Extremal Graph Colouring and Tiling Problems

Jan Corsten

A thesis submitted for the degree of
Doctor of Philosophy

Department of Mathematics

The London School of Economics
and Political Science

9th July 2020

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work, with the exceptions outlined below.

I declare that this thesis consists of 46960 words.

Statement of co-authored work

The contents of Chapter 2 are well-known results of various authors not including myself.

I confirm that Chapter 3 contains the following joint work. Section 3.1 is based on [18], which is jointly co-authored with Sebastián Bustamante, Nóra Frankl, Alexey Pokrovskiy and Jozef Skokan. Section 3.2 is based on [31], which is jointly co-authored with Walner Mendonça.

I confirm that Chapter 4 contains the following joint work. Section 4.1 is based on [30], which is jointly co-authored with Louis DeBiasio, Ander Lamaison and Richard Lang. Section 4.2 is based on [29], which is jointly co-authored with Louis DeBiasio and Paul McKenney.

I confirm that Chapter 5 is based on [3], which is jointly co-authored with Peter Allen, Julia Böttcher, Ewan Davies, Matthew Jenssen, Patrick Morris, Barnaby Roberts and Jozef Skokan.

Abstract

In this thesis, we study a variety of different extremal graph colouring and tiling problems in finite and infinite graphs.

Confirming a conjecture of Gyárfás, we show that for all $k, r \in \mathbb{N}$ there is a constant $C > 0$ such that the vertices of every r-edge-coloured complete k-uniform hypergraph can be partitioned into a collection of at most C monochromatic tight cycles. We shall say that the family of tight cycles has *finite r-colour tiling number*. We further prove that, for all natural numbers k, p and r, the family of p-th powers of k-uniform tight cycles has finite r-colour tiling number. The case where $k = 2$ settles a problem of Elekes, Soukup, Soukup and Szentmiklóssy. We then show that for all natural numbers Δ, r, every family $\mathcal{F} = \{F_1, F_2, \dots\}$ of graphs with $v(F_n) = n$ and $\Delta(F_n) \leq \Delta$ for every $n \in \mathbb{N}$ has finite r-colour tiling number. This makes progress on a conjecture of Grinshpun and Sárközy.

We study Ramsey problems for infinite graphs and prove that in every 2-edge-colouring of $K_{\mathbb{N}}$, the countably infinite complete graph, there exists a monochromatic infinite path P such that $V(P)$ has upper density at least $(12 + \sqrt{8})/17 \approx 0.87226$ and further show that this is best possible. This settles a problem of Erdős and Galvin. We study similar problems for many other graphs including trees and graphs of bounded degree or degeneracy and prove analogues of many results concerning graphs with linear Ramsey number in finite Ramsey theory.

We also study a different sort of tiling problem which combines classical problems from extremal and probabilistic graph theory, the Corrádi–Hajnal theorem and (a special case of) the Johansson–Kahn–Vu theorem. We prove that there is some constant $C > 0$ such that the following is true for every $n \in 3\mathbb{N}$ and every $p \geq Cn^{-2/3}(\log n)^{1/3}$. If G is a graph on n vertices with minimum degree at least $2n/3$, then G_p (the random subgraph of G obtained by keeping every edge independently with probability p) contains a triangle tiling with high probability.

Acknowledgements

Firstly, I would like to thank my supervisors Peter and Julia, and Jozef, who acted like a third supervisor. I would like to express my gratitude for their guidance throughout my PhD, for all the discussions (mathematical and non-mathematical), and for encouraging me to participate in workshops, conferences and summer schools around the world.

I would like to thank my fellow PhD students and good friends Attila and Nóra for our collaboration and for their unlimited supply of fun puzzles. I would like also like to thank all other PhD students for making the PhD office a great work environment, and for all the fun times we had at LSE and around the world. I would like to thank Kate, Enfale, Becca, Sarah and Ed for their administrative support and creating such a nice atmosphere in the department. I would like to thank Jozef and Jan for keeping us well hydrated.

Finally, I would like to thank my parents and my sister for their constant support throughout my life.

Contents

1 Introduction

Extremal graph theory seeks to answer questions of the following form: what is the largest or smallest value of some graph parameter among all graphs of a given class? The possibly first result in this area determines how many edges a triangle-free graph can have.

Theorem 1.0.1 (Mantel [94]). *Every graph with n vertices and more than $n^2/4$ edges contains a triangle. Furthermore, the complete bipartite graph with parts of size $\lceil n/2 \rceil$ and $\lfloor n/2 \rfloor$ contains $\lfloor n^2/4 \rfloor$ edges and is triangle-free.*

Another classical example of extremal graph theory is the study of Ramsey numbers. The *Ramsey number* of a graph H, denoted by $\mathrm{R}(H)$, is the smallest integer N such that in every two-colouring of the edges of the complete graph on N vertices K_N, we can find a monochromatic copy of H (that is a subgraph of K_N which is isomorphic to H and all its edges receive the same colour). Ramsey numbers are named after Ramsey, who proved the following result.

Theorem 1.0.2 (Ramsey [100]). $\mathrm{R}(K_n) < \infty$ *for every* $n \in \mathbb{N}$.

It immediately follows from Ramsey's theorem that Ramsey numbers of all finite graphs are finite and there has been a lot of work in determining and estimating them. One topic of interest is to classify which graphs have small, i.e. linear, Ramsey number. To be more precise, define a *sequence of graphs* to be a family of graphs $\mathcal{F} = \{F_1, F_2, \ldots\}$ where F_n has n vertices for every $n \in \mathbb{N}$. Such $\mathcal{F}$ is said to have linear Ramsey number if there exists a constant $C > 0$ such that $\mathrm{R}(F_n) \leq Cn$ for every $n \in \mathbb{N}$. Looking at the problem from the inverse perspective, it follows directly from the definition that if $\mathcal{F}$ has linear Ramsey number, it is possible to cover a linearly large part of every 2-edge-coloured K_n using only one monochromatic copy from $\mathcal{F}$.

Observation 1.0.3. *For every sequence of graphs $\mathcal{F}$ with linear Ramsey number, there is a constant $c = c(\mathcal{F}) > 0$, such that every two-edge-coloured K_n contains a monochromatic copy of some $F \in \mathcal{F}$ with at least cn vertices.*

By greedily applying Observation 1.0.3 to the set of remaining vertices, we can actually cover most of K_n using only "few" monochromatic copies from $\mathcal{F}$.[1]

Observation 1.0.4. *For every $\varepsilon > 0$ and every sequence of graphs $\mathcal{F}$ with linear Ramsey number, there is a constant $C = C(\mathcal{F}, \varepsilon)$ such that the following is true for every two-edge-coloured K_n. There is a collection of at most C vertex-disjoint monochromatic copies from $\mathcal{F}$ whose union covers at least $(1 - \varepsilon)n$ vertices.*

In other words, Observation 1.0.4 asserts that it is possible to cover almost all vertices of every two-edge coloured K_n by vertex-disjoint monochromatic copies from $\mathcal{F}$ whose average size is linearly large. It is natural to ask for which families $\mathcal{F}$ we can extend Observation 1.0.4 to cover all the vertices of K_n. Such problems are called *graph tiling problems* and we shall discuss some of them in Chapter 3.

Ramsey's theorem further implies that every two-edge-coloured $K_{\mathbb{N}}$ (the countably infinite complete graph) contains a monochromatic copy of $K_{\mathbb{N}}$.[2] Therefore, there is no direct extensions of Ramsey numbers to infinite graphs. It is possible however to generalise the concept of graphs with linear Ramsey number to infinite graphs using a density formulation similar as in Observation 1.0.3. The *upper density* of a set $A \subseteq \mathbb{N}$ is defined as $\overline{\mathrm{d}}(A) := \limsup_{n\to\infty} |A \cap [n]| / n$. It is not hard to see that there exist 2-edge-colourings of $K_{\mathbb{N}}$ such that the vertex-set of every monochromatic copy of $K_{\mathbb{N}}$ has upper density 0 (e.g., a random colouring). In Chapter 4 we will discuss infinite graphs for which this is not the case and prove infinite analogues of many results about graphs with linear Ramsey number.

[1]This was first observed by Erdős, Gyárfás and Pyber [42] to the best of the author's knowledge.
[2]In fact, this is how the theorem was originally stated.

We will also study a different sort of graph tiling problems in random graphs. The binomial random graph $G(n, p)$ is the random subgraph of K_n in which every edge of K_n is present independently with probably p. This random graph was first introduced by Gilbert [56], while Erdős and Rényi [44] independently introduced the related random graph $G(n, m)$. Erdős and Rényi then initiated the systematic study of both types of random graphs in a series of papers [45, 46, 47, 48] and it has been a very active research topic ever since. One of the central questions in the theory of random graphs is to determine threshold probabilities for certain graph properties. A classical example of this is the following theorem, which determines when $G(n, p)$ is connected.

Theorem 1.0.5 (Gilbert [56]). *Let C_n denote the set of all connected graphs on n vertices. Then, for every $\varepsilon > 0$, we have*

$$\lim_{n\to\infty} \mathbb{P}\left[G(n, p(n)) \in C_n\right] = \begin{cases} 0 & \text{if } p(n) \leq (1-\varepsilon)\log(n)/n \text{ for all } n \in \mathbb{N}, \\ 1 & \text{if } p(n) \geq (1+\varepsilon)\log(n)/n \text{ for all } n \in \mathbb{N}. \end{cases}$$

We will study similar questions regarding the existence of a triangle tiling in random subgraphs of certain deterministic graphs in Chapter 5.

1.1 Notation

1.1.1 Elementary Notation

We denote by $\mathbb{R}$ the set of real numbers, by $\mathbb{R}_{\geq 0}$ the set of non-negative real numbers and by $\mathbb{R}_+$ the set of positive real numbers. We denote by $\mathbb{N} := \{1, 2, \dots\}$ the set of positive integers and by $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ the set of non-negative integers.

Given a set V, we denote by 2^V the set of all subsets of V and, given $k \in \mathbb{N}_0$, we denote by $\binom{V}{k}$ the set of all k-element subsets of V. For integers $0 \leq t \leq n$, we define $n!_t$ to be the number of ways to select a list of t distinct numbers from $[n]$. That is, $n!_t := \frac{n!}{(n-t)!} = n \cdot (n-1) \cdots (n-t+1)$.

Given $k \in \mathbb{N}_0$, we denote by $[k] := \{1, \dots, k\}$ the set of the first k positive integers. Given $k \in \mathbb{N}_0$, we denote by $[k]_0 := \{0, \dots, k\}$ the set of the first $k + 1$

non-negative integers. Given $k \leq \ell \in \mathbb{N}_0$, we denote by $[k, \ell] := \{k, \dots, \ell\}$ the set of all integers z with $k \leq z \leq \ell$. Given $a \leq b \in \mathbb{R}$, we denote by $[a, b]_{\mathbb{R}}$ the set of all reals x with $a \leq x \leq b$. Given $a, b \in \mathbb{R}$, we denote by $(a, b) = (a, b)_{\mathbb{R}}$ the set of all reals x with $a < x < b$.

We denote by exp the exponential function with base e and, for $k \in \mathbb{N}$, by $\exp^k$ the k-th composition of the exponential function. We denote by log the natural logarithm, that is, the logarithm with base e.

In this thesis we will frequently break long proofs into smaller claims. We denote by ■ the end of the proof of such a claim and by □ the end of the main proof.

1.1.2 O-Notation

Let $f, g : \mathbb{N} \to \mathbb{R}$ be functions taking non-negative values. We say that

- $f = O(g)$ if there is some $C > 0$ such that $f(n) \leq C \cdot g(n)$ for all sufficiently large $n \in \mathbb{N}$.
- $f = \Omega(g)$ if $g = O(f)$.
- $f = \Theta(g)$ if $f = O(g)$ and $g = O(f)$.
- $f = o(g)$ if for all $\varepsilon > 0$, there is some $n_0 > 0$ so that $f(n) < \varepsilon \cdot g(n)$ for all $n \geq n_0$.
- $f = \omega(g)$ if $g = o(f)$.

If the functions f, g depend on more than one variable, we shall indicate the parameter tending to infinity in the subscript and consider all other variables as constant. For example, $n^2 + m^3 = o_n(n^3 + m^2)$ and $n^2 + m^3 = \omega_m(n^3 + m^2)$.

Furthermore, given real numbers a, b, c, we say $a = b \pm c$ if $|a - b| \leq |c|$, and we will frequently use the following slightly informal notation: Given reals $\varepsilon, \delta > 0$, we say $\delta \ll \varepsilon$ if δ can be chosen arbitrarily small in terms of ε. This shall replace the quantification "for all $\varepsilon > 0$ there exists some $\delta_0 > 0$ such that for all $0 < \delta \leq \delta_0$ ..." when we have a long list of such quantifications.

1.1.3 Graphs

We use standard notation from graph theory. A *graph* $G = (V, E)$ is a tuple consisting of a set V, called the *vertices* of G, and a set $E \subseteq \binom{V}{2}$, called the *edges* of G. We write $V(G)$ to refer to the vertex-set of G and $E(G)$ to refer to the edge-set of V. We further denote the sizes of $V(G)$ and $E(G)$ by $v(G) := |V(G)|$ and $e(G) := |E(G)|$. Note that $V(G)$ and $E(G)$ are possibly infinite. We will often abbreviate an edge $\{u, v\} \in E(G)$ as uv.

Definition 1.1.1 (Neighbourhoods and degrees)**.** Given a graph G, a vertex $v \in V(G)$ and a set $U \subseteq V(G)$, we define the neighbourhood of v in U by $N_G(v, U) := \{u \in U : uv \in E(G)\}$. Furthermore, the *degree* of v in U is given by $\deg_G(v, U) := |N_G(v, U)|$. Given a set of vertices $V \subseteq V(G)$, we define the *common neighbourhood* of V in U by $N_G(V; U) := \bigcap_{v \in V} N_G(v, U)$ and $\deg_G(V; U) := |N_G(V; U)|$. Given two vertices $v_1, v_2 \in V(G)$, we write $N_G(v_1, v_2; U)$ for $N_G(\{v_1, v_2\}; U)$, and similarly we write $\deg_G(v_1, v_2; U)$ for $\deg_G(\{v_1, v_2\}; U)$. If $U = V(G)$, we simply write $N_G(v)$, $\deg_G(v)$, $N_G(V)$ and $\deg_G(V)$ and if G is clear from context we drop the subscript.

Definition 1.1.2 (Families and sequences of graphs)**.** A *family of graphs* $\mathcal{F}$ is an infinite set of graphs. A *sequence of graphs* is a family of graphs $\mathcal{F} = \{F_1, F_2, \ldots\}$ with $v(F_n) = n$ for all $n \in \mathbb{N}$.

Definition 1.1.3 (Maximum and minimum degree)**.** The *maximum degree* of G is defined by $\Delta(G) := \sup_{v \in V(G)} \deg(v)$ and the *minimum degree* is defined by $\delta(G) := \min_{v \in V(G)} \deg(v)$. The *maximum degree* of a family of graphs $\mathcal{F}$ is given by $\Delta(\mathcal{F}) := \sup_{F \in \mathcal{F}} \Delta(F)$.

Definition 1.1.4 (Degeneracy)**.** A graph G is called *d-degenerate* for some $d \in \mathbb{N}$ if there is a linear order $<$ of $V(G)$ such that $|N(v) \cap \{u \in V(G) : u < v\}| \leq d$ for every $v \in V(G)$. We denote by $\mathrm{degen}(G)$ the smallest $d \in \mathbb{N}$ such that G is d-degenerate. The *degeneracy* of a family of graphs $\mathcal{F}$ is given by $\mathrm{degen}(\mathcal{F}) := \sup_{F \in \mathcal{F}} \mathrm{degen}(F)$.

Note that we have $\mathrm{degen}(G) \leq \Delta(G)$ and $\chi(G) \leq \mathrm{degen}(G) + 1$ for every graph G. Furthermore, there are families with finite degeneracy but infinite maximum degree (for example, the family of all stars).

Definition 1.1.5 (Subgraphs)**.** A graph H is called a *subgraph* of a graph G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. We call H *spanning* if $V(H) = V(G)$. If $U \subseteq V(G)$, we define $G[U]$ to be the subgraph of G induced by U, i.e. $V(G[U]) = U$ and $E(G[U]) = \{e \in E(G) : e \subseteq U\}$.

Definition 1.1.6 (Independent sets)**.** A set $I \subseteq V(G)$ is called *independent* if there is no edge $e \in E(G)$ with $e \subseteq I$. The *independence number* of G is given by $\alpha(G) := \sup\{|I| : I \in \mathcal{I}(G)\}$, where $\mathcal{I}(G)$ denotes the set of all independent sets in G. Given a partition $\mathcal{P}$ of $V(G)$, we say that G is $\mathcal{P}$-partite if every $U \in \mathcal{P}$ is an independent set. A graph G is t-partite if it is $\mathcal{P}$-partite for some partition $\mathcal{P}$ of $V(G)$ with t parts. The minimum $t \in \mathbb{N}$ such that G is t-partite is called the *chromatic number* of G and is denoted by $\chi(G)$. If G is not $\mathcal{P}$-partite for any partition $\mathcal{P}$ into finitely many parts, we set $\chi(G) := \infty$.

Definition 1.1.7 (Complete (partite) graphs)**.** We define $K(V) := (V, \binom{V}{2})$ to be the complete graph on the vertex-set V and we write $K_n := K([n])$. Given a collection disjoint sets $\mathcal{P} = \{V_1, \ldots, V_t\}$, we denote by $K(\mathcal{P}) = K(V_1, \ldots, V_t)$ the complete t-partite graph with parts $V_1, \ldots, V_t$, i.e. the graph with vertex-set $V(G) = V_1 \cup \ldots \cup V_t$ and edge-set $E(G) = \{e \in \binom{V}{2} : |e \cap V_i| \leq 1 \text{ for all } i \in [t]\}$. Given a graph G and disjoint sets $U_1, \ldots, U_t$, we denote by $G[U_1, \ldots, U_t]$ the t-partite subgraph of G with vertex-set $U_1 \cup \ldots \cup U_t$ and edge-set $E = \{e \in E(G) : |e \cap V_i| \leq 1 \text{ for all } i \in [t]\}$.

Definition 1.1.8 ((Partial) Embeddings)**.** A *partial function* $f : X \to Y$ is a function $f' : X' \to Y$ for some $X' \subseteq X$. We define the *range* of f by $\operatorname{rg} f := \{y \in Y : f(x) = y \text{ for some } x \in X\}$ and the domain of f, denoted by $\operatorname{dom} f$, as the set of all $x \in X$ for which $f(x)$ is defined. A *(partial) embedding* of a graph H into another graph G is a (partial) injective function $f : V(G) \to V(H)$ such that $f(u)f(v) \in E(H)$ whenever $uv \in E(G)$.

Definition 1.1.9 (Orderings)**.** A *linear order* on a set X is a binary relation $\leq$ on X such that, for all $x, y \in X$, we have

(i) $x \leq y$ or $y \leq x$,

(ii) $x \leq y$ and $y \leq x$ if and only if $x = y$, and

(iii) if $x \leq y$ and $y \leq z$, then $x \leq z$.

Observe that every linear order on a finite set X induces an enumeration $X = \{x_1, x_2, \ldots\}$ so that $x_1 < x_2 < \ldots$, where we write $x < y$ if $x \leq y$ and $x \neq y$.

We will now define the graphs which will appear frequently in this thesis.

Definition 1.1.10 (Paths)**.** For $n \geq 0$, we define the *path* P_n as the graph with $n+1$ vertices $v_1, \ldots, v_{n+1}$ and n edges $v_1v_2, v_2v_3, \ldots, v_nv_{n+1}$.

Definition 1.1.11 (Cycles)**.** For $n \geq 3$, we define the *cycle* C_n as the graph obtained from P_{n-1} by adding the edge v_nv_1. For technical reasons, we shall consider a single vertex and a single edge as degenerate cycles C_1 and C_2.

Definition 1.1.12 (Trees)**.** A graph T is called a *tree* if it does not contain any finite cycles of length at least 3 as subgraphs. A *rooted tree* is a tuple (T, r) where T is a tree and $r \in V(T)$. Given a rooted tree (T, r) and $s, t \in V(T)$, we say that t is a *child* of t (and s is the *parent* of t) if $st \in E(T)$ and s is on the unique path from t to r. Observe that every vertex other than the root has a unique parent but possibly many or no children.

Definition 1.1.13 (Powers of graphs)**.** The *distance* of two vertices u, v in a graph H, is the length of the shortest path from u to v in H. If no such path exists, the distance of u and v is infinite. The *k-th (distance) power* of a graph H, denoted by H^k, is the graph obtained from H by adding an edge between any two vertices of distance at most k in H. Popular examples of this are powers of paths and powers of cycles.

1.1.4 Hypergraphs

Given an integer $k \geq 2$, a *k-uniform hypergraph* (for short k-graph) $G = (V, E)$ is a tuple consisting of a set V, called the *vertices* of G, and a set $E \subseteq \binom{V}{k}$, called the *edges* of G. We write $V(G)$ to refer to the vertex-set of G and $E(G)$ to refer to the edge-set of G. We further denote the sizes of $V(G)$ and $E(G)$ by $v(G) := |V(G)|$ and $e(G) := |E(G)|$.

Definition 1.1.14 (Degree)**.** Given some $e \subseteq V(G)$ with $|e| \in [k-1]$, the *degree* of e is given by $\deg_G(e) := |f \in E(H) : e \subseteq f|$. If $e = \{v\}$ for some $v \in V(H)$ we simply write $\deg_G(v)$ for $\deg_G(\{v\})$ and if $|e| = k-1$ we also call $\deg_G(e)$ *co-degree*. If G is clear from context, we drop the subscript.

Definition 1.1.15 (Subgraphs)**.** A k-graph H is called a *sub-k-graph* of a k-graph G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. We call H *spanning* if $V(H) = V(G)$. If $U \subseteq V(G)$, we define $G[U]$ to be the subgraph of G induced by U, i.e. $V(G[U]) = U$ and $E(G[U]) = \{e \in E(G) : e \subseteq U\}$.

Definition 1.1.16 (Independent sets)**.** A set $I \subseteq V(G)$ is called *independent* if there is no edge $e \in E(G)$ with $e \subseteq I$. The *independence number* of G, denoted by $\alpha(G)$, is the size of the largest independent set in G. Given a k-graph G and a partition $\mathcal{P} = \{V_1, \ldots, V_t\}$ of $V(G)$, we say that G is $\mathcal{P}$-partite if $|e \cap V_i| \leq 1$ for every $e \in E(H)$ and every $i \in [t]$. G is called t-partite if it is $\mathcal{P}$-partite for some partition $\mathcal{P}$ of V with t parts.

Definition 1.1.17 (Complete (partite) graphs)**.** We define $K^{(k)}(V) := (V, \binom{V}{k})$ to be the complete k-graph on the vertex-set V and we write $K_n^{(k)} := K^{(k)}([n])$. Given a collection disjoint sets $\mathcal{P} = \{V_1, \ldots, V_t\}$, we denote by $K^{(k)}(\mathcal{P}) = K^{(k)}(V_1, \ldots, V_t)$ the complete t-partite k-graph with parts $V_1, \ldots, V_t$, i.e. the k-graph with vertex-set $V(G) = V_1 \cup \ldots \cup V_t$ and edge-set $E(G) = \{e \in \binom{V}{k} : |e \cap V_i| \leq 1 \text{ for all } i \in [t]\}$. Given a k-graph G and disjoint sets $U_1, \ldots, U_t$, we denote by $G[U_1, \ldots, U_t]$ the t-partite subgraph of G with vertex-set $U_1 \cup \ldots \cup U_t$ and edge-set $E = \{e \in E(G) : |e \cap V_i| \leq 1 \text{ for all } i \in [t]\}$. Given some $2 \leq j \leq k-1$ and a j-graph H, we define $K^{(k)}(H)$ to be the set of all k-cliques in $H^{(j)}$, seen as a k-graph on V.

Definition 1.1.18 (Link graphs and (more) degrees)**.** Given a k-graph G and some $e = \{v_1, \ldots, v_\ell\} \subseteq V(G)$ with $\ell \in [k-1]$, we define the *link-graph* $\mathrm{Lk}_G(e)$ as the $(k-\ell)$-graph with vertex-set $V(G)$ and edge-set $\{f \in \binom{V(G)}{k-\ell} : e \cup f \in E(G)\}$. Note that $\deg_G(e) = |\mathrm{Lk}_G(e)|$. If in addition we are given disjoint sets $V_1, \ldots, V_{k-\ell} \subseteq V(G) \setminus e$, we denote by $\mathrm{Lk}_G(v_1, \ldots, v_\ell; V_1, \ldots, V_{k-\ell})$ the $(k-\ell)$-partite $(k-\ell)$-graph with parts $V_1, \ldots, V_{k-\ell}$ and edge-set $\{e \in K^{(k-\ell)}(V_1, \ldots, V_{k-\ell}) : e \cup \{v_1, \ldots, v_\ell\} \in E(G)\}$. Furthermore, we define $\deg_G(v_1, \ldots, v_\ell; V_1, \ldots, V_{k-\ell}) := |\mathrm{Lk}_G(v_1, \ldots, v_\ell; V_1, \ldots, V_{k-\ell})|$. If G is clear from context, we drop the subscript.

Definition 1.1.19 (Loose path)**.** The *loose k-uniform path* of length $m \geq 1$ is the k-graph consisting of $m(k-1)+1$ distinct linearly ordered vertices and m edges, each of which is formed of k consecutive vertices so that consecutive edges intersect in exactly one vertex. More precisely, its vertex-set is $\{v_1, \ldots, v_{m(k-1)+1}\}$ and its edge-set is $\{v_1, \ldots, v_k\}, \{v_k, \ldots, v_{2k-1}\}, \ldots, \{v_{(m-1)(k-1)+1}, \ldots, v_{m(k-1)+1}\}$. We consider a single vertex as a loose path of length 0.

Definition 1.1.20 (Loose cycle)**.** The *loose k-uniform cycle* of length $m \geq 3$ is the k-graph consisting of $m(k-1)$ distinct cyclically ordered vertices and m edges, each of which formed of k consecutive vertices so that consecutive edges intersect in exactly one vertex. More precisely, its vertex-set is $\{v_1, \ldots, v_{m(k-1)}\}$ and its edge-set is $\{v_1, \ldots, v_k\}, \{v_k, \ldots, v_{2k-1}\}, \ldots, \{v_{(m-1)(k-1)+1}, \ldots, v_1\}$. We consider a single vertex as a loose cycle of length 1 and a single edge as a loose cycle of length 2.

Definition 1.1.21 (Tight path)**.** The *tight k-uniform path* of length $m \geq 1$ is the k-graph with $m+k-1$ distinct linearly ordered vertices in which any k consecutive vertices form an edge. More precisely, its vertex-set is $\{v_1, \ldots, v_{m+k-1}\}$ and its edges are $\{v_1, \ldots, v_k\}$, $\{v_2, \ldots, v_{k+1}\}$, $\ldots$, $\{v_m, \ldots, v_{m+k-1}\}$. For $i \in [k-1]$, we consider an independent set of size i a tight path (all of length 0).

Definition 1.1.22 (Tight cycle)**.** The *tight k-uniform cycle* of length $m \geq 1$ is the k-graph with m distinct cyclically ordered vertices in which any k consecutive vertices form an edge. More precisely, its vertex-set is $\{v_1, \ldots, v_m\}$ and its edges are $\{v_1, \ldots, v_k\}$, $\{v_2, \ldots, v_{k+1}\}$, $\ldots$, $\{v_m, \ldots, v_{k-1}\}$. Note that, for $m \in [k-1]$, the tight cycle of length m has no edges and the tight cycle of length k is a single edge.

Definition 1.1.23 (Powers of paths and cycles)**.** The p-th power of a k-uniform tight path (cycle) is the k-graph obtained by replacing every edge of the $(k+p-1)$-uniform tight path (cycle) by the complete k-graph on $k+p-1$ vertices (on the same vertex-set). For $k=2$ this coincides with Definition 1.1.13 for paths (cycles).

1.1.5 Edge Colourings and Ramsey Numbers

Since a graph is precisely a 2-graph, we will only define the following concepts for hypergraphs. Given a k-graph G and a positive integer $r \in \mathbb{N}$, an r-edge-colouring of G is an assignment of colours from a list of r colours to each edge of G. More formally, an r-edge-colouring of G is a function $\chi : E(G) \to S$, where S is a set of size r (usually $S = \{\text{red}, \text{blue}\}$ if $r = 2$ and $S = [r]$ if $r \geq 3$). An r-edge-coloured k-graph is a tuple (G, χ) of a k-graph G and an r-edge-colouring χ of G. We will frequently abuse notation and write r-edge-coloured k-graph G.

Two k-graphs H and G are isomorphic if there is a bijection $f : V(H) \to V(G)$ such that $e \in E(H)$ if and only if $f(e) := \{f(v) : v \in e\} \in E(G)$. Given an r-edge-coloured k-graph G and a k-graph H, a monochromatic copy of H in G is a (not necessarily induced) subgraph H' of G such that H' is isomorphic to H and every edge of H' receives the same colour.

Definition 1.1.24. The r-colour k-uniform *Ramsey number* of a k-graph G, denoted by $\mathrm{R}_r^{(k)}(G)$, is the smallest integer N such that every r-coloured $K_N^{(k)}$ contains a monochromatic copy of G. If $k = 2$, we drop the superscript (k) and simply say r-colour Ramsey number of G. If $r = 2$, we drop the subscript r and simply say k-uniform Ramsey number of G. If $r = k = 2$, we drop both superscript and subscript and simply say Ramsey number of G.

We will make the following definition only for ordinary graphs (i.e. for k=2).

Definition 1.1.25 (Coloured subgraphs, neighbourhoods and degrees)**.** Given an edge-coloured graph G and a colour i, we define G_i to be the spanning subgraph of G with all edges of colour i. If G is clear from context and we are given a vertex $v \in V(G)$ and a set $U \subseteq V(G)$, we write $N_i(v, U) := N_{G_i}(v, U)$ and $\deg_i(v, U) := \deg_{G_i}(v, U)$. If $U = V(G)$, we simply write $N_i(v)$ and $\deg_i(v)$.

1.1.6 Infinite Graph Theory

Definition 1.1.26 (Upper and lower density)**.** The *upper density* of a set $A \subseteq \mathbb{N}$ is defined as

$$\overline{\mathrm{d}}(A) = \limsup_{t \to \infty} \frac{|A \cap [t]|}{t}$$

and the upper density of a graph G with $V(G) \subseteq \mathbb{N}$ is given by the upper density of its vertex-set, i.e. $\overline{\mathrm{d}}(G) := \overline{\mathrm{d}}(V(G))$. The *lower density* $\underline{\mathrm{d}}(A)$ (or $\underline{\mathrm{d}}(G)$) is defined similarly in terms of the infimum and we speak of the *density*, whenever lower and upper density coincide.

Definition 1.1.27 (Ramsey upper density)**.** The *r-colour Ramsey upper density* of G, denoted by $\overline{\mathrm{Rd}}_r(G)$, is the supremum over all $d \geq 0$ such that every r-edge-coloured $K_{\mathbb{N}}$ contains a monochromatic copy of G with $\overline{\mathrm{d}}(G) \geq d$. If $r = 2$, we drop the subscript.

We will now define the infinite graphs which will appear frequently in this thesis.

Definition 1.1.28 (Infinite path)**.** The *infinite path* P_∞ is the graph with vertex-set $V(P) = \{v_i : i \in \mathbb{N}\}$ and edge-set $E(P) = \{v_i v_{i+1} : i \in \mathbb{N}\}$. P_∞ is sometimes called one-way infinite path. The *two-way infinite path* C_∞ is the graph with vertex-set $V(P) = \{v_i : i \in \mathbb{Z}\}$ and edge-set $E(P) = \{v_i v_{i+1} : i \in \mathbb{Z}\}$. C_∞ can be seen as the infinite analogue of a cycle.

1.1.7 Random Graphs

A *(finite) probability space* is a tuple $(\Omega, \mathbb{P})$ of a finite set Ω and a probability function, that is a function $\mathbb{P} : 2^\Omega \to [0,1]_{\mathbb{R}}$ with $\mathbb{P}[\emptyset] = 0$, $\mathbb{P}[\Omega] = 1$ and $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$ for all disjoint sets $A, B \subseteq \Omega$. Note that $\mathbb{P}$ is uniquely determined by the values of $\mathbb{P}[x] := \mathbb{P}[\{x\}]$ for all $x \in \Omega$.

Definition 1.1.29 (Conditional Probability)**.** Given a finite probability space $(\Omega, \mathbb{P})$ and events $A, B \subseteq \Omega$ with $\mathbb{P}[B] > 0$, we define the *conditional probability* of A given B as $\mathbb{P}[A|B] := \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$.

Definition 1.1.30 (With high probability)**.** For each $n \in N$, let E_n be an event in some probability space $(\Omega_n, \mathbb{P}_n)$. We say that the event $E = (E_n)_{n \in \mathbb{N}}$ holds *with high probability (w.h.p.)* if $\lim_{n\to\infty} \mathbb{P}[E_n] = 1$.

A *random variable* is a function $X : \Omega \to S$ for some set S. If $S \subseteq \mathbb{R}$, we denote the *expected value* of X by

$$\mathbb{E}[X] := \sum_{x \in \Omega} \mathbb{P}[x] \cdot X(x) = \sum_{y \in S} \mathbb{P}[X = y] \cdot y,$$

where we denote $\mathbb{P}[X = y] := \mathbb{P}[\{x \in \Omega : X(x) = y\}]$. The *distribution* of a random variable is the probability function $\mathbb{P}_X : 2^S \to [0,1]_{\mathbb{R}}$, where $\mathbb{P}_X$ is induced by $\mathbb{P}_X(y) := \mathbb{P}[X = y]$. On the other hand, given a probability space $(S, \mathbb{P})$, the random variable $X : S \to S$ defined by $X(y) = y$ has distribution $\mathbb{P}$. Therefore, when the original probability space is not important, we often identify a random variable with its distribution. A *random graph* is a random variable taking values in a set of graphs. We will usually only be interested in its distribution. We will be dealing with the following random graphs in this thesis. Here, given a (hyper-)graph G, we denote by $\mathcal{G}_G$ the set of all spanning subgraphs of G and we write $\mathcal{G}_n := \mathcal{G}_{K_n}$ and $\mathcal{G}_n^{(k)} := \mathcal{G}_{K_n^{(k)}}$, and we denote by $\mathcal{G}_{n,m}$ the set of all spanning subgraphs of K_n with exactly m edges.

Definition 1.1.31 (Random graphs)**.** Let $n \in \mathbb{N}$ and let $p \in [0,1]_{\mathbb{R}}$.

- $G(n,p)$ is the random spanning subgraph of K_n in which every edge of K_n is present independently with probability p. Its distribution $\mathbb{P} : 2^{\mathcal{G}_n} \to [0,1]_{\mathbb{R}}$ is induced by $\mathbb{P}[G(n,p) = H] = p^{e(H)}(1-p)^{e(K_n)-e(H)}$ for $H \in \mathcal{G}_n$.

- Given $m \in \mathbb{N}$ with $0 \leq m \leq \binom{n}{2}$, we define $G(n,m)$ to be the random spanning subgraph of K_n which is chosen uniformly from all spanning subgraphs of K_n with exactly m edges. Its distribution $\mathbb{P} : 2^{\mathcal{G}_{n,m}} \to [0,1]_{\mathbb{R}}$ is induced by $\mathbb{P}[G(n,m) = H] = 1/\binom{e(K_n)}{m}$ for $H \in \mathcal{G}_{n,m}$.

- Given $k \in \mathbb{N}$, we define $G^{(k)}(n,p)$ to be the random spanning subgraph of $K_n^{(k)}$ in which every edge of $K_n^{(k)}$ is present independently with probability p. Its distribution $\mathbb{P} : 2^{\mathcal{G}_n^{(k)}} \to [0,1]_{\mathbb{R}}$ is induced by $\mathbb{P}\left[G^{(k)}(n,p) = H\right] = p^{e(H)}(1-p)^{e(K_n^{(k)})-e(H)}$ for $H \in \mathcal{G}_n^{(k)}$.

- Given a graph G, we define G_p to be the random spanning subgraph of G in which every edge of G is present independently with probability p. Its distribution $\mathbb{P} : 2^{\mathcal{G}_G} \to [0,1]_{\mathbb{R}}$ is induced by $\mathbb{P}\left[G_p = H\right] = p^{e(H)}(1-p)^{e(G)-e(H)}$ for $H \in \mathcal{G}_G$.

Definition 1.1.32 (Graph properties and threshold probabilities)**.** A *graph property* $\mathcal{P}$ is a set of graphs. We say that a graph G has property $\mathcal{P}$ if $G \in \mathcal{P}$. Let now $\mathcal{P}$

be a graph property whose members are subgraphs of K_n for some $n \in \mathbb{N}$ and let $p_0 : \mathbb{N} \to [0,1]_{\mathbb{R}}$ be a function.

- p_0 is called a *coarse threshold probability* for $\mathcal{P}$ if

$$\lim_{n\to\infty} \mathbb{P}\left[G(n, p(n)) \in \mathcal{P}\right] = \begin{cases} 0 & \text{if } p = o(p_0), \\ 1 & \text{if } p = \omega(p_0). \end{cases}$$

- p_0 is called a *sharp threshold probability* for $\mathcal{P}$ if for every $\varepsilon > 0$, we have

$$\lim_{n\to\infty} \mathbb{P}\left[G(n, p(n)) \in \mathcal{P}\right] = \begin{cases} 0 & \text{if } p(n) \leq (1-\varepsilon)p_0(n) \text{ for all } n \in \mathbb{N}, \\ 1 & \text{if } p(n) \geq (1+\varepsilon)p_0(n) \text{ for all } n \in \mathbb{N}. \end{cases}$$

Note that threshold probabilities do not necessarily exist for all graph properties. However, Bollobás and Thomason [12] proved that coarse thresholds exist for all *monotone* graph properties ($\mathcal{P}$ is monotone if $H \in \mathcal{P}$ and $H \subseteq G$ implies $G \in \mathcal{P}$).

1.2 Graphs with Linear Ramsey Number

Investigating graphs with linear Ramsey number has been one of the most studied topics in Ramsey Theory, motivated particularly by a series of conjectures of Burr and Erdős [16, 17]. A lot of work presented in this thesis was inspired by results from this area, and hence we will start with a brief introduction to it.

A sequence of graphs $\mathcal{F} = \{F_1, F_2, \ldots\}$[3] is said to have *linear r-colour Ramsey number* if we have $\mathrm{R}_r(F_n) = O(n)$. If $r = 2$, we will simply say $\mathcal{F}$ has linear Ramsey number. The following observation shows that graphs with linear Ramsey number must be quite "sparse".

Observation 1.2.1. *If $\mathcal{F} = \{F_1, F_2, \ldots\}$ is a sequence of graphs with linear Ramsey number, then we have $e(F_n) = O(n\log(n))$.*

[3] Recall that $\mathcal{F} = \{F_1, F_2, \ldots\}$ is a sequence of graphs if F_n is a graph with n vertices for every $n \in \mathbb{N}$.

Proof. Since $\mathcal{F}$ has linear Ramsey number, there is some constant $C \in \mathbb{N}$ such that $\mathrm{R}(F_n) \leq Cn$ for all $n \in \mathbb{N}$. We will show that $d(F_n) \leq 2\log_2(Cn) + 2/n$. Fix some $n \in \mathbb{N}$ and let $N := Cn$. Colour the edges of K_N uniformly at random with 2 colours and let E_n be the event that there is a monochromatic copy of F_n. Since $N \geq \mathrm{R}(F_n)$, we have $\mathbb{P}[E_n] = 1$. On the other hand we have $\mathbb{P}[E_n] \leq 2 \cdot N^n \cdot 2^{-e(F_n)}$. We conclude that $e(F_n) \leq \log_2(2N^n) = 1 + n\log_2(Cn)$, concluding the proof. □

1.2.1 Paths and Cycles

Paths and cycles are two of the most elementary examples of graphs with linear Ramsey number. While it is not hard to see that paths have linear Ramsey number, there has been a lot of work in trying to determine their Ramsey numbers exactly. For two colours, Gerencsér and Gyárfás [55] proved the following result.

