say that $(x, y) \in V \times V$ is* β-connectable in H *if the set*

$$Z_{xy} = \{z \in V : xyz \in E(H) \text{ and } d(y, z) \geqslant \beta|V|\},$$

has size at least $\beta|V|$. Moreover, we say that an (a, b)-(c, d)-path is β-connectable *if the pairs (b, a) and (c, d) are β-connectable.*

Observe that the starting pair of the path is asked to be β-connectable in the inverse direction that as it appears in the path.

The proof of Theorem 1.3.3 splits into three lemmas. Let $H = (V, E)$ be a $(\rho, 1/4 + \varepsilon, \therefore)$-dense hypergraph on n vertices, with $1/n \ll \rho \ll \varepsilon$. First we prove that such hypergraphs can be almost covered by a collection of 'few' tight paths. We remark that this is even true under the weaker assumption of non-vanishing $\therefore$-density. A straight forward proof is presented in Section 4.3.

Lemma 4.1.2 (Almost Covering Lemma)**.** *For all $d, \gamma \in (0, 1]$ there exist $\rho, \beta > 0$, and n_0 such that in every $(\rho, d, \therefore)$-dense hypergraph H on $n \geqslant n_0$ vertices there exists a collection of at most $1/\beta$ disjoint β-connectable paths, that cover all but at most $\gamma^2 n$ vertices of H.*

Next we discuss how to find an absorbing path, which contains a collection of several smaller structures, called *absorbers*. For $v \in V$, we call $A_v \subseteq H$ an absorber for v if both A_v and $A_v \cup \{v\}$ span tight paths with same ends (we say that A_v absorbs v). The main difficulty is to define the absorbers in such a way that we can prove that every vertex is contained in many of them. In Section 4.5 we see that the absobers considered here are in fact more complicated and absorb sets of three vertices instead of one. This leads to a divisibility issue which we consider separately in Lemma 4.13. Going further, we can find a relatively small collection of tight paths which can absorb any sufficiently small given set of vertices. After finding this collection we connect them together to form one tight path with the absortion property described in the following lemma.

Lemma 4.1.3 (Absorbing Path Lemma)**.** *For every $\varepsilon > 0$ there exist $\rho, \beta, \gamma' > 0$ and n_0 such that the following is true for every positive $\gamma \leqslant \gamma'$ and every $(\rho, 1/4 + \varepsilon, \therefore)$-dense hypergraph $H = (V, E)$ on $n \geqslant n_0$ vertices with $\delta_1(H) \geqslant \varepsilon n^2$.*

For every $R \subseteq V$, with $|R| \leqslant 2\gamma^2 n$, there exists a tight β-connectable path A satisfying $V(A) \subseteq V \setminus R$ and $|V(A)| \leqslant \gamma n$, such that for every $U \subseteq V(H) \setminus A$ with $|U| \leqslant 3\gamma^2 n$, the hypergraph $H[V(A) \cup U]$ has a tight path with the same ends as A.

The set of vertices R in Lemma 4.1.3 will act as a reservoir of vertices that will be used later for connecting the tight paths mentioned in Lemmas 4.1.2 and 4.1.3, without interfering with the vertices already used by those tight paths.

The next lemma justifies Definition 4.1.1 and shows that between every two β-connectable pairs there exist several short tight paths connecting them. As it was said before, this is used for connecting the absorbers in the proof of Lemma 4.1.3. Moreover, observe that all tight paths mentioned in Lemma 4.1.2 and 4.1.3 are β-connectable. This allows us to connect them together into an almost spanning cycle and the absorbing path in this cycle will absorb all the remaining vertices to complete the Hamilton cycle.

Lemma 4.1.4 (Connecting Lemma). *For every $\varepsilon, \beta > 0$ there exist $\rho, \alpha > 0$ and n_0 such that for every $(\rho, 1/4 + \varepsilon, \therefore)$-dense hypergraph H on $n \geqslant n_0$ vertices the following holds.*

For every pair of disjoint ordered β-connectable pairs of vertices $(x, y), (w, z) \in V \times V$ there exists an integer $\ell \leqslant 15$ such that the number of (x, y)-(z, w)-paths with ℓ inner vertices is at least αn^{ℓ}

In view of the construction given in Example 1.3.2, one can see that the 1/4 in the $\therefore$-density assumption in Lemma 4.1.4 cannot be dropped. In that example, there are two classes of pairs that cannot be connected by a tight path (namely the pairs in G and in $\overline{G}$), although they are β-connectable. Hence, $\therefore$-density of at least 1/4 is required for Lemma 4.1.4.

Also Lemma 4.1.3 requires $\therefore$-density bigger than 1/4. In the proof of Lemma 4.1.3 this assumption will be crucial for connecting the so-called absorbers to a tight path, which makes use of Lemma 4.1.4. Moreover, the type of absorbers used here, leads to a 'divisibility issue'. It is addressed in Lemma 4.13 for which we also employ the same density assumption.

We now deduce Theorem 1.3.3 from Lemmata 4.1.2 − 4.1.4.

Proof of Theorem 1.3.3. Given $\varepsilon > 0$ we apply Lemma 4.1.3 and obtain ρ_1, β_1 and γ'. Lemma 4.1.2 applied with $d = 1/4$ and $\gamma = \min\{\gamma', \varepsilon/2\}$ yields ρ_2 and β_2. Applying Lemma 4.1.4 with ε and

$$\beta = \frac{1}{8}\min\{\beta_1, \beta_2\},$$

reveals α and ρ_3. Finally we set

$$\rho = \min\{\rho_1, \rho_2/8, \rho_3\},$$

and n be sufficiently large. Having fixed all constants, let H be a $(\rho, 1/4 + \varepsilon, \therefore)$-dense hypergraph on n vertices.

We consider a random set $R \subseteq V$, in which each vertex is present independently with probability γ^2. For every positive integer $\ell \leqslant 15$ consider two pairs $(x, y), (w, z) \in V \times V$ between which there are at least αn^{ℓ} paths with ℓ inner vertices. Let $Y = Y(\ell, (x, y), (z, w))$ count the number of such paths whose inner vertices are contained in R. We point out that Y is a function determined by n independent random variables, each of which can influence the value of Y by at most $n^{\ell-1}$. Therefore a standard application of Azuma's inequality (see [40]*Section 2.4) implies that

$$\mathbb{P}\left(Y \leqslant \frac{\gamma^{2\ell}}{2} \cdot \alpha n^{\ell}\right) = \exp(-\Omega(n)) < \frac{1}{2} \cdot \frac{1}{15n^4}, \tag{4.1}$$

for any fixed ℓ, (x,y), and (w,z). Moreover, by Markov's inequality we have that

$$\mathbb{P}\big(|R| \geqslant 2\gamma^2 n\big) \leqslant \frac{1}{2}. \tag{4.2}$$

Therefore there exists a realisation of R, which from now on will take over the name R, that is not in the event considered in (4.2) and in any of the events considered in (4.1) (all 4-tuples of vertices and values of ℓ). Since $\gamma' < \gamma$, $\rho < \rho_1$, and $|R| < 2\gamma^2 n$, Lemma 4.1.3 ensures that we can find a β_1-connectable absorbing path A of size smaller than γn and which does not intersect R.

Let $V' = V \smallsetminus (V(A) \cup R)$. Since $|V(A) \cup R| \leqslant 3\gamma n \leqslant n/2$, the induced hypergraph $H[V']$ is $(8\rho, 1/4+\epsilon, \therefore)$-dense. In particular, $H[V']$ is $(8\delta, 1/4+\varepsilon, \therefore)$-dense and since $8\rho \leqslant \rho_2$, Lemma 4.1.2 implies that there exists a collection of at most $1/\beta_2$ paths with β_2-connectable ends in $H[V']$ that cover all but at most $\gamma^2 n$ vertices.

Set $t = \lfloor 1/\beta_2 + 1 \rfloor$ and let $(P_i)_{i\in[t]}$ be any cyclic ordering of such paths together with the absorbing path. Assume that we were able to find connections in R between the paths $P_1, P_2, \ldots, P_i$, using inner vertices from R only. Moreover, we make sure that each connection is made with at most 15 inner vertices. Let C_i be the path that begins with P_1 and ends in P_i using those connections. Therefore

$$|V(C_i) \cap R| \leqslant t \cdot 15 = o(n).$$

Now, we want to show that we can connect P_i with P_{i+1} to construct C_{i+1}. Observe that all the tight paths from $(P)_{i\in[t]}$ are β-connectable. This follows from the choice $\beta \leqslant \beta_1$ for the absorbing path A. From the paths given by Lemma 4.1.2 we know that they are β_2-connectable in $H[V']$. Owing to $\beta \leqslant \beta_2/2$ and $|V'| \geqslant n/2$ the β-connectibility follows.

Let (x_i, y_i) be the ending pair of P_i and (z_i, w_i) the starting pair P_{i+1}. Lemma 4.1.4 implies that, for some $\ell_i \leqslant 15$, there exist at least αn^{ℓ_i} tight (x_i, y_i)-(z_i, w_i) paths, each with ℓ_i inner vertices. By the choice of R, the number of (x_i, y_i)-(z_i, w_i) paths of length ℓ_i+2 whose inner vertices lie in R is at least $\gamma^2 \alpha n^{\ell_i}/2$. Since at most $|V(C_i) \cap R| n^{\ell_i - 1} = o(n^{\ell_i})$ such paths contain a vertex from C_i, for sufficiently large n large enough we can find one tight path disjoint from C_i.

Finally, consider C_t the final cycle obtained in this process, by connecting P_t to P_1. Then, as C_t includes all the tight paths in the almost covering the number of vertices not covered by C_t is at most

$$|V \smallsetminus V(C_t)| \leqslant |R| + \gamma^2 n \leqslant 3\gamma^2 n.$$

This finishes the proof, since A can absorb these vertices into a new path with the same endings. $\square$

4.2 Preliminary results and basic definitions

In this section we collect some preliminary results and introduce the necessary notation. Given $\eta, d \in [0,1]$ and a bipartite graph $G = (V_1 \cup V_2, E)$ we say that G is (η, d)-*regular* if for every two sets of vertices $X \subseteq V_1$ and $Y \subseteq V_2$ we have

$$|e(X,Y) - d|X||Y|| \leqslant \eta |V_1||V_2|\,.$$

It is easy to see that every dense graph contains a linear sized bipartite regular subgraph, with almost the same density. That can be proved by a simple application of Szemerédi's Regularity Lemma or alternatively by a more direct density increment argument (see [55]).