Theorem 1.2.2 (Gerencsér–Gyárfás [55])**.** *We have* $\mathrm{R}(P_{n-1}) = \lfloor 3n/2 - 1 \rfloor$ *for every* $n \geq 2$.

Confirming a conjecture of Faudree and Schelp [50] for large n, Gyárfás, Ruszinkó, Sárközy and Szemerédi [65] proved that

$$\mathrm{R}_3(P_{n-1}) = \begin{cases} 2n-2 & \text{if } n \text{ is even} \\ 2n-1 & \text{if } n \text{ is odd.} \end{cases}$$

for every large enough $n \in \mathbb{N}$. For more than three colours, less is known. It easily follows from a result of Erdős and Gallai [38] that $\mathrm{R}_r(P_{n-1}) \leq rn$ and Bierbrauer and Gyárfás [10] proved that $\mathrm{R}_r(P_{n-1}) \geq (r - 1 - o(1))n$. Recently, there has been a lot of work on this problem and, after progress by Sárközy [109], and Davies, Jenssen and Roberts [32], the currently best known upper bound is due to Knierim and Su [79], who proved that $\mathrm{R}_r(P_{n-1}) \leq (r - 1/2 + o(1))n$.

Clearly we have $\mathrm{R}_r(C_n) \geq \mathrm{R}_r(P_{n-1})$ for all $n, r \geq 2$ since C_n contains P_{n-1} and it turns out the two numbers are very similar for even cycles. Faudree and Schelp [49] and independently Rosta [103] showed that $\mathrm{R}(C_n) = 3n/2 - 1$ for all even $n \geq 6$ and Benevides and Skokan [8] showed that $\mathrm{R}_3(C_n) = 2n$ for all sufficiently large even n. For multiple colours, Łuczak, Simonovits and Skokan [93] proved that

$R_r(C_n) \leq (r + o(1))n$ and all recent result about paths above ([32, 79, 109]) came with extensions to even cycles, in particular we have $R_r(C_n) \leq (r - 1/2 + o(1))n$.

Interestingly, the problem is very different for odd cycles. Bondy and Erdős [13] showed that $R(C_n) = 2n-1$ for all odd $n \geq 5$ and, confirming a conjecture of Bondy and Erdős, Jenssen and Skokan [74] showed that $R_r(C_n) = 2^{r-1}(n-1) + 1$ for all sufficiently large odd n.

1.2.2 Trees

Trees are another well-studied family of graphs in Ramsey theory. It is again easy to show that trees have linear r-colour Ramsey number. Indeed, let us prove that $R_r(T) \leq 2r(n-1)+1$ for every tree T with n vertices. Let $N = 2r(n-1)+1$ and let K_N be edge-coloured with r colours. By the pigeonhole principle, one colour (say blue) appears at least $\binom{N}{2}/r$ times and hence the average degree of the blue subgraph is at least $(N-1)/r$. Therefore, there is a blue subgraph with minimum degree at least $(N-1)/(2r) \geq n-1$.[4] It is now easy to greedily embed T into this subgraph.

A famous conjecture of Burr and Erdős [17], which was solved for sufficiently large n by Zhao [117], states that this can be improved roughly by a factor of 2 in the case of two colours.

Conjecture 1.2.3 (Burr–Erdős [17])**.** *For every tree T on n vertices, we have* $R(T) \leq 2n-2$ *if n is even and* $R(T) \leq 2n-3$ *if n is odd.*

1.2.3 Graphs with Bounded Degree

Burr and Erdős also studied Ramsey numbers of graphs with bounded degree and made the following conjecture, which states that every sequence of graphs with bounded degree has linear Ramsey number.

Conjecture 1.2.4 (Burr–Erdős [16])**.** *For all $\Delta \geq 1$, there exists some $c = c(\Delta) > 0$ such that* $R(G) \leq cn$ *for every graph G on n vertices with maximum degree at most Δ.*

[4]It is a well-known combinatorial fact that every graph of average degree d has a subgraph of minimum degree at least $d/2$.

Conjecture 1.2.4 was solved by Chvatál, Rödl, Szemerédi, Trotter [21] in an early application of the regularity lemma. Since then, there has been many improvements to the constant $c(\Delta)$. After an improvement of Eaton [36], Graham, Rödl and Ruciński [57] came close to settling the problem by showing that the constant can be chosen to be $c(\Delta) = 2^{O(\Delta \log^2(\Delta))}$ and providing the following lower bound.

Theorem 1.2.5 (Graham–Rödl–Ruciński [57])**.** *There is a sequence of bipartite graphs $\mathcal{F} = \{F_1, F_2, \ldots\}$ with $\mathrm{R}(F_n) = 2^{\Omega(\Delta)} n$.*

The currently best known upper bound is due to Conlon, Fox and Sudakov [26].

Theorem 1.2.6 (Conlon–Fox–Sudakov [26])**.** *There is some constant $c > 0$ such that $\mathrm{R}(G) \leq 2^{c\Delta \log(\Delta)} n$ for every graph G on n vertices with maximum degree at most Δ.*

Furthermore, Conlon [23], and independently Fox and Sudakov [52] were able to remove the extra $\log(\Delta)$-factor in the exponent for bipartite graphs with maximum degree at most Δ. Allen, Brightwell and Skokan [5] showed that the constant in Theorem 1.2.6 can be significantly improved to for a wide range of graphs. The *bandwidth* of a graph G, denoted by $\mathrm{bw}(G)$, is the smallest integer k for which there is a bijection $f : V(G) \to [v(G)]$ such that $|f(u) - f(v)| \leq k$ for every edge $uv \in E(G)$.

Theorem 1.2.7 (Allen–Brightwell–Skokan [5])**.** *For every $\Delta \in \mathbb{N}$, there are constants $\beta > 0$ and $n_0 \in \mathbb{N}$ such that $\mathrm{R}(G) \leq (2\chi(G) + 4)n$ for every graph G with $n \geq n_0$ vertices, maximum degree at most Δ and bandwidth at most βn.*

Finally, let us mention another famous conjecture of Burr and Erdős regarding graphs with bounded degeneracy.

Conjecture 1.2.8 (Burr–Erdős [16])**.** *For all $d \geq 1$, there exists some $C = C(d) > 0$ such that $\mathrm{R}(G) \leq Cn$ for every d-degenerate graph G on n vertices.*

Conjecture 1.2.8 was recently confirmed in a breakthrough result of Lee [90].

1.3 Monochromatic Graph Tiling Problems

1.3.1 History

The first result in this area is due to Gerencsér and Gyárfás [55]. Recall that single vertices and edges are considered (degenerate) paths and cycles in this thesis.

Theorem 1.3.1 (Gerencsér–Gyárfás [55]). *The vertices of every 2-edge-coloured complete graph on n vertices can be partitioned into two monochromatic paths, one of each colour.*

A conjecture of Lehel (which first appeared in the PhD thesis of Ayel [7] in 1979, where it was also proved for some special types of colourings of K_n) states that the same should be true for cycles instead of paths. Almost 20 years later, in 1998, Łuczak, Rödl and Szemerédi [92] proved Lehel's conjecture for all sufficiently large n using the regularity method. In [2], Allen gave an alternative proof, which gave a better bound on n. Finally, Bessy and Thomassé [9] proved Lehel's conjecture for all integers $n \geq 1$.

Theorem 1.3.2 (Bessy–Thomassé [9]). *The vertices of every 2-edge-coloured complete graph on n vertices can be partitioned into two monochromatic cycles, one of each colour.*

The problem was soon extended to multiple colours.

Theorem 1.3.3 (Gyárfás [59]). *The vertices of every r-edge-coloured complete graph on n vertices can be covered by $O(r^4)$ monochromatic paths.*

Theorem 1.3.4 (Erdős–Gyárfás–Pyber [42]). *The vertices of every r-edge-coloured complete graph on n vertices can be partitioned into $O(r^2 \log r)$ monochromatic cycles.*

It is a major open problem to determine $\mathrm{cp}(r)$, the smallest number of cycles needed in Theorem 1.3.4. More precisely, $\mathrm{cp}(r)$ is the smallest integer t such that the vertices of every r-edge-coloured K_n can be partitioned into a collection of at most t vertex-disjoint monochromatic cycles. Note that (without Theorem 1.3.4) it is not clear at all whether $\mathrm{cp}(r)$ is finite (i.e. independent of n, the size of the host-graph).

Problem 1.3.5. *What is* $\mathrm{cp}(r)$*?*

It was further conjectured in [42] that $\mathrm{cp}(r) = r$ for all $r \in \mathbb{N}$. This conjecture was refuted however by Pokrovskiy [95], who showed that $\mathrm{cp}(r) > r$ for every $r \geq 3$. Gyárfás, Ruszinkó, Sárközy and Szemerédi [62] showed that $O(r \log r)$ cycles suffice for all sufficiently large n.

Pokrovskiy proposed the following slightly weaker conjecture.

Conjecture 1.3.6 (Pokrovskiy [95])**.** *For every positive integer r, there is some constant $c = c(r)$, such that in every r-edge-coloured K_n, there is a collection of r vertex-disjoint monochromatic cycles covering all but at most c vertices.*

It is still possible that r monochromatic paths suffice.

Conjecture 1.3.7 (Gyárfás [59])**.** *The vertices of every r-edge-coloured complete graph on n vertices can be partitioned into r monochromatic paths.*

1.3.2 New Results

Hypergraph Cycles

It is natural to ask if cycle partition problems like Theorem 1.3.4 can be generalised to hypergraphs. Such questions were first studied by Gyárfás and Sárközy [64] who proved the following result about loose cycles.

Theorem 1.3.8 (Gyárfás–Sárközy)**.** *For every k, $r \in \mathbb{N}$, there is some $c = c(k, r)$ such that the vertices of every r-edge-coloured complete k-graph can be partitioned into a collection of at most c loose cycles.*

Later, Sárközy [107] showed that $c(k, r)$ can be be chosen to be $50rk \log(rk)$. Gyárfás [60] conjectured that a similar result can be obtained for tight cycles (again, a single vertex is considered a tight cycle here).

Conjecture 1.3.9 (Gyárfás [60])**.** *For every $k, r \in \mathbb{N}$, there is some $c = c(k, r)$ so that the vertices of every r-edge-coloured complete k-graph can be partitioned into at most c monochromatic tight cycles.*

We prove this conjecture and a generalisation in which we allow the host-hypergraph to be any k-graph with bounded independence number (the main motivation of this generalisation is an interesting application discussed below).

Theorem 1.3.10 (Bustamante–Corsten–Frankl–Pokrovskiy–Skokan [18])**.** *For every $k, r, \alpha \in \mathbb{N}$, there is some $c = c(k, r, \alpha)$ such that the vertices of every r-edge-coloured k-graph G with independence number $\alpha(G) \leq \alpha$ can be partitioned into at most c monochromatic tight cycles.*

The proof of Theorem 1.3.10, which is based on the absorption method and the hypergraph regularity method, will be presented in Section 3.1.

Tiling Number of Sequences of Graphs

We turn our attention back to graphs now considering other "tiles" than cycles. We investigate problems in which we are given a sequence of graphs $\mathcal{F}$ and an edge-coloured complete graph K_n and our goal is to partition $V(K_n)$ into few monochromatic copies of graphs from $\mathcal{F}$. An $\mathcal{F}$*-tiling* $\mathcal{T}$ of a graph G is a collection of vertex-disjoint copies of graphs from $\mathcal{F}$ in G with $V(G) = \bigcup_{T \in \mathcal{T}} V(T)$. If G is edge-coloured, we say that $\mathcal{T}$ is *monochromatic* if every $T \in \mathcal{T}$ is monochromatic (not necessarily in the same colour). Using this notation, Theorem 1.3.4 states that in every r-edge coloured K_n, we can find a monochromatic C-tiling of size at most $O(r^2 \log r)$, where C is the family of all cycles (including a single vertex and a single edge).

Definition 1.3.11 (Tiling number)**.** Let $\tau_r(\mathcal{F}, n)$ be the minimum $t \in \mathbb{N}$ such that in every r-edge-coloured K_n there is a monochromatic $\mathcal{F}$-tiling of size at most t. We call $\tau_r(\mathcal{F}) = \sup_{n \in \mathbb{N}} \tau_r(\mathcal{F}, n)$ the r-colour *tiling number* of $\mathcal{F}$. If $r = 2$, we simply write $\tau(\mathcal{F}, n)$ and $\tau(\mathcal{F})$, and simply say *tiling number*.

The results of Pokrovskiy and of Erdős, Gyárfás and Pyber above imply that, for all $r \geq 3$,

$$r + 1 \leq \tau_r(C) \leq O(r^2 \log r).$$

Note that, in general, it is not clear at all that $\tau_r(\mathcal{F})$ is finite and it is a natural question to ask for which families this is the case. The study of such tiling problems for more general families of graphs was initiated by Grinshpun and Sárközy [58] who proved the following result. Let $\mathscr{F}_\Delta$ be the collection of all sequences of graphs $\mathcal{F}$ with $\Delta(\mathcal{F}) \leq \Delta$.

Theorem 1.3.12 (Grinshpun–Sárközy [58])**.** *We have $\tau_2(\mathcal{F}) \leq 2^{O(\Delta \log \Delta)}$ for all $\mathcal{F} \in \mathscr{F}_\Delta$. In particular, $\tau_2(\mathcal{F})$ is finite whenever $\Delta(\mathcal{F})$ is finite.*

Grinshpun and Sárközy also proved that $\tau_2(\mathcal{F}) \leq 2^{O(\Delta)}$ for every sequence of bipartite graphs $\mathcal{F}$ of maximum degree Δ and showed that this is best possible up to the implicit constant (see Theorem 1.3.19). Sárközy [108] further proved that the constant in Theorem 1.3.12 can be improved a lot for the special case of powers of cycles.

For more than two colours not much is known. Elekes, D. Soukup, L. Soukup and Szentmiklóssy [37, Problem 6.4] asked the following problem after proving a similar statement for infinite graphs.[5]

Problem 1.3.13 (Elekes et al. [37])**.** *Prove that for every $r, p \in \mathbb{N}$, there is some $c = c(r, p)$ such that the vertices of every r-edge-coloured complete graph can be partitioned into at most c monochromatic p-th powers of cycles.*

We answer this problem positively. In fact, we obtain the following generalisation to hypergraphs and host-graphs with bounded independence number as a corollary of Theorem 1.3.10.

Theorem 1.3.14 (Bustamante–Corsten–Frankl–Pokrovskiy–Skokan [18])**.** *For every $k, r, p, \alpha \in \mathbb{N}$, there is some $c = c(k, r, p, \alpha)$ such that the vertices of every r-edge-coloured k-graph G with $\alpha(G) \leq \alpha$ can be partitioned into at most c monochromatic p-th powers of tight cycles.*

Grinshpun and Sárközy [58] conjectured that their Theorem 1.3.12 should extend to r colours as well.

Conjecture 1.3.15 (Grinshpun–Sárközy [58])**.** *For every $r \in \mathbb{N}$, there is some $c_r > 0$ so that $\tau_r(\mathcal{F}) \leq 2^{\Delta^{c_r}}$ for all $\Delta \geq 1$ and all $\mathcal{F} \in \mathscr{F}_\Delta$.*

[5] The problem is phrased differently in [37] but this version is stronger, as Elekes et al. explain below the problem.

Note that it is still open to decide whether $\tau_r(\mathcal{F})$ is finite for such families. We show that this is the case and make progress towards Conjecture 1.3.15. Recall that $\exp^k$ denotes the k-th composition of the exponential function.

Theorem 1.3.16 (Corsten–Mendonça)**.** *There is an absolute constant $K > 0$ such that for all integers $r, \Delta \geq 2$ and all $\mathcal{F} \in \mathscr{F}_\Delta$, we have $\tau_r(\mathcal{F}) \leq \exp^3\left(Kr^2\Delta^3\right)$. In particular, $\tau_r(\mathcal{F}) < \infty$ whenever $\Delta(\mathcal{F}) < \infty$.*

Theorem 1.3.16 also provides a positive answer to Problem 1.3.13 and the bound is much better than the implicit bound in Theorem 1.3.14. The proof of Theorem 1.3.16, which is based on the absorption method and the regularity method, will be presented in Section 3.2.

Lower Bounds and Graphs with Linear Ramsey Number

Sequences of graphs with finite tiling number are closely related to sequences with linear Ramsey number. Observe that a sequence of graphs $\mathcal{F} = \{F_1, F_2, \ldots\}$ has linear r-colour Ramsey number if and only if there exists some $\rho > 0$ such that every r-edge-coloured K_n contains a monochromatic copy of F_m for every $m \leq \rho n$. A slightly weaker condition (in which we require only one monochromatic copy of F_m for some $m \geq \delta n$) is a necessary condition to have finite tiling number: given a sequence of graphs $\mathcal{F}$, it follows from the pigeonhole principle that every r-edge-coloured K_n contains a monochromatic copy from $\mathcal{F}$ of size at least $n/\tau_r(\mathcal{F})$. Define $\rho_r(\mathcal{F})$ to be the supremum over all $\rho \geq 0$ such that every r-edge coloured K_n contains a monochromatic copy from $\mathcal{F}$ of size at least ρn. The above application of the pigeonhole principle gives the following observation.

Observation 1.3.17. *We have $\tau_r(\mathcal{F}) \geq 1/\rho_r(\mathcal{F})$ for every sequence of graphs $\mathcal{F}$.*

Note that $\rho_r(\mathcal{F}) > 0$ for every family $\mathcal{F}$ with linear r-colour Ramsey number but the converse is not always true.[6] It is however true if $\mathcal{F}$ is an increasing family of graphs (i.e. if $F_n \subseteq F_{n+1}$ for every $n \in \mathbb{N}$), in which case we have

$$\rho_r(\mathcal{F}) = \inf_{n \in \mathbb{N}} \left(n/\mathrm{R}_r(F_n)\right) \tag{1.3.1}$$

[6] For example, consider the sequence $\mathcal{F} = \{F_1, F_2, \ldots\}$ where F_n is an independent set for all even n and a clique for all odd n.

and thus, using Observation 1.3.17,

$$\tau_r(\mathcal{F}) \geq \sup_{n\in\mathbb{N}} \left(\mathrm{R}_r(F_n)/n\right). \tag{1.3.2}$$

This relationship can be used to translate existing lower bounds for Ramsey numbers of sequences of graphs to lower bounds for their tiling number. While (1.3.2) can only be directly applied to increasing sequences of graphs, Grinshpun and Sárközy [58] observed that one can modify sequences of bounded degree to avoid this problem.

Observation 1.3.18. *For every sequence of graphs $\mathcal{F} = \{F_1, F_2, \dots\}$ with maximum degree at most Δ and $\mathrm{R}(F_n) \geq Cn$ for every $n \in \mathbb{N}$, there is an increasing sequence of graphs $\tilde{\mathcal{F}} = \{\tilde{F}_1, \tilde{F}_2, \dots\}$ with maximum degree Δ and $\mathrm{R}(\tilde{F}_n) \geq Cn/4$ for every $n \in \mathbb{N}$.*

Combining this with Theorem 1.2.5 of Graham, Rödl and Ruciński and (1.3.2) immediately gives the following lower bound, which almost matches the upper bound in Theorem 1.3.12.

Theorem 1.3.19 (Grinshpun–Sárközy [58])**.** *There is a sequence of graphs $\mathcal{F} \in \mathscr{F}_\Delta$ with $\tau_2(\mathcal{F}) \geq 2^{\Omega(\Delta)}$.*

Observation 1.3.17 further implies that $\rho_r(\mathcal{F}) > 0$ (or having linear r-colour Ramsey number if $\mathcal{F}$ is increasing) is a necessary condition for $\mathcal{F}$ to have finite tiling number. In the other direction, we have $\tau_r(\mathcal{F}, n) \leq \log(n)/\rho_r(\mathcal{F})$ (this follows by greedily taking the largest monochromatic copy from $\mathcal{F}$ among the uncovered vertices until all vertices are covered). The following example, which was provided by Alexey Pokrovskiy (personal communication), shows that this is essentially tight and therefore the above necessary condition is not sufficient. Let S_n be a star with n vertices and let $\mathcal{S} = \{S_1, S_2, \dots\}$ be the family of stars. It follows readily from the pigeonhole principle that $R_r(S_n) \leq r(n-2)+2$ for every $n \in \mathbb{N}$ and thus $\rho_r(\mathcal{S}) \geq 1/r$. However, $\tau_r(\mathcal{S}) = \infty$ for every $r \geq 2$.

Example 1.3.20. For every $r \geq 2$, we have $\tau_r(\mathcal{S}, n) \geq r \cdot \log n/8$ for all sufficiently large n.

Proof. Let $\tau = r \log n/8$ and colour $E(K_n)$ uniformly at random with r colours. Given a vertex $v \in [n]$ and a colour c, let $S_c(v)$ be the star centred at v formed by all the edges of colour c incident on v. Note that there is a monochromatic $\mathcal{S}$-tiling of size at most τ if and only if there are distinct vertices $v_1, \ldots, v_\tau$ and colours $c_1, \ldots, c_\tau \in [r]$ such that $\bigcup_{i \in [\tau]} V(S_{c_i}(v_i)) = [n]$.

Fix distinct vertices $v_1, \ldots, v_\tau \in [n]$ and colours $c_1, \ldots, c_\tau \in [r]$. Let U be the random set $U = \bigcup_{i \in [\tau]} V(S_{c_i}(v_i))$. Notice that the events $\{v \in U\}$, $v \in [n] \setminus \{v_1, \ldots, v_\tau\}$, are independent and each has probability $1-(1-1/r)^\tau$. Therefore, using $e^{-x/(1-x)} \leq 1 - x \leq e^x$ for all $x \leq 1$, we get

$$\begin{aligned}\mathbb{P}[U = [n]] = (1-(1-1/r)^\tau)^{n-\tau} &\leq \exp\left(-(n-\tau)(1-1/r)^\tau\right) \\ &\leq \exp\left(-n(1-1/r)^{\tau+1}\right) \\ &\leq \exp\left(-n\exp\left(-4\tau/r\right)\right) \\ &\leq \exp\left(-\sqrt{n}\right).\end{aligned}$$

Taking a union bound over all choices of $v_1, \ldots, v_\tau$ and $c_1, \ldots, c_\tau$, we conclude that the probability that there is a monochromatic $\mathcal{S}$-tiling of size τ is at most

$$(rn)^{-\tau} \cdot e^{-\sqrt{n}} < 1$$

for all sufficiently large n. Hence, there exists an r-colouring of $E(K_n)$ without a monochromatic $\mathcal{S}$-tiling of size at most τ, finishing the proof. □

Recall that Lee [90] proved that sequences of graphs with bounded degeneracy have linear Ramsey number. Since every star has degeneracy 1, Example 1.3.20 further shows that it is not possible to extend this to a tiling result. Nevertheless, it may be possible to allow unbounded degrees in this case.

Question 1.3.21. *Is there a function $\omega : \mathbb{N} \to \infty$ with $\lim_{n\to\infty} \omega(n) = \infty$, such that the following is true for all integers $r, d \geq 2$? If $\mathcal{F} = \{F_1, F_2, \ldots\}$ is a sequence of d-degenerate graphs with $v(F_n) = n$ and $\Delta(F_n) \leq \omega(n)$ for all $n \in \mathbb{N}$, then $\tau_r(\mathcal{F}) < \infty$.*

It is possible to obtain such a result for graphs of bounded arrangeability (a concept in between bounded degree and bounded degeneracy) as explained in Section 3.2.5.

1.4 Ramsey Problems for Infinite Graphs

1.4.1 Infinite Paths

Theorem 1.2.2 implies that that every 2-edge-coloured K_n contains a monochromatic path with $\lfloor (2n+1)/3 \rfloor$ vertices. In Section 4.1 we will investigate a similar problem concerning the infinite path P_∞, where, given a 2-edge-coloured $K_\mathbb{N}$, we try to find a monochromatic infinite path with the largest possible upper density. Since this path does not need to respect the order of $\mathbb{N}$, we cannot easily infer any bounds from the finite case. Due to the following result of Erdős and Galvin [39], we will restrict our attention to upper densities.

Theorem 1.4.1 (Erdős–Galvin [39])**.** *There is a 2-edge-colouring of $K_\mathbb{N}$ in which every monochromatic infinite path has lower density* 0*.*

Recall the definition of Ramsey upper density, Definition 1.1.27. Probably the first result in this direction is due to Rado [99].

Theorem 1.4.2 (Rado [99])**.** *Every r-edge-coloured $K_\mathbb{N}$ contains a collection of at most r vertex-disjoint monochromatic paths covering all vertices. In particular, one of them has upper density at least $1/r$.*

Interestingly, this gives a positive answer of an infinite analogue of Conjecture 1.3.7. For two colours, this further implies that we can always find a monochromatic path of upper density at least $1/2$. Erdős and Galvin [39] improved this special case.

Theorem 1.4.3 (Erdős–Galvin [39])**.** *We have* $2/3 \le \overline{\mathrm{Rd}}(P_\infty) \le 8/9$*.*

DeBiasio and McKenney [33] recently improved the lower bound to $3/4$ and conjectured the correct value to be $8/9$. Progress towards this conjecture was made by Lo, Sanhueza-Matamala and Wang [91], who raised the lower bound to $(9+\sqrt{17})/16 \approx 0.82019$. We settle the problem of determining the Ramsey upper density of the infinite path by proving that the correct value is $\overline{\mathrm{Rd}}(P_\infty) = (12+\sqrt{8})/17 \approx 0.87226$.

Theorem 1.4.4 (Corsten–DeBiasio–Lamaison–Lang [30])**.** *There exists a 2-edge-colouring of $K_{\mathbb{N}}$ such that every monochromatic path has upper density at most* $(12+\sqrt{8})/17$.

Theorem 1.4.5 (Corsten–DeBiasio–Lamaison–Lang [30])**.** *Suppose the edges of $K_{\mathbb{N}}$ are coloured with two colours. Then, there exists a monochromatic path with upper density at least* $(12+\sqrt{8})/17$.

Note that the problem of determining the r-colour Ramsey upper density remains open and it would be very interesting to make any improvement on Rado's lower bound of $1/r$ for $r \geq 3$ colours (see [33, Corollary 3.5] for the best known upper bound). In particular for three colours, the correct value is between $1/3$ and $1/2$.

Question 1.4.6. *What is $\overline{\mathrm{Rd}}_r(P_\infty)$ for $r \geq 3$?*

The proof of Theorems 1.4.4 and 1.4.5, which is based on a mix of the regularity method and novel yet elementary combinatorial arguments, will be presented in Section 4.1.

Bipartite Ramsey densities

In Section 4.2.5 we will briefly discuss a bipartite version of Theorem 1.4.5. Gyárfás and Lehel [61] and independently Faudree and Schelp [51] proved that every 2-edge-coloured $K_{n,n}$ contains a monochromatic path with at least $2\lfloor n/2 \rfloor + 1$ vertices (that is, roughly half the vertices of the graph). They further proved that this is best possible. We will prove an analogue of this for infinite graphs. Here, $K_{\mathbb{N},\mathbb{N}}$ is the infinite complete bipartite graph with one part being all even positive integers and the other part being all odd positive integers.

Theorem 1.4.7 (Corsten–DeBiasio–McKenney)**.** *Every 2-coloured $K_{\mathbb{N},\mathbb{N}}$ contains a monochromatic path of upper density at least* $1/2$.

Pokrovskiy [95] proved that the vertices of every 2-edge-coloured complete bipartite graph $K_{n,n}$ can be partitioned into three monochromatic paths. Soukup [112] proved an analogue of this for infinite graphs which holds for multiple colours: the vertices of every r-edge-coloured $K_{\mathbb{N},\mathbb{N}}$ can be partitioned into $2r-1$ monochromatic paths. He also presents an example where this is best possible. However, in his example all but finitely many vertices can be covered by r monochromatic paths. Our next result shows that this is always possible in the case of two colours.

Theorem 1.4.8 (Corsten–DeBiasio–McKenney)**.** *The vertices of every 2-edge-coloured $K_{\mathbb{N},\mathbb{N}}$ can be partitioned into a finite set and at most two monochromatic paths.*

Theorem 1.4.7 is an immediate consequence of Theorem 1.4.8. We will provide an example which demonstrates that Theorems 1.4.7 and 1.4.8 are best possible (see Example 4.2.9). We believe that a similar statement is true for multiple colours.

Conjecture 1.4.9 (Corsten–DeBiasio–McKenney)**.** *The vertices of every r-edge-coloured $K_{\mathbb{N},\mathbb{N}}$ can be partitioned into a finite set and at most r monochromatic paths.*

Example 4.2.9 further shows that Conjecture 1.4.9 is best possible, if true.

1.4.2 Infinite Trees

In Section 4.2.6 we will present an analogue of Conjecture 1.2.3 (which implies that every 2-edge-coloured K_n contains a monochromatic copy of every tree on $n/2+1$ vertices) for infinite trees.

Theorem 1.4.10 (Corsten–DeBiasio–McKenney)**.** $\overline{\mathrm{Rd}}(T) \geq 1/2$ *for every infinite tree T.*

There are a couple of examples showing that this is best possible. For example, we have $\overline{\mathrm{Rd}}(S_\infty) = 1/2$, where S_∞ is the infinite star (that is the infinite tree in which one vertex has infinitely many neighbours of degree 1). This can be seen by colouring the edges of $K_{\mathbb{N}}$ uniformly at random with red and blue (see Rado colouring Section 4.2.2). Another such example is T_∞, the infinite tree in which every vertex has infinite degree (see Example 4.2.8). Furthermore, there is a locally finite (that is, every vertex has finite degree) tree T with $\overline{\mathrm{Rd}}(T) = 1/2$ (see Example 4.2.8).

1.4.3 Infinite Graphs with "Linear Ramsey Number"

In the finite case, the study of graphs with linear Ramsey number has received a lot of attention and we will now study possible analogues for infinite graphs. Recall that a sequence of graphs $\mathcal{F} = \{F_1, F_2, \ldots\}$ has linear r-colour Ramsey number if and only if there is some $\delta > 0$ such that every r-edge-coloured K_n contains a monochromatic copy of F_m for every $m \leq \delta n$. There are two natural analogues of this for infinite graphs. One is the class of infinite graphs with positive Ramsey upper density, that is, those graphs G for which there is some $\delta > 0$ such that every 2-edge-coloured $K_{\mathbb{N}}$ contains a monochromatic copy of G with upper density at least δ. For example P_∞ has positive Ramsey upper density as seen above. Another natural analogue is the class of infinite graphs G for which every 2-edge coloured $K_{\mathbb{N}}$ contains a monochromatic copy of G with positive upper density (we call such graphs *Ramsey-dense*). Note that this density depends on the colouring and therefore, since there are infinitely many colourings, a Ramsey-dense graph does not necessarily have positive Ramsey upper density. On the other hand, every infinite graph G with positive Ramsey upper density is Ramsey-dense. The Rado graph is an example of a Ramsey-dense graph with Ramsey upper density 0 (see Corollary 1.4.17 below).

Graphs with positive Ramsey upper density

We will now investigate an analogue of Conjecture 1.2.4 (which states that sequences of graphs with bounded degree have linear Ramsey number) for infinite graphs. Given some integer $k \geq 2$, we say that an infinite graph G is *one-way k-locally finite* if there exists a partition of $V(G)$ into k independent sets $V_1, \ldots, V_k$ with $\infty = |V_1| \geq \ldots \geq |V_k|$ such that for all $1 \leq i < j \leq k$ and all $v \in V_j$, we have $d(v, V_i) < \infty$. Note that every vertex in V_k has finite degree, but it is possible for any vertex in $V_1 \cup \cdots \cup V_{k-1}$ to have infinite degree. Further note that one-way k-locally finite graphs have chromatic number at most k and, if G is locally finite with $\chi(G) < \infty$, then G is one-way $\chi(G)$-locally finite.

Theorem 1.4.11 (Corsten–DeBiasio–McKenney)**.** *Let $r, k \geq 2$ be integers and let G be a one-way k-locally finite graph.*

(i) If $k = 2$, then $\overline{\mathrm{Rd}}_r(G) \geq 1/r$.

(ii) If $k \geq 3$, then $\overline{\mathrm{Rd}}(G) \geq 1/(2(k-1))$.

(iii) If $k \geq 3$, then $\overline{\mathrm{Rd}}_r(G) \geq \left(\sum_{i=0}^{(k-2)r+1}(r-1)^i\right)^{-1} = (1+o_k(1))r^{-(k-2)r-1}$.

As a consequence, we obtain the following corollary about graphs with bounded degree. The two-colour case of this answers a question from [33] and has been independently proven by Lamaison [89].

Corollary 1.4.12 (Corsten–DeBiasio–McKenney)**.** *If G is an infinite graph with maximum-degree Δ for some $\Delta \in \mathbb{N}$, then $\overline{\mathrm{Rd}}(G) \geq 1/(2\Delta)$ and $\overline{\mathrm{Rd}}_r(G) \geq r^{-\Delta r}$ for every $r \geq 3$.*

Interestingly, the constant $1/(2\Delta)$ in the two-colour case is much better than expected: for finite graphs, we cannot hope for anything better than $2^{-\Theta(\Delta)}$ by Theorem 1.2.5. However, note that the constant is very similar to that in Theorem 1.2.7.

Recall that for finite graphs, Theorem 1.2.6 can be extended to a tiling result (Theorems 1.3.12 and 1.3.16). Our proof of Theorem 1.4.11 actually gives the following stronger result of a similar form.

Theorem 1.4.13 (Corsten–DeBiasio–McKenney)**.** *Let $r, k \geq 2$ be integers and let G be a one-way k-locally finite graph. Then, in every r-edge-coloured $K_{\mathbb{N}}$, there exists a collection of at most*

$$f(r,k) = \begin{cases} r & \text{if } k = 2 \\ 2(k-1) & \text{if } r = 2 \\ \sum_{i=0}^{(k-2)r+1}(r-1)^i & \text{if } k \geq 3 \end{cases}$$

vertex-disjoint, monochromatic copies of G whose union covers all but finitely many vertices.

This generalises a result of Elekes, D. Soukup, L. Soukup, and Szentmiklóssy [37] who proved a similar statement for powers of cycles.

We will present the proof of Theorems 1.4.11 and 1.4.13 in Section 4.2.7. It would be very interesting to prove an analogue of Conjecture 1.2.8 for infinite graphs of bounded degeneracy.

Question 1.4.14 (Corsten–DeBiasio–McKenney)**.** *Do we have* $\overline{\mathrm{Rd}}(G) > 0$ *for every infinite graph G with finite degeneracy?*

Ramsey-dense graphs

We call an infinite graph G *r-Ramsey-dense* if in every r-edge-colouring of $K_{\mathbb{N}}$ there is a monochromatic copy of G with positive upper density. If $r = 2$, we drop the prefix and just say G is *Ramsey-dense*. We will describe a simple property that is sufficient to guarantee that a graph is Ramsey-dense and then we show that every Ramsey-dense graph is not far from having this property.

A set $X \subseteq V(G)$ is called *dominating* if every vertex $v \in V(G) \setminus X$ has a neighbour in X. We call a set $X \subseteq V(G)$ *ruling* if X is finite and all but finitely many vertices $v \in V(G) \setminus X$ have a neighbour in X. We say that an infinite graph G is *t-ruled* if there are at most t disjoint minimal ruling sets. The *ruling number* of a graph G, denoted by $\mathrm{rul}(G)$, is the smallest $t \in \mathbb{N}$ such that G is t-ruled and $\mathrm{rul}(G) := \infty$ if no such t exists. Equivalently, $\mathrm{rul}(G)$ is the matching number of the hypergraph whose edges are all minimal ruling sets. Note that a graph is 0-ruled if and only if there is no finite dominating set and finitely-ruled (i.e. t-ruled for some $t \in \mathbb{N}_0$) if and only if there is a finite set $S \subseteq V(G)$ such that $G[\mathbb{N} \setminus S]$ has no finite dominating sets.

Theorem 1.4.15 (Corsten–DeBiasio–McKenney)**.** *If G is an infinite graph with finite ruling number, then G is r-Ramsey-dense for every $r \in \mathbb{N}$.*

Theorem 1.4.15 has a couple of interesting corollaries. For example, it implies that every infinite graph with finite degeneracy is Ramsey-dense.

Corollary 1.4.16 (Corsten–DeBiasio–McKenney)**.** *If G is an infinite graph with finite degeneracy, then G is r-Ramsey-dense for every $r \in \mathbb{N}$.*

Proof. By Theorem 1.4.15, it suffices to show that every d-degenerate infinite graph G is d-ruled. Suppose for contradiction, there is a d-degenerate infinite graph G with $\mathrm{rul}(G) > d$ for some $d \in \mathbb{N}$. Let $S_1, \ldots, S_{d+1}$ be disjoint minimal ruling sets and let $S_0 \subseteq V(G) \setminus (S_1 \cup \ldots \cup S_{d+1})$ be the set of vertices which do not have a neighbour in some S_i. Note that $S := S_0 \cup S_1 \cup \ldots \cup S_{d+1}$ is finite. Therefore, there is a vertex $u \in \mathbb{N} \setminus S$ which comes after all vertices in S in a d-degenerate ordering of $V(G)$ and hence $\deg(u, S) \leq d$. However, by construction, u has a neighbour in each of $S_1, \ldots, S_{d+1}$, a contradiction. □

Another corollary concerns the well-studied Rado graph. The *Rado graph* $\mathcal{R}$ is the graph with vertex-set $\mathbb{N}$ defined by placing an edge between $m < n$ if and only if the mth digit in the binary expansion of n is 1 (starting with the least significant bit). It is easy to verify that the Rado graph does not have any finite dominating sets and hence $\mathrm{rul}(\mathcal{R}) = 0$.

Corollary 1.4.17 (Corsten–DeBiasio–McKenney)**.** *The Rado graph $\mathcal{R}$ is r-Ramsey-dense for every $r \in \mathbb{N}$.*

We will see later that the Rado graph has Ramsey upper density 0 (Corollary 4.2.7), showing that not every Ramsey-dense graph has positive Ramsey upper density.

We do not know if the converse of Theorem 1.4.15 holds:

Question 1.4.18 (Corsten–DeBiasio–McKenney)**.** *Is there a Ramsey-dense graph G with infinite ruling number?*

If the answer is no, then we have a complete characterisation of Ramsey-dense graphs. The following result shows that if G has infinite ruling number and additionally the sizes of the minimal ruling sets do not grow too fast, then G is not Ramsey-dense.

Theorem 1.4.19 (Corsten–DeBiasio–McKenney)**.** *Let G be an infinite graph with disjoint ruling sets $F_1, F_2, \ldots$ satisfying $|F_n| \leq \log_2(n)$ for all sufficiently large n. Then G is not Ramsey-dense.*

The proof of Theorems 1.4.15 and 1.4.19 will be presented in Section 4.2.8.

1.5 Robust Triangle Tilings in Random Graphs

1.5.1 Shamir's Problem

A *perfect matching* in a k-graph G is a collection of disjoint edges covering all vertices. The following classical result of Erdős and Rényi [46] determines the threshold probability for the existence of a perfect matching in $G(n,p)$.