Lemma 4.2.1. *For all η, $d > 0$ there exists some $\mu > 0$ such that for every n-vertex graph G with $e(G) \geqslant dn^2/2$, there exist disjoint subsets V_1, $V_2 \subseteq V(G)$, with $|V_1| = |V_2| = \lceil \mu n \rceil$ such that the bipartite induced subgraph $G[V_1, V_2]$ is (η, d')-regular for some $d' \geqslant d$.* □

For a hypergraph $H = (V, E)$ recall its *shadow* ∂H is the subset of $V^{(2)}$ of those pairs that are contained in some edge of H. For disjoint sets of vertices V_1, $V_2 \subseteq V$ with a slight abuse of notation we write $\partial H[V_1, V_2]$ for the set of ordered pairs in $V_1 \times V_2$ that correspond to unordered pairs in the shadow, i.e.,

$$\partial H[V_1, V_2] = \big\{(v_1, v_2) \in V_1 \times V_2 \colon \{v_1, v_2\} \in \partial H\big\}\,.$$

Given $\rho, d > 0$, a set of ordered pairs of vertices $P \in V^2$, and a subset $X \subseteq V$ we say that H is $(\rho, d, \therefore)$-*dense over* (X, P) if for every subset of vertices $X' \subseteq X$ and every subset of pairs $P' \subseteq P$ we have

$$e(X', P') \geqslant d\,|X'||P'| - \rho\,|X||P|\,,$$

which is a version of $\therefore$-density restricted to P and X. For the next lemma we also need the following concept of restricted vertex neighbourhood. Given a vertex $v \in V$ and a set of ordered pairs $P \in V^2$ we define its *neighbourhood restricted to* P by

$$N(v, P) = \{(x, y) \in P \colon vxy \in E\}\,.$$

Lemma 4.2.2. *Let $H = (V, E)$ be a hypergraph, $X \subseteq V$ be a set of vertices, and $P \subseteq V^2$. If H is $(\rho, d, \therefore)$-dense over (X, P) for some constants ρ, $d > 0$, then*

$$\big|\big\{x \in X \colon |N(x, P)| < (d - \sqrt{\rho})|P|\big\}\big| < \sqrt{\rho}\,|X|\,.$$

Proof. Let $X' \subseteq X$ be the vertices with less than $(d-\sqrt{\rho})|P|$ neighbour pairs in P. The definition of X' and the $(\rho, d, \therefore)$-density of H over (X, P) provide the following upper and lower bounds on $e(X', P)$

$$d|X'||P| - \rho|X||P| \leqslant e(X', P) \leqslant (d-\sqrt{\rho})|P| \cdot |X'|$$

and the desired bound on $|X'|$ follows. $\square$

The following result asserts that hypergraph contains subhypergraph with almost the same density and such that every pair of vertices with positive codegree has at least $\Omega(|V|)$ neighbours. This fact can be proved by removing iteratively the edges which contain a pair with small codegree and we omit the details.

Lemma 4.2.3. *For every $\beta > 0$ and every n-vertex hypergraph H there is a hypergraph $H_\beta \subseteq H$ on the same vertex set with $e(H_\beta) \geqslant e(H) - \beta n^3$ such that for every pair of vertices x, y either $d_{H_\beta}(x, y) = 0$ or $d_{H_\beta}(x, y) \geqslant \beta n$. In particular, if we have $d_{H_\beta}(x, y) > 0$, then (x, y) is β-connectable in H.* $\square$

Let F and F' be two hypergraphs. We say that F *contains a homomorphic copy of* F' if there is a function ϕ from $V(F')$ to $V(F)$ such that for every edge $xyz \in E(F')$ we have that $\phi(x)\phi(y)\phi(z) \in E(F)$. We denote this fact as $F' \xrightarrow{hom} F$ and we recall the following well known consequence from Erdős [22].

Lemma 4.2.4. *For every $\xi > 0$ and k, $\ell \in \mathbb{N}$ there is $\zeta > 0$ and $n_0 \in \mathbb{N}$ such that the following holds. Let F and F' be hypergraphs such that $|V(F)| = k$ and $|V(F')| = \ell$ and $F' \xrightarrow{hom} F$. If a hypergraph H on $n > n_0$ vertices contains at least ξn^k copies of F, then H contains ζn^ℓ copies of F'.* $\square$

We denote the hypergraph with four vertices and three edges by $K_4^{(3)-}$. We refer to the vertex of degree three as the *apex*. Glebov, Král, and Volec [35] showed that $\therefore$-density bigger than $1/4$ yields the existence of a, in fact of many copies of, $K_4^{(3)-}$.

Theorem 4.2.5 (Glebov, Král & Volec, 2016). *For every $\epsilon > 0$ there exist ρ and $\xi > 0$ such that every sufficiently large $(\rho, 1/4 + \epsilon, \therefore)$-dense n-vertex hypergraph contains ξn^4 copies of $K_4^{(3)-}$.* $\square$

4.3 Almost covering

In this section we present a very straightforward proof of Lemma 4.1.2.

Proof of Lemma 4.1.2. Given $d, \gamma > 0$ take β and ρ such that

$$\beta = \rho = \frac{d\gamma^6}{13}.$$

We show that a maximal collection of β-connectable tight paths, each of which having at least βn vertices, must cover all but at most $\gamma^2 n$ vertices. We do that by showing that in every set $X \subseteq V(H)$ with at least $\gamma^2 n$ vertices there exists a β-connectable tight path of size βn. Indeed, the $(\rho, d, \therefore)$-density implies that in such a set X, we have

$$e(X) \geqslant \frac{d|X|^3}{6} - \rho n^3,$$

where we discounted the ordering of triples. In $H[X]$ we remove, iteratively, every edge that contains an (unordered) pair of vertices with codegree smaller than βn. In this way, we remove at most βn^3 edges and get a hypergraph with at least

$$\begin{aligned} e(X) - \beta n^3 &\geqslant \frac{d|X|^3}{6} - \rho n^3 - \beta n^3 \\ &\geqslant \left(\frac{d\gamma^6}{6} - \rho - \beta\right) n^3, \end{aligned}$$

edges. Owing to the choice of β and γ this hypergraph is not empty. Now a tight path with βn vertices can be found in a greedy manner. Moreover, if (x, y) is a pair contained in such path, then we have that the set

$$Z_{xy} = \{z \in V : xyz \in E \text{ and } d(y, z) \geqslant \beta n\}$$

has at least βn vertices. □

4.4 Connecting Lemma

We dedicate this section to prove the Connecting Lemma (Lemma 4.1.4). The proof splits into several lemmata. The Connecting Lemma asserts that every ordered connectable pair can be connected to any other ordered connectable pair. In a first step in Lemmata 4.4.1 and 4.4.3 we show that there are many connections between large sets of unordered pairs (without specifying the order of the ending pairs). In fact, these connection can be achieved by paths consisting of only two edges, which we refer to as *cherries* (see Definition 4.4.2 below). On the price of extending the length by at most two, in Lemma 4.4.4 we establish that one can even fix the order of one of the sets of given pairs. On the other hand, this is complemented by Lemma 4.4.7 showing that there are many pairs of unordered pairs that can be connected in any orientation. We call such pairs of pairs *turnable* (see Definition 4.4.5 below).

For the proof of the Connecting Lemma we can now start with any given connectable pair (x, y) and move to its second neighbourhood, which is a large set of ordered pairs. From that set we shall reach many turnable pairs. Similarly, from any given ending pair (z, w) we also reach many turnable pairs. These paths give the turnable pairs an orientation, but since the turnable pairs can be connected in any orientation, we find the desired tight (x, y)-(z, w)-paths. The detailed presentation of this argument renders the proof of the Connecting Lemma, which we defer to the end of this section.

Lemma 4.4.1. *For all ξ, $\epsilon \in (0, 1]$ there exist η, $\rho > 0$ such that the following holds for sufficiently large m.*

Suppose V_1, V_2, V_3 are pairwise disjoint sets of size m and suppose $G = (V_1 \cup V_2, P)$ is an (η, ξ)-regular bipartite graph. If $H = (V_1 \cup V_2 \cup V_3, E)$ is a 3-partite hypergraph that is $(\rho, 1/4 + \epsilon, \therefore)$-dense over (V_3, P), then

$$\left|\partial H[V_1, V_3]\right| + \left|\partial H[V_2, V_3]\right| \geqslant (1 + \varepsilon)\, m^2 .$$

Proof. Given ξ and ε we set

$$\rho = \left(\frac{\varepsilon}{21}\right)^2 \quad \text{and} \quad \eta \leqslant \frac{\xi\varepsilon}{36} .$$

Let $G = (V_1 \cup V_2, P)$ and $H = (V_1 \cup V_2 \cup V_3, E)$ be given. Since G is bipartite we may view P as a subset of $V_1 \times V_2$ and, hence, as a set of ordered pairs. Lemma 4.2.2 applied to V_3 and P ensures for the hypergraph H that there are at most $\sqrt{\rho}m$ vertices in V_3 with less than $(1/4 + \epsilon - \sqrt{\rho})|P|$ neighbour pairs in P. We remove such vertices from V_3 and let V_3' be the resulting subset of V_3.

Consider a fixed vertex $v_3 \in V_3'$. By the definition of V_3', we have

$$|N(v_3, P)| \geqslant \left(\frac{1}{4} + \epsilon - \sqrt{\rho}\right)|P| \geqslant \left(\frac{1}{4} + \frac{15}{16}\epsilon\right)|P| . \tag{4.3}$$

For $i = 1$, 2 we consider the neighbourhood of v_3 in $\partial H[V_i, V_3]$ defined by

$$N_i(v_3) = \big\{v_i \in V_i \colon (v_i, v_3) \in \partial H[V_i, V_3]\big\}$$

and note that

$$|N(v_3, P)| \leqslant e_G\big(N_1(v_3), N_2(v_3)\big) .$$

Consequently, the (η, ξ)-regularity of G yields

$$|N(v_3, P)| \leqslant \xi|N_1(v_3)||N_2(v_3)| + \eta m^2 . \tag{4.4}$$

Combining (4.3) and (4.4) with the upper bound on $|P|$ provided by the regularity of G we obtain

$$4\xi|N_1(v_3)||N_2(v_3)| \geqslant \Big(1+\frac{15}{4}\epsilon\Big)|P| - 4\eta m^2 \geqslant \Big(1+\frac{15}{4}\epsilon\Big)(\xi-\eta)m^2 - 4\eta m^2 \geqslant \Big(1+\frac{7}{2}\epsilon\Big)\xi m^2 ,$$

where the last inequality makes use of the choice of η. Hence, the AM-GM inequality tells us

$$\big(|N_1(v_3)| + |N_2(v_3)|\big)^2 \geqslant 4\,\big|N_1(v_3)\big|\big|N_2(v_3)\big| \geqslant \Big(1 + \frac{7}{2}\epsilon\Big)m^2$$

and, consequently, we arrive at

$$|N_1(v_3)| + |N_2(v_3)| \geqslant \Big(1 + \frac{7}{2}\epsilon\Big)^{1/2} m \geqslant \Big(1 + \frac{11}{10}\epsilon\Big)m\,.$$

Finally, summing for all vertices $v_3 \in V_3'$ we obtain the desired lower bound

$$\begin{aligned}
\big|\partial H[V_1, V_3]\big| + \big|\partial H[V_2, V_3]\big| &\geqslant \sum_{v_3 \in V_3'} \big(|N_1(v_3)| + |N_2(v_3)|\big) \\
&\geqslant \Big(1 + \frac{11}{10}\varepsilon\Big)m \cdot |V_3'| \\
&\geqslant \Big(1 + \frac{11}{10}\varepsilon\Big)(1 - \sqrt{\rho})m^2 \\
&\geqslant (1 + \varepsilon)m^2\,,
\end{aligned}$$

where we used the choice of ρ for last inequality. $\square$

Tight paths of length two will play a special rôle in our proof and the following notation will be useful.