Theorem 1.5.1 (Erdős–Rényi [46])**.** *Let $\varepsilon > 0$, $n \geq 2$ be an even integer and $\mathcal{M}_n$ be the set of all n-vertex graphs containing a perfect matching. Then,*

$$\lim_{n\to\infty} \mathbb{P}\left[G(n,p) \in \mathcal{M}_n\right] = \begin{cases} 0 & \text{if } p \leq (1-\varepsilon)\log(n)/n, \\ 1 & \text{if } p \geq (1+\varepsilon)\log(n)/n. \end{cases}$$

Shamir's problem, which first appeared in [41] (where Erdős attributes it to Shamir) asks the analogous question for hypergraphs.

Problem 1.5.2 (Shamir's problem)**.** *For $k \geq 3$, what is the threshold probability for the existence of a perfect matching in $G^{(k)}(n,p)$ (when n is divisible by k)?*

Given two graphs H and G, an H-tiling in G is a collection of vertex-disjoint copies of H in G covering all vertices of G. In this thesis, we shall be mostly interested in the following tiling version of Shamir's problem (which was first raised by Ruciński [105] for $k = 3$).

Problem 1.5.3 (Shamir's problem - tiling version)**.** *For $k \geq 3$, what is the threshold probability for the existence of a K_k-tiling in $G(n,p)$ (when n is divisible by k)?*

An obvious necessary condition for Problem 1.5.2 is that there are no isolated vertices (that is, vertices of degree 0). It is not hard to see that the (sharp) threshold for $G^{(k)}(n,p)$ not to have isolated vertices is $p_m(n,k) := (\log n)/\binom{n-1}{k-1}$ (that is, when the expected degree of each vertex is $\log n$). The analogous necessary condition for Problem 1.5.3 is that every vertex must be in at least one copy of K_k. It is well known (see [73, Theorem 3.22]) that the (sharp) threshold probability for this is

$$p_t(n,k) := (\log n)^{1/\binom{k}{2}} \binom{n-1}{k-1}^{-1/\binom{k}{2}}$$

(that is, when we expect exactly $\log n$ copies of K_k containing a given vertex).

Problems 1.5.2 and 1.5.3 have received a lot of attention, but proved to be very challenging. After some progress in [6, 53, 78, 85, 104, 110], Johansson, Kahn and Vu [75] completely solved both problems in an unexpected breakthrough by showing that p_m and p_t are already the (coarse) thresholds for Shamir's problem. In fact, their result does not only hold for K_k-tilings but many more graphs and hypergraphs (so-called strictly balanced graphs). We will state the theorem only in the case of Problem 1.5.3.

Theorem 1.5.4 (Johansson–Kahn–Vu [75])**.** *For every $k \in \mathbb{N}$, there is a constant $C > 0$ such that for all $n \in \mathbb{N}$ divisible by k and $p \geq C(\log n)^{1/\binom{k}{2}} n^{-2/k}$, $G(n, p)$ has a K_k-tiling w.h.p.*

By the above discussion, Theorem 1.5.4 is best possible up to the constant C. Recently, Kahn [76] determined the asymptotics completely for Problem 1.5.2 by showing that $p_m(n)$ is a sharp threshold as well. Further recent results of Riordan [101] and Heckel [68] which 'couple' the two problems show that this implies that p_t is a sharp threshold for Problem 1.5.3 as well.

1.5.2 Robustness

A *Hamilton cycle* in a graph G is a cycle in G spanning all its vertices and a graph is called *Hamiltonian* if it has a Hamilton cycle. A classical theorem of Dirac [34] gives a sufficient condition for a graph to be Hamiltonian.

Theorem 1.5.5 (Dirac [34])**.** *If G is an n-vertex graph with $\delta(G) \geq n/2$, then G is Hamiltonian.*

Considering the complete bipartite graph with parts of size $\lfloor n/2 \rfloor + 1$ and $\lceil n/2 \rceil - 1$ shows that this result is best possible. Another classical result, independently proven by Korshunov [84] and by Pósa [96], determines the probability threshold for $G(n, p)$ to be Hamiltonian.

Theorem 1.5.6 (Korshunov [84], Pósa [96])**.** *There is a constant $C > 0$ such that for all $n \in \mathbb{N}$ and all $p \geq C\log(n)/n$, $G(n, p)$ is Hamiltonian w.h.p.*

Since every Hamiltonian graph is connected, Theorem 1.0.5 implies that this is best possible up to the implicit constant. A more exact window for the transition phase in Theorem 1.5.6 was proved independently by Bollobás [11] and by Komlós and Szemerédi [83]. Krivelevich, Lee and Sudakov [86] proved a remarkable common generalisation of Theorems 1.5.5 and 1.5.6.

Theorem 1.5.7 (Krivelevich–Lee–Sudakov [86]). *There is a constant $C > 0$ such that for all $n \in \mathbb{N}$, $p \geq C\log(n)/n$ and all n-vertex graphs G with $\delta(G) \geq n/2$, G_p is Hamiltonian w.h.p.*

Observe that, by choosing $p = 1$, we obtain Theorem 1.5.5, and by choosing $G = K_n$ we obtain Theorem 1.5.6. Therefore, both the minimum degree condition and the lower bound on p (up to the constant C) are best possible. This can be seen as a *robust* version of Theorem 1.5.5: the conclusion even holds if we randomly remove most edges.

New Results

We shall prove a similar result for triangle tilings. The classical Corrádi–Hajnal theorem [28] gives a sufficient condition for the existence of a triangle tiling in terms of the minimum degree.

Theorem 1.5.8 (Corrádi–Hajnal [28]). *If G is an n-vertex graph for some n divisible by 3 with $\delta(G) \geq 2n/3$, then G contains a triangle tiling.*

We will prove a common generalisation of Theorem 1.5.8 and of Theorem 1.5.4 for $k = 3$.

Theorem 1.5.9 (Allen et al. [3]). *There is a constant $C > 0$ such that for all $n \in \mathbb{N}$ divisible by 3, all $p \geq C(\log n)^{1/3}n^{-2/3}$ and all n-vertex graphs G with $\delta(G) \geq 2n/3$, G_p has a triangle tiling w.h.p.*

Similarly as above, we obtain Theorem 1.5.8 by choosing $p = 1$, and Theorem 1.5.4 for $k = 3$ by choosing $G = K_n$. Therefore, both the minimum degree condition and the lower bound on p (up to the constant C) are best possible. We will present the proof of Theorem 1.5.9 in Chapter 5. The main tool in our proof is the entropy method and, while many ideas are similar to those in [75], the proof strategy actually follows an alternative proof of Problem 1.5.2, which was given by Allen, Böttcher, Davies, Jenssen, Kohayakawa and Roberts [4].

2 Preliminaries

2.1 The Absorption Method

The absorption method, introduced by Erdős, Gyárfás and Pyber in [42] to prove Theorem 1.3.4 below, has become a standard tool in graph tiling problems and has been applied to many problems in the area. We shall make use of it in the proofs of Theorems 1.3.10 and 1.3.16 and we will briefly sketch the proof of Theorem 1.3.4 in this section in order to introduce the method.

Theorem 1.3.4 (Erdős–Gyárfás–Pyber [42])**.** *The vertices of every r-edge-coloured complete graph on n vertices can be partitioned into $O(r^2 \log r)$ monochromatic cycles.*

For sake of clarity, we will not make an effort to calculate the exact constants and only show that $O_n(1)$ cycles suffice (i.e. the number of cycles will be a function independent of n). At the heart of the proof are so called absorbers.

Definition 2.1.1. A pair (H, A) of a graph H and a set $A \subseteq V(H)$ is called an *absorber* if $H[V(H) \setminus X]$ contains a Hamilton cycle for every $X \subseteq A$.

Note that this definition is specific to cycles and will be different for other graphs. It is vital for the proof that we can find large absorbers.

Lemma 2.1.2. *For every $r \geq 2$, there is some constant $c = c(r) \in (0, 1/4)$, such that every r-edge-coloured K_n contains a monochromatic absorber (H, A) with $|A| = |V(H) \setminus A| = cn$.*

We will further need the following approximate version of the theorem.

Lemma 2.1.3. *For every $r \geq 2$ and $\varepsilon > 0$, there is some constant $C = C(r, \varepsilon)$ such that every r-edge-colouring of K_n contains a collection of at most C vertex-disjoint monochromatic cycles covering all but at most εn vertices.*

Lemma 2.1.3 follows from the fact the cycles have linear r-colour Ramsey number (see Section 1.2.1) by greedily taking out the largest monochromatic cycle disjoint from all previous cycles, one at a time. The key part of the proof is the following Absorption Lemma.

Lemma 2.1.4 (Absorption Lemma for cycles)**.** *For every $r \geq 2$, there exist constants $\varepsilon = \varepsilon(r) > 0$ and $C = C(r) > 0$ such that the following holds. Let V_1, V_2 be sets with $|V_1| \leq \varepsilon |V_2|$ and let G be an r-edge-coloured complete bipartite graph with parts V_1, V_2. Then, there is a collection of at most C vertex-disjoint monochromatic cycles in G covering V_1.*

The proofs of Lemmas 2.1.2 to 2.1.4 are not essential to understand the method and therefore we will not present it here (see [42] for the proofs).

Proof of Theorem 1.3.4. Fix $r, n \in \mathbb{N}$ and an r-edge-coloured K_n for the rest of this proof. Let c_1 be the constant from Lemma 2.1.2, $C_2(\varepsilon)$ be the constant from Lemma 2.1.3, and C_3 and ε_3 be the constants from Lemma 2.1.4.

By Lemma 2.1.2, there is a monochromatic absorber (H, A) with $|A| = |V(H) \setminus A| = c_1 \cdot n$. We apply now Lemma 2.1.3 with input r and $\varepsilon_3 \cdot c_1$ to find a collection of at most $C_2(\varepsilon_3 \cdot c_1)$ vertex-disjoint monochromatic cycles in $V(K_n) \setminus V(H)$ such that the set of leftover vertices R satisfies

$$|R| \leq \varepsilon_3 \cdot c_1(n - |V(H)|) \leq \varepsilon_3 \cdot |A| \,.$$

Hence, we can apply Lemma 2.1.4 with input $X_1 = R$ and $X_2 = A$ to find a collection of at most C_3 vertex-disjoint cycles covering R and a set $R' \subseteq A$.

We have covered $V(K_n) \setminus (V(H) \setminus R')$ so far, and we can cover $V(H) \setminus R'$ with one monochromatic cycle using the absorption property of H. We thus covered all vertices with $f(r) = C_2(\varepsilon_3 \cdot c_1) + C_3 + 1$ monochromatic cycles.[1] □

[1] Carrying out the exact calculations leads to $f(r) = O(r^2 \log(r))$, see [42].

2.2 Graph Regularity

The graph regularity method has become one of the most powerful tools in extremal graph theory since its first use in Szemerédi's theorem [114, 116].[2] It is centred around Szemerédi's Regularity Lemma, which was first explicitly stated in its current form in [115] and there has been a lot of work developing this method ever since, see [82] for a survey on the topic. We will use results from this area in every chapter of thesis and therefore give a brief introduction, and then list all applications and related results we shall make use of in this thesis.

2.2.1 Definitions and Basic Properties

Let $G = (V, E)$ be a graph and A and B be non-empty, disjoint subsets of V. We write $e_G(A, B)$ to denote the number of edges in G with one vertex in A and one in B and we define the *density* of the pair (A, B) to be $d_G(A, B) = e_G(A, B)/(|A||B|)$. If G is clear from context, we drop the subscript. The pair (A, B) is called ε*-regular* (in G) if we have $|d_G(A', B') - d_G(A, B)| \leq \varepsilon$ for all $A' \subseteq A$ with $|A'| \geq \varepsilon|A|$ and $B' \subseteq B$ with $|B'| \geq \varepsilon|B|$. The pair (A, B) is called (ε, d)*-regular* if it is ε-regular and $d_G(A, B) = d$.

We will often require an even stronger condition than ε-regularity in which we have control over the degrees of all vertices. A pair (A, B) of disjoint subsets of $V(G)$ is called (ε, d, δ)*-super-regular* (in G) if

(i) (A, B) is (ε, d)-regular in G,

(ii) $\deg(v, V_{3-i}) \geq \delta|V_i|$ for all $v \in V_i$ and $i = 1, 2$.

We call (A, B) $(\varepsilon, d^+, \delta)$*-super-regular* if it is $(\varepsilon, d', \delta)$-super-regular for some $d' \geq d$ and (ε, d^+)*-super-regular* if it is (ε, d^+, d)-super-regular.

We begin with some simple facts about (super-)regular pairs. The first one is known as slicing lemma and roughly says that if we shrink a dense regular pair we still get a dense regular pair. Its proof is straight forward from the definition of regular pair.

[2] Szemerédi's theorem states that every set $A \subseteq \mathbb{N}$ with positive upper density contains infinitely many arithmetic progressions of length k for every $k \in \mathbb{N}$.

Lemma 2.2.1 (Slicing lemma). *Let $\beta > \varepsilon > 0$, and let (V_1, V_2) be an ε-regular pair. Then any pair (U_1, U_2) with $|U_i| \geq \beta|V_i|$ and $U_i \subseteq V_i$, $i = 1, 2$, is ε'-regular with $\varepsilon' = \max\{\varepsilon/\beta, 2\varepsilon\}$.*

The following lemma essentially says that after removing few vertices from a super-regular pair and adding few new vertices with large degree, we still have a super-regular pair.

Lemma 2.2.2. *Let $0 < \varepsilon < 1/2$ and let $d, \delta \in [0, 1]$ so that $\delta \geq 4\varepsilon$. Let (V_1, V_2) be an $(\varepsilon, d^+, \delta)$-super-regular pair in a graph G. Let $X_i \subseteq V_i$ for $i \in [2]$, and let Y_1, Y_2 be disjoint subsets of $V(G) \setminus (V_1 \cup V_2)$. Suppose that for each $i \in [2]$ we have $|X_i|, |Y_i| \leq \varepsilon^2|V_i|$ and $\deg(y, V_i) \geq \delta|V_i|$ for every $y \in Y_{3-i}$. Then the pair $((V_1 \setminus X_1) \cup Y_1, (V_2 \setminus X_2) \cup Y_2)$ is $(8\varepsilon, (d - 8\varepsilon)^+, \delta/2)$-super-regular.*

Proof. Let $U_i = (V_i \setminus X_i) \cup Y_i$ for $i \in [2]$. We will show that (U_1, U_2) is $(8\varepsilon, d - 8\varepsilon, \delta/2)$-super-regular. Let now $Z_i \subseteq U_i$ with $|Z_i| \geq 8\varepsilon|U_i|$, and let $Z_i' = Z_i \setminus Y_i$ and $Z_i'' = Z_i \cap Y_i$ for $i \in [2]$. Note that we have

$$\begin{aligned}
&|Z_i| \geq 8\varepsilon|U_i| \geq \varepsilon|V_i|,\\
&|Z_i''| \leq |Y_i| \leq \varepsilon^2|V_i| \leq \varepsilon|Z_i| \text{ and}\\
&|Z_i'| = |Z_i| - |Z_i''| \geq (1 - \varepsilon)|Z_i|
\end{aligned}$$

for both $i \in [2]$. We therefore have

$$e(Z_1, Z_2) \leq e(Z_1', Z_2') + e(Z_1'', Z_2) + e(Z_1, Z_2'') \leq e(Z_1', Z_2') + 2\varepsilon|Z_1||Z_2|$$

and thus

$$d(Z_1, Z_2) \leq d(Z_1', Z_2') + 2\varepsilon.$$

On the other hand, we have

$$\begin{aligned}
d(Z_1, Z_2) = \frac{e(Z_1, Z_2)}{|Z_1||Z_2|} &\geq \frac{e(Z_1', Z_2')}{|Z_1'||Z_2'|} \cdot \frac{|Z_1'||Z_2'|}{|Z_1||Z_2|}\\
&\geq d(Z_1', Z_2')(1 - \varepsilon)^2 \geq d(Z_1', Z_2') - 2\varepsilon
\end{aligned}$$

and hence $d(Z_1, Z_2) = d(Z'_1, Z'_2) \pm 2\varepsilon$. Furthermore, by ε-regularity of (V_1, V_2), we have $d(Z'_1, Z'_2) = d(V_1, V_2) \pm \varepsilon$ and we conclude

$$d(Z_1, Z_2) = d(V_1, V_2) \pm 3\varepsilon.$$

This holds in particular for $Z_1 = U_1$ and $Z_2 = U_2$ and therefore the pair (U_1, U_2) is $(8\varepsilon, (d - 8\varepsilon)^+, 0)$-super-regular.

Let $u_1 \in U_1$ now. By assumption, we have $\deg(u_1, V_2) \geq \delta|V_2|$ and therefore

$$\begin{aligned}\deg(u_1, U_2) \geq \deg(u_1, V_2 \setminus X_2) &\geq (\delta - \varepsilon^2)|V_2| \\ &\geq (\delta - \varepsilon^2)|U_2| \geq \delta/2 \cdot |U_2|.\end{aligned}$$

A similar statement is true for every $u_2 \in U_2$ finishing the proof. □

Given disjoint sets of vertices $V_1, \dots, V_k \subseteq V(G)$, we call $Z = (V_1, \dots, V_k)$ a k-cylinder and often identify it with the induced k-partite subgraph $G[V_1, \dots, V_k]$. We write $V_i(Z) = V_i$ for every $i \in [k]$. We say that Z is *ε-balanced* if

$$\max_{i\in[k]} |V_i(Z)| \leq (1 + \varepsilon) \min_{i\in[k]} |V_i(Z)|$$

and *balanced* if it is 0-balanced. Furthermore, we say that Z is ε-regular if all the $\binom{k}{2}$ pairs (V_i, V_j) are ε-regular and we define $(\varepsilon, d^+, \delta)$-super-regular cylinders and (ε, d^+)-super-regular cylinders similarly.

Lemma 2.2.3. *Let k be a positive integer and $d, \varepsilon > 0$ with $\varepsilon \leq 1/(2k)$. If $Z = (V_1, \dots, V_k)$ is an ε-regular k-cylinder and $d(V_i, V_j) \geq d$ for all $1 \leq i < j \leq k$, then there is some $\gamma \geq 1 - k\varepsilon$ and sets $\tilde{V}_1 \subseteq V_1, \dots, \tilde{V}_k \subseteq V_k$ with $|\tilde{V}_i| = \lceil (1-\gamma)|V_i| \rceil$ for all $i \in [k]$ so that the k-cylinder $\tilde{Z} = (\tilde{V}_1, \dots, \tilde{V}_k)$ is $(2\varepsilon, (d - k\varepsilon)^+)$-super-regular.*

Proof. For $i \neq j \in [k]$, let $A_{i,j} := \{v \in V_i : \deg(v, V_j) < (d-\varepsilon)|V_j|\}$. By definition of ε-regularity, we have $\left|A_{i,j}\right| < \varepsilon|V_i|$ for every $i \neq j \in [k]$. For each $i \in [k]$, let $A_i = \bigcup_{j\in[k]\setminus\{i\}} A_{i,j}$. Clearly $|A_i| < (k - 1)\varepsilon|V_i|$ for every $i \in [k]$, so we can add arbitrary vertices from $V_i \setminus A_i$ to A_i until $|A_i| = \lfloor (k - 1)\varepsilon|V_i| \rfloor$ for every $i \in [k]$.

Let now $\tilde{V}_i = V_i \setminus \tilde{A}_i$ for every $i \in [k]$ and let $\tilde{Z} = (\tilde{V}_1, \dots, \tilde{V}_k)$. Observe that $|\tilde{V}_i| = (1-\gamma)|V_i|$ for all $i \in [k]$, where $\gamma = (k-1)\varepsilon$. It follows from Lemma 2.2.1 and definition of A_i that $\tilde{Z}$ is $(2\varepsilon, (d-\varepsilon)^+, d-k\varepsilon)$-super-regular. □

The following lemma further allows us to control the exact density of a super-regular pair by deleting edges if necessary.

Lemma 2.2.4. *For all $\varepsilon > 0$, there is some $n_0 > 0$, such that the following is true for every $n \geq n_0$ and every bipartite graph Γ with parts V_1, V_2 of size n. Suppose that (V_1, V_2) is $(\varepsilon^2, d, \delta)$-super-regular for some $d \geq \delta \geq 4\varepsilon$. Then there is a spanning subgraph $\Gamma' \subseteq \Gamma$ so that (V_1, V_2) is $(4\varepsilon, \delta, \delta - \varepsilon)$-super-regular in Γ'.*

For the proof, we will need the hypergeometric distribution: A random variable $X : \Omega \to \mathbb{N}_0$ is *hypergeometrically distributed* with parameters $N \in \mathbb{N}$, $K \in [0, N]$ and $t \in \mathbb{N}_0$ if $\mathbb{P}[X = k]$ is the probability that when drawing t balls from a set of N balls (K of which are blue and $N - K$ red) without replacement, exactly k are blue. That is,

$$\mathbb{P}[X = k] = \frac{\binom{K}{k}\binom{N-K}{t-k}}{\binom{N}{t}}.$$

We will use the following concentration inequality, which Chvátal [22] deduced from the so-called Hoeffding's inequality [69].

Lemma 2.2.5. *Let X be hypergeometrically distributed with parameters $N \in \mathbb{N}$, $K \in [0, N]$ and $t \in \mathbb{N}_0$ and let $\mu := \mathbb{E}[X] = tK/N$. Then, for all $\lambda > 0$, we have*

$$\mathbb{P}[|X - \mu| > \lambda\mu] \leq 2e^{-2\lambda^2 (K/N)\cdot\mu}.$$

Proof of Lemma 2.2.4. For $i \in [2]$, let $Y_i := \{v \in V_i : \deg(v, V_{3-i}) \leq (d-\varepsilon)n\}$ and observe that by ε^2-regularity of (V_1, V_2), we have $|Y_i| \leq \varepsilon^2 n$ for both $i \in [2]$. Let $E_Y \subseteq E(\Gamma)$ be the set of edges with at least one vertex in $Y := Y_1 \cup Y_2$ and let $E := E(\Gamma) \setminus E_Y$. Let $m := |E_Y| \leq 2\varepsilon^2 n^2$. Let $p := \frac{\delta n^2 - m}{|E(\Gamma)|} = \frac{\delta \pm \varepsilon^2}{d}$. Let E' be a uniformly random subset of E of size exactly $p|E|$ and let Γ' be the spanning subgraph of Γ with edge-set $E' \cup E_Y$. By construction, we have $d_{\Gamma'}(V_1, V_2) = \delta$; we will show that (V_1, V_2) is $(4\varepsilon, \delta, \delta - \varepsilon)$-super-regular in Γ' w.h.p.

Let $A_i \subseteq V_i$ with $A_i \geq 4\varepsilon n$, and let $B_i = A_i \setminus Y_i$ and $B'_i = A_i \setminus B_i$ for both $i \in [2]$. By ε^2-regularity in Γ, we have $Y := |E_\Gamma(B_1, B_2)| = (d \pm \varepsilon^2)|B_1||B_2|$. Let now $X := |E_{\Gamma'}(B_1, B_2)|$. Then X is hypergeometrically distributed with parameters $N = |E|, K = Y, t = p|E|$ and thus $\mu := \mathbb{E}[X] = p \cdot Y = (\delta \pm 2\varepsilon)|B_1||B_2|$. Since $\mu \geq 8\varepsilon^3 n^2$, it follows from Lemma 2.2.5 that

$$\mathbb{P}[|X - \mu| > \varepsilon \cdot \mu] \leq 2e^{-2\varepsilon^2 K/N \cdot \mu} \leq 2e^{-\varepsilon^8 n^2}.$$

In particular, we have $\mathbb{P}[d(A_1, A_2) = \delta \pm 4\varepsilon] \geq 1 - 2e^{-\varepsilon^8 n^2}$. By taking a union bound over all choices of X_1, X_2, we deduce that (V_1, V_2) is 2ε-regular with probability at least $1 - 2e^{2n-\varepsilon^8 n^2}$. Similarly, we deduce that $\deg_{\Gamma'}(v_i, V_{3-i}) \geq (\delta - \varepsilon)n$ for each $i \in [2]$ and $v_i \in V_i$ with probability at least $1 - 4ne^{-\varepsilon^8 n}$. (Note that this is automatically true for all $v \in Y$ as we fixed their neighbours.) Hence, taking another union bound, it follows that (V_1, V_2) is $(4\varepsilon, \delta, \delta - \varepsilon)$-super-regular in Γ' w.h.p. Therefore, for all large enough n, there is a suitable choice for E'. □

The following lemma states that every regular k-cylinder contains roughly the 'right amount' of k-cliques. Given a graph G, let $K^{(k)}(G)$ be the set of cliques of size k.

Lemma 2.2.6. *Let $k \in \mathbb{N}$, $\varepsilon > 0$ and let Γ be a k-partite graph with parts $V_1, \ldots, V_k$, such that, for all $1 \leq i < j \leq k$, (V_i, V_j) is $(\varepsilon, d_{i,j})$-regular for some $d_{i,j} > 0$. Then, for all $X_1 \subseteq V_1, \ldots X_k \subseteq V_k$, we have*

$$\left|K^{(k)}(\Gamma[X_1, \ldots, X_k])\right| = \left(\prod_{1 \leq i < j \leq k} d_{i,j}\right) |X_1| \cdots |X_k| \pm k^2 \varepsilon \cdot n^k.$$

Proof. We will only prove the lower bound as the upper bound is similar. We proceed by induction on k. The statement is trivial for $k = 1$ (here the empty product is 1). So let $k \geq 2$ and assume the statement is true for $k - 1$. First note that the statement is trivially true if $X_i < \varepsilon n$ for some $i \in [k]$; hence we may assume that $|X_i| \geq \varepsilon n$ for all $i \in [k]$. For $i \in [2, k]$, let $A_i := \{v_1 \in V_1 : \deg(v_1, X_i) < (d_{1,i} - \varepsilon)|X_i|\}$. By $(\varepsilon, d_{1,i})$-regularity of (V_1, V_i), we have $|A_i| \leq \varepsilon n$ for all $i \in [2, k]$ and thus $|A| := |\bigcup_{i \in [2,k]} A_i| \leq (k - 1)\varepsilon n$.

Fix some $x_1 \in X_1 \setminus A$ now and let $B_i := N(x_1, X_i)$ for all $i \in [2, k]$. By choice of x_1 we have $|B_i| \geq (d_{1,i} - \varepsilon)|X_i|$ for all $i \in [2, k]$. Using the induction hypothesis with $B_2, \ldots, B_k$ in the cylinder $(V_2, \ldots, V_k)$, we deduce

$$\begin{aligned}\left|K^{(k-1)}\left(\Gamma[B_2, \ldots, B_k]\right)\right| &\geq \left(\prod\nolimits_{2 \leq i < j \leq k} d_{i,j}\right) |B_2| \cdots |B_k| \pm (k-1)^2 \varepsilon n^{k-1} \\ &\geq \left(\prod\nolimits_{1 \leq i < j \leq k} d_{i,j}\right) |X_2| \cdots |X_k| \pm k(k-1) \varepsilon n^{k-1}.\end{aligned}$$

Finally, since every such clique extends to a clique in $K^{(k)}(\Gamma[X_1, \ldots, X_k])$ with x_1 and since $|X_1 \setminus A| \geq |X_1| - (k-1)\varepsilon n$, we deduce

$$\begin{aligned}\left|K^{(k)}\left(\Gamma[X_1, \ldots, X_k]\right)\right| &\geq (|X_1| - (k-1)\varepsilon n) \left|K^{(k-1)}\left(\Gamma[B_2, \ldots, B_k]\right)\right| \\ &\geq \left(\prod\nolimits_{1 \leq i < j \leq k} d_{i,j}\right) |X_1| \cdots |X_k| - k^2 \varepsilon \cdot n^k,\end{aligned}$$

as claimed. □

We will frequently use the following immediate corollary for (ε, d)-regular 3-cylinders.

Lemma 2.2.7. *Let $d, \varepsilon > 0$ and let Γ be a 3-partite graph with parts V_1, V_2, V_3 such that (V_1, V_2, V_3) is (ε, d)-regular. Then, for all $X_1 \subseteq V_1, X_2 \subseteq V_2, X_3 \subseteq V_3$, we have*

$$\left|K^{(3)}\left(\Gamma[X_1, X_2, X_3]\right)\right| = d^3 |X_1||X_2||X_3| \pm 10\varepsilon \cdot n^3.$$

2.2.2 The Regularity Lemma

In order to state the regularity lemma, we will first define regular partitions.

Definition 2.2.8 (ε-regular partition)**.** Given a graph G, a partition $V(G) = V_0 \cup V_1 \cup \ldots \cup V_r$ is called an *ε-regular partition* if

- $|V_1| = \ldots = |V_r|$,
- $|V_0| \leq \varepsilon n$, and
- All but at most εr^2 pairs (V_i, V_j) with $1 \leq i < j \leq r$ are ε-regular.

We call $V_0, V_1, \ldots, V_r$ clusters and V_0 the *garbage set*.

Theorem 2.2.9 (Regularity Lemma [115])**.** *For all $m, \varepsilon > 0$ there exists $M > 0$ such that the following holds. Every graph G with at least M vertices has an ε-regular partition $V(G) = V_0 \cup V_1 \cup \cdots \cup V_r$ for some $m \leq r \leq M$.*

Note that some of the regular pairs in a regular partition can have very small density (even 0), which can make certain applications difficult. We will define a graph with vertex-set $\{V_1, \ldots, V_r\}$ whose edges are regular pairs that behave nicely.

Definition 2.2.10 (Reduced graph)**.** Let $d, \varepsilon > 0$, G be a graph and $V(G) = V_0 \cup V_1 \cup \ldots \cup V_r$ be an ε-regular partition. The (ε, d)*-reduced graph* R is the graph on $\{V_1, \ldots, V_r\}$ with $V_iV_j \in E(R)$ if (V_i, V_j) is an ε-regular pair with $d(V_i, V_j) \geq d$.

The following version of the regularity lemma (see [88, Proposition 9]) shows that R inherits minimum degree properties of G.

Theorem 2.2.11 (Regularity Lemma - degree version)**.** *For all $m_0, \varepsilon, d, \gamma$ with $0 < \varepsilon < d < \gamma < 1$ there is some $M > 0$ such that every graph G on $n \geq M$ vertices with minimum degree $\delta(G) \geq \gamma n$ has an ε-regular partition $V(G) = V_0 \cup V_1 \cup \ldots \cup V_r$ for some $m \leq r \leq M$ and the (ε, d)-reduced graph R satisfies $\delta(R) \geq (\gamma - d - \varepsilon)r$.*

In Chapter 4, we will make use of a version of the regularity lemma for edge-coloured graphs which works for all colour classes simultaneously (see [82, Theorem 1.18]). Additionally, we require more control over the produced partition. A family of disjoint subsets $\{V_i\}_{i \in [r]}$ of a set V is said to *refine* a partition $\{W_j\}_{j \in [\ell]}$ of V if, for all $i \in [r]$, there is some $j \in [\ell]$ with $V_i \subseteq W_j$. For convenience later on, we will state this version without reduced graphs and ε-regular partitions, and instead consider a subgraph of G in which we remove edges between pairs of cluster which do not behave nicely.

Theorem 2.2.12 (Regularity Lemma - coloured version)**.** *For all $d \geq \varepsilon > 0$ and $m_0, \ell \geq 1$ there exists some $M > 0$ such that the following holds. Let G be a graph on $n \geq M$ vertices whose edges are coloured in red and blue and let $\{W_i\}_{i \in [\ell]}$ be a partition of $V(G)$. Then there exists a partition $V(G) = V_0 \cup V_1 \cup \ldots \cup V_m$ and a subgraph H of G with vertex-set $V(G) \setminus V_0$ such that the following holds:*

(i) $m_0 \leq m \leq M$,

(ii) $\{V_i\}_{i\in[m]}$ *refines* $\{W_i \cap V(H)\}_{i\in[\ell]}$,

(iii) $|V_0| \le \varepsilon n$ *and* $|V_1| = \cdots = |V_m| \le \lceil \varepsilon n \rceil$,

(iv) $\deg_H(v) \ge \deg_G(v) - (d+\varepsilon)n$ *for each* $v \in V(G) \setminus V_0$,

(v) $H[V_i]$ *has no edges for* $i \in [m]$,

(vi) *all pairs* (V_i, V_j) *are* ε*-regular with density either 0 or at least* d *in each colour in* H.

2.2.3 Finding Large Regular Cylinders

In Section 3.2 we will use super-regular cylinders as absorbers and only require one such cylinder instead of a partition as in the regularity lemma. This allows us to find larger clusters than guaranteed by the regularity lemma.

Given k disjoint sets $V_1, \ldots, V_k$, we call a cylinder $(U_1, \ldots, U_k)$ *relatively balanced* (w.r.t. $(V_1, \ldots, V_k)$) if there exists some $\gamma > 0$ such that $U_i \subseteq V_i$ with $|U_i| = \lfloor \gamma |V_i| \rfloor$ for every $i \in [k]$. We say that a partition $\mathcal{K}$ of $V_1 \times \cdots \times V_k$ is *cylindrical* if each partition class is of the form $W_1 \times \cdots \times W_k$ (which we associate with the k-cylinder $Z = (W_1, \ldots, W_k)$) with $W_j \subseteq V_j$ for every $j \in [k]$. Finally, we say that $\mathcal{K} = \{Z_1, \ldots, Z_N\}$ is ε-regular if

(i) $\mathcal{K}$ is a cylindrical partition,

(ii) Z_i is a relatively balanced for every $i \in [k]$, and

(iii) all but an ε-fraction of the k-tuples $(v_1, \ldots, v_k) \in V_1 \times \cdots \times V_k$ are in ε-regular cylinders.

For technical reasons, we will allow that some of the V_i are empty. In this case $(A, \emptyset)$ is considered ε-regular for every set A and $\varepsilon > 0$. If G is an r-edge-coloured graph and $i \in [r]$, we say that a cylinder Z is ε-regular in colour i if is ε-regular in G_i (the graph on $V(G)$ with all edges of colour i). We define (ε, d, δ)-super-regular in colour i and ε-regular partitions in colour i similarly.

In [24], Conlon and Fox used the weak regularity lemma of Duke, Lefmann and Rödl [35] to find a reasonably large balanced k-cylinder in a k-partite graph. In order to prove a coloured version of Conlon and Fox's result, we will need the following coloured version of the weak regularity lemma of Duke, Lefmann and Rödl, which further allows for parts of not necessarily equal sizes.

Theorem 2.2.13 (Duke–Lefmann–Rödl [35])**.** *Let $0 < \varepsilon < 1/2$, $k, r \in \mathbb{N}$ and let $\beta = \varepsilon^{rk^2\varepsilon^{-5}}$. Let G be an r-edge-coloured k-partite graph with parts $V_1, \dots, V_k$. Then there exist some $N \leq \beta^{-k}$, sets $R_1 \subseteq V_1, \dots, R_k \subseteq V_k$ with $|R_i| \leq \beta^{-1}$ and a partition $\mathcal{K} = \{Z_1, \dots, Z_N\}$ of $(V_1 \setminus R_1) \times \cdots \times (V_k \setminus R_k)$ so that $\mathcal{K}$ is ε-regular in every colour and $V_i(Z_j) \geq \lfloor \beta |V_i| \rfloor$ for every $i \in [k]$ and $j \in [N]$.*

The proof of Duke, Lefmann and Rödl [35, Proposition 2.1] only has to be adapted slightly to obtain this version, so we will not include it here. As a corollary, we obtain the following lemma, which extends Lemma 2 in [58] to multiple colours.

Lemma 2.2.14. *Let $k, r \geq 2$, $0 < \varepsilon < 1/(rk)$ and $\gamma = \varepsilon^{r^{8rk}\varepsilon^{-5}}$. Then every r-edge-coloured complete graph on $n \geq 1/\gamma$ vertices contains, in one of the colours, a balanced $(\varepsilon, (1/2r)^+)$-super-regular k-cylinder $Z = (U_1, \dots, U_k)$ with parts of size at least γn.*

Proof. Let $k, r \geq 2$, $0 < \varepsilon < 1/(rk)$ and $\gamma = \varepsilon^{r^{8rk}\varepsilon^{-5}}$. Let $n \geq 1/\gamma$ and suppose we are given an r-edge coloured K_n. Let $\tilde{k} = r^{rk}$ and let $V_1, \dots, V_{\tilde{k}} \subseteq [n]$ be disjoint sets of size $\lfloor n/\tilde{k} \rfloor$ and let G be the $\tilde{k}$-partite subgraph of K_n induced by $V_1, \dots, V_{\tilde{k}}$ (inheriting the colouring). Let $\tilde{\varepsilon} = \varepsilon/2$ and $\beta = \tilde{\varepsilon}^{r^{2rk+1}\tilde{\varepsilon}^{-5}}$. We apply Theorem 2.2.13 to get some $N \leq \beta^{-\tilde{k}}$, sets $R_1 \subseteq V_1, \dots, R_{\tilde{k}} \subseteq V_{\tilde{k}}$ each of which of size at most β^{-1} and a partition $\mathcal{K} = \{Z_1, \dots, Z_N\}$ of $(V_1 \setminus R_i) \times \cdots \times (V_{\tilde{k}} \setminus R_{\tilde{k}})$ which is $\tilde{\varepsilon}$-regular in every colour, and with $V_i(Z_j) \geq \lfloor \beta |V_i| \rfloor \geq 2\gamma n$ for every $i \in [\tilde{k}]$ and $j \in [N]$. Note that one of the cylinders (say Z_1) must be $\tilde{\varepsilon}$-regular in every colour and, since $(V_1, \dots, V_k)$ is balanced, so is Z_1. We consider now the complete graph with vertex-set $\{V_1(Z_1), \dots, V_{\tilde{k}}(Z_1)\}$ and colour every edge $V_i(Z_1)V_j(Z_1)$, $1 \leq i < j \leq \tilde{k}$, with a colour $c \in [r]$ so that the density of the pair $(V_i(Z_1), V_j(Z_1))$ in colour c is at least $1/r$. By Ramsey's theorem [40, 100], there is a colour, say 1, and k parts (say $V_1(Z_1), \dots, V_k(Z_1)$) so that the cylinder $(V_1(Z_1), \dots, V_k(Z_1))$ is $(\tilde{\varepsilon}, (1/r)^+, 0)$-regular in colour 1. By Lemma 2.2.3, there is an $(\varepsilon, (1/(2r))^+)$-super-regular balanced subcylinder $\tilde{Z}_1$ with parts of size at least γn. □

The following lemma further guarantees that this remains possible as long as the host-graph has many k-cliques.

Lemma 2.2.15. *Let $k \geq 2$, and let $0 < \varepsilon < 1/2$ and $2k\varepsilon \leq d \leq 1$. Let $\gamma = \varepsilon^{k^2\varepsilon^{-12}}$. Suppose that G is a k-partite graph with parts $V_1, \dots, V_k$ with at least $d|V_1|\cdots|V_k|$ cliques of size k. Then there is some $\gamma' \geq \gamma$ and an $(\varepsilon, (d/2)^+)$-super-regular k-cylinder $Z = (U_1, \dots, U_k)$ in G with $U_i \subseteq V_i$ and $|U_i| = \lfloor \gamma'|V_i| \rfloor$ for every $i \in [k]$.*

Proof. Let $k \geq 2$, and let $d, \varepsilon > 0$ with $2k\varepsilon \leq d \leq 1$. Let $\gamma = \varepsilon^{k^2\varepsilon^{-12}}$ and let G be a k-partite graph with parts $V_1, \dots, V_k$. Let $\tilde{\varepsilon} = \varepsilon/2$ and $\beta = \tilde{\varepsilon}^{k^2\tilde{\varepsilon}^{-5}}$. We may assume that $|V_i| \geq 1/\gamma$ for every $i \in [k]$ (otherwise we set $U_i := \emptyset$ for all $i \in [k]$ with $|V_i| < 1/\gamma$). In particular, we have $|V_i| \geq k/(\tilde{\varepsilon}\beta)$ for all $i \in [k]$.

We apply Theorem 2.2.13 (with $r = 1$) to get some $N \leq \beta^{-k}$, sets $R_1 \subseteq V_1, \dots, R_k \subseteq V_k$, each of which of size at most β^{-1}, and an $\tilde{\varepsilon}$-regular partition $\mathcal{K} = \{Z_1, \dots, Z_N\}$ of $(V_1 \setminus R_1) \times \cdots \times (V_k \setminus R_k)$ with $V_i(Z_j) \geq \lfloor \beta|V_i| \rfloor$ for every $i \in [k]$ and $j \in [N]$.

Note that the number of cliques of size k incident to $R = R_1 \cup \ldots \cup R_k$ is at most

$$\sum_{i=1}^{k} \beta^{-1} \prod_{j \in [k] \setminus \{i\}} |V_j| \leq \tilde{\varepsilon}|V_1| \cdots |V_k|.$$

Furthermore, there are at most $\tilde{\varepsilon}|V_1|\cdots|V_k|$ cliques of size k in cylinders of $\mathcal{K}$ which are not ε-regular. By the pigeonhole principle, there is a cylinder $\tilde{Z}$ among the remaining cylinders which contains at least $(d-2\tilde{\varepsilon})|V_1(\tilde{Z})|\cdots|V_k(\tilde{Z})|$ cliques of size k. In particular, $\tilde{Z}$ is $(\tilde{\varepsilon}, (d-2\tilde{\varepsilon})^+, 0)$-super-regular and relatively balanced with parts of size at least $\lfloor \beta|V_i| \rfloor$. Finally, we apply Lemma 2.2.3 (and possibly delete a single vertex from some parts) to get a relatively balanced $(\varepsilon, (d-(k+2)\tilde{\varepsilon})^+)$-super-regular k-cylinder Z with parts of size at least $\frac{\beta}{2}|V_i| \geq \gamma|V_i|$. This completes the proof since $(k+2)\tilde{\varepsilon} \leq k\varepsilon \leq d/2$. □

2.2.4 The Blow-up Lemma

The blow-up lemma [80, 81] of Komlós, Sárközy and Szemerédi is very useful to find spanning copies of graphs with bounded degree in regular cylinders. It is the key tool that allows us to use super-regular cylinders as absorbers in Section 3.2. We will need the following quantitative version of Sárközy [106].

Theorem 2.2.16 (Blow-up lemma, [106])**.** *There is a constant $C_{BL} > 0$, such that for all $0 < d, \delta \leq 1/2$, all $\Delta, r \in \mathbb{N}$ and all $0 < \varepsilon < (\delta d^\Delta / r)^{C_{BL}}$, the following is true. Let $N \in \mathbb{N}$ and let $V_1, \ldots, V_r$ disjoint sets of size N. Let R be a graph on $\{V_1, \ldots, V_r\}$. We will define two graphs $G_d(N)$ and $G_s(N)$ on $V_1 \cup \ldots \cup V_r$ as follows. The graph $G_d(N)$ is obtained by replacing each edge of R with a copy of $K_{N,N}$ and $G_s(N)$ is obtained by replacing each edge of R by an $(\varepsilon, d^+, \delta)$-super-regular pair. Then, if a graph H with $\Delta(H) \leq \Delta$ is embeddable into $G_d(N)$, it is already embeddable into $G_s(N)$.*

Grinshpun and Sárközy [58, Lemma 6] used of the Hajnal–Szemerédi theorem [66] below in order to turn the blow-up lemma into a tiling result which qualifies ε-balanced super-regular cylinders as absorbers.

Theorem 2.2.17 (Hajnal–Szemerédi)**.** *Every graph G with $\Delta(G) \leq \Delta$ is $(\Delta + 1)$-partite and the parts can be chosen so that their sizes differ by at most 1.*

Theorem 2.2.18 (Grinshpun–Sárközy [58])**.** *There is a constant K, such that for all $0 \leq \delta \leq d \leq 1/2$, $\Delta \in \mathbb{N}$, $k = \Delta + 2$, $0 < \varepsilon \leq (\delta d^\Delta)^K$, and $\mathcal{F} \in \mathcal{F}_\Delta$, the following is true for every $(\varepsilon, d^+, \delta)$-super-regular k-cylinder $Z = (V_1, \ldots, V_k)$.*

(i) If Z is ε-balanced, then its vertices can be partitioned into at most $\Delta+2$ copies of graphs from $\mathcal{F}$.

(ii) If $|V_i| \geq |V_1|$ for all $i = 2, \ldots, k$, then there is a copy of a graph from $\mathcal{F}$ covering V_1 and at most $|V_1|$ vertices of each of $V_2, \ldots, V_k$.

The proof of Theorem 2.2.18 is simple: First we fix a partition of every $F \in \mathcal{F}$ into $k - 1$ parts as guaranteed by Theorem 2.2.17. For (i), we apply Theorem 2.2.16 a total of $\Delta + 2$ times. In the first application, we work in the cylinder induced by all parts but the second-largest and find a copy $F \in \mathcal{F}$ which has the correct size to make the first two parts have equal sizes. We iterate this process until all parts have equal sizes and then apply Theorem 2.2.16 a final time. For (ii), we only need to apply Theorem 2.2.16 once with the graph from $\mathcal{F}$ with $(k-1)|V_1|$ vertices (in which all parts have size $|V_1|$).

A graph G is called *a-arrangeable* for some $a \in \mathbb{N}$ if its vertices can be ordered in such a way that the neighbours to the right of any vertex $v \in V(G)$ have at most a neighbours to the left of v in total. Böttcher, Kohayakawa and Taraz [14] proved an extension of the blow-up lemma to graphs H of bounded arrangeability with $\Delta(H) \leq \sqrt{n}/\log(n)$. As a consequence we obtain the following result similar to Theorem 2.2.18.

Theorem 2.2.19. *For all $d, \delta > 0$ and all $a \in \mathbb{N}$, there is some $\varepsilon > 0$ such that the following is true for all sequences of a-arrangeable graphs $\mathcal{F} = \{F_1, F_2, \dots\}$ such that, for all $n \in \mathbb{N}$, $\Delta(F_n) \leq \sqrt{n}/\log(n)$ and F_n is $(a+2)$-partite and the parts can be chosen so that their sizes differ by at most 1. Let $k = a+3$ and let $Z = (V_1, \dots, V_k)$ be an $(\varepsilon, d^+, \delta)$-super-regular k-cylinder.*

(i) If Z is ε-balanced, then its vertices can be partitioned into at most $a+3$ copies of graphs from $\mathcal{F}$.

(ii) If $|V_i| \geq |V_1|$ for all $i = 2, \dots, k$, then there is a copy of a graph from $\mathcal{F}$ covering V_1 and at most $|V_1|$ vertices of each of $V_2, \dots, V_k$.

The vertices of every ε-balanced, $(\varepsilon, d^+, \delta)$-super-regular $(a+3)$-cylinder Z can be partitioned into at most $a+3$ copies from $\mathcal{F}$.

Note here that we cannot use the Hajnal-Szemeredi theorem here and therefore had to adjust the statement accordingly.

2.3 Hypergraph Regularity

The goal of this section is to prove the following Lemma 2.3.1, which allows us to find in any dense k-graph G, a dense subgraph $H \subseteq G$ in which any two non-isolated $(k-1)$-sets are connected by a short path of a given prescribed length. The lemma follows easily from the hypergraph regularity method, which we shall introduce before proving the lemma. The reader may use Lemma 2.3.1 as a black box if they would like to avoid hypergraph regularity.

Before stating the lemma, we need to introduce some notation. Fix some $k \geq 2$ and a partition $\mathcal{P} = \{V_1, \ldots, V_k\}$ of a set V. We call a tight path in $K^{(k)}(\mathcal{P})$ *positively oriented* if its vertex sequence $(u_1, \ldots, u_m)$ travels through $\mathcal{P}$ in cyclic order, i.e. there is some $j \in [k]$ such that $u_i \in V_{i+j}$ for every $i \in [m]$, where we identify $k + 1 \equiv 1$. In this section, we will only consider positively oriented tight cycles. In particular, given some $e \in K^{(k-1)}(\mathcal{P})$, the ordering of e in a tight path starting at e is uniquely determined.

Lemma 2.3.1. *For every $d > 0$, there are constants $\delta = \delta(d) > 0$ and $\gamma = \gamma(d) > 0$, such that the following is true for every partition $\mathcal{P} = \{V_1, \ldots, V_k\}$ and every $\mathcal{P}$-partite k-graph G of density at least d. There is a $\mathcal{P}$-partite sub-k-graph $H \subseteq G$ of density at least δ such that for every set $S = S_1 \cup \ldots \cup S_k$ with $S_i \subseteq V_i$ and $|S_i| \leq \gamma |V_i|$ and any two $e, f \in K^{(k-1)}(\mathcal{P})$ with positive co-degree, there is a positively oriented tight path of length $\ell \in \{k+2, \ldots, 2k+1\}$ in H which starts at e, ends at f and avoids $S \setminus (e \cup f)$.*

Note that the length of the cycle in Lemma 2.3.1 is uniquely determined by the types of e and f. The *type* of $e \in K^{(k-1)}(\mathcal{P})$, denoted by $\mathrm{tp}(e)$, is the unique index $i \in [k]$ such that $e \cap V_i = \emptyset$. Given two $(k-1)$-sets $e, f \in K^{(k-1)}(\mathcal{P})$, the type of (e, f) is given by $\mathrm{tp}(e, f) := \mathrm{tp}(f) - \mathrm{tp}(e) \pmod k$. It is easy to see that every tight path in $K^{(k)}(\mathcal{P})$ which starts at e and ends at f has length $\ell k + \mathrm{tp}(e, f)$ for some $\ell \geq 0$. In particular, in Lemma 2.3.1, we have $\ell = k + \mathrm{tp}(e, f)$ if $\mathrm{tp}(e, f) \geq 2$ and $\ell = 2k + \mathrm{tp}(e, f)$ otherwise.

We will now introduce the basic concepts of hypergraph regularity in order to state a simple consequence of the strong hypergraph regularity lemma which guarantees a dense regular complex in every large enough k-graph.

For technical reasons, we define a 1-graph on some vertex-set V as a partition of V in what follows. In particular, given a 1-graph $H^{(1)}$ and some $k \geq 2$, $K^{(k)}(H^{(1)})$ is the complete $H^{(1)}$-partite k-graph. We call $\mathcal{H}^{(k)} = (H^{(1)}, \ldots, H^{(k)})$ a k-complex if $H^{(j)}$ is a j-graph for every $j \in [k]$ and $H^{(j)}$ underlies $H^{(j+1)}$, i.e. $H^{(j+1)} \subseteq K^{(j+1)}(H^{(j)})$ for every $j \in [k-1]$. Note that, in particular, $H^{(j)}$ is $H^{(1)}$-partite for every $j \in [k]$. We call $\mathcal{H}^{(k)}$ s-partite if $H^{(1)}$ consists of s parts.

Now, given some j-graph $H^{(j)}$ and some underlying $(j-1)$-graph $H^{(j-1)}$, we define the *density* of $H^{(j)}$ w.r.t. $H^{(j-1)}$ by

$$\mathrm{d}\left(H^{(j)}|H^{(j-1)}\right) = \frac{\left|H^{(j)} \cap K^{(j)}(H^{(j-1)})\right|}{\left|K^{(j)}(H^{(j-1)})\right|}.$$

We are now ready to define regularity.

Definition 2.3.2.

- Let $r, j \in \mathbb{N}$ with $j \geq 2$, $\varepsilon, d_j > 0$, and $H^{(j)}$ be a j-partite j-graph and $H^{(j-1)}$ be an underlying (j-partite) $(j-1)$-graph. We call $H^{(j)}$ (ε, d_j, r)-regular w.r.t. $H^{(j-1)}$ if for all $Q_1^{(j-1)}, \dots, Q_r^{(j-1)} \subseteq E(H^{(j-1)})$, we have

$$\left|\bigcup_{i\in[r]} K^{(j)}\left(Q_i^{(j-1)}\right)\right| \geq \varepsilon \left|K^{(j)}\left(H^{(j-1)}\right)\right|$$
$$\implies \left|\mathrm{d}\left(H^{(j)}\middle|\bigcup_{i\in[r]} Q_i^{(j-1)}\right) - d_j\right| \leq \varepsilon.$$

  We say $(\varepsilon, *, r)$-regular for $(\varepsilon, \mathrm{d}\left(H^{(j)}|H^{(j-1)}\right), r)$-regular and (ε, d)-regular for $(\varepsilon, d, 1)$-regular.

- Let $j, s \in \mathbb{N}$ with $s \geq j \geq 2$, $\varepsilon, d_j > 0$, and $H^{(j)}$ be an s-partite j-graph and $H^{(j-1)}$ be an underlying (s-partite) $(j-1)$-graph. The graph $H^{(j)}$ is called (ε, d_j)-regular w.r.t. $H^{(j-1)}$ if $H^{(j)}[V_1, \dots, V_j]$ is (ε, d_j)-regular w.r.t. $H^{(j-1)}[V_{i_1}, \dots, V_{i_j}]$ for all $1 \leq i_1 < \dots < i_j \leq s$, where $\{V_1, \dots, V_s\}$ is the vertex partition of $V(H^{(j)})$.

- Let $k, r \in \mathbb{N}, \varepsilon, \varepsilon_k, d_2, \dots, d_k > 0$, and $\mathcal{H}^{(k)} = (H_1, \dots, H_k)$ be a k-partite k-complex. We call $\mathcal{H}^{(k)}$ $(d_2, \dots, d_k, \varepsilon, \varepsilon_k, r)$-regular, if $H^{(j)}$ is (ε, d_j)-regular with respect to $H^{(j-1)}$ for every $j = 2, \dots, k-1$ and $H^{(k)}$ is (ε_k, d_k, r)-regular w.r.t. $H^{(k-1)}$.

The following theorem is a direct consequence of the strong hypergraph regularity lemma as stated in [102] (with the exception that we allow for an initial partition of not necessarily equal sizes).

Theorem 2.3.3. *For all integers $k \geq 2$, constants $\varepsilon_k > 0$, and functions ε : $(0,1) \to (0,1)$ and $r : (0,1) \to \mathbb{N}$, there exists some $\delta = \delta(k, \varepsilon, \varepsilon_k, r) > 0$ such* *that the following is true. For every partition $\mathcal{P} = \{V_1, \ldots, V_k\}$ of some set V and every $\mathcal{P}$-partite k-graph $G^{(k)}$ of density $d \geq 2\varepsilon_k$, there are sets $U_i \subseteq V_i$ with $|U_i| \geq \delta |V_i|$ for every $i \in [k]$ and constants $d_2, \ldots, d_{k-1} \geq \delta$ and $d_k \geq d/2$ for which there exists some $(d_2, \ldots, d_k, \varepsilon(\delta), \varepsilon_k, r(\delta))$-regular k-complex $\mathcal{H}^{(k)}$, so that $H^{(1)} = \{U_1, \ldots, U_k\}$.*

We will use the following special case of the extension lemma in [27, Lemma 5] to find short tight paths between almost any two $(k-1)$-sets in a regular complex. Fix a $(d_2, \ldots, d_k, \varepsilon, \varepsilon_k)$-regular complex $H^{(k)} = (\mathcal{P}, H^{(2)}, \ldots, H^{(k)})$, where $\mathcal{P} = \{V_1, \ldots, V_k\}$. Let $H_i^{(k-1)} \subseteq H^{(k-1)}$ denote the edges of type i and note that the dense counting lemma for complexes [27, Lemma 6] implies that

$$\left|H_{i_0}^{(k-1)}\right| = (1 \pm \varepsilon) \prod_{j=2}^{k-1} d_j^{\binom{k-1}{j}} \prod_{i \in [k] \setminus i_0} |V_i| \,.$$

Given some $\beta > 0$, we call a pair $(e, f) \in H_{i_1}^{(k-1)} \times H_{i_2}^{(k-1)}$ β-typical for $\mathcal{H}^{(k)}$ if the number of tight paths of length $\ell := k + \operatorname{tp}(i_1, i_2)$ in $H^{(k)}$ which start at e and end at f is

$$(1 \pm \beta) \prod_{j=2}^{k} d_j^{\ell\binom{k-1}{j-1} - \binom{k-1}{j}} \prod_{i \in \{i_1, \ldots, i_2\}} |V_i| \,,$$

where $\{i_1, \ldots, i_2\}$ is understood in cyclic ordering. Note here that the number of j-sets in a k-uniform tight path of length ℓ which are contained in an edge is $\ell\binom{k-1}{j-1} + \binom{k-1}{j}$. However, $2\binom{k-1}{j}$ of these are contained in e (the first $k-1$ vertices) or f (the last $k-1$ vertices), which are already fixed in our example.

Lemma 2.3.4. *Let $k, r, n_0 \in \mathbb{N}$, $\beta, d_2, \ldots, d_k, \varepsilon, \varepsilon_k > 0$ and suppose that*

$$1/n_0 \ll 1/r, \varepsilon \ll \min\{\varepsilon_k, d_2, \ldots, d_{k-1}\} \leq \varepsilon_k \ll \beta, d_k, 1/k.$$

Then the following is true for all integers $n \geq n_0$, for all indices $i_1, i_2 \in [k]$ and every $(d_2, \ldots, d_k, \varepsilon, \varepsilon_k, r)$-regular complex $\mathcal{H}^{(k)} = \left(H^{(1)}, \ldots, H^{(k)}\right)$ with $|V_i| \geq n_0$ for all $i \in [k]$, where $H^{(1)} = \{V_1, \ldots, V_k\}$. All but at most $\beta \left|H_{i_1}^{(k-1)}\right| \left|H_{i_2}^{(k-1)}\right|$ pairs $(e, f) \in H_{i_1}^{(k-1)} \times H_{i_2}^{(k-1)}$ are β-typical for $\mathcal{H}^{(k)}$.

Combining Theorem 2.3.3 and Lemma 2.3.4 gives Lemma 2.3.1.

Proof-sketch of Lemma 2.3.1. Apply Theorem 2.3.3 with suitable constants and delete all $e \in H^{(k-1)}$ of small co-degree. Let $e \in H^{(k-1)}_{i_1}$ and $f \in H^{(k-1)}_{i_2}$ for some $i_1, i_2 \in [k]$ and define

$$X = \left\{g^{(k-1)} \in H^{(k-1)}_{i_1+1} : e \cup g^{(k-1)} \in H^{(k)}\right\} \text{ and}$$
$$Y = \left\{g^{(k-1)} \in H^{(k-1)}_{i_2-1} : f \cup g^{(k-1)} \in H^{(k)}\right\}.$$

Let $\tilde{X} \subseteq X$ and $\tilde{Y} \subseteq Y$ be the sets of all those edges in X and Y avoiding S. By Lemma 2.3.4 at least one pair in $\tilde{X} \times \tilde{Y}$ must be typical and by a counting argument not all of the promised paths can touch S. □

2.4 Probabilistic Tools

In Chapter 5, we will frequently use the following concentration inequalities for random variables. The first such inequality, Chernoff's inequality [20], deals with the case of binomial random variables.

Lemma 2.4.1 (Chernoff's concentration inequality)**.** *Let $X_1, \ldots, X_n$ be independent random variables, each of which is 1 with probability p and 0 otherwise. Let $X = X_1 + \ldots + X_n$ and let $\lambda := \mathbb{E}[X] = np$. Then, for every $\varepsilon \in (0, 1)$, we have*

$$\mathbb{P}[|X - \lambda| \geq \varepsilon\lambda] \leq \exp\left(-\varepsilon^2\lambda/3\right).$$

Given a subgraph $F \subseteq E(G)$ of G (given by its edge-set), we denote by I_F the indicator random variables which is 1 if F is present in G_p and 0 otherwise. Chernoff's inequality is particularly useful to give sharp bounds on random variables of the form $X = \sum_{F \in \mathcal{F}} I_F$, where $\mathcal{F} \subseteq 2^{E(G)}$ is a collection of edge-disjoint subgraphs of G. However, when $\mathcal{F}$ consists of not-necessarily edge disjoint subgraphs of G, the situation becomes more complicated, since the random variables $\{I_F\}_{F \in \mathcal{F}}$ are not necessarily independent. Janson's inequality [72] provides a bound for the lower tail in this case.

Lemma 2.4.2 (Janson's concentration inequality)**.** *Let G be a graph and $\mathcal{F} \subseteq 2^{E(G)}$ be a collection of subgraphs of G and let $p \in [0,1]$. Let $X = \sum_{F\in\mathcal{F}} I_F$, let $\lambda = \mathbb{E}[X]$ and let*

$$\bar{\Delta} = \sum_{(F,F')\in\mathcal{F}^2:\ F\cap F'\neq\emptyset} \mathbb{E}[I_F I_{F'}].$$

Then, for every $\varepsilon \in (0,1)$, we have

$$\mathbb{P}[X \leq (1-\varepsilon)\lambda] \leq \exp\left(-\frac{\varepsilon^2\lambda^2}{2\bar{\Delta}}\right).$$

If we additionally require a bound for the upper tail, we will use the Kim-Vu inequality [77]. Let $X = \sum_{F\in\mathcal{F}} I_F$ as above. Given an edge $e \in E(G)$, we write t_e for $I_{\{e\}}$. Then, X is as a polynomial with variables t_e, $X = \sum_{F\in\mathcal{F}} \prod_{e\in F} t_e$. Given some $A \subseteq E(G)$, we obtain X_A from X by deleting all summands corresponding to $F \in \mathcal{F}$ which are disjoint from A and replacing every t_e with $e \in A$ by 1. That is,

$$X_A = \sum_{F\in\mathcal{F}:\ F\cap A\neq\emptyset} \prod_{e\in F\setminus A} t_e.$$

In other words, X_A is the the number of $F \in \mathcal{F}$ which are not disjoint from A and are present in G_p.

Lemma 2.4.3 (Kim-Vu polynomial concentration)**.** *For every $k \in \mathbb{N}$, there is a constant $c = c(k) > 0$ such that the following is true. Let G be a graph and $\mathcal{F} \subseteq 2^{E(G)}$ be a collection of subgraphs of G, each with at most k edges. Let $X = \sum_{F\in\mathcal{F}} I_F$ and $\lambda = \mathbb{E}[X]$ as above. For $i \in [k]$, define $E_i := \max\{\mathbb{E}[X_A] :\ A \subseteq E(G),\ |A| = i\}$. Further define $E' := \max_{i\in[k]} E_i$ and $E = \max\{\lambda, E'\}$. Then, for every $\mu > 0$, we have*

$$\mathbb{P}\left[|Y-\lambda| > c\cdot(EE')^{1/2}\mu^k\right] \leq ce(G)^{k-1}e^{-\mu}.$$

2.5 Entropy

In this section, we will explain basic definitions and properties of the entropy function, which will play a central role in Chapter 5. We will be following closely the notes of Galvin [54] and all proofs we do not include here can be found in [54]. Throughout this subsection, we will fix a finite probability space $(\Omega, \mathbb{P})$.

Let $X : \Omega \to S$ be a random variable. Given $x \in S$, we denote $p(x) := \mathbb{P}[X = x]$. We define the *entropy* of X by

$$h(X) := \sum_{x \in S} -p(x) \log p(x).$$

Entropy can be interpreted as the uncertainty of a random variable, or as how much information is gained by revealing X. The following lemma shows that the entropy is maximised when X is uniform. Intuitively, this makes sense as the outcome of a uniform random variable is most uncertain to us. Define the range of X as the set of values which X takes with positive probability, that is $\mathrm{rg}(X) = \{x \in S : p(x) > 0\}$.

Lemma 2.5.1 (Maximality of the uniform)**.** *For every random variable $X : \Omega \to S$, we have $h(X) \leq \log(|\mathrm{rg}(X)|) \leq \log(|S|)$ with equality if and only if $p(x) = 1/|S|$ for all $x \in S$.*

Given random variables $X_i : \Omega \to S_i$ for $i \in [n]$, we denote the entropy of the random vector $(X_1, \ldots, X_n) : \Omega \to S_1 \times \ldots \times S_n$ by $h(X_1, \ldots, X_n) := h((X_1, \ldots, X_n))$. The entropy function has the following subadditivity property.

Lemma 2.5.2 (Subadditivity)**.** *Given random variables $X_i : \Omega \to S_i$, $i \in [n]$, we have*

$$h(X_1, \ldots, X_n) \leq \sum_{i=1}^{n} h(X_i)$$

with equality if and only if the X_i are jointly independent.

Intuitively, this means that revealing a random vector cannot give us more information than revealing each component separately. If $E \subseteq \Omega$ is an event which has positive probability, we define the conditional entropy given the event as

$$h(X|E) := \sum_{x \in S} -p(x|E) \log p(x|E),$$

where $p(x|E) = \mathbb{P}[\{X = x\}|E]$. Note that $h(X|E)$ is the entropy of the random variable $Z = X|E : E \to S$ whose distribution is given by $\mathbb{P}[Z = x] = \mathbb{P}[X = x|E]$.

Given two random variables $X : \Omega \to S_X$ and $Y : \Omega \to S_Y$, the conditional entropy of X given Y is defined as

$$h(X|Y) := \mathbb{E}_Y[h(X|Y = y)] = \sum_{y \in S_Y} p(y) h(X|Y = y) \tag{2.5.1}$$

$$= \sum_{\omega \in \Omega} \mathbb{P}[\omega]\, h(X|Y = Y(\omega)), \tag{2.5.2}$$

where $p(y) = \mathbb{P}[\{Y = y\}]$. As conditioning on an event or another random variable only gives us more information, we have the following inequality. Here, we say a random variable $Y : \Omega \to S_Y$ determines another random variable $Y' : \Omega \to S_{Y'}$ if the outcome of Y' is completely determined by Y. For example if Y is the outcome of rolling a regular six-sided die and Y' is 1 if this outcome is even, and 0 otherwise, then Y determines Y'.

Lemma 2.5.3 (Dropping conditioning)**.** *Given random variables $X : \Omega \to S_X$ and $Y : \Omega \to S_Y$, and an event $E \subseteq \Omega$ which has positive probability, we have*

$$h(X|Y) \leq h(X) \quad \text{and} \quad h(X) \geq \mathbb{P}[E] \cdot h(X|E).$$

Furthermore, if $Y' : \Omega \to S_{Y'}$ is another random variable and Y determines Y', then

$$h(X|Y) \leq h(X|Y').$$

The following chain rule strengthens Lemma 2.5.2.

Lemma 2.5.4 (Chain rule)**.** *Given random variables $X : \Omega \to S_X$ and $Y : \Omega \to S_Y$, we have*

$$h(X, Y) = h(X) + h(Y|X)$$

and more generally, for random variables $X_i : \Omega \to S_i$, $i \in [n]$, we have

$$h(X_1, \ldots, X_n) = \sum_{i=1}^{n} h(X_i|X_1, \ldots, X_{i-1}).$$

If $X : \Omega \to S_X$ and $Y : \Omega \to S_Y$ are random variables, we say that X *determines* Y if there is a function $f : S_X \to S_Y$ such that $Y(\omega) = f(X(\omega))$ for all $\omega \in \Omega$. If X

determines Y, then Y does not bring any new information after revealing X. This is formalised in the following lemma.

Lemma 2.5.5 (Redundancy)**.** *If $X : \Omega \to S_X$ and $Y : \Omega \to S_Y$ are random variables and X determines Y, then*

$$h(X) = h(X, Y).$$

Lemmas 2.5.1, 2.5.2 and 2.5.4 have the following conditional versions. Given a random variable $X : \Omega \to S_X$ and an event $E \subseteq \Omega$, we define the conditional range of X given E by $\mathrm{rg}(X|E) = \{x \in S_X : p(x|E) > 0\}$.

Lemma 2.5.6 (Conditional maximality of the uniform)**.** *For every random variable $X : \Omega \to S$, we have*

$$h(X|E) \leq \log\left(|\mathrm{rg}(X|E)|\right).$$

Lemma 2.5.7 (Conditional subadditivity)**.** *Given random variables $X_i : \Omega \to S_i$, $i \in [n]$, and $Y : \Omega \to S_Y$, we have*

$$h(X_1, \ldots, X_n|Y) \leq \sum_{i=1}^{n} h(X_i|Y)$$

with equality if and only if the X_i are jointly independent conditioned on Y.

Lemma 2.5.8 (Conditional chain rule)**.** *Given random variables $X_i : \Omega \to S_i$, $i \in [n]$, and $Y : \Omega \to S_Y$, we have*

$$h(X_1, \ldots, X_n|Y) = \sum_{i=1}^{n} h(X_i|X_1, \ldots, X_{i-1}, Y).$$

The following lemma will play an essential role in the main proof. It states that if a random variable has almost maximal entropy, then it must be close to uniform. This can be seen as a stability result for Lemma 2.5.1. Given a fucntion $\zeta : S \to [0, \infty)$ and a random variable $X : \Omega \to S$, we say that the distribution of X follows ζ if

$$\mathbb{P}\left[X = x\right] = \frac{\zeta(x)}{\sum_{y \in S} \zeta(y)}$$

for all $x \in S$. We shall call ζ a weight function.

Lemma 2.5.9. *For all $\beta > 0$, there is some $\beta' > 0$ such that the following is true for every finite set S with weight function $\zeta : S \to [0, \infty)$ and every random variable $X : \Omega \to S$ whose distribution follows ζ. If $h(X) \geq \log S - \beta'$, then there is some $a \geq 0$, such that for $J = \zeta^{-1}([(1-\beta)a, (1+\beta)a])$, we have*

$$|J| \geq (1-\beta)|S| \quad \text{and} \quad \sum_{y \in J} \zeta(y) \geq (1-\beta) \sum_{y \in S} \zeta(y). \tag{2.5.3}$$

Proof. We may assume that $\beta < 1/10$. Furthermore, it suffices to prove the lemma for all functions $\zeta : S \to [0, \infty)$ with $\sum_{y \in S} \zeta(y) = 1$ (to see this, divide all necessary parameters by $\sum_{y \in S} \zeta(y)$). Let now $D := \{\underline{\zeta} \in [0,1]^S : \sum_{y \in S} \zeta_y = 1\}$ be the set of all such functions (seen as vectors) and define $h : D \to [0, \infty)$ by

$$h(\underline{\zeta}) = h((\zeta_y)_{y \in S}) := \sum_{y \in S} -\zeta_y \log(\zeta_y).$$

Note that $h(\underline{\zeta}) = h(X)$ if X is a random variable whose distribution follows $\underline{\zeta}$. We make the following easy but important observation.

Observation 2.5.10. *If $x \neq x' \in S$ and $\underline{\zeta}, \underline{\tilde{\zeta}} \in D$ with $\zeta_y = \tilde{\zeta}_y$ for all $y \in S \setminus \{x, x'\}$ and $\zeta_x \leq \tilde{\zeta}_x \leq \tilde{\zeta}_{x'} \leq \zeta_{x'}$, then we have $h(\underline{\zeta}) \leq h(\underline{\tilde{\zeta}})$.*

In words, if we push two parameters closer together, this will increase the entropy. To prove Observation 2.5.10, fix all but two parameters and take the first derivative with respect to one of the remaining parameters (noting that the last one is determined from that).

Fix now some $\underline{\zeta} \in D$ and let $a := \frac{1}{|S|}$ be its average. Let $\beta > 0$ be given and let $\beta' = \beta^3/1000$. Let $X : \Omega \to S$ be a random variable whose distribution follows $\underline{\zeta}$. Assume that $h(X) \geq \log(|S|) - \beta'$. Let $J^+ = \{y \in S : \zeta_y > (1 + \beta/4)a\}$, $J^- = \{y \in S : \zeta_y < (1 - \beta/4)a\}$ and $J = \zeta^{-1}([(1-\beta)a, (1+\beta)a])$. Note that $|J| \geq |S| - (|J^+| + |J^-|)$.

Claim 2.1. *We have $|J^+| \leq \beta/4|S|$.*

Proof. Choose $\gamma \leq \beta/4$ so that $\gamma|S| = \lfloor \beta/4|S| \rfloor$. Assume for contradiction that

$|J^+| > \gamma|S|$ and let $\tilde{J}^+ \subseteq J^+$ be a set of size exactly $\gamma|S|$. Define $\underline{\zeta}^+$ by

$$\zeta_y^+ = \begin{cases} (1+\gamma)a & \text{if } s \in \tilde{J}^+ \\ (1-\xi)a & \text{if } s \notin \tilde{J}^+, \end{cases}$$

where $\xi = \frac{\gamma^2}{1-\gamma}$ is chosen so that $\sum_{y \in S} \zeta_y^+ = 1$. Let $X^+ : \Omega \to S$ be a random variable whose distribution follows $\underline{\zeta}^+$. Let $Y = 1$ if $X^+ \in \tilde{J}^+$ and 0 otherwise. It follows from (multiple applications of) Observation 2.5.10 that

$$h(X) \le h(X^+, Y) = h(X^+|Y=1)\mathbb{P}[Y=1] + h(X^+|Y=0)\mathbb{P}[Y=0] + h(Y),$$

where we used Lemma 2.5.5 and the chain rule (Lemma 2.5.4) and the definition of conditional entropy. Note that $\mathbb{P}[Y=1] = \gamma(1+\gamma)$ and

$$h(Y) = -\gamma(1+\gamma)\log(\gamma(1+\gamma)) - (1-\gamma(1+\gamma))\log(1-(\gamma(1+\gamma))).$$

Therefore (and using Lemma 2.5.6), we get

$$\begin{aligned} h(X) &\le \log(\gamma|S|)\,\gamma(1+\gamma) + \log((1-\gamma)|S|)\,(1-\gamma(1+\gamma)) + h(Y) \\ &= \log(|S|) + \log(\gamma)\gamma(1+\gamma) + \log(1-\gamma)(1-\gamma(1+\gamma)) + h(Y) \\ &= \log(|S|) + \gamma(1+\gamma)\left(\log(\gamma) - \log(\gamma(1+\gamma))\right) \\ &\qquad\qquad + (1-\gamma(1+\gamma))\left(\log(1-\gamma) - \log(1-\gamma(1+\gamma))\right) \\ &= \log(|S|) - (\gamma+\gamma^2)\log(1+\gamma) + (1-\gamma-\gamma^2)\log\left(\frac{1-\gamma}{1-\gamma-\gamma^2}\right) \\ &\overset{(*)}{\le} \log(|S|) - \gamma^2(1+\gamma)(1-\gamma/2) + (1-\gamma-\gamma^2)\frac{\gamma^2}{1-\gamma-\gamma^2} \\ &= \log(|S|) - \gamma^2 - \gamma^3/2 + \gamma^4 + \gamma^2 \\ &\le \log(|S|) - \gamma^3/4 \\ &\le \log(|S|) - \beta', \end{aligned}$$

a contradiction. Here we used the approximation $x - x^2/2 \le \log(1+x) \le x$ for all $x \in (0,1)$ for $(*)$, which gives $\log(1+\gamma) \ge \gamma(1-\gamma/2)$ and $\log\left(\frac{1-\gamma}{1-\gamma-\gamma^2}\right) =$

$\log\left(1+\frac{\gamma^2}{1-\gamma-\gamma^2}\right) \leq \frac{\gamma^2}{1-\gamma-\gamma^2}$. ■

Similarly, we can show that $|J^-| \leq \beta/4 \cdot |S|$ and we conclude that $|J| \geq |S| - (|J^+| + |J^-|) \geq (1-\beta)|S|$. Furthermore, we have

$$\begin{aligned} \sum\nolimits_{y \in J} \zeta(y) &\geq (1-\beta/2)|S|(1-\beta/4)a \\ &\geq (1-\beta) \\ &= (1-\beta) \sum_{y \in S} \zeta(y). \end{aligned}$$

This completes the proof. □

3

Monochromatic Graph Tiling Problems

3.1 Tiling Coloured Hypergraphs with Tight Cycles

3.1.1 Overview

In this section, we are going to prove the following two results mentioned in the introduction.

Theorem 1.3.10 (Bustamante–Corsten–Frankl–Pokrovskiy–Skokan [18])**.** *For every $k, r, \alpha \in \mathbb{N}$, there is some $c = c(k, r, \alpha)$ such that the vertices of every r-edge-coloured k-graph G with independence number $\alpha(G) \leq \alpha$ can be partitioned into at most c monochromatic tight cycles.*

Theorem 1.3.14 (Bustamante–Corsten–Frankl–Pokrovskiy–Skokan [18])**.** *For every $k, r, p, \alpha \in \mathbb{N}$, there is some $c = c(k, r, p, \alpha)$ such that the vertices of every r-edge-coloured k-graph G with $\alpha(G) \leq \alpha$ can be partitioned into at most c monochromatic p-th powers of tight cycles.*

Since Theorem 1.3.14 follows easily from Theorem 1.3.10, we present its short proof here.

Proof of Theorem 1.3.14. Let $f(k, r, \alpha)$ be the smallest c for which Theorem 1.3.10 is true and let $g(k, r, p, \alpha)$ be the smallest c for which Theorem 1.3.14 is true. We will show that $g(k, r, p, \alpha) \leq f(k + p - 1, r, \tilde{\alpha})$, where $\tilde{\alpha} =$

$R^{(k)}_{r+1}(k+p-1,\dots,k+p-1,\alpha+1)-1$. (Recall that $R^{(k)}_r(s_1,\dots,s_r)$ is the r-colour Ramsey number for k-graphs, i.e. the smallest positive integer n, so that in every r-colouring of the complete k-graph on n vertices, there is some $i \in [r]$ and s_i distinct vertices which induce a monochromatic clique in colour i.) Suppose now we are given an r-edge-coloured k-graph G with $\alpha(G) \leq \alpha$. Define a $(k+p-1)$-graph H on the same vertex-set whose edges are the monochromatic cliques of size $k+p-1$ in G. By construction we have $\alpha(H) \leq \tilde{\alpha}$ and thus, by Theorem 1.3.10, there are at most $f(k+p-1, r, \tilde{\alpha})$ monochromatic tight cycles partitioning $V(H)$. To conclude, note that a tight cycle in H corresponds to a p-th power of a tight cycle in G. □

The proof of Theorem 1.3.10 combines the absorption method and the hypergraph regularity method. If the host k-graph G is complete, the proof of Theorem 1.3.10 can be summarised as follows.

First, we find a monochromatic k-graph $H_0 \subseteq G$ with the following special property: There is some $B \subseteq V(H_0)$, so that for every $B' \subseteq B$ there is a tight cycle in H_0 with vertices $V(H_0) \setminus B'$. This is explained in Section 3.1.2. We then greedily remove vertex-disjoint monochromatic tight cycles from $V(G) \setminus V(H_0)$ until the set of leftover vertices R is very small in comparison to B. Finally, in Section 3.1.3, we show that the leftover vertices can be absorbed by H_0. More precisely, we show that there are constantly many vertex-disjoint tight cycles with vertices in $R \cup B$ which cover all of R.

In order to prove the main theorem for host k-graphs with bounded independence number, we need to iterate the above process a few times. Here the main difficulty is to show that the iteration process stops after constantly many steps. This will be shown in Section 3.1.4.

3.1.2 Absorption Method for Hypergraphs

The idea of the absorption method is to first cover almost every vertex by vertex-disjoint monochromatic tight cycles and then absorb the leftover using a suitable absorption lemma.

Lemma 3.1.1. *For all $k, r, \alpha \in \mathbb{N}$ and every $\gamma > 0$, there is some $c = c(k, r, \alpha, \gamma)$ so that the following is true for every r-coloured k-graph G on n vertices with*

$\alpha(G) \leq \alpha$. There is a collection of at most c vertex-disjoint monochromatic tight cycles whose vertices cover all but at most γn vertices.