Definition 4.4.2. *Given a hypergraph $H = (V, E)$ and disjoint sets p, $q \in V^{(2)}$, we say that the edges xyz, $yzw \in E$ form a (p, q)-*cherry*, if $p = \{x, y\}$ and $q = \{z, w\}$.*

*Moreover, given two sets P, $Q \subseteq V^{(2)}$, we say that edges e, $e' \in E$ form a (P, Q)-*cherry*, if they form a (p, q)-cherry for some disjoint sets $p \in P$ and $q \in Q$.*

The next lemma asserts that in $\therefore$-dense hypergraphs with density larger than $1/4$ large sets of pairs induce many cherries.

Lemma 4.4.3. *For every ξ, $\epsilon \in (0, 1]$ there exist ρ, $\nu > 0$ such that for every sufficiently large $(\rho, 1/4 + \varepsilon, \therefore)$-dense hypergraph $H = (V, E)$ the following holds. For all sets P, $Q \subseteq V^{(2)}$ of size at least $3\xi n^2$ there are at least νn^4 (P, Q)-cherries.*

Proof. Given ξ and ϵ we apply Lemma 4.4.1 and we obtain η and ρ'. Without loss of generality we may assume that $\eta \leqslant \xi/2$. Moreover, Lemma 4.2.1 applied with η and $d = \xi$ yields some $\mu > 0$ and we fix the desired constants ρ and ν by

$$\rho = \frac{\mu^3 \xi}{56}\rho' \quad \text{and} \quad \nu = 9\rho^2\mu^4\varepsilon\,.$$

Let $H = (V, E)$ and P, $Q \subseteq V^{(2)}$ satisfy the assumptions of the lemma.

We consider a random balanced bipartition of $A \cup B$ of V and let $P_A = \{p \in P \colon p \subseteq A\}$ and $Q_B = \{q \in Q \colon q \subseteq B\}$. A standard application of Chebyshev's inequality shows that there exists a balanced partition of V such that $|P_A|$, $|Q_B| \geqslant \xi n^2/2$. We apply Lemma 4.2.1 separately to the graphs (A, P_A) and (B, Q_B) and obtain four pairwise disjoint vertex sets $A_1, A_2 \subseteq A$ and $B_1, B_2 \subseteq B$ each of size $m \geqslant \mu n/2$ such that the induced bipartite graphs $P[A_1, A_2]$ and $Q[B_1, B_2]$ are both η-regular with density at least ξ.

Next for $i = 1$, 2 we consider the 3-partite subhypergraph $H[B_i, P[A_1, A_2]]$ on $A_1 \cup A_2 \cup B_i$ with the edge set

$$\big\{\{x, y, z\} \in V^{(3)} \colon x \in B_i \text{ and } \{y, z\} \in E(P[A_1, A_2])\big\}\,.$$

Lemma 4.2.3 applied to $H[B_i, P[A_1, A_2]]$ with $\beta = \rho$ yields a subhypergraph $H_\rho^{i,P}$. Since our choice of ρ guarantees

$$\rho n^3 + \rho(3m)^3 \leqslant 28\rho n^3 \leqslant \rho' \cdot |B_i| \cdot e(P[A_1, A_2])$$

it follows from the $\therefore$-density of H, that $H_\rho^{i,P}$ is $(\rho', 1/4 + \varepsilon, \therefore)$-dense over $(B_i, P[A_1, A_2])$. Similarly, for $i = 1$, 2 we also define the 3-partite hypergraph $H_\rho^{i,Q}$ with vertex partition $B_1 \cup B_2 \cup A_i$ and note that it is $(\rho', 1/4 + \varepsilon, \therefore)$-dense over $(A_i, Q[B_1, B_2])$.

Applying Lemma 4.4.1 to the bipartite graph $P[A_1, A_2]$ and the 3-partite hypergraph $H_\rho^{1,P}$ implies

$$\big|\partial H_\rho^{1,P}[A_1, B_1]\big| + \big|\partial H_\rho^{1,P}[A_2, B_1]\big| \geqslant (1 + \varepsilon)m^2.$$

Moreover, three further applications of Lemma 4.4.1 to $P[A_1, A_2]$ with $H_\rho^{2,P}$ and to $Q[B_1, B_2]$ with $H_\rho^{1,Q}$ and with $H_\rho^{2,Q}$ show that

$$\sum_{i=1}^{2}\Big(\big|\partial H_\rho^{i,P}[A_1, B_i]\big| + \big|\partial H_\rho^{i,P}[A_2, B_i]\big|\Big) + \sum_{i=1}^{2}\Big(\big|\partial H_\rho^{i,Q}[B_1, A_i]\big| + \big|\partial H_\rho^{i,Q}[B_2, A_i]\big|\Big) \geqslant 4(1+\varepsilon)m^2.$$

In particular, rearranging the terms shows that

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\Big(\big|\partial H_\rho^{j,P}[A_i, B_j]\big| + \big|\partial H_\rho^{i,Q}[B_j, A_i]\big|\Big) \geqslant 4(1 + \varepsilon)m^2$$

and, hence, there are some indices i_0, $j_0 \in \{1, 2\}$ such that

$$\big|\partial H_\rho^{j_0,P}[A_{i_0}, B_{j_0}]\big| + \big|\partial H_\rho^{i_0,Q}[B_{j_0}, A_{i_0}]\big| \geqslant (1 + \varepsilon)m^2\,.$$

Consequently, set of ordered pairs

$$R = \big\{\{y, z\} \in V^{(2)} \colon (y, z) \in \partial H_\rho^{j_0,P}[A_{i_0}, B_{j_0}] \text{ and } (z, y) \in \partial H_\rho^{i_0,Q}[B_{j_0}, A_{i_0}]\big\}$$

has size at least εm^2.

Finally, we note that every $\{y, z\} \in R$ has positive degree in both hypergraphs $H_\rho^{j_0,P}$ and $H_\rho^{i_0,Q}$ and, hence, these degrees are at least $3\rho m$. Therefore, there are at least $9\rho^2 m^2$ distinct vertices $x \in A_{3-i_0}$ and $w \in B_{3-j_0}$ such that xyz and yzw form a (P, Q)-cherry. Summing over all pairs in R yields at least

$$\varepsilon m^2 \cdot 9\rho^2 m^2 \geqslant \nu n^4$$

(P, Q)-cherries in H. □

The following corollary allows us to find many connections between a large sets of unordered and a large set of ordered pairs.

Lemma 4.4.4. *For every ξ, $\epsilon \in (0, 1]$ there exist ζ, $\rho > 0$ such that for every sufficiently large $(\rho, 1/4 + \varepsilon, \therefore)$-dense n-vertex hypergraph $H = (V, E)$ the following holds.*

Let $P \subseteq V \times V$ be a set of ordered pairs and let $Q \subseteq V^{(2)}$ be a set of unordered pairs, each of size at least ξn^2. There is an $\ell \in \{2, 4\}$ such that there are at least $\zeta n^{\ell+2}$ tight paths of length ℓ which start with an ordered pair from P and ends in (some ordering of) with a pair from Q.

Proof. Given ξ and ϵ we apply Lemma 4.4.3 with $\xi/6$ and ε and obtain ρ and ν. Lemma 4.2.4 applied for $\nu/2$, 4, and 6 (in place of ξ, k, and ℓ in Lemma 4.2.4) yields the promised constant $\zeta > 0$. With out loss of generality we may assume that $\zeta < \nu/2$ and let n be sufficiently large.

For a given set of ordered pairs $P \subseteq V \times V$ let $\overline{P}$ be the set of unordered pairs obtained from P by ignoring the order. In particular, $|\overline{P}| \geqslant |P|/2 \geqslant \xi n^2/2$ and Lemma 4.4.3 asserts that there are νn^4 different $(\overline{P}, Q)$-cherries. That is to say there are νn^4 tight paths on four vertices of the form $xyzw$ where $\{x, y\} \in \overline{P}$ and $\{z, w\} \in Q$.

If for ζn^4 of those cherries we have that $(x, y) \in P$, then the lemma follows with $\ell = 2$. Hence, we may assume that for at least $(\nu - \zeta)n^4 \geqslant \nu n^4/2$ of those tight paths we (only) have $(y, x) \in P$. Consequently, Lemma 4.2.4 yields ζn^6 blowups of these two edge paths where the vertices y and z are doubled, i.e., H contains ζn^6 6-tuples of distinct vertices $(x, y_1, y_2, z_1, z_2, w)$ such that for every i, $j \in \{1, 2\}$ we have

$$(y_i, x) \in P\,, \quad \{z_j, w\} \in Q\,, \quad \text{and} \quad xy_iz_jw \text{ is a tight path with two edges.}$$

In particular, every such 6-tuple induces a tight path $y_1xz_1y_2wz_2$ which starts with an ordered pair from P and ends in an unordered pair from Q and this concludes the proof of the lemma. □

For establishing the Connecting Lemma (Lemma 4.1.4) we shall extend Lemma 4.4.4 in such a way that we can connect large sets P and Q, where both of them consist of ordered pairs. For that certain blowups of $K_4^{(3)-}$s will be useful and we introduce the following notation.

Definition 4.4.5. *We say a 7-tuple of distinct vertices $(a_1, a_2, a_3, b_1, b_2, c, d) \in V^7$ is a* turn *in a hypergraph $H = (V, E)$ if for every $i \in \{1, 2, 3\}$ and $j \in \{1, 2\}$ the set $\{a_i, b_j, c, d\}$ spans a copy of a $K_4^{(3)-}$ in H with a_i being the apex.*

Combining Theorem 4.2.5 and Lemma 4.2.4 shows that the hypergraphs with $\therefore$-density bigger than 1/4 contain many turns. Moreover, we observe that in a turn T the tight paths

$$a_1b_1ca_2b_2\,, \quad a_1b_1ca_3db_2a_2\,, \quad b_1a_1cda_2b_2\,, \quad \text{and} \quad b_1a_1cb_2a_2 \tag{4.5}$$

with at most 3 inner vertices connect the pairs $\{a_1, b_1\}$ and $\{a_2, b_2\}$ in all four possible orientations. This motivates the following definition.

Definition 4.4.6. *For a hypergraph $H = (V, E)$ we say two disjoint unordered pairs q, $q' \in V^{(2)}$ are* (θ, L)-turnable*, if for every ordering (q_1, q_2) of q and every ordering (q_1', q_2') of q' there exists some positive integer $\ell \leqslant L$ such that the number of tight (q_1, q_2)-(q_1', q_2')-paths in H with ℓ inner vertices is at least $\theta|V|^\ell$.*

It follows from (4.5) that pairs $\{a_1, b_1\}$ and $\{a_2, b_2\}$ that are contained in $\Omega(|V|^3)$ turns are $(\theta, 3)$-turnable for some sufficiently small $\theta > 0$. The following variation of this fact, will be useful in the proof of the Connecting Lemma.