Definition 3.1.2. Let G be a hypergraph, χ be a colouring of $E(G)$ and $A, B \subseteq V(G)$ disjoint subsets. Then A is called an *absorber* for B if there is a monochromatic tight cycle with vertices $A \cup B'$ for every $B' \subseteq B$.

Lemma 3.1.3. *For every $k, r, \alpha \in \mathbb{N}$, there is some $\beta = \beta(k, r, \alpha) > 0$ such that the following is true for every k-graph G with $\alpha(G) \leq \alpha$. In every r-colouring of $E(G)$ there are disjoint sets $A, B \subseteq V(G)$ with $|B| \geq \beta|V(G)|$ such that A absorbs B.*

The following hypergraph will function as our absorber. A very similar hypergraph was used by Gyárfás and Sárközy to absorb loose cycles [63, 64]. See Figure 3.1 for an example.

Definition 3.1.4. The (k-uniform) crown of order t, $T_t^{(k)}$, is a tight cycle with $n = t(k-1)$ vertices $v_0, \dots, v_{n-1}$ (the base) and additional vertices $u_0, \dots, u_{t-1}$ (the rim). Furthermore, for each $i = 0, \dots, t-1$, we add the k edges $\{u_i, v_{(k-1)i+j}, \dots, v_{(k-1)i+j+k-2}\}$, $j = 0, \dots, k-1$.

It is easy to see that the base of a crown is an absorber for the rim. To prove Lemma 3.1.3, we therefore only need to show that we can always find monochromatic crowns of linear size. Both this and Lemma 3.1.1 are consequences of the following theorem of Cooley, Fountoulakis, Kühn, and Osthus [27] (see also [71] and [25]).

Theorem 3.1.5. *For every $r, k, \Delta \in \mathbb{N}$, there is some $C = C(r, k, \Delta) > 0$ so that the following is true for all k-graphs $H_1, \dots, H_r$ with at most n vertices and maximum degree at most Δ, and every $N \geq Cn$. In every edge-colouring of $K_N^{(k)}$ with colours $c_1, \dots, c_r$, there is some $i \in [r]$ for which there is a c_i-monochromatic copy of H_i.*

Proof of Lemma 3.1.3. Suppose k, r, α and G are given as in the theorem and that $E(G)$ is coloured with r colours. Let $N = |V(G)|$, $\Delta := \max\left\{2k, \binom{\alpha}{k-1}\right\}$ and $c = 1/((k-1)C)$ where $C = C(r+1, k, \Delta)$ is given by Theorem 3.1.5. Furthermore, let $n = \left|V(T_{cN}^{(k)})\right| = N/C$. Consider now the $(r+1)$-colouring of $E\left(K_N^{(k)}\right)$ in which every edge in $E(G)$ receives the same colour as in G and every other edge receives colour $r+1$. Let $H_{r+1} = K_{\alpha+1}^{(k)}$ and $H_i = T_{cN}^{(k)}$ for all $i \in [t]$, and note that $\Delta(H_i) \leq \Delta$

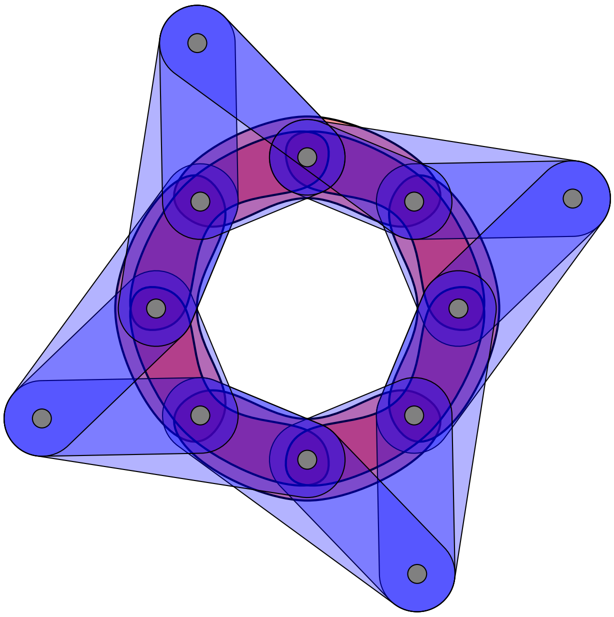


Figure 3.1: A 3-uniform crown of order 4. The edges of the tight cycle are red and the remaining edges are blue.

for all $i \in [r+1]$. By choice of Δ, there is no monochromatic H_{r+1} in colour $r+1$ and hence, since $N \geq Cn$, there is a monochromatic copy of H_i for some $i \in [r]$. Therefore, there is a monochromatic crown of size $c|V(G)|$ and its base is an absorber for its rim. □

Proof of Lemma 3.1.1. Applying Theorem 3.1.5 with $r+1$ colours, uniformity k, $\Delta = \max\{k, \binom{\alpha}{k-1}\}$, and $H_1 = \ldots = H_r$ being tight cycles on $n/(C_{Thm\ 3.1.5}(r+1,k,\Delta))$ vertices and $H_{r+1} = K^{(k)}_{\alpha+1}$ gives the following. There exist some $\varepsilon = \varepsilon(r,k,\alpha)$ so that in every r-coloured k-graph G on n vertices with $\alpha(G) \leq \alpha$, there is a monochromatic tight cycle on at least εn vertices.[1] By iterating this process i times, we find i vertex-disjoint monochromatic tight cycles covering all but $(1-\varepsilon)^i n$ vertices. This finishes the proof, since $(1-\varepsilon)^i \to 0$ as $i \to \infty$. □

3.1.3 Absorption Lemma

In this section we prove a suitable absorption lemma for our approach.

[1]Here, we treat non-edges as colour $r+1$ again.

Lemma 3.1.6. *For every $\varepsilon > 0$ and $k, r \in \mathbb{N}$, there is some $\gamma = \gamma(k, r, \varepsilon) > 0$ and some $c = c(k, r, \varepsilon)$ such that the following is true. Let H be a k-partite k-graph with parts $B_1, \ldots, B_k$ such that $|B_1| \geq \ldots \geq |B_{k-1}| \geq |B_k|/\gamma$ and $|\operatorname{Lk}(v; B_1, \ldots, B_{k-1})| \geq \varepsilon |B_1| \cdots |B_{k-1}|$ for every $v \in B_k$. Then, in every r-colouring of $E(H)$, there are c vertex-disjoint monochromatic tight cycles covering B_k.*

Note that it is enough to cover all but a bounded number of vertices, since we allow single vertices as tight cycles. We will make use of this in the proof and frequently remove few vertices.

We will use the following theorem of Pósa [97].

Theorem 3.1.7 (Pósa)**.** *In every graph G, there is a collection of at most $\alpha(G)$ cycles whose vertices partition $V(G)$.*

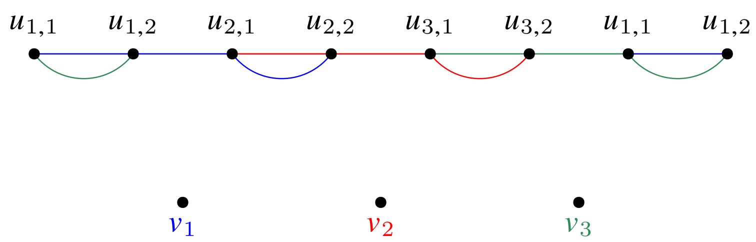
We further need the following simple but quite technical lemma, which states that, given a ground set X and a collection $\mathcal{F}$ of subsets of X of linear size, we can group almost all of these subsets into groups of size 4 which have a large common intersection. We will apply this lemma when X is the edge-set of a hypergraph G and $\mathcal{F}$ is a collection of subgraphs of G.

Lemma 3.1.8. *For every $\varepsilon > 0$ there are $\delta = \delta(\varepsilon) > 0$ and $C = C(\varepsilon) > 0$ such that the following is true for every $m \in \mathbb{N}$. Let X be set of size m and $\mathcal{F} \subseteq 2^X$ be a family of subsets such that $|F| \geq \varepsilon m$ for every $F \in \mathcal{F}$. Then there is some $\mathcal{G} \subseteq \mathcal{F}$ of size $|\mathcal{G}| \leq C$ and a partition $\mathcal{P}$ of $\mathcal{F} \setminus \mathcal{G}$ into sets of size 4 such that $\left|\bigcap_{i=1}^4 B_i\right| \geq \delta m$ for every $\{B_1, B_2, B_3, B_4\} \in \mathcal{P}$.*

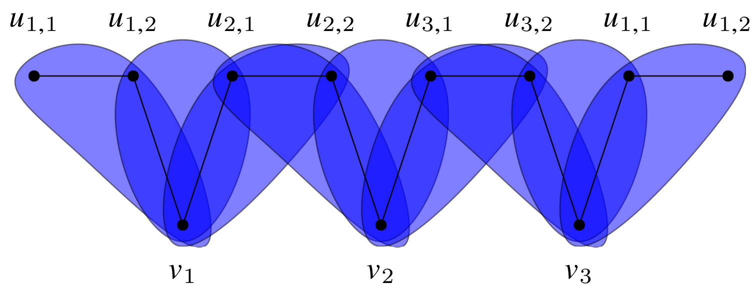
We will prove the lemma with $\delta(\varepsilon) = e^4/2^6$ and $C(\varepsilon) = 8/\varepsilon^2 + 2/\varepsilon$.

Proof. Define a graph G on $\mathcal{F}$ by $\{F_1, F_2\} \in E(G)$ if and only if $|F_1 \cap F_2| \geq (\varepsilon/2)^2 m$. We claim that $\alpha(G) \leq 2/\varepsilon$. Suppose for contradiction that there is an independent set I of size $2/\varepsilon + 1$. Then we have $|F_0 \setminus \bigcup_{F \in I \setminus \{F_0\}} F| \geq \varepsilon m/2$ for every $F_0 \in I$ and hence $|\bigcup_{F \in I} F| > m$, a contradiction.

Since every graph has a matching of size at least $\nu(G) - \alpha(G)$, we find a matching $\mathcal{P}_1$ in G of all but at most $2/\varepsilon$ vertices of G (i.e. $F \in \mathcal{F}$). Let $\mathcal{G}_1 = \mathcal{F} \setminus V(\mathcal{P}_1)$ and note that $\mathcal{P}_1$ is a partition of $\mathcal{F} \setminus \mathcal{G}_1$ into sets of size 2. Let $\mathcal{F}_1 = \{F_1 \cap F_2 : \{F_1, F_2\} \in \mathcal{P}_1\}$ and iterate the process once more. □



(a) Link graphs



(b) Tight Cycle

Figure 3.2: A sketch of Observation 3.1.9 for $k = t = 3$. Figure (a) shows the link graphs of v_1 (blue), v_2 (red) and v_3 (green). The colours are for demonstration purposes only and are not related to the given edge-colouring. Figure (b) shows the resulting tight cycle. In both figures, we identify the ends ($u_{1,1}$ and $u_{1,2}$) to simplify the drawing.

Proof of Lemma 3.1.6. By choosing γ sufficiently small, we may assume that $|B_1|, \ldots, |B_{k-1}|$ are sufficiently large for the following arguments. First we claim that it suffices to prove the lemma for $r = 1$. Indeed, partition $B_k = B_{k,1} \cup \ldots \cup B_{k,r}$ so that for each $i \in [r]$ and $v \in B_{k,i}$, we have $|\operatorname{Lk}_i(v; B_1, \ldots, B_{k-1})| \geq \varepsilon/r \cdot |B_1| \cdots |B_{k-1}|$ and delete all edges containing v whose colour is not i. (Here we denote by $\operatorname{Lk}_i(\cdot)$ the link graph with respect to G_i, the graph with all edges of colour i.) Next, for each $j \in [k-1]$, partition $B_j = B_{j,1} \cup \ldots \cup B_{j,r}$ into sets of equal sizes so that $|\operatorname{Lk}_i(v; B_{1,i}, \ldots, B_{k-1,i})| \geq \varepsilon/(2r) \cdot |B_{1,i}| \cdots |B_{k-1,i}|$. Such a partition can be found for example by choosing one uniformly at random and applying the probabilistic method). Finally, we can apply the one-colour result (with $\varepsilon' = \varepsilon/(2r)$) for each $i \in [r]$.

Fix $\varepsilon > 0$, $k \geq 2$ and a k-partite k-graph H with parts $B_1, \ldots, B_k$ as in the statement of the lemma. Choose constants $\gamma, \delta_1, \delta_2, \delta_3 > 0$ so that $0 < \gamma \ll \delta_3 \ll \delta_2 \ll \delta_1 \ll \varepsilon, 1/k$. We begin with a simple but important observation.

Observation 3.1.9. *Let $v_1, \ldots, v_t \in B_k$ be distinct vertices and C be a tight cycle in $K^{(k-1)}(B_1, \ldots, B_{k-1})$ with vertex-sequence $(u_{1,1}, \ldots, u_{1,k-1}, \ldots, u_{t,1}, \ldots, u_{t,k-1})$. Denote by $e_{s,i}$ the edge in C starting at $u_{s,i}$ and suppose that*

(i) $e_{s,i} \in \mathrm{Lk}(v_s; B_1, \ldots, B_{k-1})$ *for every $s \in [t]$ and every $i \in [k-1]$ and*

(ii) $e_{s,1} \in \mathrm{Lk}(v_{s-1}; B_1, \ldots, B_{k-1})$ *for every $s \in [t]$ (here $v_0 := v_t$).*

Then, $(u_{1,1}, \ldots, u_{1,k-1}, v_1, \ldots, u_{t,1}, \ldots, u_{t,k-1}, v_t)$ is the vertex-sequence of a tight cycle in H.

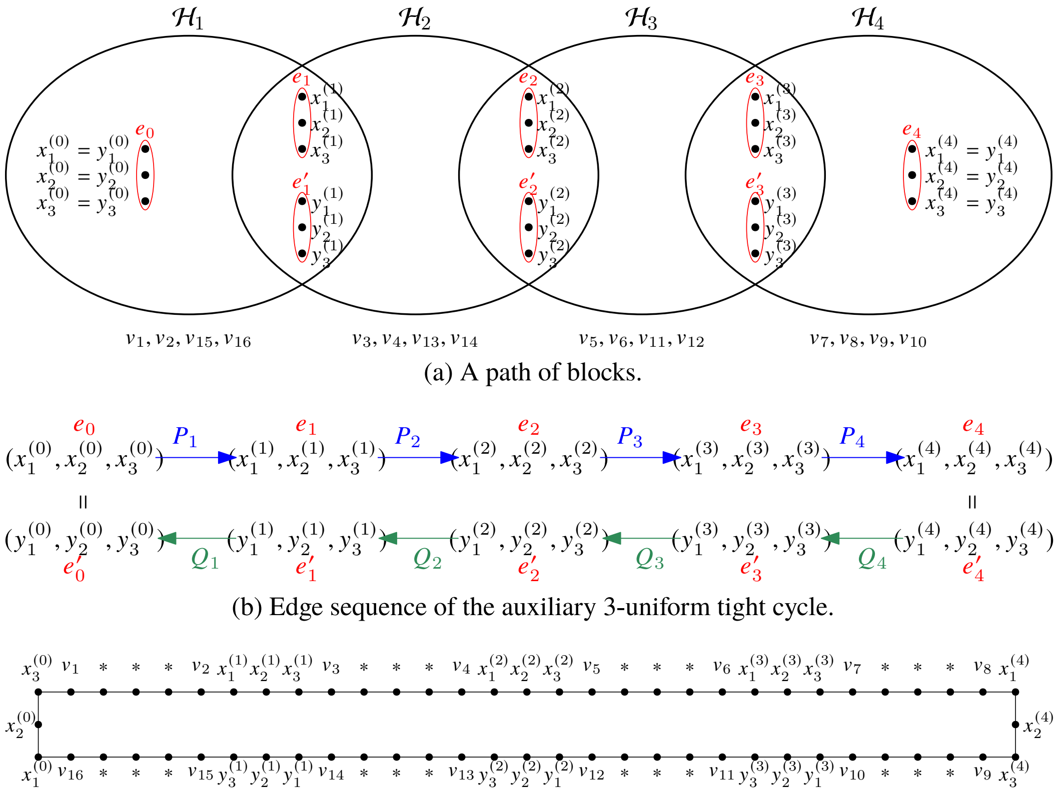
The proof of Observation 3.1.9 follows readily from the definition of the link graphs. See Figure 3.2 for an overview. We will now proceed in three steps. For simplicity, we write $H_v := \mathrm{Lk}_H(v; B_1, \ldots, B_{k-1})$ for $v \in B_k$.

Step 1 (Divide into blocks). By Lemma 3.1.8, there is some $C = C(\varepsilon) \in \mathbb{N}$ and a partition $\mathcal{P}$ of all but C graphs from $\{H_v : v \in B_k\}$ into *blocks* $\mathcal{H}$ of size 4 with $e(\mathcal{H}) := |\bigcap_{H \in \mathcal{H}} E(H)| \geq \delta_1 |B_1| \cdots |B_{k-1}|$ for every $\mathcal{H} \in \mathcal{P}$. Remove the C leftover vertices from B_k.

Think of every block $\mathcal{H}$ now as a $(k-1)$-graph with edges $E(\mathcal{H}) := \bigcap_{H \in \mathcal{H}} E(H)$. By applying Lemma 2.3.1 (with $k-1$ instead of k), for each $\mathcal{H} \in \mathcal{P}$, we find a subgraph $\mathcal{H}' \subseteq \mathcal{H}$ such that $e(\mathcal{H}') \geq \delta_2 |B_1| \cdots |B_{k-1}|$ with the same property as in Lemma 2.3.1. By deleting all the edges of $\mathcal{H} \setminus \mathcal{H}'$ we may assume that $\mathcal{H}$ itself has this property.

Step 2 (Cover blocks by paths). Define an auxiliary graph G with $V(G) = \mathcal{P}$ and $\{\mathcal{H}_1, \mathcal{H}_2\} \in E(G)$ if and only if $e(\mathcal{H}_1 \cap \mathcal{H}_2) \geq \delta_3 |B_1| \cdots |B_{k-1}|$. Similarly as in the proof of Lemma 3.1.8, we conclude that $\alpha(G) \leq 2/\delta_2$, and hence $V(G)$ can be covered by $2/\delta_2$ vertex-disjoint paths using Theorem 3.1.7.

Step 3 (Lift to tight cycles). This step is the crucial part of the argument. To make it easier to follow the proof, Figure 3.3 provides an example for $k = t = 4$.



(a) A path of blocks.

(b) Edge sequence of the auxiliary 3-uniform tight cycle.

(c) Vertex sequence of the resulting tight cycle.

Figure 3.3: Finding a tight cycle in a path of blocks when $k = t = 4$. In Figure (c), $*$ represents an internal vertex of a some path P_i or Q_i.

We will show now how to find in each path of blocks an auxiliary tight cycle in $K^{(k-1)}(B_1, \ldots, B_{k-1})$ of the desired form to apply Observation 3.1.9. Let $\mathbf{P} = (\mathcal{H}_1, \ldots, \mathcal{H}_t)$ be one of the paths. Choose disjoint edges $e_0 = \left\{x_1^{(0)}, \ldots, x_{k-1}^{(0)}\right\} \in E(\mathcal{H}_1)$ and $e_t = \left\{x_1^{(t)}, \ldots, x_{k-1}^{(t)}\right\} \in E(\mathcal{H}_t)$. For each $s \in [t-1]$, further choose two edges $e_s = \left\{x_1^{(s)}, \ldots, x_{k-1}^{(s)}\right\} \in E(\mathcal{H}_s) \cap E(\mathcal{H}_{s+1})$ and $e'_s = \left\{y_1^{(s)}, \ldots, y_{k-1}^{(s)}\right\} \in E(\mathcal{H}_s) \cap E(\mathcal{H}_{s+1})$ so that all chosen edges are pairwise disjoint. We identify $x_i^{(0)} = y_i^{(0)}$ and $x_i^{(s)} = y_i^{(s)}$ for every $i \in [k-1]$, and $e_0 = e'_0$ and $e_t = e'_t$. Assume without loss of generality, that $x_i^{(s)} \in B_i$ for every $i \in [k-1]$ and all $s = 0, \ldots, t$.

By construction, every block $\mathcal{H}$ has the property guaranteed in Lemma 2.3.1. Therefore, for every $s \in [t]$, there is a tight path $P_s \subseteq \mathcal{H}_s$ of length $2k-3$ which starts at $(x_2^{(s-1)}, \ldots, x_{k-1}^{(s-1)})$, ends at $(x_1^{(s)}, \ldots, x_{k-2}^{(s)})$ and (internally) avoids all previously used vertices. Indeed, we can choose these paths one at a time, applying Lemma 2.3.1 with $S_j \subseteq V_j$ being the set of previously used vertices in V_j. (Since every tight cycle in G uses the same number of vertices from each part, we have $|S_j| \leq |V_k| \leq \gamma|V_j|$ for every $j \in [k-1]$.) Similarly, for every $s \in [t]$, there is a tight path $Q_s \subseteq \mathcal{H}_s$ of length $2k-3$ which starts at and $(y_1^{(s)}, \ldots, y_{k-2}^{(s)})$, ends at $(y_2^{(s-1)}, \ldots, y_{k-1}^{(s-1)})$ and (internally) avoids all previously used vertices.

Let $U \subseteq B_k$ be the set of vertices v for which $H_v \in \mathcal{H}_i$ for some $i \in [t]$. To finish the proof, we want to apply Observation 3.1.9 to cover U. Label $U = \{v_1, \ldots, v_{4t}\}$ so that $H_{v_{2i+1}}, H_{v_{2i+2}}, H_{v_{4t-2i}}, H_{v_{4t-2i-1}} \in \mathcal{H}_i$ for all $i = 0, \ldots, t-1$. Consider now the tight cycle C in $K^{k-1}(B_1, \ldots, B_{k-1})$ with edge sequence

$$e'_0 = e_0, P_1, e_1, P_2, e_2, \ldots, P_t, e_t = e'_t, Q_t, \ldots, e_1, Q_1, e'_0 = e_0 \tag{3.1.1}$$

and relabel $V(C)$ so that it's vertex sequence is

$$(u_{1,1}, \ldots, u_{1,k-1}, \ldots, u_{t,1}, \ldots, u_{4t,k-1})$$

(i.e. $u_{1,i} = x_i^{(0)}$ for $i \in [k-1]$, $u_{2,1}, \ldots, u_{2,k-1}$ are the internal vertices of P_1[2], $u_{3,i} = x_i^{(1)}$ for all $i \in [3]$ and so on). By construction, C has the desired properties to apply Observation 3.1.9, finishing the proof. Note that it is important here that every block $\mathcal{H}_i$ has size 4 since we cover 2 vertices of every block "going forwards" and 2 vertices "going backwards". □

3.1.4 Proof of Theorem 1.3.10.

Fix $\alpha, r, n \in \mathbb{N}$ and a k-graph G with $\alpha(G) \leq \alpha$. Choose constants $0 < \beta, \gamma, \varepsilon \ll \max\{\alpha, r, k\}^{-1}$ so that $\gamma = \gamma(r, \varepsilon)$ works for Lemma 3.1.6 and $\beta = \beta(\alpha, r)$ works for Lemma 3.1.3. The proof proceeds in α steps (the initial $k-1$ steps are done at the same time). No effort is made to calculate the exact number of cycles we use, we only care that it is bounded (i.e. independent of n).

Step 1, ..., k-1. By Lemma 3.1.3, there is some $B \subseteq [n]$ of size βn with an absorber $A_{k-1} \subseteq [n]$. Partition B into $k-1$ sets $B_1^{(k-1)}, \ldots, B_{k-1}^{(k-1)}$ of equal sizes.

[2]Note that all P_i and Q_i have $3k-5$ vertices and hence $k-1$ internal vertices.

By Lemma 3.1.1, there is a bounded number of vertex-disjoint monochromatic tight cycles in $[n] \setminus (B \cup A_{k-1})$ so that the set R_{k-1} of uncovered vertices in $[n] \setminus (B \cup A_{k-1})$ satisfies $|R_{k-1}| \leq \gamma|B_1^{(k-1)}|$. Let $R'_{k-1} \subseteq R_{k-1}$ be the set of vertices v with $|\operatorname{Lk}(v; B_1^{(k-1)}, \ldots, B_{k-1}^{(k-1)})| < \varepsilon|B_1^{(k-1)}| \cdots |B_{k-1}^{(k-1)}|$ and let $R''_{k-1} = R_{k-1} \setminus R'_{k-1}$. By Lemma 3.1.6 we can find a bounded number of vertex-disjoint cycles in $B_1^{(k-1)} \cup \ldots \cup B_{k-1}^{(k-1)} \cup R''_{k-1}$ covering R''_{k-1}. Remove them and let $B_i^{(k)} \subseteq B_i^{(k-1)}$ be the set of leftover vertices for every $i \in [k-1]$.

Step j ($j = k, \ldots, \alpha$)**.** In each step j, we will define disjoint sets $B_1^{(j+1)}, \ldots, B_j^{(j+1)}, R'_{j+1}, A_j$. Fix some $j \in \{k, \ldots, \alpha\}$ now and suppose we have built disjoint sets $B_1^{(j)}, \ldots, B_{j-1}^{(j)}, R'_j$ and absorbers $A_2, \ldots, A_{j-1}$. By Lemma 3.1.3 there is some $B_j^{(j)} \subseteq R'_j$ of size $\beta|R'_j|$ with an absorber $A_j \subseteq R'_j$. By Lemma 3.1.1, there is a bounded number of monochromatic tight cycles in $R'_j \setminus (A_j \cup B_j^{(j)})$ so that the set R_{j+1} of uncovered vertices in $R'_j \setminus (A_j \cup B_j^{(j)})$ satisfies $|R_{j+1}| \leq \gamma|B_j^{(j)}|$. Let $R'_{j+1} \subseteq R_{j+1}$ be the set of vertices v with

$$|\operatorname{Lk}(v; B_{t_1}^{(j)}, \ldots, B_{t_{k-1}}^{(j)})| < \varepsilon \left|B_{t_1}^{(j)}\right| \cdots \left|B_{t_{k-1}}^{(j)}\right|$$

for all $1 \leq t_1 < \ldots < t_{k-1} \leq j$ and let $R''_{j+1} = R_{j+1} \setminus R'_{j+1}$. By ($\binom{j}{k}$ applications of) Lemma 3.1.6 we can find a bounded number of vertex-disjoint cycles in $B_1^{(j)} \cup \ldots \cup B_j^{(j)} \cup R''_j$ covering R''_j. Remove them and let $B_i^{(j+1)} \subseteq B_i^{(j)}$ be the set of leftover vertices for every $i \in [j]$.

In the end we have disjoint sets $B_1 := B_1^{(\alpha+1)}, \ldots, B_\alpha := B_\alpha^{(\alpha+1)}, B_{\alpha+1} := R'_{\alpha+1}$ and corresponding absorbers $A_{k-1}, \ldots, A_\alpha$ (A_{k-1} absorbs $B_1^{(\alpha+1)}, \ldots, B_{k-1}^{(\alpha+1)}$). All other vertices are covered by a bounded number of cycles.

We will show now that $R'_{\alpha+1} = \emptyset$, which finishes the proof. In order to do so, we assume the contrary and find an independent set of size $\alpha + 1$. Note that every vertex in $B_j^{(j)} \setminus B_j$ must be part of a tight cycle of our disjoint collection of tight cycles with one part in R_{j+1} and hence $\left|B_j^{(j)} \setminus B_j\right| \leq \left|R_{j+1}\right| \leq \gamma \left|B_j^{(j)}\right|$. It follows that $\left|B_j\right| \geq (1-\gamma)\left|B_j^{(i)}\right|$ for every $1 \leq j \leq i \leq \alpha$ and we conclude

$$\begin{aligned}\mathrm{Lk}\left(v; B_{i_1}, \ldots, B_{i_{k-1}}\right) &\leq \mathrm{Lk}\left(v; B_{i_1}^{(i-1)}, \ldots, B_{i_{k-1}}^{(i-1)}\right) \\ &\leq \varepsilon\left|B_{i_1}^{(i-1)}\right| \cdots \left|B_{i_{k-1}}^{(i-1)}\right| \\ &\leq \varepsilon(1-\gamma)^{-(k-1)}\left|B_{i_1}\right| \cdots \left|B_{i_{k-1}}\right| \\ &\leq 2\varepsilon\left|B_{i_1}\right| \cdots \left|B_{i_{k-1}}\right|\end{aligned}$$

for every $i \in \{k, \ldots, \alpha+1\}$, $1 \leq i_1 < \ldots < i_{k-1} < i$ and $v \in B_i$. By the following lemma, there is an independent set of size $\alpha + 1$, a contradiction. □

Lemma 3.1.10. *Let k and N be positive integers and let H be a k-uniform hypergraph. Suppose that $B_1, \ldots, B_N \subseteq V(H)$ are non-empty disjoint sets such that for every $1 \leq i_1 < \cdots < i_k \leq N$ we have*

$$\deg_H(v, B_{i_2}, \ldots, B_{i_k}) < \binom{N}{k}^{-1} |B_{i_2}| \cdots |B_{i_k}|$$

for all $v \in B_{i_1}$. Then, there exists an independent set $\{v_1, \ldots, v_N\}$ with $v_i \in B_i$, for each $i \in [N]$.

Instead of the original proof given in [18], we will include the more elegant proof of the same lemma given in [31].

Proof. For each $i \in [N]$, let v_i be chosen uniformly at random from B_i. Let $I = \{v_1, \ldots, v_N\}$. Then, the probability that I is an independent set is

$$\begin{aligned}&\sum_{1\leq i_1<\cdots<i_k\leq N} \mathbb{P}\left[\{v_{i_1}, \ldots, v_{i_k}\} \in E(\mathcal{H})\right] \\ &= \sum_{1\leq i_1<\cdots<i_k\leq N} \frac{1}{|B_{i_1}|} \sum_{v\in B_1} \mathbb{P}\left[\{v_{i_1}, \ldots, v_{i_k}\} \in E(\mathcal{H}) \mid v_{i_1} = v\right] \\ &= \sum_{1\leq i_1<\cdots<i_k\leq N} \frac{1}{|B_{i_1}|} \sum_{v\in B_1} \frac{\deg_{\mathcal{H}}(v, B_{i_2}, \ldots, B_{i_k})}{|B_{i_2}| \cdots |B_{i_k}|} \\ &< \sum_{1\leq i_1<\cdots<i_k\leq N} \binom{N}{k}^{-1} = 1.\end{aligned}$$

Therefore, there exists an independent set $\{v_1, \ldots, v_N\}$ with $v_i \in B_i$, for each $i \in [N]$. □

3.2 Tiling Coloured Graphs with Graphs of Bounded Degree

3.2.1 Overview

In this section we are going to prove the following result.

Theorem 1.3.16 (Corsten–Mendonça). *There is an absolute constant $K > 0$ such that for all integers $r, \Delta \geq 2$ and all $\mathcal{F} \in \mathscr{F}_\Delta$, we have $\tau_r(\mathcal{F}) \leq \exp^3\left(Kr^2\Delta^3\right)$. In particular, $\tau_r(\mathcal{F}) < \infty$ whenever $\Delta(\mathcal{F}) < \infty$.*

The proof of Theorem 1.3.16 combines ideas from the absorption method as in the original paper of Erdős, Gyárfás and Pyber [42] with some modern approaches involving the blow-up lemma and the weak regularity lemma. We will be following closely the framework of the proof of the two-colour result of Grinshpun and Sárközy [58]. Our *absorption lemma* (Lemma 3.2.4) states that if we have $k := \Delta+2$ disjoint sets of vertices $V_1, \ldots, V_k$ with $|V_i| \geq 2|V_1|$ for all $i = 2, \ldots, k$ such that every vertex in V_1 belongs to at least $\delta|V_2|\cdots|V_k|$ monochromatic k-cliques *transversal*[3] in $(V_1, \ldots, V_k)$, then it is possible to cover the vertices in V_1 with a constant number (depending on δ, r and Δ) of monochromatic vertex disjoint copies of graphs from $\mathcal{F}$. Furthermore, we can choose such a covering using no more than $|V_1|$ vertices in each $V_2, \ldots, V_k$.

To deduce Theorem 1.3.16 from the absorption lemma, we proceed similar as in Section 3.1: First, using the weak regularity lemma of Duke, Lefmann and Rödl [35], we can find $k-1$ monochromatic *super-regular cylinders* $Z_1, \ldots, Z_{k-1}$ covering a positive proportion of the vertices of K_n (see Lemma 2.2.14). We then apply a result of Fox and Sudakov [52] to cover *greedily* all but a small proportion of the vertices in $V(K_n) \setminus (Z_1 \cup \cdots \cup Z_{k-1})$ with a bounded number of disjoint monochromatic copies of graphs from $\mathcal{F}$ (Proposition 3.2.2).

[3] A k-clique is transversal in $(V_1, \ldots, V_k)$ if it contains one vertex in each one of the sets $V_1, \ldots, V_k$.

Let us denote by R the set of uncovered vertices in $V(K_n) \setminus (Z_1 \cup \cdots \cup Z_{k-1})$. We split R into two sets: the set R_1 of vertices in R belonging to at least $\delta|Z_1|\cdots|Z_{k-1}|$ monochromatic k-cliques transversal in $(R, Z_1, \ldots, Z_{k-1})$, and the set $R_2 = R \setminus R_1$. Using our absorption lemma we can cover the vertices in R_1 using no more than $|R_1|$ vertices of each of the cylinders $Z_1, \ldots, Z_{k-1}$. For each $i = 1, \ldots, k-1$, let Z_i' be the set of vertices in Z_i that has not been used to cover R_1. It follows from the slicing lemma that each Z_i' remains super-regular. Now, if the set R_2 was empty, then we would be done. Indeed, a consequence of the blow-up lemma (Theorem 2.2.16) guarantees that we can cover each of the cylinders $Z_1', \ldots, Z_{k-1}'$ with $k+1$ copies of vertex disjoint monochromatic graphs from $\mathcal{F}$.

So let us consider the case where R_2 is non-empty. In this case, we repeat the process above. This time we first find a reasonably large regular cylinder Z_k in R_2, then cover most of the vertices in $R_2 \setminus Z_k$ greedily and apply the absorption lemma to those vertices that have not yet been covered and belong to many monochromatic k-cliques transversal in R_2 and $k-1$ of the cylinders $Z_1', \ldots, Z_{k-1}', Z_k$. The set of leftover vertices, which we denote by R_3, is either empty (and in this case we are done, as above) or is non-empty, in which case we repeat the process to cover R_3. We can show using Ramsey's theorem that this process must stop after $R_r(k)$ many iterations.

In order to prove the absorption lemma, we employ a density increment argument. This is the most difficult part of the proof and the key new idea in this proof. First, we partition V_1 into r sets $V_1^{(1)}, \ldots, V_1^{(r)}$ so that for every $c \in [r]$, every $v \in V_i^{(c)}$ is incident to at least $d/r \cdot |V_2|\cdots|V_k|$ monochromatic cliques of colour c which are transversal in $(V_1, \ldots, V_k)$. We will cover each of these sets separately, making sure not to repeat vertices. Let us show how to cover $V_1^{(1)}$. We start by finding a large k-cylinder $Z = (U_1, \ldots, U_k)$ with $U_1 \subseteq V_1^{(1)}, U_2 \subseteq V_2, \ldots U_k \subseteq V_k$ which is super-regular in colour 1. We shall use Z as an *absorber* at the end of the proof to cover a small proportion of vertices in $V_1^{(1)}, V_2, \ldots, V_k$. Next, we greedily cover most of $V_1^{(1)} \setminus U_1$ by monochromatic copies of $\mathcal{F}$ until the set of uncovered vertices R has size much smaller then $|U_1|$. To cover the set R, we will find a partition $R = S_1 \cup T_2 \cup \ldots \cup T_k$, where each vertex in S_1 belongs to many monochromatic k-cliques of colour 1 which are transversal in $(S_1, U_2, \ldots, U_k)$ (allowing S_1 to be absorbed into the cylinder Z at the end of the proof) and each vertex in T_i, for $i \in \{2, \ldots, k\}$, belongs to at least $(\delta + \eta)|V_2| \cdots |V_{i-1}||U_i| \cdots |U_k|$ monochromatic k-cliques transversal in $(T_i, V_2, \ldots, V_i, U_{i+1}, \ldots, U_k)$, for some $\eta \ll \delta$.

To cover the vertices in T_i, for $i \in \{2, \ldots, k\}$, we repeat the argument with $(V_1, \ldots, V_k)$ replaced by $(T_i, V_2, \ldots, V_i, U_{i+1}, \ldots, U_k)$ and δ replaced by $\delta + \eta$. Because we are always increasing δ by at least η and the density cannot be larger than 1, we only need to repeat this argument $1/\eta$ times.[4] While covering each of the sets $T_2, \ldots, T_k$, we shall guarantee that the set of vertices $X_i \subseteq U_i$ that we use to cover them has size much smaller than $|U_i|$ for all $i = 2, \ldots, k$. This way, the cylinder $Z' = (U_1 \cup S_1, U_2 \setminus X_2, \ldots, U_k \setminus X_k)$ will be super-regular in colour 1 and thus we can cover $U_1 \cup S_1$ using the blow-up lemma. Repeating this for every colour $c \in [r]$, we get a covering of V_1 with $O_{\delta,r,\Delta}(1)$ many monochromatic disjoint copies of graphs from $\mathcal{F}$.

3.2.2 Tools for the Absorption Method

In the proof, we will use the following theorem of Fox and Sudakov [52] about r-colour Ramsey numbers of bounded-degree graphs.

Theorem 3.2.1 ([52, Theorem 4.3])**.** *Let $k, \Delta, r, n \in \mathbb{N}$ with $r \geq 2$ and let $H_1, \ldots, H_r$ be k-partite graphs with n vertices and maximum degree at most Δ. Then $R(H_1, \ldots, H_r) \leq r^{2rk\Delta} n$.*

Let $\mathscr{F}_{k,\Delta} \subseteq \mathscr{F}_\Delta$ denote the family of k-partite sequences of graphs with maximum degree at most Δ. The following consequence of the previous theorem shows that we can cover almost all the vertices of K_n with constantly many (i.e. independent of n) monochromatic copies of graphs from a family $\mathcal{F} \in \mathscr{F}_{k,\Delta}$.

Proposition 3.2.2. *Let $\Delta, k, r \in \mathbb{N}$, let $\gamma > 0$ and let $C = r^{2rk\Delta} \log(1/\gamma)$. Then, for every $\mathcal{F} \in \mathscr{F}_{k,\Delta}$ and every r-edge-coloured K_n, there are vertex-disjoint monochromatic copies $H_1, \ldots, H_C$ of graphs in $\mathcal{F}$ so that $|V(K_n) \setminus \bigcup_i V(H_i)| \leq \gamma n$.*

Proof. Let $\mathcal{F} = \{F_1, F_2, \ldots\}$, $c = r^{-2rk\delta}$ and $V_0 = [n]$. For every $i = 1, \ldots, C$, by Theorem 3.2.1, there is a monochromatic copy of $F_{c|V_{i-1}|}$ in V_{i-1}. Call it H_i and let

[4] Technically η depends on d, but we will see that this is not a problem.

$V_i = V_{i-1} \setminus V(H_i)$. We then have

$$|V_C| = (1-c)^C n \leq e^{-c \cdot C} n = \gamma n,$$

as claimed. □

In particular, by choosing $\gamma = 1/(2n)$, we get the following corollary.

Corollary 3.2.3. *Let $\Delta, k, r \in \mathbb{N}$ and let $C = 2r^{2rk\Delta} \log n$. Then, for every $\mathcal{F} \in \mathscr{F}_{k,\Delta}$ and every r-edge-coloured K_n, there is a collection of at most C vertex-disjoint copies from $\mathcal{F}$ whose vertex-sets partition $V(G)$.*

3.2.3 The Absorption Lemma

Given a graph G, a partition $V(G) = V_1 \cup \ldots \cup V_k$ into k parts and some $v \in V_1$, let

$$\deg_G(v, V_2, \ldots, V_k) = |\{(v_2, \ldots, v_k) \in V_2 \times \cdots \times V_k : \{v, v_2, \cdots, v_k\} \text{ is a clique}\}|$$

denote the *clique-degree* of v and let

$$\mathrm{cd}_G(v, V_2, \ldots, V_k) := \frac{\deg_G(v, V_2, \ldots, V_k)}{|V_2| \cdots |V_k|}.$$

denote the *clique-density* of v. If, additionally, we are given a colouring $\chi \colon E(G) \to [r]$ of $E(G)$, then we denote $\deg_{G,i}(v, V_2, \ldots, V_k) = \deg_{G_i}(v, V_2, \ldots, V_k)$, where G_i is the graph with vertex set $V(G)$ consisting of the edges of G with colour i. We define $\mathrm{cd}_{G,i}(v, V_2, \ldots, V_k)$ similarly and denote $\mathrm{cd}_{G,[r]}(v, V_2, \ldots, V_k) := \sum_{i=1}^{r} \mathrm{cd}_{G,i}(v, V_2, \ldots, V_k)$. If the graph G is clear from context, we may drop the G in the subscript.

Let $G = K(V_1) \cup K(V_1, \ldots, V_k)$ and let $\mathcal{H}$ be a collection of subgraphs of G. We denote by $\cup\mathcal{H}$ the graph with edge set $\bigcup_{H \in \mathcal{H}} E(H)$ and vertex set $V(\mathcal{H}) := \bigcup_{H \in \mathcal{H}} V(H)$. We say that $\mathcal{H}$ *canonically covers* V_1 if $V_1 \subseteq V(\mathcal{H})$ and

$$|V(\mathcal{H}) \cap V_i| \leq |V(\mathcal{H}) \cap V_1|$$

for all $i \in [2, k]$. The following lemma is the key ingredient of the main proof.

Lemma 3.2.4 (Absorption Lemma)**.** *There is some absolute constant $K > 0$, such that the following is true for all $d > 0$, all integers $\Delta, r \geq 2$, $C = \exp^2\left((r/d)^{K\Delta}\right)$ and for every $\mathcal{F} \in \mathcal{F}_\Delta$. Let $k = \Delta + 2$, let $V_1, \ldots, V_k$ be disjoint sets with $|V_i| \geq 2|V_1|$ for all $i \in [2, k]$ and let $G = K(V_1) \cup K(V_1, \ldots, V_k)$. Let $\chi : E(G) \to [r]$ be a colouring in which for every $v \in V_1$ we have $\mathrm{cd}_{[r]}(v, V_2, \ldots, V_k) \geq d$. Then, there is a collection of at most C vertex-disjoint monochromatic copies of graphs from $\mathcal{F}$ in G which canonically covers V_1.*

The edges inside V_1 will only be used to find copies from $\mathcal{F}$ which lie entirely in V_1 in order to greedily cover most vertices of V_1. The difficult part is finding monochromatic copies in $K(V_1, \ldots, V_k)$ covering the remaining vertices. To do so, we will reduce the problem to only one colour within $K(V_1, \ldots, V_k)$ and then deduce Lemma 3.2.4 from the following lemma.

Lemma 3.2.5. *There is some absolute constant $K > 0$, such that the following is true for all $d > 0$, all integers $\Delta, r \geq 2$, $C = \exp^2\left((r/d)^{K\Delta}\right)$ and for every $\mathcal{F} \in \mathcal{F}_\Delta$. Let $k = \Delta + 2$ and let $V_1, \ldots, V_k$ be disjoint sets with $|V_i| \geq 2|V_1|$ for all $i \in [k]$. Let $H \subseteq K(V_1, \ldots, V_k)$ be a graph with $\mathrm{cd}_H(v, V_2, \ldots, V_k) \geq d$ for every $v \in V_1$ and let $G = K(V_1) \cup H$. Let $\chi : E(G) \to [r]$ be an r-colouring in which every edge of H receives colour 1. Then, there is a collection of at most C vertex-disjoint monochromatic copies of graphs from $\mathcal{F}$ in G which canonically covers V_1 and at most $d/(2r) \cdot |V_i|$ vertices of V_i for each $i \in [2, k]$.*

Lemma 3.2.4 follows routinely from Lemma 3.2.5.

Proof of Lemma 3.2.4. Let K' be the absolute constant from Lemma 3.2.5 and let $C' := \exp^2\left((2r^2/d)^{K'\Delta}\right)$. Let $d' = d/(2r)$ and partition $V_1 = V_1^{(1)} \cup \ldots \cup V_1^{(r)}$ so that $\mathrm{cd}_j(v, V_2, \ldots, V_k) \geq 2d'$ for every $j \in [r]$ and $v \in V_1^{(j)}$. We apply Lemma 3.2.5 (with d') to $G_1 := G[V_1^{(1)} \cup V_2 \cup \ldots \cup V_k]$ to get a collection $\mathcal{H}_1$ of at most C' monochromatic copies from $\mathcal{F}$ in G_1 canonically covering $V_1^{(i)}$ and at most $d'/(2r) \cdot |V_i| = d/(4r^2) \cdot |V_i|$ vertices of V_i for each $i \in [2, k]$. Let $V_i' := V_i \setminus V(\mathcal{H}_1)$ for every $i \in [k]$, and observe that $|V_i| \geq 2|V_1'|$ for all $i \in [k]$ and $\mathrm{cd}_j(v, V_2', \ldots, V_k') \geq d'$ for every $j \in [r]$ and $v \in V_1^{(j)}$. We proceed like this to canonically cover each $V_1^{(j)}$ with vertex-disjoint monochromatic copies from $\mathcal{F}$. In total, we use at most $rC' \leq \exp^2\left((r/d)^{4K'\Delta}\right)$ copies, finishing the proof. $\square$

The proof of Lemma 3.2.5 is quite long and technical, see Section 3.2.1 for a sketch.

Proof of Lemma 3.2.5. Let Δ and r be given positive integers and $\mathcal{F} \in \mathscr{F}_\Delta$. For each $d > 0$, let $C(d)$ be the smallest integer $C \geq 0$ such that the following holds:

> ($\star$) Let $V_1, \ldots, V_k$ be disjoint sets with $|V_i| \geq 2|V_1|$ for all $i \in [k]$, let $H \subseteq K(V_1, \ldots, V_k)$ be a graph with $\operatorname{cd}(v, V_2, \ldots, V_k) \geq d$ for every $v \in V_1$ and $G = K(V_1) \cup H$. Let $\chi : E(G) \to [r]$ be a colouring such that every edge in $E(H)$ receives colour 1. Then, there is a collection $\mathcal{H}$ of at most C vertex-disjoint monochromatic copies of graphs from $\mathcal{F}$ contained in G that canonically covers V_1 and at most $d/(2r) \cdot |V_i|$ vertices of V_i for each $i \in [2, k]$.

Notice that this minimum is attained since it is taken over non-negative integers. Furthermore, $C(d)$ is a decreasing function in d, and $C(d) = 0$ for every $d > 1$ (($\star$) trivially holds for all $C \geq 0$ in this case). Our goal is to show that $C(d)$ is finite for every $d > 0$. We will do this by establishing a recursive upper bound (see (3.2.1)).

Let us first define all relevant constants used in the proof. Let K' be the universal constant given by Theorem 2.2.16 and fix some $0 < d \leq 1$. Define

$$\varepsilon = \left(\frac{d}{100}\right)^{2K'\Delta}, \quad \gamma = \tfrac{1}{r} \cdot \varepsilon^{k^2\varepsilon^{-12}} \quad \text{and} \quad \eta = \frac{d\gamma^k}{2}.$$

It might be of benefit for the reader to have in mind that those constants obey the following hierarchy:

$$1 \geq d \gg \varepsilon \gg \gamma \gg \eta > 0.$$

Furthermore, define

$$P(d) := r^{4rk^2} \log(2/\eta^2) + 1.$$

Note that, since r and k are fixed, η depends only on d and thus P is indeed a function of d. We will prove that for every $d' \geq d$ we have

$$C(d') \leq P(d) + kC(d' + \eta). \tag{3.2.1}$$

Since $C(d') = 0$ if $d' > 1$, it follows by iterating that $C(d) \leq (2k)^{2/\eta} P(d)$. Furthermore, we have

$$2/\eta \leq \gamma^{-2k} \leq \varepsilon^{-2rk^3\varepsilon^{-12}} \leq \exp\left(r\varepsilon^{-20}\right) \leq \exp\left((r/d)^{400K'\Delta}\right).$$

It follows that

$$C(d) \leq \exp^2\left((r/d)^{500K'\Delta}\right) P(d) \leq \exp^2\left((r/d)^{1000K'\Delta}\right)$$

concluding the proof of Lemma 3.2.5.

It remains to prove (3.2.1). Let $d' \geq d$ be fixed now and let $V_1, \ldots, V_k$, G and $\chi : E(G) \to [r]$ be as in ($\star$) (with d' playing the role of d). By assumption, there are at least $d|V_1||V_2|\cdots|V_k|$ cliques of size k in $G[V_1, V_2, \ldots, V_k]$ each of which is monochromatic in colour 1. Since $\gamma = \varepsilon^{k^2\varepsilon^{-12}}$ and $d \geq 2k\varepsilon$, we can apply Lemma 2.2.15 to get some $\gamma' \geq \gamma$ and a k-cylinder $Z = (U_1, \ldots, U_k)$ which is $(\varepsilon, (d/2)^+)$-super-regular with $U_i \subseteq V_i$ and $|U_i| = \lfloor \gamma' |V_i| \rfloor$ for every $i \in [k]$. Without loss of generality we may assume that $\gamma|V_i|$ is an integer for every $i \in [k]$ and that we have $\gamma' = \gamma$. By Proposition 3.2.2, there is a collection $\mathcal{H}_R$ of at most $r^{4rk^2} \log(2/\eta^2)$ vertex-disjoint monochromatic copies of graphs from $\mathcal{F}$ contained in $K(V_1 \setminus U_1)$ covering all vertices in $V_1 \setminus U_1$ except for a set R with $|R| \leq \eta^2 |V_1|$. We remark here that

$$|R| \leq \eta/(4k) \cdot |U_1| \leq \varepsilon^2 |U_1|. \tag{3.2.2}$$

It remains now to cover the vertices in R. For each $i \in [k]$, let

$$d_i = \frac{1 - \gamma^i}{1 - \gamma^k} \cdot d' \tag{3.2.3}$$

and notice that $(1 - \gamma)d' \leq d_1 \leq \cdots \leq d_k = d'$. For $i \in [2, k]$, let $\tilde{V}_i = V_i \setminus U_i$ and define

$$S_i = \{v \in R : \mathrm{cd}(v, V_2, \dots, V_{i-1}, V_i, U_{i+1}, \dots, U_k) \geq d_i\},$$
$$T_i = \{v \in R : \mathrm{cd}(v, V_2, \dots, V_{i-1}, \tilde{V}_i, U_{i+1}, \dots, U_k) > d' + 2\eta\}.$$

We will prove (3.2.1) using a series of claims, which we shall prove at the end.

Claim 3.1. *We have $R = S_1 \cup T_2 \cup \dots \cup T_k$.*

Without loss of generality, we may assume that $S_1, T_2, \dots, T_k$ are pairwise disjoint (more formally, we can define $T_i' := T_i \setminus (S_1 \cup T_2 \cup \dots \cup T_{i-1})$ for all $i \in [2, k]$ and continue the proof with these sets.) Our goal now is to cover each of the sets $S_1, T_2, \dots, T_k$ one by one using the following claims.

Claim 3.2. *For every $i \in [2, k]$ and every set $A \subseteq V(G) \setminus T_i$ with $|A \cap V_s| \leq |R|$ for all $s \in [2, k]$, there is a collection $\mathcal{H}_i$ of at most $C(d' + \eta)$ monochromatic disjoint copies of graphs from $\mathcal{F}$ in G, such that*

(i) $V(\mathcal{H}_i) \cap V_1 = T_i$,

(ii) $V(\mathcal{H}_i) \cap A = \emptyset$, *and*

(iii) $\left|V(\mathcal{H}_i) \cap V_j\right| \leq |T_i|$ *for all* $j \in [2, k]$.

Claim 3.3. *For every set $A \subseteq V(G) \setminus (S_1 \cup U_1)$ with $|A \cap V_s| \leq |R|$ for all $s \in [2, k]$, there is a monochromatic copy H_1 of a graph from $\mathcal{F}$ in G, such that*

(i) $V(H_1) \cap V_1 = S_1 \cup U_1$,

(ii) $V(H_1) \cap A = \emptyset$ *and*

(iii) $\left|V(H_1) \cap V_j\right| \leq |S_1 \cup U_1|$ *for all* $j \in [2, k]$.

With these claims at hand, we can now prove (3.2.1). First, we apply Claim 3.2 repeatedly to get collections $\mathcal{H}_2, \dots, \mathcal{H}_k$ of monochromatic copies from $\mathcal{F}$ as follows. Let $i \in \{2, \dots, k\}$ and suppose we have constructed $\mathcal{H}_2, \dots, \mathcal{H}_{i-1}$. Let $A_i := V(\mathcal{H}_2) \cup \dots \cup V(\mathcal{H}_{i-1})$ and note that $|A_i \cap V_s| \leq |T_2| + \dots + |T_{i-1}| \leq |R|$ for all $s \in [2, k]$. Apply now Claim 3.2 for i and $A = A_i$ to get the desired collection

$\mathcal{H}_i$. Next, we apply Claim 3.3 with $A = V(\mathcal{H}_2) \cup \ldots \cup V(\mathcal{H}_k)$ to get a copy H_1 of a graph from $\mathcal{F}$ with the desired properties. Note that, similarly as above, we have $|A \cap V_s| \leq |R|$ for all $s \in [2, k]$. By construction $V(H_1), V(\mathcal{H}_2), \ldots, V(\mathcal{H}_k)$ and $V(\mathcal{H}_R)$ are disjoint and cover V_1. Moreover, for every $s \in [2, k]$, we have

$$\begin{aligned} |(V(H_1) \cup \ldots \cup V(\mathcal{H}_k) \cup V(\mathcal{H}_R)) \cap V_s| &\leq |S_1 \cup U_1| + |R_1| + |T_2| + \cdots + |T_k| \\ &\leq |U_1 \cup R| \leq d/(2r) \cdot |V_1|. \end{aligned}$$

Hence, $\{H_1\} \cup \ldots \cup \mathcal{H}_k \cup \mathcal{H}_R$ canonically covers V_1 and at most $d/(2r) \cdot |V_i|$ vertices of every $i \in [2, k]$. Finally, we have $|\{H_1\} \cup \ldots \cup \mathcal{H}_k \cup \mathcal{H}_R| \leq P(d) + kC\big(d' + \eta\big)$, proving (3.2.1). It remains now to prove Claims 3.1 to 3.3.

Proof of Claim 3.1. Since $S_k = R$, it suffices to show $S_i \subseteq S_{i-1} \cup T_i$ for each $i \in [2, k]$. Let $i \in [2, k]$ and let $v \in S_i \setminus S_{i-1}$. We have

$$\begin{aligned} \deg(v, V_2, \ldots, V_{i-1}, \tilde{V}_i, U_{i+1}, \ldots, U_k) =& \deg(v, V_2, \ldots, V_{i-1}, V_i, U_{i+1}, \ldots, U_k) \\ &- \deg(v, V_2, \ldots, V_{i-1}, U_i, U_{i+1}, \ldots, U_k). \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{cd}(v, V_2, \ldots, V_{i-1}, \tilde{V}_i, U_{i+1}, \ldots, U_k) =\; & \mathrm{cd}(v, V_2, \ldots, V_{i-1}, V_i, U_{i+1}, \ldots, U_k)\frac{|V_i|}{|\tilde{V}_i|} \\ & - \mathrm{cd}(v, V_2, \ldots, V_{i-1}, U_i, U_{i+1}, \ldots, U_k)\frac{|U_i|}{|\tilde{V}_i|} \\ >\; & d_i\frac{|V_i|}{|\tilde{V}_i|} - d_{i-1}\frac{|U_i|}{|\tilde{V}_i|} \\ =\; & \frac{d_i - \gamma d_{i-1}}{1-\gamma} \\ =\; & \frac{(1-\gamma^i)d' - \gamma(1-\gamma^{i-1})d'}{(1-\gamma)(1-\gamma^k)} \\ =\; & \frac{d'}{1-\gamma^k} \geq d' + 2\eta, \end{aligned}$$

where we use (3.2.3) and the definition of η to obtain the last identities. Thus $v \in T_i$ and hence $S_i \subseteq S_{i-1} \cup T_i$. ∎

Proof of Claim 3.2. Let $V'_s := V_s \setminus A$ for all $s \in [2, i-1]$, $\tilde{V}'_i := \tilde{V}_i \setminus A$ and $U'_s := U_s \setminus A$ for all $s \in [i+1, k]$. Then, by (3.2.2), we have

$$\begin{aligned} |V'_s| \geq |V_s| - |R| \geq \left(1 - \tfrac{\eta}{4k}\right)|V_s| \geq \frac{|V_s|}{2}, \text{ for } s = 2, \ldots, i-1, \\ |\tilde{V}'_i| \geq |\tilde{V}_i| - |R| \geq \left(1 - \tfrac{\eta}{4k}\right)|\tilde{V}_i| \geq \frac{|\tilde{V}_i|}{2}, \text{ and} \\ |U'_s| \geq |U_s| - |R| \geq \left(1 - \tfrac{\eta}{4k}\right)|U_s| \geq \frac{|U_j|}{2}, \text{ for } s = i+1, \ldots, k. \end{aligned}$$

In particular, it follows that

$$\begin{aligned} |V_s \setminus V'_s| \leq |R| \leq \frac{\eta}{4k}|V_s| \leq \frac{\eta}{2k}|V'_s|, \text{ for } s = 2, \ldots, i-1, \\ |V_i \setminus V'_i| \leq |R| \leq \frac{\eta}{4k}|V_i| \leq \frac{\eta}{2k}|V'_i|, \text{ and} \\ |U_s \setminus U'_s| \leq |R| \leq \frac{\eta}{4k}|U_s| \leq \frac{\eta}{2k}|U'_s|, \text{ for } s = i+1, \ldots, k. \end{aligned}$$

Therefore, for every $v \in T_i$, we have

$$\begin{aligned} &\operatorname{cd}(v, V'_2, \ldots, V'_{i-1}, \tilde{V}'_i, U'_{i+1}, \ldots, U'_k) \\ &\geq d' + 2\eta - \sum_{s=2}^{i-1} \frac{|V_s \setminus V'_s|}{|V'_s|} - \frac{|\tilde{V}_i \setminus \tilde{V}'_i|}{|\tilde{V}'_i|} - \sum_{s=i+1}^{k} \frac{|U_s \setminus U'_s|}{|U'_s|} \\ &\geq d' + 2\eta - (k-1)\frac{\eta}{2k} \geq d' + \eta. \end{aligned}$$

Hence, by definition of $C(d' + \eta)$ (see ($\star$)), there exists a collection $\mathcal{H}_i$ of at most $C(d' + \eta)$ monochromatic copies of graphs from $\mathcal{F}$ which canonically covers T_i in the graph $K(T_i) \cup K(T_i, V'_2, \ldots, V'_{i-1}, \tilde{V}'_i, U'_{i+1}, \ldots, U'_k)$. By construction, $\mathcal{H}_i$ satisfies the requirements of the claim ((*iii*) holds since $\mathcal{H}_i$ is a canonical covering). ■

Proof of Claim 3.3. Let $Y_1 = S_1$ and, for each $i \in [2, k]$, let $X_i = U_i \cap A$. Observe that $|Y_1| \leq |R| \leq \varepsilon^2 |U_1|$ and $|X_i| \leq |R| \leq \varepsilon^2 |U_i|$ for all $i \in [2, k]$. Let $U'_1 = U_1 \cup Y_1$ and, for each $i \in [2, k]$, let $U'_i := U_i \setminus X_i$. We now consider the cylinder $Z' := (U'_1, \ldots, U'_k)$. By definition of S_1, we have $\operatorname{cd}(v, U_2, \ldots, U_k) \geq d_1 \geq d/2$ and in particular $\deg(v, U_i) \geq d/2 \cdot |U_i|$ for all $v \in Y_1$ and $i \in [2, k]$.

Hence, by Lemma 2.2.2, Z' is $(8\varepsilon, (d/4)^+)$-super-regular. Furthermore, we have

$|U_1'| \leq |U_i'|$ for all $i \in [k]$. Thus, by Theorem 2.2.16, there is a monochromatic copy H_1 of a graph from $\mathcal{F}$ in Z that covers $U_1' = U_1 \cup S_1$ and at most $|U_1'|$ vertices from each of $U_2', \ldots, U_k'$. By construction, this copy satisfies the requirements of the claim. ■

This finishes the proof of Lemma 3.2.5. □

3.2.4 Proof of Theorem 1.3.16

In this section, we will use Lemma 3.2.4 to finish the proof of Theorem 1.3.16. We will follow the same proof technique as in Section 3.1.4. In particular, we will make use of Lemma 3.1.10 which we will restate here for convenience.

Lemma 3.1.10. *Let k and N be positive integers and let H be a k-uniform hypergraph. Suppose that $B_1, \ldots, B_N \subseteq V(H)$ are non-empty disjoint sets such that for every $1 \leq i_1 < \cdots < i_k \leq N$ we have*

$$\deg_H(v, B_{i_2}, \ldots, B_{i_k}) < \binom{N}{k}^{-1} |B_{i_2}| \cdots |B_{i_k}|$$

for all $v \in B_{i_1}$. Then, there exists an independent set $\{v_1, \ldots, v_N\}$ with $v_i \in B_i$, for each $i \in [N]$.

We are now able to prove Theorem 1.3.16. The main idea is to find reasonably large cylinders that are super-regular for one of the colours, greedily cover most of the remaining vertices using Proposition 3.2.2 and then apply the Absorption Lemma (Lemma 3.2.4) to the set of remaining vertices that share many monochromatic cliques with the cylinders. We then iterate this process until no vertices remain. Using Lemma 3.1.10, we will show that a bounded number of iterations suffices.

Proof of Theorem 1.3.16. Fix $r, \Delta \geq 2$, $\mathcal{F} \in \mathcal{F}_\Delta$. Let G be an r-edge-coloured complete graph on n vertices. Let

$$k = \Delta + 2, \quad N = r^{rk}, \quad \delta = N^{-k} \quad \text{and} \quad d = \frac{1}{2r}.$$

In order to use Theorem 2.2.16 and Lemma 2.2.14, respectively, consider the constants

$$\varepsilon = (\delta d^{\Delta})^{2K'} \quad \text{and} \quad \gamma = \varepsilon^{r^{8rk}\varepsilon^{-5}},$$

where K' is the universal constant given by Theorem 2.2.16. Consider also the constants

$$\alpha = \varepsilon^2 \quad \text{and} \quad C_1 = r^{2rk\Delta} \log\left(\frac{4}{\alpha\gamma}\right)$$

in order to use Proposition 3.2.2. Finally, let

$$C_2 = \exp^2((2r/\delta)^{\tilde{K}\Delta}) \leq \exp^2\left(r^{16\tilde{K}r\Delta^3}\right) \leq \exp^3\left(16\tilde{K}r^2\Delta^3\right),$$

where $\tilde{K}$ is the universal constant from Lemma 3.2.4, and let $K = 20\tilde{K}$.

We will build a framework consisting of many k-cylinders working as absorbers and small sets which can be absorbed by them. More precisely, our goal is to define sets with the following properties.

Framework. There are sets $Z_1, \ldots, Z_N, S_{k-1}, \ldots, S_N, R_k, \ldots, R_{N+1}, R'_k, \ldots, R'_{N+1}$ with the following properties.

(F.1) $V(G) = \bigcup_{i=1}^{N} Z_i \cup \bigcup_{i=k-1}^{N} S_i \cup \bigcup_{i=k}^{N+1} R'_i$ is a partition.

(F.2) $Z_1, \ldots, Z_N$ are the vertex-sets of k-cylinders which are (ε, d^+)-super-regular in one of the colours (or empty).

(F.3) $S_{k-1}, \ldots, S_N$ are sets of vertices which we will cover greedily by monochromatic copies of graphs from $\mathcal{F}$.

(F.4) For each $i \in [k, N+1]$, R'_i can be partitioned into sets $R'_{i,I}$ for all $I \in \binom{[i-1]}{k-1}$, such that, for each $I = \{i_1, \ldots, i_{k-1}\} \subseteq [i]$, we have $\mathrm{cd}_{[r]}(u, Z_{i_1}, \ldots, Z_{i_{k-1}}) \geq \delta$ for all $u \in R'_{i+1,I}$.

(F.5) For each $k \leq i < j \leq N+1$, we have $S_j, Z_j, R'_j \subseteq R_i$ and $|R_i| \leq \alpha|Z_{i-1}|$.

Figure 3.4 should help the reader to understand the structure of those sets as we define them. First, if $n < 1/4\gamma$, then Corollary 3.2.3 gives a covering with at most C_2 monochromatic vertex-disjoint copies of graphs from $\mathcal{F}$. Therefore we may

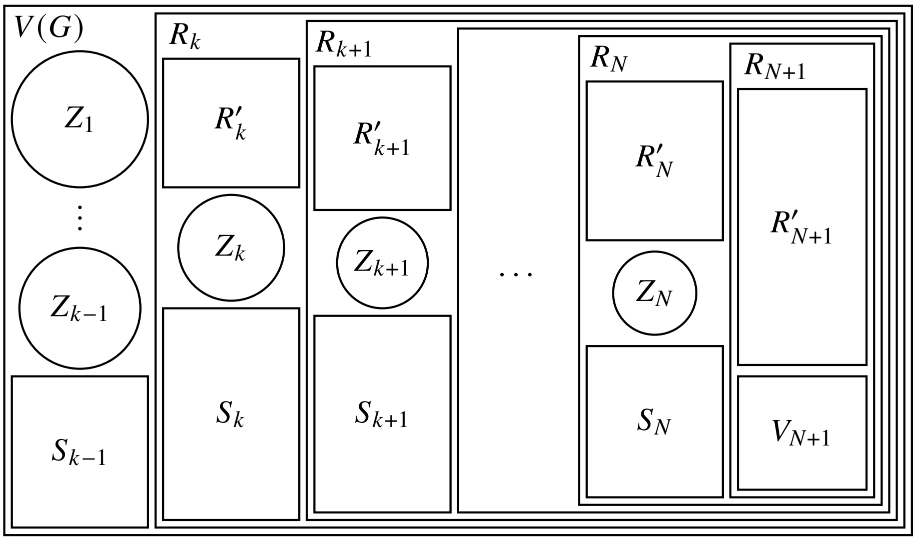


Figure 3.4: A partition of $V(G)$. Each set in the picture is much smaller than the closest cylinder Z_i to the left.

assume that $n \geq 1/4\gamma$. Hence, by applying Lemma 2.2.14 multiple times, we find $k-1$ vertex-disjoint k-cylinders $Z_1, \dots, Z_{k-1}$ such that each of them is (ε, d^+)-super-regular in some colour (not necessarily the same) and $|Z_1| \geq \cdots \geq |Z_{k-1}| \geq \gamma n/2$. Let $V_{k-1} = V(G) \setminus (Z_1 \cup \cdots \cup Z_{k-1})$. By Proposition 3.2.2, there is a collection of at most C_1 monochromatic vertex-disjoint copies from $\mathcal{F}$ in V_{k-1} covering a set S_{k-1} such that the leftover vertices $R_k = V_{k-1} \setminus S_{k-1}$ satisfies $|R_k| \leq \alpha\gamma n/2 \leq \alpha|Z_{k-1}|$. Let $R'_k \subseteq R_k$ be the set of vertices $u \in R_k$ with $\mathrm{cd}_{[r]}(u, Z_1, \dots, Z_{k-1}) \geq \delta$. Let $R'_{k,[k-1]} = R'_k$ and $V_k = R_k \setminus R'_k$.

Inductively, for each $i = k, \dots, N$, we do the following. If $|V_i| < 1/4\gamma$, we use Corollary 3.2.3 to cover V_i using at most C_2 monochromatic vertex-disjoint copies from $\mathcal{F}$ and let $Z_i = S_i = R_{i+1} = R'_{i+1} = V_{i+1} = \emptyset$. Otherwise, we apply Lemma 2.2.14 to find a monochromatic (ε, d^+)-super-regular k-cylinder Z_i contained in V_i with $|Z_i| \geq \gamma\,|V_i|$. By Proposition 3.2.2, there is a collection of at most C_1 monochromatic, vertex-disjoint copies from $\mathcal{F}$ in $V_i \setminus Z_i$ covering a set $S_i \subseteq V_i$, so that the set of leftover vertices $R_{i+1} = V_i \setminus S_i$ has size at most $\alpha\gamma|V_i| \leq \alpha\,|Z_i|$.

Let R'_{i+1} be the set of vertices u in R_{i+1} for which there is a set $I = \{i_1, \dots, i_{k-1}\} \subseteq$

$[i]$ such that $\mathrm{cd}_{[r]}(u, Z_{i_1}, \ldots, Z_{i_{k-1}}) \geq \delta$. Let

$$R'_{i+1} = \bigcup_{I \in \binom{[i]}{k-1}} R'_{i+1,I}$$

be a partition of R'_{i+1} so that, for each $I = \{i_1, \ldots, i_{k-1}\} \subseteq [i]$, we have $\mathrm{cd}_{[r]}(u, Z_{i_1}, \ldots, Z_{i_{k-1}}) \geq \delta$ for all $u \in R'_{i+1,I}$. Finally, let $V_{i+1} = R_{i+1} \setminus R'_{i+1}$. The following claim implies (F.1).

Claim 3.4. *The set V_{N+1} is empty.*

Proof. Define a k-uniform hypergraph $\mathcal{H}$ with vertex set $U = Z_1 \cup \ldots \cup Z_N \cup V_{N+1}$ and hyperedges corresponding to monochromatic k-cliques in $G[U]$. If V_{N+1} is non-empty, then so are $Z_1, \ldots, Z_N$. Since for each $i = k, \ldots, N$ we have $Z_i \subseteq R_i \setminus R'_i$, it follows that $\mathcal{H}$ satisfies the hypothesis of Lemma 3.1.10. Therefore, there is an independent set $\{v_1, \ldots, v_{N+1}\}$ in $\mathcal{H}$ of size $N + 1$. On the other hand, since $N \geq R_r(K_k)$, it follows that the set $\{v_1, \ldots, v_{N+1}\}$ has a monochromatic k-clique in $G[U]$, which is a contradiction. ∎

The vertices in $S_{k-1} \cup \cdots \cup S_N$ are already covered by monochromatic copies of graphs from $\mathcal{F}$. Our goal now is to cover the sets $R'_k, \ldots, R'_{N+1}$ using Lemma 3.2.4 without using too many vertices from the cylinders $Z_1, \ldots, Z_N$. This way, we can cover the remaining vertices in $Z_1 \cup \cdots \cup Z_N$ using Theorem 2.2.16.

Claim 3.5. *Let $i \in \{k, \ldots, N+1\}$ and $I = \{i_2, \ldots, i_k\} \subseteq [i-1]$. Let $A \subseteq V(G) \setminus R_{i,I}$ be a set with $|A \cap Z_j| \leq \alpha |Z_j|$ for each $j \in I$. Then there is a collection of at most C_2 monochromatic vertex-disjoint copies of graphs from $\mathcal{F}$ in*

$$G' = K(R'_{i,I}) \cup K(R'_{i,I}, Z_{i_2}, \ldots, Z_{i_k})$$

which are disjoint from A and canonically cover $R'_{i,I}$.

Proof. Let $\tilde{V}_1 = R'_{i,I}$ and for $j \in [k] \setminus \{1\}$, let $\tilde{V}_j = Z_{i_j} \setminus A$. Note that $|\tilde{V}_j| \geq 2|\tilde{V}_1|$

for every $j \in [k] \setminus \{1\}$ and

$$\begin{aligned}\deg_{[r]}(v, \tilde{V}_2, \dots, \tilde{V}_k) &\geq \deg_{[r]}(v, Z_{i_2}, \dots, Z_{i_k}) - k\alpha |Z_{i_2}| \cdots |Z_{i_k}| \\ &\geq (\delta - k\alpha)|Z_{i_2}| \cdots |Z_{i_k}| \\ &\geq \delta/2 \cdot |Z_{i_2}| \cdots |Z_{i_k}|\end{aligned}$$

for every $v \in \tilde{V}_1$. Hence, by Lemma 3.2.4, there is a collection of at most C_2 vertex-disjoint copies from $\mathcal{F}$ in $\tilde{V}_1 \cup \ldots \cup \tilde{V}_k$ which canonically covers $\tilde{V}_1$, finishing the proof. ■

We will use Claim 3.5 now to cover $\bigcup_{i=k}^{N+1} R'_i$. Let $\prec$ be a linear order on $\mathcal{I} := \left\{(i, I) : i \in [k, N+1], I \in \binom{[i-1]}{k-1}\right\}$. Let $(i, I) \in \mathcal{I}$ and suppose that, for all $(i', I') \in \mathcal{I}$ with $(i', I') \prec (i, I)$, we have already constructed a family $\mathcal{H}_{i',I'}$ of monochromatic copies of graphs from $\mathcal{F}$ which canonically covers $R'_{i',I'}$ in $K(R'_{i',I'}) \cup K(R'_{i',I'}, Z_{i'_2}, \dots, Z_{i'_k})$, where $I' = \{i'_2, \dots, i'_k\}$.

Let $A = \bigcup_{(i',I') \prec (i,I)} V(\mathcal{H}_{i',I'})$ be the set of already covered vertices. We claim that

$$\left|A \cap Z_j\right| \leq |R_{j+1}| \leq \alpha |Z_j| \tag{3.2.4}$$

for each $j \in [N]$. Indeed, given some $j \in [N]$, we have $V(\mathcal{H}_{i',I'}) \cap Z_j = \emptyset$ for all $(i', I') \in \mathcal{I}$ with $i' \leq j$, and $\left|V(\mathcal{H}_{i',I'}) \cap Z_j\right| \leq |R'_{i',I'}|$ for all $(i', I') \in \mathcal{I}$ with $i > j$ since $\mathcal{H}_{i,I}$ is canonical. This implies (3.2.4), since the sets $\{R'_{i',I'} : (i', I') \in \mathcal{I}, i > j\}$ are disjoint subsets of R_{j+1}. In particular, by Claim 3.5, there is a collection $\mathcal{H}_{i,I}$ of monochromatic copies of graphs from $\mathcal{F}$ which canonically covers $R'_{i,I}$ in $K(R'_{i,I}) \cup K(R'_{i,I}, Z_{i_2}, \dots, Z_{i_k})$, where $I = \{i_2, \dots, i_k\}$.

It remains to cover $\bigcup_{i=1}^{N} Z_i$. Let $A := \bigcup_{(i,I) \in \mathcal{I}} V(\mathcal{H}_{i,I})$ be the set of vertices covered in the previous step. Note that, similarly as in (3.2.4), we have $|A \cap Z_j| \leq |R_j| \leq \alpha |Z_j|$ for all $j \in [N]$. Therefore, by Lemma 2.2.2, the cylinder $\tilde{Z}_j$ obtained from Z_j by removing all vertices in A is $(8\varepsilon, (d/2)^+)$-super-regular and ε-balanced for every $j \in [N]$. It follows from Theorem 2.2.16 that, for every $j \in [N]$, there is a collection $\mathcal{H}_j$ of at most $\Delta + 3$ monochromatic vertex-disjoint copies of graphs from $\mathcal{F}$ contained in Z_j covering $V(Z_j)$.

In total, the number of monochromatic copies we used to cover $V(G)$ is at most

$$
\begin{aligned}
N \cdot C_1 + N^k \cdot C_2 + N \cdot (\Delta + 3) &\leq 2N^k C_2 \\
&\leq 2r^{rk^2} \cdot \exp^3\left(16\tilde{K}r^2\Delta^3\right) \\
&\leq \exp^3\left(Kr^2\Delta^3\right).
\end{aligned}
$$

This concludes the proof of Theorem 1.3.16. □

3.2.5 Concluding Remarks

Recall that a graph G is called *a-arrangeable* for some $a \in \mathbb{N}$ if its vertices can be ordered in such a way that the neighbours to the right of any vertex $v \in V(G)$ have at most a neighbours to the left of v in total. Using an extension of the blow-up lemma by Böttcher, Kohayakawa and Taraz [14], we obtain the following theorem.

Theorem 3.2.6. *For all integers $r, a \geq 2$ and all sequences of a-arrangeable graphs $\mathcal{F} = \{F_1, F_2, \dots\}$ with $\Delta(F_n) \leq \sqrt{n}/\log(n)$ for all $n \in \mathbb{N}$, we have $\tau_r(\mathcal{F}) < \infty$.*

The proof is almost identical, with the following two differences. First, instead of Theorem 2.2.18, we will use Theorem 2.2.19. In order to do so, we show that it suffices to prove Theorem 3.2.6 for graphs with balanced $(a+2)$-partitions. Indeed, given a sequence $\mathcal{F} = \{F_1, F_2, \dots\}$ of a-arrangeable graphs with $\Delta(F_n) \leq \sqrt{n}/\log(n)$ for every $n \in \mathbb{N}$, we define another sequence of graphs $\tilde{\mathcal{F}} = \{\tilde{F}_1, \tilde{F}_2, \dots\}$ as follows. Since every a-arrangeable graph is $(a+2)$-colourable, we can fix a partition of $V(F_n) = V_1(F_n) \cup \dots \cup V_k(F_n)$ into independent sets, where $k = a + 2$. Then, for every $j \in \mathbb{N}$, we define $\tilde{F}_{jk}$ to be the disjoint union of k copies of F_j. Note that each $\tilde{F}_{jk}$ has a k-partition into parts of equal sizes (by rotating each copy around). Finally, for each $j \in \mathbb{N} \cup \{0\}$ and every $i \in [k-1]$, we define $\tilde{F}_{jk+i}$ to be the disjoint union of $\tilde{F}_{jk}$ and i isolated vertices (here $\tilde{F}_0$ is the empty graph). Observe that all $\tilde{F}_n$ have k-partitions into parts of almost equal sizes. Furthermore, every $\tilde{\mathcal{F}}$-tiling $\mathcal{T}$ corresponds to an $\mathcal{F}$-tiling $\tilde{\mathcal{T}}$ of size at most $(2k-1)|\mathcal{T}|$.

Second, we need to replace Theorem 3.2.1 with a similar theorem for a-arrangeable graphs G with $\Delta(G) \leq \sqrt{n}/\log(n)$, where $n = v(G)$. For two colours, such a theorem was proved by Chen and Schelp [19]. For more than two colours, this was (to the best of the author's knowledge) never explicitly stated, but is easy to obtain using modern tools (for example, by applying the above mentioned blow-up lemma for a-arrangeable graphs).

4

Ramsey Problems for Infinite Graphs

4.1 Ramsey Upper Density of Paths

4.1.1 Overview

In this section, we are going to prove the following two results.

Theorem 1.4.4 (Corsten–DeBiasio–Lamaison–Lang [30])**.** *There exists a 2-edge-colouring of $K_{\mathbb{N}}$ such that every monochromatic path has upper density at most* $(12 + \sqrt{8})/17$.

Theorem 1.4.5 (Corsten–DeBiasio–Lamaison–Lang [30])**.** *Suppose the edges of $K_{\mathbb{N}}$ are coloured with two colours. Then, there exists a monochromatic path with upper density at least* $(12 + \sqrt{8})/17$.