Lemma 4.4.7. *For every $\epsilon \in (0, 1]$ there exist θ, $\rho > 0$ such that for every sufficiently large $(\rho, 1/4 + \varepsilon, \therefore)$-dense hypergraph $H = (V, E)$ the following holds.*

There exists a set $Q \subseteq V^{(2)}$ of size at least $\theta|V|^2$ such that for every $q \in Q$ there exists a set $Q'(q) \subseteq V^{(2)}$ of size at least $\theta|V|^2$ such that q and q' are $(\theta, 3)$-turnable for every $q' \in Q'(q)$.

Proof. Let $H = (V, E)$ be a sufficiently large $(\rho, 1/4+\epsilon, \therefore)$-dense hypergraph on n vertices. A combined application of Theorem 4.2.5 and Lemma 4.2.4 yields a set $\mathcal{T} \subseteq V^7$ of at least ζn^7 turns $(a_1, a_2, a_3, b_1, b_2, c, d)$ in H for some sufficiently small $\zeta = \zeta(\varepsilon) > 0$ and we shall deduce the conclusion of the lemma for

$$\theta = \frac{\zeta}{8}\,.$$

For every pair $(a, b) \in V \times V$ and $i \in \{1, 2\}$ let $\mathcal{T}_i(a, b)$ be the set of such turns where a and b play the rôles of a_i and b_i, respectively. We consider the set

$$\mathcal{T}^\star = \big\{(a, a', a_3, b, b', c, d) \in \mathcal{T} \colon |\mathcal{T}_1(a, b) \cap \mathcal{T}_2(a', b')| \geqslant \zeta n^3/2\big\}$$

and note that $|\mathcal{T}^\star| \geqslant \zeta n^7/2$. By a standard averaging argument there are at least $\zeta n^2/4$ pairs $(a,b) \in V \times V$ for which we have

$$|\mathcal{T}_1(a,b) \cap \mathcal{T}^\star| \geqslant \frac{\zeta}{4} n^5$$

and we denote the set of these ordered pairs by R. Note that for every pair $(a,b) \in R$ there is a set $R'(a,b) \subseteq V \times V$ with

$$|R'(a,b)| \geqslant \frac{\zeta}{4} n^2 \text{ such that } \left|\mathcal{T}_1(a,b) \cap \mathcal{T}_2(a',b')\right| \geqslant \frac{\zeta}{2} n^3 \tag{4.6}$$

for every $(a',b') \in R'(a,b)$. Finally, let Q be the set of unordered pairs derived from R, i.e.,

$$Q = \left\{\{q_1,q_2\} \in V^{(2)} \colon (q_1,q_2) \in R\right\}$$

and for every $q = \{q_1,q_2\}$ set

$$Q'(q) = \left\{\{q_1',q_2'\} \in V^{(2)} \colon (q_1',q_2') \in R'(q_1,q_2) \cup R'(q_2,q_1)\right\}.$$

Clearly,

$$|Q| \geqslant \frac{|R|}{2} \geqslant \frac{\zeta}{8} n^2 = \theta n^2 \quad \text{and} \quad Q'(q) \overset{(4.6)}{\geqslant} \frac{\zeta}{8} n^2 = \theta n^2$$

and the required number of tight paths for every orientation of $q \in Q$ and $q' \in Q'(q)$ follows from (4.5) and (4.6). $\square$

Roughly speaking, the proof of Lemma 4.1.4 follows from Lemmata 4.4.4 and 4.4.7. The definition of connectable pairs allows us to move from the given ordered pairs (x,y) and (w,z), that need to be connected, to large sets of ordered pairs P, P', by considering their second neighbourhoods. Moreover, Lemma 4.4.7 yields sets $Q \subseteq V^{(2)}$ and $Q'(q) \subseteq V^{(2)}$ for every $q \in Q$ of turnable pairs. Applying Lemma 4.1.4 first to P and Q and then to P' and $Q'(q)$ for all $q \in Q$ leads to the desired tight (x,y)-(z,w)-paths.

Proof of Lemma 4.1.4. For given ϵ, $\beta > 0$ let θ and ρ_1 be the constants provided by Lemma 4.4.7. We set

$$\xi = \min\{\theta, \beta^2\}$$

and Lemma 4.4.4 applied with ξ and ε yields ζ and ρ_2. Finally, we define the promised constants

$$\rho = \min\{\rho_1, \rho_2\} \quad \text{and} \quad \alpha = \frac{\zeta^2 \theta}{13}.$$

Let $H = (V,E)$ be a sufficiently large $(\rho, 1/4+\epsilon, \therefore)$-dense hypergraph on n vertices and let (x,y), (w,z) be two disjoint β-connectable pairs. Consider the second neighbourhoods of these pairs defined by

$$P = \{(u,v) \in V \times V \colon xyu,\ yuv \in E\} \quad \text{and} \quad P' = \{(u',v') \in V \times V \colon wzu',\ zu'v' \in E\}. \tag{4.7}$$

Owing to the β-connectability, both sets P and P' have size at least $\beta^2 n^2 \geqslant \xi n^2$.

Next, let $Q \subseteq V^{(2)}$ and $Q'(q) \subseteq V^{(2)}$ for every $q \in Q$ be the sets of size at least $\theta n^2 \geqslant \xi n^2$ provided by Lemma 4.4.7. For every $q \in Q$ we denote by $P_4(q)$ (resp. $P_6(q)$) the number of tight (u,v)-(q_1,q_2)-paths having 4 (resp. 6) vertices and $(u,v) \in P$ and $\{q_1,q_2\} = q$. Moreover, we normalise these numbers by

$$\eta_P(q) = \max\left\{\frac{P_4(q)}{n^4}, \frac{P_6(q)}{n^6}\right\}$$

and note that Lemma 4.4.4 applied to P and Q ensures

$$\sum_{q \in Q} \eta_P(q) \geqslant \zeta\,. \tag{4.8}$$

Analogously, we define $P_4'(q')$, $P_6'(q')$, and $\eta_{P'}(q')$ for every $q' \in \bigcup_{q\in Q} Q'(q)$ and Lemma 4.4.4 applied to P' and $Q'(q)$ implies

$$\sum_{q' \in Q'(q)} \eta_{P'}(q') \geqslant \zeta\,. \tag{4.9}$$

for every $q \in Q$. Recall, that the paths accounted for in (4.8) and (4.9) induce an ordering of the vertices in q and in q'. However, by Lemma 4.4.7 the pairs q and q' are $(\theta,3)$-turnable for every $q \in Q$ and $q' \in Q'(q)$, which means that these pairs can be connected for any possible orientation. Consequently, there is some ℓ with

$$5 \leqslant \ell \leqslant \max\{4,6\} + \max\{1,2,3\} + \max\{4,6\} = 15$$

such that the number of (x,y)-(z,w)-walks in H is at least

$$\frac{n^\ell}{12} \cdot \sum_{q\in Q} \eta_P(q) \cdot \theta \cdot \sum_{q' \in Q'(q)} \eta_{P'}(q') \overset{(4.9)}{\geqslant} \frac{n^\ell}{12} \cdot \sum_{q \in Q} \eta_P(q) \cdot \theta \cdot \zeta \overset{(4.8)}{\geqslant} \frac{\zeta^2\theta}{12} n^\ell\,.$$

At most $O(n^{\ell-1})$ of these walks might not be a path and, hence, the lemma follows for sufficiently large n. □

4.5 Absorbing path

We dedicate this section to the proof of Lemma 4.1.3. Similarly as in [59] the absorbers we consider here have two parts. Moreover, we use an idea of Polcyn and Reiher [56], which reduces the abundant existence of absorbers to a degenerate Turán problem on the price that we can only absorb exactly three vertices at each time.

Consider first the complete 3-partite hypergraph $K^{(3)}_{3,3,3}$ with parts $A_i = \{x_i, y_i, z_i\}$, for every $i = 1,2,3$. Note that this hypergraph contains the tight paths

$$x_1x_2x_3y_1y_2y_3z_1z_2z_3\,, \tag{4.10}$$

and

$$x_1x_2x_3z_1z_2z_3\,. \tag{4.11}$$

This means that from every copy of $K^{(3)}_{3,3,3}$, ordered as a tight path like in (4.10), we may remove the three inner vertices y_1, y_2, y_3 to obtain a tight path with the same ends. Since we only consider dense hypergraphs, we can guarantee that many copies $K^{(3)}_{3,3,3}$ exist. In other words, in such a situation the tight path $x_1x_2x_3z_1z_2z_3$ could absorb the three vertices y_1, y_2, and y_3. However, not every triple might be contained in a $K^{(3)}_{3,3,3}$ and this will be addressed by the second part of the absorbers used here.

Suppose we want to absorb some arbitrary vertices v_1, v_2, and v_3. The idea, similarly as in [59], is to exchange v_i with y_i contained in some $K^{(3)}_{3,3,3}$. Suppose we have found a $K^{(3)}_{3,3,3}$ as described above, but additionally we find a path (as a graph) on four vertices with edges from $N_H(v_i) \cap N_H(y_i)$ disjointly for each $i = 1, 2, 3$. We argue that this whole structure can absorb v_1, v_2, v_3. Indeed, if $a_ib_ic_id_i$ is a path on four vertices with edges from $N_H(v_i) \cap N_H(y_i)$, then both $P(v_i) = a_ib_iv_ic_id_i$ and $P(y_i) = a_ib_iy_ic_id_i$ are tight paths in the hypergraph and with the same endings. Moreover, the minimum degree and the uniform density imply that for each vertex $v \in V$, most vertices of V have $\Omega(n^2)$ common neighbours with v, which is enough to find such paths.

Therefore, if we choose to absorb v_1, v_2, v_3, we will consider the tight paths $P(v_1)$, $P(v_2)$, and $P(v_3)$ and the tight path of $K^{(3)}_{3,3,3}$ as in (4.10). On the other hand, if we choose not to absorb them, then we consider the tight paths $P(y_1)$, $P(y_2)$, and $P(y_3)$ and the tight path of $K^{(3)}_{3,3,3}$ as in (4.11). We will also show that for each triple of vertices, we can find many of these configurations, so that we can choose a small amount of them that still can absorb every triple and also connect them into a single tight path. Observe that this absorbing path can only absorb sets of vertices with size divisible by three, an issue with which we deal later. First we prove that for every triple there are many absorbers.

Definition 4.5.1. *Let $H = (V, E)$ be a hypergraph and $(v_1, v_2, v_3) \in V^3$. We say*

$$A = (K, P_1, P_2, P_3) \in V^9 \times V^4 \times V^4 \times V^4\,,$$

with $K = (x_1, x_2, x_3, y_1, y_2, y_3, z_1, z_2, z_3)$ and $P_i = (a_i, b_i, c_i, d_i)$ is an absorber *for (v_1, v_2, v_3) if the ordered sets*

(i) $x_1x_2x_3y_1y_2y_3z_1z_2z_3$, $x_1x_2x_3z_1z_2z_3$,

(ii) $a_ib_iv_ic_id_i$ and $a_ib_iy_ic_id_i$ for $i = 1, 2, 3$

induce tight paths in H. All hyperedges of those paths that do not include a vertices from $\{v_1, v_2, v_3\}$ are called internal edges *of the absorber A.*

Formally absorbers are defined to be four tuples. However, sometimes it will be convenient to view them as 21-tuples of vertices.

Lemma 4.5.2. *For all d, $\epsilon \in (0,1]$ there exist ρ, $\xi > 0$ such that for sufficiently large n the following holds.*

For every $(\rho, d, \therefore)$-dense hypergraph $H = (V, E)$ on n vertices with $\delta_1(H) \geqslant \epsilon n^2$ and every triple $T = (v_1, v_2, v_3) \in V^3$ of distinct vertices there are at least ξn^{21} absorbers for T.

Proof. Given d and ϵ we define some auxiliary constant $\zeta = (d/2)^{27}/3$ and set

$$\rho = \frac{1}{36}\left(\frac{d}{2}\right)^{54} \quad \text{and} \quad \xi = \frac{\zeta d^9 \varepsilon^9}{2^{11}}\,.$$

Let $H = (V, E)$ be a $(\rho, d, \therefore)$-dense hypergraph on n vertices and consider some triple of vertices $T = (v_1, v_2, v_3) \in V^3$.

Three applications of Lemma 4.2.2 each with $X = V$ and for $i \in [3]$ with the set of ordered pairs

$$\big\{(u, w)\colon \{u, w\} \in N_H(v_i)\big\}$$

tells us, that there are at most $3\sqrt{\rho}n$ *bad vertices* $v \in V$ that may fail to satisfy

$$\big|N_H(v) \cap N_H(v_i)\big| \geqslant (d - \sqrt{\rho})\big|N_H(v_i)\big| \geqslant (d - \sqrt{\rho})\delta_1(H) \geqslant \frac{d}{2}\varepsilon n^2 \tag{4.12}$$

for some $i \in [3]$. Moreover, the $(\rho, d, \therefore)$-density of H implies that the edge density of H is at least $d - 2\rho > d/2$ and since the extremal number of any fixed 3-partite hypergraph is $o(n^3)$ we have $K^{(3)}_{3,3,3} \subseteq H$ for sufficiently large n. In fact, the standard proof of this fact from [22] yields at least $((d/2)^{27} - o(1))n^9$ such copies. Consequently, for sufficiently large n there are at least

$$\left(\left(\frac{d}{2}\right)^{27} - o(1)\right) n^9 - 3\sqrt{\rho}n \cdot n^8 \geqslant \zeta n^9$$

copies of $K^{(3)}_{3,3,3}$ in H that contain no bad vertex. Let $\mathcal{K} = \mathcal{K}_T \subseteq V^9$ be the set of these $K^{(3)}_{3,3,3}$ in H.

Consider some $K = (x_1, x_2, x_3, y_1, y_2, y_3, z_1, z_2, z_3) \in \mathcal{K}$. Since none of the vertices of K is bad, for every vertex v from K inequality (4.12) holds for every $i \in [3]$. In particular, for every $i \in [3]$ we have $|N_H(y_i) \cap N_H(v_i)| \geqslant d\varepsilon n^2/2$ and it follows from [10] that there exist at least $((d\varepsilon/2)^3 - o(1))n^4$ paths on four vertices with edges from $N_H(y_i) \cap N_H(v_i)$. Consequently, for sufficiently large n, there exist at least

$$|\mathcal{K}| \cdot \left(\left(\frac{d^3\varepsilon^3}{8} - o(1)\right)n^4\right)^3 \geqslant \zeta n^9 \cdot \frac{d^9\varepsilon^9}{2^{10}} n^{12} \geqslant 2\xi n^{21}$$

4-tuples $A = (K, P_1, P_2, P_3) \in V^9 \times V^4 \times V^4 \times V^4$ with P_i inducing a path in $N_H(y_i) \cap N_H(v_i)$ for $i = [3]$. Such an A may only fail to be an absorber for T, if it contains some vertex from T itself or if its 21 vertices are not distinct. However, since there are at most $O(n^{20})$ such "degenerate" A's the lemma follows for sufficiently large n. $\square$

Note that for the proof of Lemma 4.5.2 positive $\therefore$-density was sufficient. However, to address the aforementioned divisibility issue, we will show that the hypergraphs H considered here contain a copy of $C_8(4)$, the 4-blow-up of the tight cycle on 8 vertices. For the proof of that, we make use of the assumption that the $\therefore$-density of H is bigger than $1/4$.

The $C_8(4)$ is formed by 8 cyclicly ordered independent sets $\{e_i, f_i, g_i, h_i\}_{i\in[8]}$ such that the only edges are the ones with vertices from three consecutive such sets. Note that $C_8(4)$ contains the tight path

$$e_1e_2 \dots e_8 f_1 f_2 \dots f_8 g_1 g_2 \dots g_8 h_1 h_2 \dots h_8. \tag{4.13}$$

Moreover, by removing the sets $\{f_i\}_{i\in[8]}$ or $\{f_i, g_i\}_{i\in[8]}$ from the path in (4.13) leads to tight paths with the same ends in $C_8(4)$ with 24 or 16 vertices, respectively. We also remark that 16, 24 and 32 are congruent to 1, 0 and 2 modulo 3, respectively. Therefore, if we connect such tight path to the absorbing path, we can decide to remove some of the vertices so that the size of the leftover set is divisible by 3.

Lemma 4.5.3. *For all $\epsilon > 0$ there exist ρ, $\theta > 0$ such that every sufficiently large $(\rho, 1/4 + \epsilon, \therefore)$-dense hypergraph $H = (V, E)$ contains $\theta|V|^{32}$ copies of $C_8(4)$.*

Proof. Given $\epsilon > 0$ we apply Theorem 4.2.5 to obtain ρ_1 and ξ. Then, the application of Lemma 4.4.3 to $\xi/6$ and ε yields ρ_2 and ν. Set $\rho = \min\{\rho_1, \rho_2\}$ and let n be sufficiently large.

Let $H = (V, E)$ be a $(\rho, 1/4 + \epsilon, \therefore)$-dense hypergraph on n vertices. In view of Lemma 4.2.4 it suffices to show that H contains ζn^8 copies of C_8 for some $\zeta > 0$.

Theorem 4.2.5 implies that H contains at least ξn^4 copies of $K_4^{(3)-}$. Let R be the set of ordered pairs (a, x) such that both vertices are contained in at least $\xi n^2/2$ of these $K_4^{(3)-}$ with a being the apex. By double counting we infer $|R| \geqslant \xi n^2/2$.

For every $(a, x) \in R$, let $P_{a,x} \subseteq V^{(2)}$ be those pairs $\{y, z\}$ that span such a copy of $K_4^{(3)-}$ together with a and x. We apply Lemma 4.4.3 to $P = Q = P_{a,x}$ and infer that there are at least νn^4 (P, Q)-cherries, i.e., tight paths with 4 vertices starting and ending at a pair from $P_{a,x}$.

Let F be the hypergraph with vertex set $\{a, x, y, y', z, z'\}$ such that the following holds. The vertices $\{a, x, y, z\}$ and $\{a, x, y', z'\}$ span copies of $K_4^{(3)-}$ with apex a and F contains a $(\{y, z\}, \{y', z'\})$-cherry. Observe that since y and z (resp. y' and z') play a symmetric role in $K_4^{(3)-}$, regardless of the orientation of the pairs $\{y, z\}$ and $\{y', z'\}$ in the cherry the resulting hypergraph is isomorphic. Without loss of generality we will assume that the cherry is a tight path of the form $yzy'z'$. By the reasoning above, H contains at least

$$|R| \cdot \nu n^4 \geqslant \frac{\xi}{2}\nu n^6$$

copies of F. We argue that there is a homomorphism of C_8 in F. Indeed, if we consider the vertices of F in the following cyclic ordering

$$xayzy'z'ay'$$

one can check that every consecutive triple forms an edge in F. Since there are at least $\Omega(n^6)$ copies of F in H, then by Lemma 4.2.4 and taking ζ small enough, we have that there are at least ζn^8 copies of C_8. □

We are now ready to prove Lemma 4.1.3.

Proof of Lemma 4.1.3. Given $\varepsilon > 0$ the constants appearing in this proof will satisfy the following hierarchy

$$1 > \varepsilon \gg \xi\,, \theta \gg \beta \gg \rho\,, \alpha \gg \gamma' \geqslant \gamma \gg \frac{1}{n}\,, \tag{4.14}$$

where the auxiliary constants ξ, θ, and α are provided by Lemmata 4.5.2, 4.5.3, and 4.1.4 and it is easy to check that (4.14) complies with the quantification of these lemmata. Let H be a $(\rho, 1/4+\epsilon, \therefore)$-dense hypergraph with $\delta_1(H) \geqslant \epsilon n^2$ and let R be a subset of V with at most $2\gamma^2 n$ vertices. Fix the subhypergraph $H_\beta \subseteq H$ provided by Lemma 4.2.3.

For $T \in V^3$, let $\mathcal{A}_T$ be the set of those absorbers for T in H that have no vertex in R and all its 36 internal edges from H_β. It follows from Lemma 4.5.2 applied with $d = 1/4 + \varepsilon$ and ε that

$$|\mathcal{A}_T| \geqslant \xi n^{21} - 21\,|R| n^{20} - 6\cdot 36\big(e(H) - e(H_\beta)\big)n^{18} \geqslant \xi n^{21} - 42\,\gamma^2 n^{21} - 216\,\beta n^{21} \overset{(4.14)}{\geqslant} \frac{\xi}{2} n^{21}\,.$$

Let $\mathcal{A} = \bigcup_T \mathcal{A}_T$ be the union over all triples $T \in V^3$ and consider a random collection of absorbers $\mathcal{C} \subseteq \mathcal{A}$ in which each element of $\mathcal{A}$ is present independently with probability

$$p = \frac{\gamma^{4/3} n}{2|\mathcal{A}|}\,.$$

Since $\mathbb{E}|\mathcal{A}| = p|\mathcal{A}|$, Markov's inequality ensures that

$$\mathbb{P}\big(|\mathcal{C}| \geqslant \gamma^{4/3} n\big) \leqslant \frac{1}{2}. \tag{4.15}$$

Moreover, for every $T \in V^3$ we have

$$\mathbb{E}|\mathcal{C} \cap \mathcal{A}_T| = p|\mathcal{A}_T| \geqslant \frac{\gamma^{4/3}n}{2|\mathcal{A}|} \cdot \frac{\xi n^{21}}{2} \geqslant \frac{\gamma^{4/3}\xi n}{4} \overset{(4.14)}{\geqslant} 4\gamma^2 n\,,$$

Chernoff's inequality combined with the union bound over all triples yields

$$\mathbb{P}\big(\exists T \in V^3 \colon |\mathcal{C} \cap \mathcal{A}_T| < 3\gamma^2 n\big) \leqslant o(1)\,. \tag{4.16}$$