Erdős and Galvin [39] constructed a 2-edge-colouring of $K_{\mathbb{N}}$ in which every monochromatic path has upper density at most $8/9$. It turns out that this can be improved by using the very same colouring but reordering the vertices. The proof of Theorem 1.4.4 is based on this observation.

In their proof of Theorem 1.4.3, Erdős and Galvin introduced an auxiliary vertex-colouring on top of the given edge-colouring and reduced the problem to finding a monochromatic path forest (i.e. a graph F with $\Delta(F) \leq 2$ which does not contain any finite cycles and all edges, leaves and isolated vertices receive the same colour). In order to prove Theorem 1.4.5, we use the regularity method to further reduce the problem to finding what we call a monochromatic simple forest (a union of edges and isolated vertices in which every edge, every isolated vertex and at least one of the endpoints of each edge receive the same colour). The key part of the proof is finding such a monochromatic simple forest of large upper density.

Throughout the proof, given a function f and a set $X \subseteq \operatorname{dom} f$, we denote $f(X) := \{f(x) : x \in X\}$. Furthermore, given a graph G, a matching $M \subseteq G$ and a set of vertices $X \subseteq V(G)$, we say that M *saturates* X if $X \subseteq V(M) := \bigcup M$.

4.1.2 Proof of Theorem 1.4.4

Let $q > 1$ be a real number, whose exact value will be chosen later on. We start by defining a colouring of the edges of the infinite complete graph. Let $A_0, A_1, \ldots$ be a partition of $\mathbb{N}$, such that every element of A_i precedes every element of A_{i+1} and $|A_i| = \lfloor q^i \rfloor$. We colour the edges of $G = K_{\mathbb{N}}$ such that every edge uv with $u \in A_i$ and $v \in A_j$ is red if $\min\{i, j\}$ is odd, and blue if it is even. A straightforward calculation shows that for $q = 2$, every monochromatic path P in G satisfies $\bar{d}(P) \leq 8/9$ (see Theorem 1.5 in [39]). We will improve this bound by reordering the vertices of G and then optimizing the value of q.

For convenience, we will say that the vertex $v \in A_i$ is red if i is odd and blue if i is even. We also denote by B the set of blue vertices and by R be the set of red vertices. Let b_i and r_i denote the i-th blue vertex and the i-th red vertex, respectively. We define a monochromatic red matching M_r by forming a matching between A_{2i-1} and the first $|A_{2i-1}|$ vertices of A_{2i} for each $i \geq 1$. Similarly, we define a monochromatic blue matching M_b by forming a matching between A_{2i} and the first $|A_{2i}|$ vertices of A_{2i+1} for each $i \geq 0$.

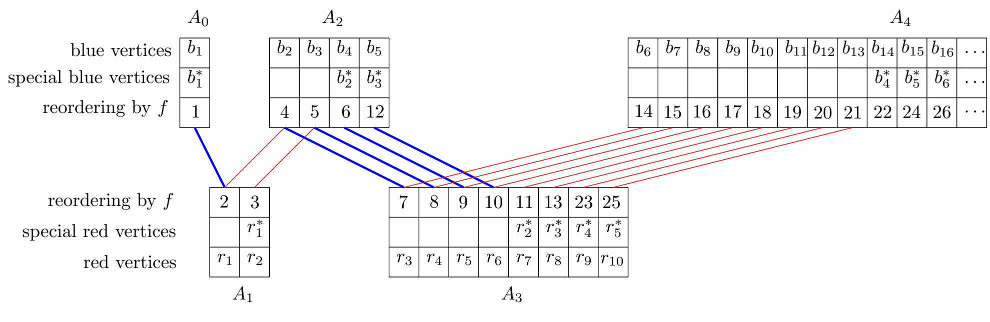


Figure 4.1: The colouring for $q = 2$ and the reordering by f.

Next, let us define a bijection $f : \mathbb{N} \to V(G)$, which will serve as a reordering of

G. Let r_t^* denote the t-th red vertex not in M_b, and b_t^* denote the t-th blue vertex not in M_r. The function f is defined as follows. We start enumerating blue vertices, in their order, until we reach b_1^*. Then we enumerate red vertices, in their order, until we reach r_1^*. Then we enumerate blue vertices again until we reach b_2^*. We continue enumerating vertices in this way, changing colours whenever we find an r_t^* or a b_t^*. (See Figure 4.1.) Finally, for every $H \subseteq G$, we define

$$\bar{d}(H; f) = \limsup_{t\to\infty} \frac{|V(H) \cap f([t])|}{t}.$$

Note that $\bar{d}(H; f)$ is the upper density of H in the reordered graph $f^{-1}(G)$.

Claim 4.1. *Let P_r and P_b be infinite monochromatic red and blue paths in G, respectively. Then $\bar{d}(P_r; f) \le \bar{d}(M_r; f)$ and $\bar{d}(P_b; f) \le \bar{d}(M_b; f)$.*

Claim 4.2. *We have*

$$\bar{d}(M_r; f),\ \bar{d}(M_b; f) \le \frac{q^2 + 2q - 1}{q^2 + 3q - 2}.$$

We can easily derive Theorem 1.4.4 from these two claims. Note that the rational function in Claim 4.2 evaluates to $(12 + \sqrt{8})/17$ at $q := \sqrt{2} + 1$. It then follows from Claim 4.1 and 4.2, that every monochromatic path P in G satisfies $\bar{d}(P; f) \le (12 + \sqrt{8})/17$. Thus we can define the desired colouring of $K_{\mathbb{N}}$, by colouring each edge ij with the colour of the edge $f(i)f(j)$ in G.

It remains to prove Claim 4.1 and 4.2. The intuition behind Claim 4.1 is that in every monochromatic red path P_r there is a red matching with the same vertex set, and that M_r has the largest upper density among all red matchings, as it contains every red vertex and has the largest possible upper density of blue vertices. Note that the proof of Claim 4.1 only uses the property that f preserves the order of the vertices inside R and inside B.

Proof of Claim 4.1. We will show $\bar{d}(P_r; f) \le \bar{d}(M_r; f)$. (The other case is analogous.) We prove that, for every positive integer k, we have $|V(P_r) \cap f([k])| \le |V(M_r) \cap f([k])|$. Assume, for contradiction, that this is not the case and let k be the minimum positive integer for which the inequality does not hold. Every red vertex is saturated by M_r, so $|V(P_r) \cap f([k]) \cap B| > |V(M_r) \cap f([k]) \cap B|$. By the minimality of k, $f(k)$ must be in P_r but not in M_r, and in particular it must be blue.

Let $f(k) \in A_{2i}$. Since $f(k) \notin M_r$, we know that $f(k)$ is not among the first $|A_{2i-1}|$ vertices of A_{2i}. Therefore, since f preserves the order of the vertices inside B, the first $|A_{2i-1}|$ blue vertices in A_{2i} are contained in $f([k])$, and hence

$$|V(P_r) \cap f([k]) \cap B| > |V(M_r) \cap f([k]) \cap B| = \sum_{j=1}^{i} |A_{2j-1}|. \tag{4.1.1}$$

On the other hand, every edge between two blue vertices is blue, so the successor of every blue vertex in P_r is red, and in particular there is a red matching between $V(P_r) \cap B$ and R saturating $V(P_r) \cap B$. By (4.1.1), this implies that the number of red neighbours of $V(P_r) \cap f([k]) \cap B$ is at least $|V(P_r) \cap f([k]) \cap B| > \sum_{j=1}^{i} |A_{2j-1}|$. Observe that by the definition of f, we have $V(P_r) \cap f([k]) \cap B \subseteq \bigcup_{j=0}^{i} A_{2j}$. Hence the red neighbourhood of $V(P_r) \cap f([k]) \cap B$ is contained in $\bigcup_{j=1}^{i} A_{2j-1}$, a contradiction. □

Proof of Claim 4.2. Let $\ell_r(t)$ and $\ell_b(t)$ denote the position of r_t^* among the red vertices and of b_t^* among the blue vertices, respectively. In other words, let $\ell_r(t) = i$ where $r_t^* = r_i$ and $\ell_b(t) = j$ where $b_t^* = b_j$ (so for example in Figure 4.1, $\ell_r(4) = 9$ and $\ell_b(4) = 14$). We claim that $f(\ell_b(t) + \ell_r(t)) = r_t^*$. Indeed, by definition of f, at $x = \ell_b(t) + \ell_r(t)$ we switched exactly t times from enumerating blue vertices to red vertices and $t-1$ times vice-versa. Hence, for $\ell_b(t-1) + \ell_r(t-1) \le k \le \ell_b(t) + \ell_r(t) - 1$, the set $f([k])$ contains exactly $t-1$ vertices outside of M_b and at least $t-1$ vertices outside of M_r. As a consequence, we obtain

$$\begin{aligned} \bar{d}(M_r; f),\ \bar{d}(M_b; f) &\le \limsup_{k\to\infty} (1 - h(k)) \\ &= \limsup_{t\to\infty} \left(1 - \frac{t-1}{\ell_r(t) + \ell_b(t) - 1}\right), \end{aligned} \tag{4.1.2}$$

where $h(k) = (t-1)/k$ if $\ell_b(t-1) + \ell_r(t-1) \le k \le \ell_b(t) + \ell_r(t) - 1$. It is easy to see that

$$\ell_r(t) = t + \sum_{j=0}^{i} |A_{2j}| \text{ if } \sum_{j=0}^{i-1} |A_{2j+1}| - |A_{2j}| < t \le \sum_{j=0}^{i} |A_{2j+1}| - |A_{2j}|, \text{ and}$$

$$\ell_b(t) = t + \sum_{j=1}^{i} |A_{2j-1}| \text{ if } \sum_{j=1}^{i-1} |A_{2j}| - |A_{2j-1}| < t - |A_0| \le \sum_{j=1}^{i} |A_{2j}| - |A_{2j-1}|.$$

Note that $\ell_r(t) - t$ and $\ell_b(t) - t$ are piecewise constant and non-decreasing. We claim that, in order to compute the right hand side of (4.1.2), it suffices to consider values of t for which $\ell_r(t) - t > \ell_r(t-1) - (t-1)$ or $\ell_b(t) - t > \ell_b(t-1) - (t-1)$. This is because we can write

$$1 - \frac{t-1}{\ell_r(t) + \ell_b(t) - 1} = \frac{1}{2} + \frac{(\ell_r(t) - t) + (\ell_b(t) - t) + 1}{2(\ell_r(t) + \ell_b(t) - 1)}.$$

In this expression, the second fraction has a positive, piecewise constant numerator and a positive increasing denominator. Therefore, the local maxima are attained precisely at the values for which the numerator increases. We will do the calculations for the case when $\ell_r(t) - t > \ell_r(t-1) - (t-1)$ (the other case is similar), in which we have

$$t = 1 + \sum_{j=0}^{i-1} (|A_{2j+1}| - |A_{2j}|) = 1 + \sum_{j=0}^{i-1} (1 + o(1)) q^{2j}(q-1) = (1 + o(1)) \tfrac{q^{2i}}{q+1},$$

$$\ell_r(t) = t + \sum_{j=0}^{i} |A_{2j}| = (1 + o(1)) \left(\tfrac{q^{2i}}{q+1} + \sum_{j=0}^{i} q^{2j} \right) = (1 + o(1)) \tfrac{(q^2+q-1)q^{2i}}{q^2-1},$$

$$\ell_b(t) = t + \sum_{j=1}^{i} |A_{2j-1}| = (1 + o(1)) \left(\tfrac{q^{2i}}{q+1} + \sum_{j=1}^{i} q^{2j-1} \right) = (1 + o(1)) \tfrac{(2q-1)q^{2i}}{q^2-1}.$$

Plugging this into (4.1.2) gives the desired result. □

4.1.3 Proof of Theorem 1.4.5

This section is dedicated to the proof of Theorem 1.4.5. A *total colouring* of a graph G is a colouring of the vertices and edges of G. Due to an argument of Erdős and

Galvin, the problem of bounding the upper density of monochromatic paths in edge coloured graphs can be reduced to the problem of bounding the upper density of monochromatic path forests in totally coloured graphs.

Definition 4.1.1 (Monochromatic path forest)**.** Given a totally coloured graph G, a forest $F \subseteq G$ is said to be a *monochromatic path forest* if $\Delta(F) \leq 2$ and there is a colour c such that all leaves, isolated vertices, and edges of F receive colour c.

Lemma 4.1.2. *For every $\gamma > 0$ and $k \in \mathbb{N}$, there is some $n_0 = n_0(k, \gamma)$ so that the following is true for every $n \geq n_0$. For every total 2-colouring of K_n, there is an integer $t \in [k, n]$ and a monochromatic path forest F with $d(F, t) \geq (12 + \sqrt{8})/17 - \gamma$.*

Some standard machinery related to Szemerédi's regularity lemma, adapted to the ordered setting, will allow us to reduce the problem of bounding the upper density of monochromatic path forests to the problem of bounding the upper density of monochromatic simple forests.

Definition 4.1.3 (Monochromatic simple forest)**.** Given a totally coloured graph G, a forest $F \subseteq G$ is said to be a *monochromatic simple forest* if $\Delta(F) \leq 1$ and there is a colour c such that all edges and isolated vertices of F receive colour c and at least one endpoint of each edge of F receives colour c.

Lemma 4.1.4. *For every $\gamma > 0$, there exists $k_0, N \in \mathbb{N}$ and $\alpha > 0$ such that the following holds for every integer $k \geq k_0$. Let G be a totally 2-coloured graph on kN vertices with minimum degree at least $(1 - \alpha)kN$. Then there exists an integer $t \in [k/8, kN]$ and a monochromatic simple forest F such that $d(F, t) \geq (12 + \sqrt{8})/17 - \gamma$.*

The heart of the proof is Lemma 4.1.4, which we shall prove in Section 4.1.6. But first, in the next two sections, we show how to deduce Theorem 1.4.5 from Lemmas 4.1.2 and 4.1.4.

4.1.4 From Path Forests to Paths

In this section we use Lemma 4.1.2 to prove Theorem 1.4.5. Our exposition follows that of Theorem 1.6 in [33].

Proof of Theorem 1.4.5. Fix a 2-colouring of the edges of $K_{\mathbb{N}}$ in red and blue. We define a 2-colouring of the vertices by colouring $n \in \mathbb{N}$ red if there are infinitely many $m \in \mathbb{N}$ such that the edge nm is red and blue otherwise.

Case 1. Suppose there are vertices x and y of the same colour, say red, and a finite set $S \subseteq \mathbb{N}$ such that there is no red path disjoint from S which connects x to y.

We partition $\mathbb{N} \setminus S$ into sets X, Y, Z, where $x' \in X$ if and only if there is a red path, disjoint from S, which connects x' to x and $y' \in Y$ if and only if there is a red path disjoint from S which connects y to y'. Note that every edge from $X \cup Y$ to Z is blue. Since x and y are coloured red, both X and Y are infinite, and by choice of x and y all edges in the bipartite graph between X and $Y \cup Z$ are blue. Hence there is a blue path with vertex set $X \cup Y \cup Z = \mathbb{N} \setminus S$.

Case 2. Suppose that for every pair of vertices x and y of the same colour c, and every finite set $S \subseteq \mathbb{N}$, there is a path from x to y of colour c which is disjoint from S.

Let γ_n be a sequence of positive reals tending to zero, and let a_n and k_n be increasing sequences of integers such that

$$a_n \geq n_0(k_n, \gamma_n) \text{ and } k_n/(a_1 + \cdots + a_{n-1} + k_n) \longrightarrow 1,$$

where $n_0(k, \gamma)$ is as in Lemma 4.1.2. Let $\mathbb{N} = (A_i)$ be a partition of $\mathbb{N}$ into consecutive intervals with $|A_n| = a_n$. By Lemma 4.1.2 there are monochromatic path forests F_n with $V(F_n) \subseteq A_n$ and initial segments $I_n \subseteq A_n$ of length at least k_n such that

$$|V(F_n) \cap I_n| \geq \left(\frac{12 + \sqrt{8}}{17} - \gamma_n \right) |I_n|.$$

It follows that for any $G \subseteq K_{\mathbb{N}}$ containing infinitely many F_n's we have

$$\bar{d}(G) \geq \limsup_{n \to \infty} \frac{|V(F_n) \cap I_n|}{a_1 + \cdots + a_{n-1} + |I_n|} \geq \limsup_{n \to \infty} \frac{12 + \sqrt{8}}{17} - \gamma_n = \frac{12 + \sqrt{8}}{17}.$$

By the pigeonhole principle, there are infinitely many F_n's of the same colour, say blue. We will recursively construct a blue path P which contains infinitely many of these F_n's. To see how this is done, suppose we have constructed a finite initial segment p of P. We will assume as an inductive hypothesis that p ends at a blue vertex v. Let n be large enough that $\min(A_n)$ is greater than every vertex in p, and F_n is blue. Let $F_n = \{P_1, \ldots, P_s\}$ for some $s \in \mathbb{N}$ and let w_i, w_i' be the endpoints of the path P_i (note that w_i and w_i' could be equal) for every $i \in [s]$. By the case assumption, there is a blue path q_1 connecting v to w_1, such that q_1 is disjoint from $A_1 \cup \cdots \cup A_n$. Similarly, there is a blue path q_2 connecting w_1' to w_2, such that q_2 is disjoint from $A_1 \cup \cdots \cup A_n \cup \{q_1\}$. Continuing in this fashion, we find disjoint blue paths $q_3, \ldots, q_s$ such that q_i connects w_{i-1}' to w_i. Hence, we can extend p to a path p' which contains all of the vertices of F_n and ends at a blue vertex. □

4.1.5 From Simple Forests to Path Forests

In this section we use Lemma 4.1.4 to prove Lemma 4.1.2. The proof is based on Szemerédi's Regularity Lemma, which we already introduced in Section 2.2. The main difference to standard applications of the Regularity Lemma is that we have to define an ordering of the reduced graph, which approximately preserves densities. This is done by choosing a suitable initial partition. It is well-known (see for instance [67]) that dense regular pairs contain almost spanning paths. We include a proof of this fact for completeness.

Lemma 4.1.5. *For $0 < \varepsilon < 1/4$ and $d \geq 2\sqrt{\varepsilon} + \varepsilon$, every ε-regular pair (A, B) with $|A| = |B|$ and density at least d contains a path with both endpoints in A and covering all but at most $2\sqrt{\varepsilon}|A|$ vertices of $A \cup B$.*

Proof. We will construct a path $P_k = (a_1 b_1 \ldots a_k)$ for every $k = 1, \ldots, \lceil (1 - \sqrt{\varepsilon})|A| \rceil$ such that $B_k := N(a_k) \setminus V(P_k)$ has size at least $\varepsilon|B|$. As $d \geq \varepsilon$, this is easy for $k = 1$. Assume now that we have constructed P_k for some $1 \leq k < (1 - \sqrt{\varepsilon})|A|$. We will show how to extend P_k to P_{k+1}. By ε-regularity of (A, B), the set $\bigcup_{b \in B_k} N(b)$ has size at least $(1 - \varepsilon)|A|$. So $A' := \bigcup_{b \in B_k} N(b) \setminus V(P_k)$ has size at least $(\sqrt{\varepsilon} - \varepsilon)|A| \geq \varepsilon|A|$. Let $B' = B \setminus V(P_k)$ and note that $|B'| \geq \sqrt{\varepsilon}|B|$ as $k < (1 - \sqrt{\varepsilon})|A|$ and $|A| = |B|$. By ε-regularity of (A, B), there exists $a_{k+1} \in A'$ with at least $(d - \varepsilon)|B'| \geq 2\varepsilon|B|$ neighbours in B'. Thus we can define $P_{k+1} = (a_1 b_1 \ldots a_k b_k a_{k+1})$, where $b_k \in B_k \cap N(a_{k+1})$. □

Before we start with the proof of Lemma 4.1.2, we will briefly describe the setup and proof strategy. Consider a totally 2-coloured complete graph $G = K_n$. Denote the sets of red and blue vertices by R and B, respectively. For $\ell \geq 4$, let $\{W_j\}_{j\in[\ell]}$ be a partition of $[n]$ such that each W_j consists of at most $\lceil n/\ell \rceil$ subsequent vertices. The partition $\{W'_j\}_{j\in[2\ell]}$, with parts of the form $W_i \cap R$ and $W_i \cap B$, refines both $\{W_j\}_{j\in[\ell]}$ and $\{R, B\}$. Suppose that $V_0 \cup \cdots \cup V_m$ is a partition obtained from Lemma 2.2.12 applied to G and $\{W'_j\}_{j\in[2\ell]}$ with parameters ε, m_0, 2ℓ and d. We define the (ε, d)-*reduced graph* G' to be the graph with vertex set $V(G') = [m]$ where ij is an edge of G' if and only if if (V_i, V_j) is an ε-regular pair of density at least d in the red subgraph of H or in the blue subgraph of H. Furthermore, we colour ij red if (V_i, V_j) is an ε-regular pair of density at least d in the red subgraph of H, otherwise we colour ij blue. As $\{V_i\}_{i\in[m]}$ refines $\{R, B\}$, we can extend this to a total 2-colouring of G' by colouring each vertex i red, if $V_i \subseteq R$, and blue otherwise. By relabelling the clusters, we can furthermore assume that $i < j$ if and only if $\max(V_i) < \max(V_j)$. Note that, by choice of $\{W_j\}_{j\in[\ell]}$, any two vertices in V_i differ by at most n/ℓ. Moreover, a simple calculation (see [87, Proposition 42]) shows that G' has minimum degree at least $(1 - d - 3\varepsilon)m$.

Given this setup, our strategy to prove Lemma 4.1.2 goes as follows. First, we apply Lemma 4.1.4 to obtain $t' \in [m]$ and a, red say, simple forest $F' \subseteq G'$ with $d(F', t') \approx (12 + \sqrt{8})/17$. Next, we turn F' into a red path forest $F \subseteq G$. For every isolated vertex $i \in V(F')$, this is straightforward as $V_i \subseteq R$ by the refinement property. For every edge $ij \in E(F')$ with $i \in R$, we apply Lemma 4.1.5 to obtain a red path that almost spans (V_i, V_j) and has both ends in V_i. So the union F' of these paths and vertices is indeed a red path forest. Since the vertices in each V_i do not differ too much, it will follow that $d(F, t) \approx (12 + \sqrt{8})/17$ for $t = \max(V_{t'})$.

Proof of Lemma 4.1.2. Suppose we are given $\gamma > 0$ and $k \in \mathbb{N}$ as input. Let $k_0, N \in \mathbb{N}$ and $\alpha > 0$ be as in Lemma 4.1.4 with input $\gamma/4$. We choose constants $d, \varepsilon > 0$ and $\ell, m_0 \in \mathbb{N}$ satisfying

$$2\sqrt{\varepsilon} + \varepsilon \leq 1/\ell, d \leq \alpha/8 \text{ and } m_0 \geq 4N/d, 2k_0N.$$

We obtain M from Lemma 2.2.12 with input ε, m_0 and 2ℓ. Finally, set $n_0 = 16k\ell MN$.

Now let $n \geq n_0$ and suppose that K_n is an ordered complete graph on vertex set $[n]$ and with a total 2-colouring in red and blue. We have to show that there is an integer $t \in [k, n]$ and a monochromatic path forest F such that $|V(F) \cap [t]| \geq ((12 + \sqrt{8})/17 - \gamma)t$.

Denote the red and blue vertices by R and B, respectively. Let $\{W'_j\}_{j \in [\ell]}$ refine $\{R, B\}$ as explained in the above setting. Let $\{V_0, \ldots, V_m\}$ be a partition of $[n]$ with respect to $G = K_n$ and $\{W'_j\}_{j \in [\ell]}$ as detailed in Lemma 2.2.12 with totally 2-coloured (ε, d)-reduced graph G'' of minimum degree $\delta(G'') \geq (1 - 4d)m$. Set $k' = \lfloor m/N \rfloor \geq k_0$ and observe that the subgraph G' induced by G'' in $[k'N]$ satisfies $\delta(G') \geq (1 - 8d)m \geq (1 - \alpha)m$ as $m \geq 4N/d$. Thus we can apply Lemma 4.1.4 with input G', k', $\gamma/4$ to obtain an integer $t' \in [k'/8, k'N]$ and a monochromatic (say red) simple forest $F' \subseteq G'$ such that $d(F', t') \geq (12 + \sqrt{8})/17 - \gamma/4$.

Set $t = \max(V_{t'})$. We have that $V_{t'} \subseteq W_j$ for some $j \in [\ell]$. Recall that $i < j$ if and only if $\max(V_i) < \max(V_j)$ for any $i, j \in [m]$. It follows that $V_i \subseteq [t]$ for all $i \leq t'$. Hence

$$t \geq t'|V_1| \geq \frac{k'}{8}|V_1| \geq \left\lfloor \frac{m}{N} \right\rfloor \frac{(1-\varepsilon)n}{8m} \geq \frac{n}{16N}. \tag{4.1.3}$$

This implies $t \geq k$ by choice of n_0. Since $[t]$ is covered by $V_0 \cup W_j \cup \bigcup_{i \in [t']} V_i$, it follows that

$$\begin{aligned} t'|V_1| &\geq t - |V_0| - |W_j| \\ &\geq \left(1 - \varepsilon\frac{n}{t} - \frac{4}{\ell}\frac{n}{t}\right)t \\ &\geq \left(1 - 16\varepsilon N - \frac{64N}{\ell}\right)t \qquad \text{(by (4.1.3))} \\ &\geq \left(1 - \frac{\gamma}{2}\right)t. \end{aligned} \tag{4.1.4}$$

For every edge $ij \in E(F')$ with $V_i \subseteq R$, we apply Lemma 4.1.5 to choose a path P_{ij} which starts and ends in V_i and covers all but at most $2\sqrt{\varepsilon}|V_1|$ vertices of each V_i and V_j. We denote the isolated vertices of F' by I'. For each $i \in I'$ we have $V_i \subseteq R$. Hence the red path forest $F := \bigcup_{i \in I'} V_i \cup \bigcup_{ij \in E(F')} P_{ij} \subseteq K_n$ satisfies

$$
\begin{aligned}
|V(F) \cap [t]| &= \sum_{i \in I'} |V_i \cap [t]| + \sum_{ij \in E(F')} |V(P_{ij}) \cap [t]| \\
&\geq \sum_{i \in I' \cap [t']} |V_i| + \sum_{i \in V(F'-I') \cap [t'],} (|V_i| - 2\sqrt{\varepsilon}|V_1|) \\
&\geq (1 - 2\sqrt{\varepsilon})|V_1||V(F') \cap [t']| \\
&\geq (1 - 2\sqrt{\varepsilon}) \left(\frac{12 + \sqrt{8}}{17} - \frac{\gamma}{4} \right) t'|V_1| \\
&\overset{(4.1.4)}{\geq} \left(\frac{12 + \sqrt{8}}{17} - \gamma \right) t
\end{aligned}
$$

as desired. □

4.1.6 Upper Density of Simple Forests

In this section we prove Lemma 4.1.4. For a better overview, we shall define all necessary constants here. Suppose we are given $\gamma' > 0$ as input and set $\gamma = \gamma'/4$. Fix a positive integer $N = N(\gamma)$ and let $0 < \alpha \leq \gamma/(8N)$. The exact value of N will be determined later on. Let $k_0 = \lceil 8/\gamma \rceil$ and fix a positive integer $k \geq k_0$. Consider a totally 2-coloured graph G' on $n = kN$ vertices with minimum degree at least $(1 - \alpha)n$.

Denote the sets of red and blue vertices by R and B, respectively. As it turns out, we will not need the edges inside R and B. So let G be the spanning bipartite subgraph, obtained from G' by deleting all edges within R and B. For each red vertex v, let $d_b(v)$ be the number of blue edges incident to v in G. Let $a_1 \leq \cdots \leq a_{|R|}$ denote the degree sequence taken by $\{d_b(v) : v \in R\}$. The whole proof of Lemma 4.1.4 revolves around analysing this sequence.

Fix an integer $t = t(\gamma, N, k)$ and subsets $R' \subseteq R$, $B' \subseteq B$. The value of t and nature of R', B' will be determined later (R' and B' will be chosen as initial segments of R and B). The following two observations explain our interest in the sequence $a_1 \leq \cdots \leq a_{|R|}$.

Claim 4.3. *If $a_j > j - t$ for all $1 \leq j \leq |R'| - 1$, then there is a blue simple forest covering all but at most t vertices of $R' \cup B$.*

Proof. We write $R' = \{v_1, \ldots, v_{|R'|}\}$ so that $d_b(v_i) \leq d_b(v_j)$ for every $1 \leq i \leq j \leq |R'|$. By assumption, we have $d_b(v_j) \geq a_j > j - t$ for all $1 \leq j \leq |R'| - 1$. Thus we can greedily select a blue matching containing $\{v_t, v_{t+1}, \ldots, v_{|R'|-1}\}$, which covers all but t vertices of R'. Together with the rest of B, this forms the desired blue simple forest. □

Claim 4.4. *If $a_i < i + t$ for all $1 \leq i \leq |B'| - t$, then there is a red simple forest covering all but at most $t + \alpha n$ vertices of $R \cup B'$.*

Proof. Let X' be a minimum vertex cover of the red edges in the subgraph of G induced by $R \cup B'$. If $|X'| \geq |B'| - t - \alpha n$, then by König's theorem there exists a red matching covering at least $|B'| - t - \alpha n$ vertices of B'. This together with the vertices in R yields the desired red simple forest.

Suppose now that $|X'| < |B'| - t - \alpha n$. Since every edge between $R \setminus (X' \cap R)$ and $B' \setminus (X' \cap B')$ is blue, we have for every vertex v in $R \setminus (X' \cap R)$,

$$d_b(v) \geq |B'| - |X' \cap B'| - \alpha n = |X' \cap R| + |B'| - |X'| - \alpha n > |X' \cap R| + t.$$

Here, we used that $|X' \cap B'| = |X'| - |X' \cap R|$ in the second step and $|B'| - |X'| - \alpha n > t$ in the last step. In particular, this implies $a_i \geq i + t$ for $i = |X' \cap R| + 1$. Since

$$|X' \cap R| + 1 \leq |X'| + 1 < |B'| - t - \alpha n + 1,$$

we have $|X' \cap R| + 1 \leq |B'| - t$, contradicting the assumption. □

Motivated by this, we introduce the following definitions.

Definition 4.1.6 (Oscillation, $\ell^+(t)$, $\ell^-(t)$)**.** Let $a_1, \ldots, a_n$ be a non-decreasing sequence of non-negative real numbers. We define its *oscillation* as the maximum value T, for which there exist indices $i, j \in [n]$ with $a_i - i \geq T$ and $j - a_j \geq T$. For all $0 < t \le T$, set

$$\ell^+(t) = \min\{i \in [n]:\ a_i \ge i + t\},$$
$$\ell^-(t) = \min\{j \in [n]:\ a_j \le j - t\}.$$

Suppose that the degree sequence $a_1, \ldots, a_{|R|}$ has oscillation T and fix some positive integer $t \le T$. We define $\ell = \ell(t)$ and $\lambda = \lambda(t)$ by

$$\ell = \ell^+(t) + \ell^-(t) = \lambda t. \tag{4.1.5}$$

The next claim combines Claims 4.3 and 4.4 into a density bound of a monochromatic simple forest in terms of the ratio $\ell/t = \lambda$. (Note that, in practice, the term αn will be of negligible size.)

Claim 4.5. *There is a monochromatic simple forest $F \subseteq G$ with*

$$d(F, \ell + t) \ge \frac{\ell - \alpha n}{\ell + t} = \frac{\lambda t - \alpha n}{(1 + \lambda)t}.$$

Proof. Let $R' = R \cap [\ell + t]$ and $B' = B \cap [\ell + t]$ so that $\ell^+(t) + \ell^-(t) = \ell = |R'| + |B'| - t$. Thus we have either $\ell^-(t) \ge |R'|$ or $\ell^+(t) > |B'| - t$. If $\ell^-(t) \ge |R'|$, then $a_j > j - t$ for every $1 \le j \le |R'| - 1$. Thus Claim 4.3 provides a blue simple forest F covering all but at most t vertices of $[\ell + t]$. On the other hand, if $\ell^+(t) > |B'| - t$, then $a_i < i + t$ for every $1 \le i \le |B'| - t$. In this case Claim 4.4 yields a red simple forest F covering all but at most $t + \alpha n$ vertices of $[\ell + t]$. □

Claim 4.5 essentially reduces the problem of finding a dense simple forest to a problem about bounding the ratio ℓ/t in integer sequences. It is, for instance, not hard to see that we always have $\ell \ge 2t$ (which, together with the methods of the previous two subsections, would imply the bound $\bar{d}(P) \ge 2/3$ of Erdős and Galvin). The following lemma provides an essentially optimal lower bound on $\ell/t = \lambda$. Note that for $\lambda = 4 + \sqrt{8}$, we have $\frac{\lambda}{\lambda+1} = (12 + \sqrt{8})/17$.

Lemma 4.1.7. *For all $\gamma \in \mathbb{R}^+$, there exists $N \in \mathbb{N}$ such that, for all $k \in \mathbb{R}^+$ and all* *sequences with oscillation at least kN, there exists a real number $t \in [k, kN]$ with*

$$\ell := \ell^+(t) + \ell^-(t) \geq \left(4 + \sqrt{8} - \gamma\right) t.$$

The proof of Lemma 4.1.7 is deferred to the next subsection. We now finish the proof of Lemma 4.1.4. Set $N = N(\gamma)$ to be the integer returned by Lemma 4.1.7 with input γ (recall that $\gamma = \gamma'/4$). In order to use Lemma 4.1.7, we have to bound the oscillation of $a_1, \dots, a_{|R|}$:

Claim 4.6. *The degree sequence $a_1, \dots, a_{|R|}$ has oscillation $T \geq kN/8$ or there is a monochromatic simple forest $F \subseteq G$ with $d(F, n) \geq (12 + \sqrt{8})/17 - \gamma$.*

Before we prove Claim 4.6, let us see how this implies Lemma 4.1.4.

Proof of Lemma 4.1.4. By Claim 4.6, we may assume that the sequence $a_1, \dots, a_{|R|}$ has oscillation at least $kN/8$. By Lemma 4.1.7, there is a real number $t' \in [k/8, kN/8]$ with

$$\ell = \ell^+(t') + \ell^-(t') \geq (4 + \sqrt{8} - \gamma)t'.$$

Let $t = t(\gamma, N, k) = \lceil t' \rceil$. Since the a_i's are all integers, we have $\ell^+(t) = \ell^+(t')$ and $\ell^-(t) = \ell^-(t')$. Let $F \subseteq G$ be the monochromatic simple forest obtained from Claim 4.5. As $n = kN$, $\ell \geq t' \geq k/8 \geq 1/\gamma$, $\alpha \leq \gamma/(8N)$, and by (4.1.5), it follows that

$$\begin{aligned}
d(F, \ell + t) &\geq \frac{\ell - \alpha n}{\ell + t} = \frac{1 - \alpha n/\ell}{1 + \frac{t}{\ell}} \geq \frac{1 - 8\alpha N}{1 + \frac{t'}{\ell} + \frac{1}{\ell}} \geq \frac{1}{1 + \frac{t'}{\ell}} - 2\gamma \\
&\geq \frac{1}{1 + \frac{1}{4 + \sqrt{8} - \gamma}} - 2\gamma = \frac{4 + \sqrt{8} - \gamma}{5 + \sqrt{8} - \gamma} - 2\gamma \\
&\geq \frac{4 + \sqrt{8}}{5 + \sqrt{8}} - 4\gamma = \frac{12 + \sqrt{8}}{17} - \gamma',
\end{aligned}$$

as desired. □

To finish, it remains to show Claim 4.6. The proof uses König's theorem and is similar to the proof of Claim 4.4.

Proof of Claim 4.6. Let X be a minimum vertex cover of the red edges. If $|X| \geq |B| - (1/8 + \alpha)n$, then König's theorem implies that there is a red matching covering all but at most $(1/8 + \alpha)n$ blue vertices. Thus adding the red vertices, we obtain a red simple forest F with $d(F, kN) \geq 7/8 - \alpha \geq (12 + \sqrt{8})/17 - \gamma$. Therefore, we may assume that $|X| < |B| - (1/8 + \alpha)n$. Every edge between $R \setminus (X \cap R)$ and $B \setminus (X \cap B)$ is blue. So there are at least $|R| - |X \cap R|$ red vertices v with

$$d_b(v) \geq |B| - |X \cap B| - \alpha n = |X \cap R| + |B| - |X| - \alpha n > |X \cap R| + n/8.$$

This implies that $a_i \geq i + n/8$ for $i = |X \cap R| + 1$. (See Figure 4.2.)

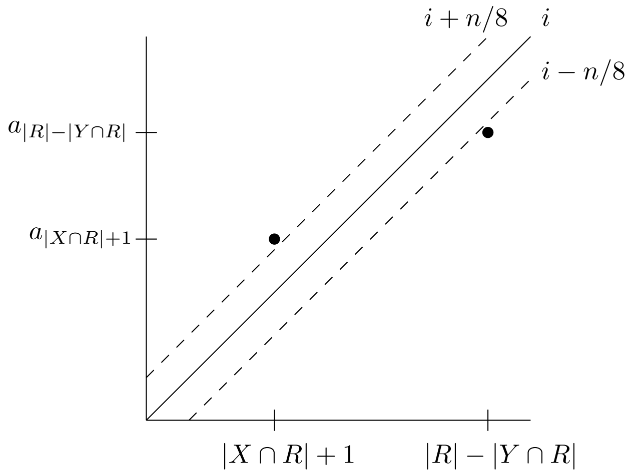


Figure 4.2: The sequence $a_1, \ldots, a_{|R|}$ has oscillation at least $kN/8$.

Let Y be a minimum vertex cover of the blue edges. Using König's theorem as above, we can assume that $|Y| \leq |R| - n/8$. Every edge between $R \setminus (Y \cap R)$ and $B \setminus (Y \cap B)$ is red. It follows that there are at least $|R| - |Y \cap R|$ red vertices v with

$$d_b(v) \leq |Y \cap B| = |Y| - |Y \cap R| \leq |R| - |Y \cap R| - \frac{n}{8}.$$

This implies that $a_j \leq j - n/8$ for $j = |R| - |Y \cap R|$. Thus $a_1, \ldots, a_{|R|}$ has oscillation at least $n/8 = kN/8$. □

4.1.7 Sequences and Oscillation

We now present the quite technical proof of Lemma 4.1.7. We will use the following definition and related lemma in order to describe the oscillation from the diagonal.

Definition 4.1.8 (k-good, $u_{\mathrm{o}}(k)$, $u_{\mathrm{e}}(k)$)**.** Let $a_1, \ldots, a_n$ be a sequence of non-negative real numbers and let k be a positive real number. We say that the sequence is k-*good* if there exists an odd i and an even j such that $a_i \geq k$ and $a_j \geq k$. If the sequence is k-good, we define for all $0 < t \leq k$

$$\begin{aligned} u_{\mathrm{o}}(t) &= a_1 + \cdots + a_{i_o - 1} \quad \text{where } i_o = \min\{i\colon\ a_i \geq t, i \text{ odd}\}, \\ u_{\mathrm{e}}(t) &= a_1 + \cdots + a_{i_e - 1} \quad \text{where } i_e = \min\{i\colon\ a_i \geq t, i \text{ even}\}. \end{aligned}$$

Lemma 4.1.9. *For all $\gamma \in \mathbb{R}^+$ there exists $N \in \mathbb{N}$ such that for all $k \in \mathbb{R}^+$ and all (kN)-good sequences, there exists a real number $t \in [k, kN]$ with*

$$u_o(t) + u_e(t) \geq \left(3 + \sqrt{8} - \gamma\right) t.$$

First we use Lemma 4.1.9 to prove Lemma 4.1.7.