Letting Y be the number of pairs of distinct absorbers A, $A' \in \mathcal{C}$ that share a vertex we note

$$\mathbb{E}Y = p^2 \cdot n^{21} \cdot 21^2 \cdot n^{20} = \frac{\gamma^{8/3}n^2}{4|\mathcal{A}|^2} \cdot 441 n^{41} \leqslant \frac{441\gamma^{8/3}n}{\xi^2} \overset{(4.14)}{\leqslant} \frac{\gamma^2 n}{4}$$

and by Markov's inequality, we have

$$\mathbb{P}(Y \geqslant \gamma^2 n) \leqslant \frac{1}{4}\,. \tag{4.17}$$

Consequently, with positive probability none of the bad events from (4.15), (4.16), and (4.17) happen. In particular, there exists a realisation of $\mathcal{C}$ such that

$$|\mathcal{C}| < \gamma^{4/3}n\,, \qquad |\mathcal{C} \cap \mathcal{A}_T| \geqslant 3\gamma^2 n \text{ for every } T \in V^3, \quad \text{and} \quad |Y(\mathcal{C})| < \gamma^2 n\,.$$

For every pair of absorbers accounted in $Y(\mathcal{C})$ we remove one of the involved absorbers in an arbitrary way and obtain a subset $\mathcal{B} \subseteq \mathcal{C}$ of pairwise vertex disjoint absorbers satisfying

$$|\mathcal{B}| \leqslant |\mathcal{C}| < \gamma^{4/3}n \quad \text{and} \quad |\mathcal{B} \cap \mathcal{A}_T| > |\mathcal{C} \cap \mathcal{A}_T| - \gamma^2 n \geqslant 2\gamma^2 n \text{ for every } T \in V^3.$$

Recall that if the absorbing path would only contain the absorbers from $\mathcal{B}$, then it could only absorb sets U with a cardinality that is divisible by 3. We address this divisibility issue by adding a copy of $C_8(4)$ to the path. Lemma 4.5.3 guarantees at least θn^{32} copies of $C_8(4)$ in H. Similarly, as for the estimate of $\mathcal{A}_T$, we infer that there is one such $C_8(4)$ which is vertex disjoint from the set R and from all absorbers from $\mathcal{B}$ and which only contains edges from H_β. In fact, this follows from

$$\begin{aligned} \theta n^{32} - 32\,|R|n^{31} - 21\,|\mathcal{B}|n^{31} - 6 \cdot e(C_8(4))\big(e(H) - e(H_\beta)\big)n^{29} \\ \geqslant \theta n^{32} - 64\,\gamma^2 n^{32} - 21\,\gamma^{4/3}n^{32} - 3072\,\beta n^{32} \overset{(4.14)}{>} 0\,. \end{aligned}$$

Fix an ordering of the vertices of such a $C_8(4)$ that induces a tight path (see, e.g., (4.13)) and denote this path by P_C.

In order to obtain the final absorbing path, each absorber $(K, P_1, P_2, P_3) \in \mathcal{B}$ will be viewed as a collection of four tight paths: $x_1x_2x_3z_1z_2z_3$ and $a_ib_iy_ic_id_i$, for $i = 1, 2, 3$, as in Definition 4.5.1. Therefore, together with joining P_C we have to connect $t = 4|\mathcal{B}| + 1$ tight paths to build the promised absorbing path A. For each of the connections we will appeal to Lemma 4.1.4 and each application will require to add up at most 15 inner vertices.

Let $(P_i)_{i\in[t]}$ be an arbitrary enumeration of all these tight paths that need to be connected. We continue in an inductive manner starting with $A_1 = P_1$, let A_j be the already constructed tight path containing P_i for every $i \leqslant j$. Since every connection requires at most 15 inner vertices and the longest path in $(P_i)_{i\in[t]}$ has 32 vertices we have

$$|V(A_j)| + \sum_{i=j+1}^{t} |V(P_i)| \leqslant 15(j-1) + 32t \leqslant 47t \leqslant 47\big(4|\mathcal{B}| + 1\big) \leqslant 47\big(4\gamma^{4/3}n + 1\big) \leqslant \gamma n\,. \tag{4.18}$$

Suppose now that we want to connect P_j, which ends in (x, y), to P_{j+1}, which starts at (z, w). Since all tight paths P_i with $i \in [t]$ have its edges in H_β, by Lemma 4.2.3 they are β-connectable. Therefore, Lemma 4.1.4 implies that there are at least αn^ℓ tight paths, with $\ell \leqslant 15$ inner vertices, connecting (x, y) with (z, w) in H. Consequently, in view of (4.18) and $|R| \leqslant 2\gamma^2 n$ our choice of γ in (4.14) shows that at least one of such connecting paths must be vertex disjoint from

$$V(A_j) \cup \bigcup_{i=j+1}^{t} V(P_i) \cup R\,,$$

which concludes the inductive step and proves the existence of the tight path A_{j+1}.

Finally, let $A = A_t$ be the final tight path and let $U \subseteq V \smallsetminus V(A)$ with $|U| \leqslant 3\gamma^2 n$. First we remove 0, 8 or 16 vertices from P_C in A and reallocate them to U to get a set U' with size divisible by three. Moreover $|U'| \leqslant 3\gamma^2 n + 16 \leqslant 3(\gamma^2 n + 6)$ and, hence, U' can be split into at most $\gamma^2 n + 6$ disjoint triples. Since each triple has at least $2\gamma^2 n > \gamma^2 n + 6$ absorbers in A, we can greedily assign one for each and absorb all of them into A. $\square$

4.6 Proof of Theorem 1.3.4

In this section we discuss the few modifications necessary in the proof of Theorem 1.3.3 in order to prove Theorem 1.3.4. Recall that both theorems have the same minimum vertex degree assumption. However, where Theorem 1.3.4 requires the given hypergraph H to be $\wedge$-dense for some positive density, Theorem 1.3.3 requires $\therefore$-density bigger than 1/4. In other words, the uniform density assumptions of both theorems are incomparable.

The proof of Theorem 1.3.3 consist of three main parts, namely Lemmata 4.1.2 – 4.1.4. Observe that Lemma 4.1.2 can be applied directly under the conditions of Theorem 1.3.4, but for Lemmata 4.1.3 and 4.1.4 we have the assumption of $\therefore$-density at least 1/4 which is not provided by Theorem 1.3.4. We start with the discussion of the Connecting Lemma in the context of Theorem 1.3.4 in the next section and defer the discussion of the adjustments for the Absorbing Path Lemma (Lemma 4.1.3) to Section 4.6.2.

4.6.1 Connecting Lemma for Theorem 1.3.4

The following lemma will play the rôle of Lemma 4.1.4 in the proof of Theorem 1.3.3.

Lemma 4.6.1 (Connecting Lemma for $\curlywedge$-density conditions)**.** *For every d, $\beta > 0$ there exist ρ, $\alpha > 0$ and an n_0 such that for every $(\rho, d, \curlywedge)$-dense hypergraph H on $n \geqslant n_0$ vertices the following holds.*

For every $\ell \in \{5, 6, 7\}$ and for every pair of disjoint ordered β-connectable pairs (x, y), $(w, z) \in V \times V$, the number of (x, y)-(z, w)-paths with ℓ inner vertices is at least αn^ℓ.

Proof of Lemma 4.6.1 (sketch). We begin with the following observation. Let P, $P' \subseteq V \times V$ each of size at least $\Omega(n^2)$ we show that

$$\text{there are at least } \Omega(n^5) \ p\text{-}p'\text{-paths with one inner vertex and } p \in P,\ p' \in P'. \tag{4.19}$$

Note that every $(\rho, d, \curlywedge)$-dense hypergraph is $(\rho, d, \therefore)$-dense and in view of Lemma 4.2.2 applied to P and V there is a set $X \subseteq V$ such that $|X| = \Omega(n)$ and for every $x \in X$ we have $|N(x, P)| = \Omega(n^2)$. Similarly, another application of Lemma 4.2.2 to P' and X yields a set $X' \subseteq X$ of size $\Omega(n)$ such that

$$|N(x, P)| = \Omega(n^2) \quad \text{and} \quad |N(x, Q)| = \Omega(n^2)$$

for every $x \in X'$. Consequently, a standard averaging argument tells us that each of the sets

$$Q = \big\{(p_2, x) \in V \times X' \colon |\{p_1 \in V \colon (p_1, p_2) \in P \text{ and } p_1 p_2 x \in E\}| = \Omega(n)\big\}$$

and

$$Q' = \big\{(x, p_1') \in X' \times V \colon |\{p_2' \in V \colon (p_1', p_2') \in P' \text{ and } x p_1' p_2' \in E\}| = \Omega(n)\big\}$$

has size $\Omega(n^2)$. Finally, the $\therefore$-density of H applied to Q and Q' yields $\Omega(n^5)$ p-p'-paths starting in P and ending in P' with an inner vertex from X, i.e., it establishes (4.19).

For given connectable pairs (x, y) and (w, z) letting P and P' be their second neighbourhoods as defined in (4.7), yields the conclusion of Lemma 4.6.1 for $\ell = 5$.

For $\ell = 6$ we note that $\therefore$-density implies that there are $\Omega(n^2)$ β'-connectable pairs (y, y') with $xyy' \in E$ for sufficiently small $\beta' = \beta'(d) > 0$. Applying the same argument as above for every such pair (y, y') proves the case $\ell = 6$. Finally, for $\ell = 7$ the same reasoning applied to the connectable pairs (y', y'') with xyy', $yy'y'' \in E$ concludes the proof. $\square$

4.6.2 Absorbing Path Lemma for Theorem 1.3.4

Recall that the proof of Lemma 4.1.3 required $\underline{\therefore}$-density bigger than 1/4 in only two places:

(i) for the connection of the absorbers to a tight path and

(ii) in Lemma 4.5.3 for addressing the divisibility issue of the size of the absorbable sets,

while for the abundant existence of the absorbers $\underline{\therefore}$-density d for any $d > 0$ is sufficient (see Lemma 4.5.2). As shown in Section 4.6.1 for the connecting lemma positive $\wedge$-density suffices, which addresses i. Moreover, in Lemma 4.6.1 we are even free to choose the length of the connecting paths, which renders the divisibility issue from ii in this context.

4.7 Concluding remarks

We briefly discuss a few open problems for 3-uniform hypergraphs and possible generalisations of Theorems 1.3.3 and 1.3.4 to k-uniform hypergraphs.

4.7.1 Problems for 3-uniform hypergraphs

Theorems 1.3.3 and 1.3.4 concern asymptotically optimal assumptions for uniformly dense hypergraphs that guarantee the existence of Hamilton cycles. The following notation will be useful for the further discussion.

Definition 4.7.1. *Given* $\star \in \{\therefore, \underline{\therefore}, \wedge\}$ *and* $a \in \{1, 2\}$. *We say a pair of reals* (d, α) *is* $(\star, a)$-Hamilton *if the following assertion holds:*

For every $\varepsilon > 0$ *there exist* $\rho > 0$ *and* n_0 *such that every* $(\rho, d+\varepsilon, \star)$*-dense hypergraph* $H = (V, E)$ *with* $|V| = n \geqslant n_0$ *and* $\delta_a(H) \geqslant (\alpha+\varepsilon)\binom{n}{3-a}$ *contains a tight Hamilton cycle.*

We remark that we can restrict our attention to tight Hamilton cycles, since the result of Lenz, Mubayi, and Mycroft [50] asserts that already $(0, 0)$ would be $(\star, a)$-Hamilton for loose cycles for every choice of $\star \in \{\therefore, \underline{\therefore}, \wedge\}$ and $a \in \{1, 2\}$. For tight Hamilton cycles Aigner-Horev and Levy [2] showed that $(0, 0)$ is $(\wedge, a)$-Hamilton for $a = 2$ and this was extended by Gan and Han [33] and by Theorem 1.3.4 to $a = 1$. It remains to characterise the minimal pairs (d, α) that are $(\star, a)$-Hamilton for the four combinations $\star \in \{\therefore, \underline{\therefore}\}$ and $a \in \{1, 2\}$.

Example 1.3.2 shows that for (d, α) being $(\underline{\therefore}, 1)$-Hamilton we must have

$$\max\{d, \alpha\} \geqslant \frac{1}{4}\,. \tag{4.20}$$

On the other hand, Theorem 1.3.3 asserts that for $d = 1/4$ already $\alpha = 0$ suffices. It would be interesting to determine the smallest value $\alpha_{\dot{-},1}$ such that $d = 0$ suffices. In view of (4.20) we have $\alpha_{\dot{-},1} \geqslant 1/4$ and the result from [59] bounds $\alpha_{\dot{-},1}$ by $5/9$. Since all known lower bound constructions for that result are lacking to be $\dot{-}$-dense it seems plausible that $\alpha_{\dot{-},1} < 5/9$.

Similarly, let $\alpha_{\dot{-},2}$ be the infimum over all $\alpha \geqslant 0$ such that $(0, \alpha)$ is $(\dot{-}, 2)$-Hamilton. Here it follows from [64] that $\alpha_{\dot{-},2} \leqslant 1/2$. Moreover, Example 1.3.2 yields a hypergraph with minimum codegree $(1/4 - o(1))n$ that fails to contain a tight Hamilton cycle. Therefore, we have $\alpha_{\dot{-},2} \geqslant 1/4$ and at this point we are not aware of any reason that excludes the possibility that $\alpha_{\dot{-},2}$ matches this lower bound.

Problem: Determine $\alpha_{\dot{-},1}$ and $\alpha_{\dot{-},2}$.

For tight Hamilton cycles in $\therefore$-dense hypergraphs the problem appears to be more delicate as the following unbalanced version of Example 1.3.2 shows. Instead of a uniformly chosen bipartition of $E(K_{n-2})$ we may colour the edges independently red with probability p and blue with probability $1 - p$. Let H_p be the resulting hypergraph, where the rest of the construction is carried out in the same way as in Example 1.3.2. By symmetry we may assume $p \geqslant 1/2$ and for the same reasons as in Example 1.3.2 the hypergraph H_p contains no tight Hamilton cycle. Moreover, for every fixed $\rho > 0$ we have with high probability that

$$\delta_1(H_p) = \big(\min\{1 - p\,, p^3 + (1 - p)^3\} - \rho\big)\tbinom{n}{2} \quad \text{and} \quad \delta_2(H_p) = \big((1 - p)^2 - \rho\big)n$$

and that H_p is $(\rho, p^3 + (1 - p)^3, \therefore)$-dense. For p close to 1 this shows that there is no $d < 1$ such that $(d, 0)$ is $(\therefore, a)$-Hamilton for $a \in \{1, 2\}$. In particular, there is no straightforward analogue of Theorem 1.3.3 in this setting.

It would be intriguing if this construction is essentially optimal for every $p \geqslant 1/2$. In such an event it would imply a resolution of the following problems, where the lower bound would be obtained from H_p for $p = 2/3$ and $p = 1/2$. **Problem:** Is it true that

(i) $(1/3, 1/3)$ is $(\therefore, 1)$-Hamilton?

(ii) $(1/4, 1/4)$ is $(\therefore, 2)$-Hamilton?

4.7.2 Possible generalisations to k-uniform hypergraphs

The notion of tight Hamilton cycles straight forwardly extends to k-uniform hypergraphs. Moreover, the definition of uniformly dense hypergraphs is inspired from the theory of quasirandom hypergraphs (see, e.g., [1, 75] and the references therein). Below we briefly recall the generalisation of Definition 1.3.1 for general k-uniform hypergraphs, where we follow the presentation from [57].

Given a nonnegative integer k, a finite set V, and a set $Q \subseteq [k]$ we write V^Q for the set of all functions from Q to V. It will be convenient to identify the Cartesian power V^k with $V^{[k]}$ by regarding any k-tuple $\mathbf{v} = (v_1, \dots, v_k)$ as being the function $i \longmapsto v_i$. We denote by $\mathbf{v} \longmapsto \mathbf{v} \,|\, Q$ the projection from V^k to V^Q and the preimage of any set $G_Q \subseteq V^Q$ is denoted by

$$\mathcal{K}_k(G_Q) = \big\{\mathbf{v} \in V^k \colon (\mathbf{v} \,|\, Q) \in G_Q\big\}\,.$$

We may think of $G_Q \subseteq V^Q$ as a directed hypergraph (where vertices in the directed hyperedges are also allowed to repeat). More generally, for a subset $\mathcal{Q} \subseteq \mathcal{P}([k])$ of the power set of $[k]$ and a family $\mathscr{G} = \{G_Q \colon Q \in \mathcal{Q}\}$ with $G_Q \subseteq V^Q$ for all $Q \in \mathcal{Q}$, we define

$$\mathcal{K}_k(\mathscr{G}) = \bigcap_{Q \in \mathcal{Q}} \mathcal{K}_k(G_Q)\,. \tag{4.21}$$

Moreover, if $H = (V, E)$ is a k-uniform hypergraph on V, then $e_H(\mathscr{G})$ denotes the cardinality of the set

$$E_H(\mathscr{G}) = \big\{(v_1, \dots, v_k) \in \mathcal{K}_k(\mathscr{G}) \colon \{v_1, \dots, v_k\} \in E\big\}\,.$$

Now we are ready to state the generalisation of Definition 1.3.1.

Definition 4.7.2. *Let ρ, $d \in (0,1]$, let $H = (V, E)$ be a k-uniform hypergraph on n vertices, and let $\mathcal{Q} \subseteq \mathcal{P}([k])$ be given. We say that H is* $(\rho, d, \mathcal{Q})$-dense *if for every family $\mathscr{G} = \{G_Q \colon Q \in \mathcal{Q}\}$ associating with each $Q \in \mathcal{Q}$ some $G_Q \subseteq V^Q$ we have*

$$e_H(\mathscr{G}) \geqslant d\,|\mathcal{K}_k(\mathscr{G})| - \rho n^k\,.$$

It is easy to check that for $k = 3$ the following subsets of $\mathcal{P}$‘([3])

$$\mathcal{Q}_{\therefore} = \big\{\{1\},\{2\},\{3\}\big\}\,, \quad \mathcal{Q}_{\stackrel{\cdot}{\cdot\!-\!\cdot}} = \big\{\{1\},\{2,3\}\big\}\,, \quad \text{and} \quad \mathcal{Q}_{\boldsymbol{\wedge}} = \big\{\{1,2\},\{1,3\}\big\}$$

correspond to $\therefore$-, $\stackrel{\cdot}{\cdot\!-\!\cdot}$-, and $\boldsymbol{\wedge}$-dense hypergraphs. More precisely, for every $\star \in \{\therefore, \stackrel{\cdot}{\cdot\!-\!\cdot}, \boldsymbol{\wedge}\}$ we have that a 3-uniform hypergraph is $(\rho, d, \star)$-dense if and only if it is $(\rho, d, \mathcal{Q}_\star)$-dense.

Example 1.3.2 straight forwardly extends to k-uniform hypergraphs. In fact, we may consider a random bipartition $G \cup \overline{G}$ of the $(k-1)$-element subsets of an $(n-2)$-element set and we define a k-uniform hypergraph containing only those hyperedges with the property that all of its $(k-1)$-element subsets are in the same partition class. Finally, we may add two vertices x and y such that the $(k-1)$-uniform link of x is G and the $(k-1)$-uniform link of y is $\overline{G}$. We remark that for $k = 2$ this construction leads to two disjoint cliques with $\sim n/2$ vertices, which is a lower bound construction for Dirac's theorem [21] in graphs.

It is easy to check that the resulting k-uniform hypergraph H does not contain a tight Hamilton cycle and for every fixed $\rho > 0$ it is $(\rho, 2^{1-k}, \mathcal{Q})$-dense for

$$\mathcal{Q} = \big\{Q \in [k]^{(k-2)} \colon 1 \in Q\big\} \cup \big\{\{2, \dots, k\}\big\}$$

with high probability for sufficiently large n. Note that for $k = 3$ we have $\mathcal{Q} = \mathcal{Q}_{\therefore}$ and H provides a lower bound for Theorem 1.3.3. It seems plausible that the hypergraph H is essentially optimal for $\mathcal{Q}$-dense hypergraphs also for $k > 3$, i.e., that $\mathcal{Q}$-dense k-uniform n-vertex hypergraphs with density bigger than 2^{1-k} and minimum vertex degree $\Omega(n^{k-1})$ contain a tight Hamilton cycle. This would be an interesting extension of Theorem 1.3.3 to k-uniform hypergraphs.

Moreover, one can check that for

$$\mathcal{Q}' = \big\{\{1, \dots, k-1\}, \{1, \dots, k-2, k\}\big\}$$

the hypergraph H constructed above is not $(\rho, d, \mathcal{Q}')$-dense for any fixed $d > 0$ and sufficiently small $\rho > 0$. Note that for $k = 3$ we have $\mathcal{Q}' = \mathcal{Q}_{\wedge}$ and, in fact, Theorem 1.3.4 asserts that $(\rho, d, \mathcal{Q}')$-dense hypergraphs with minimum vertex degree $\Omega(n^2)$ contain a Hamilton cycle for any $d > 0$ and sufficiently small ρ. We remark that the proof of Theorem 1.3.4 discussed in Section 4.6 extends to k-uniform $\mathcal{Q}'$-dense hypergraphs with an appropriate minimum vertex degree condition.

Bibliography

[1] Elad Aigner-Horev, David Conlon, Hiêp Hàn, Yury Person, and Mathias Schacht. Quasirandomness in hypergraphs. *Electronic Notes in Discrete Mathematics*, 61:13–19, 2017.

[2] Elad Aigner-Horev and Gil Levy. Tight Hamilton cycles in cherry quasirandom 3-uniform hypergraphs. *arXiv:1712.00186*, 2017.

[3] Peter Allen, Julia Böttcher, Jozef Skokan, and Maya Stein. Regularity inheritance in pseudorandom graphs. *Random Structures & Algorithms*, 56(2):306–338, 2020.

[4] Noga Alon. Explicit ramsey graphs and orthonormal labelings. *the electronic journal of combinatorics*, pages R12–R12, 1994.

[5] József Balogh, Béla Csaba, and Wojciech Samotij. Local resilience of almost spanning trees in random graphs. *Random Structures & Algorithms*, 38(1-2):121–139, 2011.

[6] József Balogh, Robert Morris, and Wojciech Samotij. Independent sets in hypergraphs. *Journal of the American Mathematical Society*, 28(3):669–709, 2015.

[7] Sören Berger, Joonkyung Lee, and Mathias Schacht. Odd cycles in subgraphs of sparse pseudorandom graphs. *arXiv:1906.05100*, 2019.

[8] Guido Besomi, Matías Pavez-Signé, and Maya Stein. Degree conditions for embedding trees. *SIAM J. Discrete Math.*, 33(3):1521–1555, 2019.

[9] Guido Besomi, Matías Pavez-Signé, and Maya Stein. On the Erdős-Sós conjecture for trees with bounded degree. *arXiv e-prints*, 2019.

[10] G R Blakley and Prabir Roy. A hölder type inequality for symmetric matrices with nonnegative entries. *Proc. Amer. Math. Soc.*, 16:1244–1245.

[11] Julia Böttcher. Large-scale structures in random graphs. *Surveys in Combinatorics*, 440:87, 2017.

[12] Stefan A Burr. Ramsey numbers involving graphs with long suspended paths. *J. Lond. Math. Soc.*, 2(3):405–413, 1981.

[13] Stefan A Burr and Paul Erdős. Generalizations of a Ramsey-theoretic result of Chvátal. *J. Graph Theory*, 7(1):39–51, 1983.

[14] Enno Buß, Hiệp Hàn, and Mathias Schacht. Minimum vertex degree conditions for loose Hamilton cycles in 3-uniform hypergraphs. *Journal of Combinatorial Theory, Series B*, 103(6):658–678, 2013.

[15] Fan R K Chung, Ronald L Graham, and Richard M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[16] Vasek Chvátal. Tree-complete graph Ramsey numbers. *J. Graph Theory*, 1(1):93–93, 1977.

[17] Vasek Chvátal and Paul Erdős. A note on Hamiltonian circuits. *Discrete Mathematics*, 2(2):111–113, 1972.

[18] David Conlon. A new upper bound for diagonal ramsey numbers. *Annals of Mathematics*, pages 941–960, 2009.

[19] David Conlon, Jacob Fox, and Yufei Zhao. Extremal results in sparse pseudorandom graphs. *Advances in Mathematics*, 256:206–290, 2014.

[20] David Conlon and William Timothy Gowers. Combinatorial theorems in sparse random sets. *Ann. of Math.*, pages 367–454, 2016.

[21] Gabriel Andrew Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.

[22] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel J. Math.*, 2:183–190.

[23] Paul Erdős. Extremal problems in graph theory. In *Theory of graphs and its applications, Proc. Sympos. Smolenice*, pages 29–36, 1964.

[24] Paul Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292–294, 1947.

[25] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[26] Paul Erdős, Alfréd Rényi, et al. On random graphs. *Publicationes mathematicae*, 6(26):290–297, 1959.

[27] Paul Erdős and Joel H Spencer. Imbalances in k-colorations. *Networks*, 1:379–385, 1972.

[28] Paul Erdős, Arthur H Stone, et al. On the structure of linear graphs. *Bull. Amer. Math. Soc*, 52(1087-1091):1, 1946.

[29] Paul Erdős and George Szekeres. A combinatorial problem in geometry. *Compositio mathematica*, 2:463–470, 1935.

[30] Peter Frankl and Vojtěch Rödl. Large triangle-free subgraphs in graphs withoutk 4. *Graphs and Combinatorics*, 2(1):135–144, 1986.

[31] Peter Frankl and Vojtěch Rödl. Large triangle-free subgraphs in graphs withoutk 4. *Graphs Combin.*, 2(1):135–144, 1986.

[32] Joel Friedman and Nicholas Pippenger. Expanding graphs contain all small trees. *Combinatorica*, 7(1):71–76, 1987.

[33] Luyining Gan and Jie Han. Hamiltonicity in cherry-quasirandom 3-graphs.

[34] Stefanie Gerke and Angelika Steger. *The sparse regularity lemma and its applications*, page 227–258. Surveys in Combinatorics 2005 (London Mathematical Society Lecture Note Series). Cambridge University Press, 2005.

[35] Roman Glebov, Jan Volec, et al. A problem of erdős and sós on 3-graphs. *Israel Journal of Mathematics*, 211(1):349–366, 2016.

[36] Penny E Haxell. Tree embeddings. *J. Graph Theory*, 36(3):121–130, 2001.

[37] Penny E Haxell, Yoshiharu Kohayakawa, and Tomasz Luczak. Turán' s extremal problem in random graphs: Forbidding even cycles. *Journal of Combinatorial Theory, Series B*, 64(2):273–287, 1995.

[38] Penny E Haxell, Yoshiharu Kohayakawa, and Tomasz Luczak. Turán's extremal problem in random graphs: Forbidding odd cycles. *Combinatorica*, 16(1):107–122, 1996.

[39] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*, volume 45. John Wiley & Sons, 2011.

[40] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*, volume 45. John Wiley & Sons, 2011.

[41] Gyula Y Katona and Hal A Kierstead. Hamiltonian chains in hypergraphs. *Journal of Graph Theory*, 30(3):205–212, 1999.

[42] Yoshiharu Kohayakawa. Szemerédi's regularity lemma for sparse graphs. In *Found. Comput. Math.*, pages 216–230. Springer, 1997.

[43] Yoshiharu Kohayakawa, Tomasz Łuczak, and Vojtěch Rödl. Arithmetic progressions of length three in subsets of a random set. *Acta Arithmetica*, 75(2):133–163, 1996.

[44] Yoshiharu Kohayakawa, Tomasz Łuczak, and Vojtěch Rödl. On k4-free subgraphs of random graphs. *Combinatorica*, 17(2):173–213, 1997.

[45] Yoshiharu Kohayakawa, Guilherme Oliveira Mota, and Mathias Schacht. Monochromatic trees in random graphs. *Math. Proc. Philos. Soc.*, 166(1):191–208, 2019.

[46] Yoshiharu Kohayakawa and Vojtěch Rödl. Szemerédi's regularity lemma and quasi-randomness. In *Recent advances in algorithms and combinatorics*, pages 289–351. Springer, 2003.

[47] János Komlós, Gábor N Sárközy, and Endre Szemerédi. Proof of a packing conjecture of Bollobás. *Combin. Probab. Comput.*, 4(3):241–255, 1995.

[48] Michael Krivelevich. Embedding spanning trees in random graphs. *SIAM J. Discrete Math.*, 24(4):1495–1500, 2010.

[49] Daniela Kühn and Deryk Osthus. Loose Hamilton cycles in 3-uniform hypergraphs of high minimum degree. *Journal of Combinatorial Theory, Series B*, 96(6):767–821, 2006.

[50] John Lenz, Dhruv Mubayi, and Richard Mycroft. Hamilton cycles in quasirandom hypergraphs. *Random Structures & Algorithms*, 49(2):363–378, 2016.

[51] Shoham Letzter. Path Ramsey number for random graphs. *Combin. Probab. Comput.*, 25(4):612–622, 2016.

[52] Tomasz Łuczak, Andrzej Ruciński, and Bernd Voigt. Ramsey properties of random graphs. *J. Combin. Theory Ser. B*, 56(1):55–68, 1992.

[53] Richard Montgomery. Embedding bounded degree spanning trees in random graphs. *arXiv:1405.6559*, 2014.

[54] Luiz Moreira. Ramsey goodness of paths in random graphs. *arXiv:1909.00030*, 2019.

[55] Yuejian Peng, Vojtech Rödl, and Andrzej Ruciński. Holes in graphs. *Electron. J. Combin.*, 9(1):Research Paper 1, 18.

[56] J. Polcyn. Personal communication.

[57] Chr. Reiher and Mathias Rödl, Vojtěch andSchacht. On a generalisation of mantel's theorem to uniformly dense hypergraphs. *Int. Math. Res. Not. IMRN*, (16):4899–4941.

[58] Christian Reiher. Extremal problems in uniformly dense hypergraphs. *arXiv:1901.04027*, 2019.

[59] Christian Reiher, Vojtěch Rödl, Andrzej Ruciński, Mathias Schacht, and Endre Szemerédi. Minimum vertex degree condition for tight Hamiltonian cycles in 3-uniform hypergraphs. *arXiv:1611.03118*, 2016.

[60] Christian Reiher, Vojtěch Rödl, and Mathias Schacht. Embedding tetrahedra into quasirandom hypergraphs. *Journal of Combinatorial Theory, Series B*, 121:229–247, 2016.

[61] Christian Reiher, Vojtěch Rödl, and Mathias Schacht. Some remarks on the extremal function for uniformly two-path dense hypergraphs. *arXiv:1602.02299*, 2016.

[62] Vojtěch Rödl and Andrzej Rucinski. Lower bounds on probability thresholds for Ramsey properties. *Combinatorics, Paul Erdős is eighty*, 1:317–346, 1993.

[63] Vojtěch Rödl and Andrzej Ruciński. Threshold functions for Ramsey properties. *J. Amer. Math. Soc.*, 8(4):917–942, 1995.

[64] Vojtěch Rödl, Andrzej Ruciński, and Endre Szemerédi. A Dirac-type theorem for 3-uniform hypergraphs. *Combinatorics, Probability and Computing*, 15(1-2):229–251, 2006.

[65] Klaus F Roth. On certain sets of integers. *Journal of the London Mathematical Society*, 1(1):104–109, 1953.

[66] Imre Z Ruzsa and Endre Szemerédi. Triple systems with no six points carrying three triangles. *Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bolyai*, 18:939–945, 1978.

[67] Ashwin Sah. Diagonal ramsey via effective quasirandomness. *arXiv:2005.09251*, 2020.

[68] David Saxton and Andrew Thomason. Hypergraph containers. *Inventiones mathematicae*, 201(3):925–992, 2015.

[69] Mathias Schacht. Extremal results for random discrete structures. *Ann. of Math.*, pages 333–365, 2016.

[70] Issai Schur. *Uber die Existenz unendlich vieler Primzahlen in einigen speziellen arithmetischen Progressionen.* 1912.

[71] Benny Sudakov, Tibor Szabó, and H Van Vu. A generalization of turán's theorem. *Journal of Graph Theory*, 49(3):187–195, 2005.

[72] Endre Szemeredi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith*, 27:299–345, 1975.

[73] Andrew Thomason. Pseudo-random graphs. In *North-Holland Mathematics Studies*, volume 144, pages 307–331. Elsevier, 1987.

[74] Andrew Thomason. An upper bound for some ramsey numbers. *Journal of graph theory*, 12(4):509–517, 1988.

[75] Henry Towsner. σ-algebras for quasirandom hypergraphs. *Random Structures & Algorithms*, 50(1):114–139, 2017.

[76] Paul Turán. Eine extremalaufgabe aus der graphentheorie. *Mat. Fiz. Lapok*, 48(436-452):61, 1941.

[77] Bartel van der Waerden. Beweis einer baudetschen vermutung. *Nieuw Arch. Wisk.*, 19:212–216, 1927.