Proof of Lemma 4.1.7. Given $\gamma > 0$, let N be obtained from Lemma 4.1.9. Let $k \in \mathbb{R}^+$ and $a_1, \ldots, a_n$ be a sequence with oscillation at least kN. Suppose first that $a_1 \geq 1$. Partition $[n]$ into a family of non-empty intervals $I_1, \ldots, I_r$ with the following properties:

- For every odd i and every $j \in I_i$, we have $a_j \geq j$.
- For every even i and every $j \in I_i$, we have $a_j < j$.

Define $s_i = \max\left\{|a_j - j|\colon\ j \in I_i\right\}$. Intuitively, this is saying that the values in the odd indexed intervals are "above the diagonal" and the values in the even indexed intervals are "below the diagonal" and s_i is the largest gap between sequence values and the "diagonal" in each interval.

Since $a_1, \ldots, a_n$ has oscillation at least kN, the sequence $s_1, \ldots, s_r$ is (kN)-good and thus by Lemma 4.1.9, there exists $t \in [k, kN]$ such that

$$u_{\mathrm{o}}(t) + u_{\mathrm{e}}(t) \geq \left(3 + \sqrt{8} - \gamma\right) t. \tag{4.1.6}$$

Since the sequence $a_1, a_2, \ldots, a_n$ is non-decreasing, $a_j - j$ can decrease by at most one in each step and thus we have $|I_i| \geq s_i$ for every $i \in [r-1]$. Moreover, we can find bounds on $\ell^+(t)$ and $\ell^-(t)$ in terms of the s_i:

- $\ell^+(t)$ must lie in the interval I_i with the smallest odd index i_o such that $s_{i_o} \geq t$, therefore $\ell^+(t) \geq s_1 + \cdots + s_{i_o - 1} = u_{\mathrm{o}}(t)$.

- $\ell^-(t)$ must lie in the interval I_j with the smallest even index i_e such that $s_{i_e} \geq t$. Moreover, it must be at least the t-th element in this interval, therefore $\ell^-(t) \geq s_1 + \cdots + s_{i_e - 1} + t = u_{\mathrm{e}}(t) + t$.

Combining the previous two observations with (4.1.6) gives

$$\ell^+(t) + \ell^-(t) \geq u_{\mathrm{o}}(t) + u_{\mathrm{e}}(t) + t \geq \left(4 + \sqrt{8} - \gamma\right) t,$$

as desired.

If $0 \leq a_1 < 1$, we start by partitioning $[n]$ into a family of non-empty intervals $I_1, \ldots, I_r$ with the following properties:

- For every even i and every $j \in I_i$, we have $a_j \geq j$.

- For every odd i and every $j \in I_i$, we have $a_j < j$.

From this point, the proof is analogous. □

Finally, it remains to prove Lemma 4.1.9. The proof is by contradiction and the main strategy is to find a subsequence with certain properties which force the sequence to become negative eventually.

Proof of Lemma 4.1.9. Let $\rho = 3 + \sqrt{8} - \gamma$ and let $m := m(\rho)$ be a positive integer which will be specified later. Suppose that the statement of the lemma is false for $N = 6 \cdot 4^m$ and let $a_1, \ldots, a_n$ be an (Nk)-good sequence without t as in the statement. We first show that a_i has a long strictly increasing subsequence. Set

$$I = \{i \colon\ a_i \geq k, a_i > a_j \text{ for all } j < i\},$$

denote the elements of I by $i_1 \leq i_2 \leq \cdots \leq i_r$ and let $a'_j = a_{i_j}$. Consider any $j \in [r-1]$ and suppose without loss of generality that i_{j+1} is odd. For δ small enough, this implies $u_o(a'_j + \delta) = a_1 + \cdots + a_{i_{j+1}-1} \geq a'_1 + \cdots + a'_j$, and $u_e(a'_j + \delta) \geq a_1 + \cdots + a_{i_{j+1}} \geq a'_1 + \cdots + a'_{j+1}$. By assumption we have $u_o(a'_j + \delta) + u_e(a'_j + \delta) < \rho \cdot (a'_j + \delta)$. Hence, letting $\delta \to 0$ we obtain $2\left(a'_1 + \cdots + a'_j\right) + a'_{j+1} \leq \rho a'_j$, which rearranges to

$$a'_{j+1} \leq (\rho - 2)a'_j - 2\left(a'_1 + \cdots + a'_{j-1}\right). \tag{4.1.7}$$

In particular, this implies $a'_{j+1} \leq (\rho - 2)a'_j < 4a'_j$. Moreover, we have $a'_1 \leq u_o(k)$ if i_1 is even and $a'_1 \leq u_e(k)$ if i_1 is odd. Therefore,

$$\begin{aligned} 6k \cdot 4^m = kN \leq a'_r &< 4^r \cdot a'_1 \\ &\leq 4^r \max\{u_o(k), u_e(k)\} \\ &\leq 4^r (u_o(k) + u_e(k)) \\ &< 4^r \cdot \rho k \\ &< 6k \cdot 4^r \end{aligned}$$

and thus $r \geq m$.

Finally, we show that any sequence of reals satisfying (4.1.7), will eventually become negative, but since a'_i is non-negative this will be a contradiction.

We start by defining the sequence $b_1, b_2, \ldots$ recursively by $b_1 = 1$ and $b_{i+1} = (\rho - 2)b_i - 2(b_1 + \cdots + b_{i-1})$. Note that

$$\begin{aligned} b_{i+1} &= (\rho - 2)b_i - 2(b_1 + \cdots + b_{i-1}) \\ &= (\rho - 1)b_i - b_i - 2(b_1 + \cdots + b_{i-1}) \\ &= (\rho - 1)b_i - ((\rho - 2)b_{i-1} - 2(b_1 + \cdots + b_{i-2})) - 2(b_1 + \cdots + b_{i-1}) \\ &= (\rho - 1)b_i - \rho b_{i-1} \end{aligned}$$

So equivalently the sequence is defined by,

$$b_1 = 1,\ b_2 = \rho - 2,\ \text{and } b_{i+1} = (\rho - 1)b_i - \rho b_{i-1} \text{ for } i \geq 2.$$

It is known that a second order linear recurrence relation whose characteristic polynomial has non-real roots will eventually become negative (see [15]). Indeed, the characteristic polynomial $x^2 - (\rho - 1)x + \rho$ has discriminant $\rho^2 - 6\rho + 1 < 0$ and so its roots $\alpha, \bar{\alpha}$ are non-real.[1] Hence the above recursively defined sequence has the closed form of $b_i = z\alpha^i + \bar{z}\bar{\alpha}^i = 2\mathrm{Re}\left(z\alpha^i\right)$ for some complex number z. By expressing $z\alpha^i$ in polar form we can see that $b_m < 0$ for some positive integer m. Note that the calculation of m only depends on ρ.

We will be done if we can show that $a'_j \le a'_1 b_j$ for all $j \in [m]$ and every sequence $a'_1, \dots, a'_m$ of non-negative reals satisfying (4.1.7); so suppose this is false. Let $\{a'_j\}_{j=1}^m$ be a counterexample which coincides with $\{a'_1 b_j\}_{j=1}^m$ on the longest initial subsequence among all counterexamples. By assumption, there is some $s \in [m]$ such that $a'_s > a'_1 b_s$. Let p be the minimum value such that $a'_p \ne a'_1 b_p$. Clearly $p > 1$. Applying (4.1.7) to $j = p - 1$ we see that

$$\begin{aligned} a'_p &\le (\rho - 2)a'_{p-1} - 2(a'_1 + \cdots + a'_{p-2}) \\ &= (\rho - 2)a'_1 b_{p-1} - 2(a'_1 b_1 + \cdots + a'_1 b_{p-2}) \\ &= a'_1((\rho - 2)b_{p-1} - 2(b_1 + \cdots + b_{p-2})) \\ &= a'_1 b_p \end{aligned}$$

and thus $a'_p < a'_1 b_p$.

Let $\beta = (a'_1 b_p - a'_p)/a'_1 > 0$. Now consider the sequence a''_j where $a''_j = a'_j$ for $j < p$ and $a''_j = a'_j + \beta a'_{j-p+1}$ for $j \ge p$. Then $a''_p = a'_1 b_p = a''_1 b_p$. Clearly, this new sequence satisfies (4.1.7) for every $j < p$. Furthermore, we have

$$\begin{aligned} a''_{p+j} &= a'_{p+j} + \beta a'_{j+1} \\ &\le (\rho - 2)a'_{p+j-1} - 2(a'_1 + \ldots + a'_{p+j-2}) + \beta(\rho - 2)a'_j - 2\beta(a'_1 + \ldots + a'_{j-1}) \\ &= (\rho - 2)a''_{p+j-1} - 2(a''_1 + \ldots + a''_{p+j-2}) \end{aligned}$$

for every $j \ge 0$. Hence, the whole sequence satisfies (4.1.7). We also have $a''_s \ge a'_s > a'_1 b_s = a''_1 b_s$. This contradicts the fact that a'_j is a counterexample which

[1] Note that $3 + \sqrt{8}$ is the positive root of $x^2 - 6x$, which is the reason why we chose $\rho := 3 + \sqrt{8} - \gamma$.

coincides with $a_1' b_j$ on the longest initial subsequence. □

4.2 Upper Density of Monochromatic Subgraphs

4.2.1 Overview

In this section, we are going to prove a number of different results concerning Ramsey densities of infinite graphs. In Sections 4.2.2 and 4.2.3 we start with a collection of constructions and examples. We move on to Section 4.2.4, where we introduce the concept of ultrafilters and prove a couple of results which demonstrate how they can be used to embed infinite graphs.

In Section 4.2.5 we move on to bipartite Ramsey densities and prove the following two results.

Theorem 1.4.7 (Corsten–DeBiasio–McKenney)**.** *Every 2-coloured $K_{\mathbb{N},\mathbb{N}}$ contains a monochromatic path of upper density at least* $1/2$*.*

Theorem 1.4.8 (Corsten–DeBiasio–McKenney)**.** *The vertices of every* 2*-edge-coloured $K_{\mathbb{N},\mathbb{N}}$ can be partitioned into a finite set and at most two monochromatic paths.*

Recall that Example 4.2.9 shows that both are best possible. The idea of the proof is to partition the vertices of a 2-edge-coloured graph into a finite set and two sets V_1, V_2 with the following property: for all $i \in [2]$, there is a colour c_i such that any two vertices $u, v \in V_i$ are the endpoints of infinitely many internally disjoint finite paths which are monochromatic in colour c_i. It is then easy to cover each of V_1 and V_2 with a monochromatic infinite path.

In Section 4.2.6 we discuss Ramsey densities of infinite trees and prove the following result.

Theorem 1.4.10 (Corsten–DeBiasio–McKenney)**.** $\overline{\mathrm{Rd}}(T) \geq 1/2$ *for every infinite tree T.*

We will prove Theorem 1.4.10 separately for trees with a vertex of infinite degree, and for trees with a path of infinite length. In the first case we will show that every

2-edge-coloured $K_{\mathbb{N}}$ contains a monochromatic subgraph H of density at least $1/2$ in which one vertex is adjacent to every other vertex and every vertex has infinite degree. We then describe an embedding algorithm to surjectively embed any tree with a vertex of infinite degree in H. In the second case we will show that every 2-edge-coloured $K_{\mathbb{N}}$ contains a monochromatic subgraph H of density at least $1/2$ which is infinitely connected (that is, it remains connected after deleting any finite set of vertices). We then describe an embedding algorithm to surjectively embed any tree with a path of infinite length in H.

In Section 4.2.7 we will discuss Ramsey densities of graphs with bounded chromatic number and prove the following result.

Theorem 1.4.11 (Corsten–DeBiasio–McKenney)**.** *Let $r, k \geq 2$ be integers and let G be a one-way k-locally finite graph.*

(i) If $k = 2$, then $\overline{\mathrm{Rd}}_r(G) \geq 1/r$.

(ii) If $k \geq 3$, then $\overline{\mathrm{Rd}}(G) \geq 1/(2(k-1))$.

(iii) If $k \geq 3$, then $\overline{\mathrm{Rd}}_r(G) \geq \left(\sum_{i=0}^{(k-2)r+1} (r-1)^i\right)^{-1} = (1 + o_k(1))r^{-(k-2)r-1}$.

The proof of Theorem 1.4.11 is quite technical and relies heavily on ultrafilters Section 4.2.4. Finally, in Section 4.2.8, we will discuss Ramsey-dense graphs and prove the following result.

Theorem 1.4.15 (Corsten–DeBiasio–McKenney)**.** *If G is an infinite graph with finite ruling number, then G is r-Ramsey-dense for every $r \in \mathbb{N}$.*

The proof of Theorem 1.4.15 also relies on ultrafilters but is less technical. We further use the Rado colouring (see Section 4.2.2) to prove the following partial converse.

Theorem 1.4.19 (Corsten–DeBiasio–McKenney)**.** *Let G be an infinite graph with disjoint ruling sets $F_1, F_2, \ldots$ satisfying $|F_n| \leq \log_2(n)$ for all sufficiently large n. Then G is not Ramsey-dense.*

4.2.2 Preliminaries

Properties of upper density

In this subsection, we collect some basic properties of upper densities.

Proposition 4.2.1.

(i) If $A_1, A_2 \subseteq \mathbb{N}$, *then* $\overline{\mathrm{d}}(A_1 \cup A_2) \leq \overline{\mathrm{d}}(A_1) + \overline{\mathrm{d}}(A_2)$.

(ii) If $A_1, A_2 \subseteq \mathbb{N}$ *are disjoint, then* $\overline{\mathrm{d}}(A_1 \cup A_2) \geq \underline{\mathrm{d}}(A_1) + \overline{\mathrm{d}}(A_2)$.

(iii) If $A_1, A_2 \subseteq \mathbb{N}$, *then* $\underline{\mathrm{d}}(A_1 \cup A_2) \leq \underline{\mathrm{d}}(A_1) + \overline{\mathrm{d}}(A_2)$.

(iv) If $A_1, A_2 \subseteq \mathbb{N}$ *are disjoint, then* $\underline{\mathrm{d}}(A_1 \cup A_2) \geq \underline{\mathrm{d}}(A_1) + \underline{\mathrm{d}}(A_2)$.

The Rado Graph

Recall that the *Rado graph* $\mathcal{R}$ is the graph with vertex-set $\mathbb{N}$ with $mn \in E(\mathcal{R})$ for some $m < n$ if and only if the mth digit in the binary expansion of n is 1 (starting with the least significant bit). The Rado graph first appeared in work of Ackermann [1] and was more systematically studied by Erdős and Rényi [43], and by Rado [98].

We say that an infinite G graph has the *extension property* if for every pair of disjoint finite sets $F, F' \subseteq V(G)$, there is a vertex $v \in V(G) \setminus (F \cup F')$ such that v is adjacent to every $w \in F$ and not adjacent to any $w' \in F'$. The following well-known theorem shows why this property is useful.

Theorem 4.2.2. *Any two infinite graphs satisfying the extension property are isomorphic.*

Furthermore, it is not hard to see that the Rado graph $\mathcal{R}$ and, with probability 1, the infinite random graph (the graph on $\mathbb{N}$ in which every edge is present independently with probability $1/2$) both satisfy the extension property. Hence, with probability 1, the infinite random graph is isomorphic to the Rado graph.

Recall that an infinite graph G is 0-ruled if it has no finite dominating set. Observe that G is 0-ruled if and only if G satisfies the "non-adjacency" half of the extension property above, i.e. if for every finite $F' \subseteq V(G)$ there is a vertex $v \in V(G) \setminus F'$ such that v is not adjacent to any $w' \in F'$. We will call G 0-*coruled* if G satisfies only the "adjacency" half of extension property, i.e. for every finite $F \subseteq V(G)$ there is a $v \in V(G) \setminus F$ such that v is adjacent to every $w \in F$. Using this, it is easy (and very similar to the proof of Theorem 4.2.2) to prove the following proposition. Here, given graphs G and H, we write $G \preceq H$ if G is isomorphic to a spanning subgraph of H.

Proposition 4.2.3. *G is 0-ruled if and only if $G \preceq \mathcal{R}$. On the other hand, G is 0-coruled if and only if $\mathcal{R} \preceq G$.*

We note that there exist graphs which are both 0-ruled and 0-coruled, but not isomorphic to $\mathcal{R}$ (for example, the half-graph discussed below).

The *bipartite Rado graph* $\mathcal{R}_2$ is the bipartite graph with biparts $V = \{v_0, v_1, \ldots\}$ and $U = \{u_0, u_1, \ldots\}$ with $u_i v_j, v_i u_j \in E(\mathcal{R}_2)$ for some $i < j$ if the ith digit in the binary expansion of j is 1 (starting with the least significant bit). The bipartite Rado graph has similar properties as the Rado graph. In particular, with probability 1, $\mathcal{R}_2$ is isomorphic to the random subgraph of $K_{\mathbb{N},\mathbb{N}}$ in which every edge of $K_{\mathbb{N},\mathbb{N}}$ is present independently with probability $1/2$.

The Rado Colouring

The *Rado colouring* $\rho : E(K_{\mathbb{N}}) \to \{0, 1\}$ is the 2-edge-colouring of $K_{\mathbb{N}}$ defined as follows: given $m < n$, define $\rho(\{m, n\})$ to be the value of the mth digit in the binary expansion of n (starting with the least significant bit). Note that the graphs induced by taking all edges of colour 0 or 1 are both isomorphic to the Rado graph. The Rado colouring further has the following simple but important property.

Observation 4.2.4. *For every $F \subseteq \mathbb{N}$ and $i \in \{0, 1\}$, we have*

$$d\left(\bigcap\nolimits_{v \in F} N_i(v)\right) = 2^{-|F|}, \tag{4.2.1}$$

where $N_i(v)$ denotes the set of vertices $u \in \mathbb{N}$ with $\rho(uv) = i$.

The Half-Graph

The *(infinite) half-graph* $\mathcal{H}$ is the bipartite graph on $\mathbb{N}$ (with one part being the even numbers and the other part being the odd numbers) defined as follows: Given an even number $i \in \mathbb{N}$ and an odd number $j \in \mathbb{N}$, we have $ij \in E(\mathcal{H})$ if and only if $i > j$. This graph has the very interesting property that every even vertex has finite degree and every odd vertex has cofinite degree.

Relationship of Important Graph Parameters

We will present some examples here to get a better understanding of how the different properties discussed in this section are related.

Example 4.2.5.

(i) There is an infinite graph G with $\mathrm{rul}(G) = 0$, but $\chi(G) = \infty$ (Rado graph).

(ii) There is an infinite graph G with $\chi(G) = 2$, but $\mathrm{rul}(G) = \infty$ (complete bipartite graph).

(iii) There is an infinite graph G with $\mathrm{rul}(G) = 0$ and $\chi(G) = 2$, but $\mathrm{degen}(G) = \infty$ (bipartite Rado graph, half-graph).

(iv) There is a locally finite infinite graph G with $\chi(G) = \infty$ (infinite collection of finite cliques of increasing size).

Section 4.2.2 gives an overview of these examples and some results from the introduction.

4.2.3 Constructions

In this subsection we will construct some colourings in order to give upper bounds on upper densities.

Example 4.2.6. We have $\overline{\mathrm{Rd}}(G) \leq 1/(\chi(G) - 1)$ for every connected graph G.

Proof. Suppose first that $\chi(G)$ is finite and let $k = \chi(G) - 1$. Partition $\mathbb{N}$ into k sets $V_1, \ldots, V_k$, each of which has density $1/k$ (for example, using residues modulo k). Colour an edge red if it is inside one of the sets $V_1, \ldots, V_k$ and blue otherwise. Since the blue subgraph is k-partite, there is no blue copy of G. Furthermore, since G is connected, every red copy of G must lie entirely in V_i for some $i \in [k]$ and therefore its upper density is at most $1/k$.

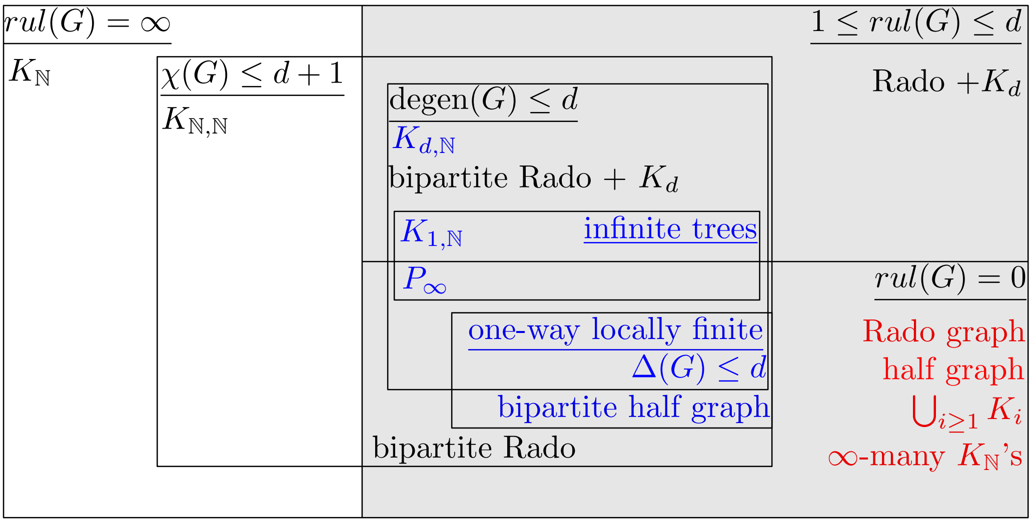


Figure 4.3: The lightly shaded area represents graphs which are Ramsey-dense. The blue text represents graphs G for which $\overline{\text{Rd}}(G) > 0$. The red text represents graphs G which are Ramsey-dense, but $\overline{\text{Rd}}(G) = 0$.

If $\chi(G)$ is infinite, then the above example shows that $\overline{\text{Rd}}(G) \leq 1/k$ for every $k \in \mathbb{N}$ and thus $\overline{\text{Rd}}(G) = 0$. □

This result implies that the Rado graph has Ramsey upper density 0. Recall however, that the Rado graph is r-Ramsey-dense for every $r \in \mathbb{N}$.

Corollary 4.2.7. *The Rado graph $\mathcal{R}$ has Ramsey upper density* 0.

Proof. By Observation 4.2.4, every infinite independent set in $\mathcal{R}$ has density 0 and therefore $\chi(\mathcal{R}) = \infty$. Therefore, the result follows from Example 4.2.6. □

Example 4.2.8. Let $r \in \mathbb{N}$.

(i) If T is a D-ary tree for $D \geq 2$, then $\overline{\text{Rd}}_r(T) \leq \frac{1}{r}(1 + \frac{1}{D})$.

(ii) We have $\overline{\text{Rd}}_r(T_\infty) \leq 1/r$.[2]

(iii) There exists a locally finite tree T with $\overline{\text{Rd}}_r(T) \leq 1/r$.

[2]Recall that T_∞ is the infinite tree in which every vertex has infinite degree.

Proof. Partition $\mathbb{N}$ by residues mod r, that is $\mathbb{N} = R_1 \cup \ldots \cup R_r$ where R_i is the set of all $n \in \mathbb{N}$ with $n \equiv i \pmod r$. We define an r-edge-colouring as follows: if $m \in R_i$ and $n > m$, colour the edge mn with colour i. Note that if $n \not\equiv i \pmod r$, then n has exactly $\lfloor (n-1)/r \rfloor$ neighbours of colour i and all of them are in $[n-1]$.

(i) Let T be a D-ary tree and suppose we have a copy of T of colour i. For all $n > 0$, let V'_n be the set of vertices in $V(T) \cap [n]$ which are not congruent to $i \pmod r$ and let $t_n = |V'_n|$. Since any vertex $m \in V'_n$ can only have neighbours (of colour i) in $R_i \cap [n-1]$, we must have $D \cdot t_n \leq (n-1)/r$. So

$$\frac{|V(T) \cap [n]|}{n} \leq \frac{\lceil \frac{n}{r} \rceil + t_n}{n} \leq \frac{\lceil \frac{n}{r} \rceil + \frac{n-1}{rD}}{n} \xrightarrow{n \to \infty} \frac{1}{r}(1 + \frac{1}{D}).$$

(ii) Suppose $U \subseteq \mathbb{N}$ is the vertex-set of a copy of T_∞ in colour $i \in [r]$. Since every vertex in $\mathbb{N} \setminus R_i$ has only finitely many neighbours in colour i, we have $U \subseteq R_i$ and thus $\overline{\mathrm{d}}(U) \leq 1/r$.

(iii) Let $0 < d_1 < d_2 < \ldots$ be an increasing sequence. Let H be a tree in which every vertex on level i has degree d_i. We can repeat the argument from case (i), except now we have $t_n/n \to 0$ as $n \to \infty$. □

Example 4.2.9. There is an r-edge-colouring of $K_{\mathbb{N},\mathbb{N}}$ in which every monochromatic path has upper density at most $1/r$. In particular, it is not possible to cover all but finitely many vertices with less than r monochromatic paths.

Proof. Let A and B be the parts of $K_{\mathbb{N},\mathbb{N}}$ (that is, A consists of all odd numbers and B consists of all even numbers) and partition both of them into r parts $A_1, \ldots, A_r$ and $B_1, \ldots, B_r$, each of density $1/(2r)$. For all $i, j \in [r]$, colour every edge between A_i and B_j by $(i - j) \mod r$. It is easy to see that every part is incident to exactly one other part of each colour and therefore, every monochromatic path can cover at most two parts, finishing the proof. □

Proposition 4.2.10. *Let K be a countably infinite complete multipartite graph.*

(i) If K has at least two infinite parts, or infinitely many vertices in finite parts, then K is not Ramsey-dense.

(ii) If K has exactly one infinite part and exactly n vertices in finite parts, then

$$\frac{1}{2^{2n-1}} \leq \overline{\mathrm{Rd}}(K) \leq \frac{1}{2^n}.$$

Proof. Colour the edges of $K_{\mathbb{N}}$ with the Rado colouring. If K has at least two infinite parts, or infinitely many vertices in finite parts, then K contains a spanning copy of $K_{\mathbb{N},\mathbb{N}}$; let (A, B) be such a spanning copy of $K_{\mathbb{N},\mathbb{N}}$. Let $a_1, a_2, \ldots$ be the elements of A. Then B is contained in the neighbourhood of $a_1, \ldots, a_n$, and hence, by Observation 4.2.4, has density at most 2^{-n}, for each n. Hence B must have density 0. Repeating the argument the other way around shows that A has density 0, too.

Suppose now that K has exactly one infinite part and exactly n vertices in finite parts. Then by Observation 4.2.4, we have $\overline{\mathrm{Rd}}(K) \leq \frac{1}{2^n}$.

To see $\overline{\mathrm{Rd}}(K) \geq \frac{1}{2^{2n-1}}$, assume that we are given an arbitrary 2-edge-colouring of $K_{\mathbb{N}}$. We will proceed similarly as in a well-known proof of Ramsey's theorem. Let $v_1 \in \mathbb{N}$ be an arbitrary vertex. By Proposition 4.2.1, there is a colour c_1 such that $V_2 := N_{c_1}(v)$ satisfies $\overline{\mathrm{d}}(N_i(v)) \geq 1/2$. Now do the following recursively for all $j \in [2, 2n-1]$. Let $v_j \in V_j$ be an arbitrary vertex. It follows from Proposition 4.2.1 that there exists a colour $c_j \in [2]$ such that $V_{j+1} := N_{c_j}(v_j; V_j)$ satisfies $\overline{\mathrm{d}}(V_{j+1}) \geq \overline{\mathrm{d}}(V_j)/2 \geq 2^{-j}$. By the pigeonhole principle, there is some a set $J = \{j_1, \ldots, j_n\} \subseteq [2n-1]$ of size n such that c_j is constant on J. Hence, $v_{j_1}, \ldots, v_{j_n}, V_{2n}$ induce a monochromatic copy of $K_{1,\ldots,1,\mathbb{N}}$ with upper density $\overline{\mathrm{d}}(V_{2n}) \geq 2^{-2n+1}$. Clearly $K_{1,\ldots,1,\mathbb{N}}$ contains a spanning copy of K, finishing the proof. □

4.2.4 Ultrafilters and Embedding

The concept of ultrafilters will play a very important role in this section.

Definition 4.2.11. Given a set X, a set system $\mathscr{U} \subseteq 2^X$ is called an *ultrafilter* if the following properties hold.

(i) $X \in \mathscr{U}$ and $\emptyset \notin \mathscr{U}$.

(ii) If $A \in \mathscr{U}$ and $A \subseteq B \subseteq X$, then $B \in \mathscr{U}$.

(iii) If $A, B \in \mathscr{U}$, then $A \cap B \in \mathscr{U}$.

(iv) For all $A \subseteq X$, either $A \in \mathscr{U}$ or $X \setminus A \in \mathscr{U}$, or

(iv)′ $\mathscr{U}$ is maximal among all families satisfying (i) - (iii).

A family satisfying (i)-(iii) is called a *filter*. Conditions (iv) and $(iv)'$ are equivalent for filters (see [70, Chapter 11, Lemma 2.3]) and we will make use whichever is more convenient for the current application. Let us list some additional properties of ultrafilters.

Proposition 4.2.12. *If $\mathscr{U}$ is an ultrafilter on X, then the following is true.*

(i) if $A_1, \ldots, A_n \in \mathscr{U}$, then $A_1 \cap \ldots \cap A_n \in \mathscr{U}$.

(ii) If $A_1 \cup \ldots \cup A_n \in \mathscr{U}$ are pairwise disjoint, then there is exactly one $i \in [n]$ with $A_i \in \mathscr{U}$.

Informally, we think of sets $A \in \mathscr{U}$ as "large" sets. A common example of an ultrafilter are the so called trivial ultrafilters $\mathscr{U}_x := \{A \subseteq X : x \in A\}$ for $x \in X$. It is not hard to see that an ultrafilter is trivial if and only if it contains a finite set.

We say that an ultrafilter $\mathscr{U}$ on N is *positive* if every set $A \in \mathscr{U}$ has positive upper density. Positive ultrafilters play a crucial role in the rest of this section.

Proposition 4.2.13. *If $X \subseteq \mathbb{N}$ is infinite, then there exists a non-trivial ultrafilter $\mathscr{U}$ on X. Furthermore, there exists a positive ultrafilter $\mathscr{U}$ on $\mathbb{N}$.*

Proof. To prove the first part of the theorem, let $\mathcal{F}_0 \subseteq 2^X$ be the set of all cofinite subsets of X and apply Zorn's lemma to

$$\{\mathcal{F} \subseteq 2^X : \mathcal{F}_0 \subseteq \mathcal{F} \text{ and } \mathcal{F} \text{ satisfies (i) - (iii) in Definition 4.2.11}\}$$

to get a maximal such family $\mathscr{U}$, which must be an ultrafilter. Finally, if A is finite, $\mathscr{U}$ contains the cofinite set A^c and hence $A \notin \mathscr{U}$.

To prove the second part, let $\mathcal{F}_1 \subseteq 2^{\mathbb{N}}$ be the set of all sets of lower density 1 and apply Zorn's lemma to

$$\{\mathcal{F} \subseteq 2^{\mathbb{N}} : \mathcal{F}_1 \subseteq \mathcal{F} \text{ and } \mathcal{F} \text{ satisfies (i) - (iii) in Definition 4.2.11}\}$$

to get a maximal such family $\mathscr{U}$, which must be an ultrafilter. Furthermore, if $A \subseteq X$ has upper density 0, then $X \setminus A$ has lower density 1 and consequently $A \notin \mathscr{U}$. □

Definition 4.2.14 (Vertex-colouring induced by $\mathscr{U}$)**.** Let $r \geq 2$ and suppose the edges of an infinite graph G are coloured with r colours. Let $\mathscr{U}$ be a non-trivial ultrafilter on $V(G)$. Define a colouring $c_{\mathscr{U}} : V(G) \to [r]$ where $c_{\mathscr{U}}(v) = i$ if and only if $N_i(v) \in \mathscr{U}$. Since $V(G) \setminus \{v\} \in \mathscr{U}$ for all $v \in V(G)$, it follows from Proposition 4.2.12 (ii) that $c_{\mathscr{U}}$ is well defined. We call $c_{\mathscr{U}}$ the vertex-colouring induced by $\mathscr{U}$.

The following propositions allow us to use ultrafilters to embed the desired subgraphs in the proof of Theorem 1.4.11 and Theorem 1.4.15.

Proposition 4.2.15. *Let $k \geq 2$, let G be a one-way k-locally finite graph and let H be graph such that $\{U_1, \dots, U_k\}$ is a partition of $V(H)$ with $|U_1| = \cdots = |U_k| = \infty$ and, for every $i \in [k]$ and every finite subset $W \subseteq U_1 \cup \cdots \cup U_{i-1}$, the set of common neighbours of W in U_i is infinite. Then, there is an embedding f of G into H with $U_1 \subseteq \operatorname{rg} f$.*

Proposition 4.2.16. *Let G be an infinite 0-ruled graph and let H be a graph having the property that for every finite set of vertices $W \subseteq V(H)$, the set of common neighbours of W is infinite. Then we can embed G surjectively into H.*

Given a k-partite graph G with parts $V_1, \dots, V_k$ and a set $S \subseteq V(G)$, the *left neighbourhood cascade* of S is the tuple $(S_1, \dots, S_k)$, where $S_k = S \cap V_k$, and for all $1 \leq i \leq k-1$, $S_i = (S \cup \bigcup_{j=i+1}^{k} N(S_j)) \cap V_i$.

Proof. Let $V_1 \cup V_2 \cup \cdots \cup V_k$ be a partition of $V(G)$ into independent sets which witness the fact that G is one-way k-locally-finite (in particular V_1 is infinite). We will assume that $V(G) = \mathbb{N}$ and that it is ordered with the natural ordering. We will construct an embedding f iteratively in finite pieces. Initially, f is the empty embedding. Then, for each $n \in \mathbb{N}$, we will proceed as follows: let

$$S_n = \{\min(V_i \setminus \operatorname{dom} f) : i \in [k] \text{ with } V_i \setminus \operatorname{dom} f \neq \emptyset\}.$$

That is, S_n contains the smallest not yet embedded vertex of each V_i which is not completely embedded yet. Let $(T_{1,n}, \dots, T_{k,n})$ be left neighbourhood cascade of S_n in G. We will now extend f to cover $\bigcup_{i\in[k]} T_{i,n}$. Observe that $T_{i,n}$ is disjoint from $\operatorname{dom} f$ for all $i \in [k]$ since we embedded the whole left neighbourhood cascade in every previous step. Since V_1 is infinite, $T_{1,n}$ is non-empty. Let $T'_{1,n} \subseteq U_1 \setminus \operatorname{rg} f$ be the set of $|T_{1,n}|$ smallest vertices in $U_1 \setminus \operatorname{rg} f$ and extend f by embedding $T_{1,n}$ into $T'_{1,n}$ arbitrarily. By assumption $T'_{1,n}$ has infinitely many common neighbours in U_2. Since $\operatorname{rg} f$ is finite, we can select a set $T'_{2,n} \subseteq (U_2 \cap N^{\cap}(T'_{1,n})) \setminus \operatorname{rg} f$ of size $|T_{2,n}|$. Extend f by embedding $T_{2,n}$ into $T'_{2,n}$ arbitrarily. Similarly, we can extend f by embedding $T_{i,n}$ into appropriate sets $T'_{i,n}$ for all $i = 3, \dots, k$.

Since we maintain a partial embedding of G into H throughout the process and every vertex of G will eventually be embedded (by choice of S_n which contains the smallest not yet embedded vertex of $V(G)$), the resulting function f defines an embedding of G into H. Since we cover the smallest not-yet covered vertex of U_1 in each step, we further have $U_1 \subseteq \operatorname{rg} f$. □

Proof of Proposition 4.2.16. Let $v_1, v_2, \dots$ be an enumeration of $V(G)$ and let $u_1, u_2, \dots$ be an enumeration of $V(H)$. Suppose we have already embedded $\{v_1, \dots, v_{n-1}\}$ (and no vertex v_i with $i \geq n$) into H for some $n \in \mathbb{N}$. Call the partial embedding f,[3] and let u_{i_n} be the vertex of smallest index in $V(H) \setminus f(\{v_1, \dots, v_{n-1}\})$. Since G is 0-ruled, there exists a vertex $v_p \in V(G)$ with $p \geq n$ such that v_p has no neighbours in $\{v_1, \dots, v_{n-1}\}$. We set $f(v_p) = u_{i_n}$ and we embed all vertices from $\{v_n, \dots, v_{p-1}\}$ into H one at a time (using the property that every finite subset of $V(G)$ has infinitely many common neighbours). Continuing in this way, we clearly obtain an embedding of G into H. Since we embed the vertex of lowest index among the remaining vertices in every step, this embedding is surjective. □

4.2.5 Bipartite Ramsey Densities

In this section we prove Theorem 1.4.8. An infinite graph G is said to be *infinitely connected* if G remains connected after removing any finite set of vertices. Note that every vertex of an infinitely connected graph has infinite degree. Given some

[3] We will update f in every step.

set of vertices $S \subseteq V(G)$, we say that S is *infinitely connected* if $G[S]$ is infinitely connected. Similarly, we call a set $S \subseteq V(G)$ *infinitely linked* if for all distinct $u, v \in S$, there are infinitely many internally vertex-disjoint paths in G from u to v (note that the internal vertices of these paths need not be contained in the set S). Note that every infinitely connected set is also infinitely linked but the converse is not true (for example, both parts of $K_{\mathbb{N},\mathbb{N}}$ are infinitely linked but not connected). Further note that if $S_1, \ldots, S_k$ are sets, each of which is infinitely linked, then there are disjoint paths $P_1, \ldots, P_k$ such that $P_1 \cup \ldots \cup P_k$ covers $S_1 \cup \ldots \cup S_k$.

If G is a coloured graph and c is a colour, we say that G is *infinitely connected in c* if G_c (the spanning subgraph of G with all edges of colour c) is infinitely connected. A set $S \subseteq V(G)$ is *infinitely connected in colour c* (*infinitely linked in colour c*) if S is infinitely connected (infinitely linked) when restricted to G_c. S is called monochromatic infinitely connected (infinitely linked) if it is infinitely connected in some colour c.

The following proposition easily implies Theorem 1.4.8.

Proposition 4.2.17. *Every* 2*-edge-coloured* $K_{\mathbb{N},\mathbb{N}}$ *can be partitioned into a finite set and two monochromatic infinitely linked sets* X *and* Y.

Proof of Proposition 4.2.17. Let V_1, V_2 be the parts of the bipartite graph and let $\mathscr{U}_1, \mathscr{U}_2$ be non-trivial ultrafilters on V_1 and V_2. For $i = 1, 2$, let $B_i \subseteq V_i$ be the blue vertices in the induced vertex-colouring and let $R_i = V_i \setminus B_i$ be the red vertices.

Case 1 ($|R_1| = |R_2| = |B_1| = |B_2| = \infty$). If there is an infinite red matching M between R_1 and R_2, then $X := R_1 \cup R_2$ is infinitely linked in red. Indeed, if $v_1, v_2 \in R_1$ or $v_1, v_2 \in R_2$, then they have infinitely many common red neighbours using properties of the ultrafilter. If $v_1 \in R_1$ and $v_2 \in R_2$, we will construct infinitely many internally disjoint paths of length 5 between $x_0 := v_1$ and $x_5 := v_2$ as follows: let $x_2x_3 \in E(M)$ so that $x_2 \in R_1$ and $x_3 \in R_2$, and let x_1 be a common red neighbour of x_0 and x_2 (of which we have infinitely many as above) and x_4 be a common neighbour of x_3 and x_5. It is clear that $x_0, \ldots, x_5$ defines a red path and that we can construct infinitely many internally disjoint paths like this. If there is no infinite red matching between R_1 and R_2, then there is a finite set S so that $X := (R_1 \cup R_2) \setminus S$ induces a complete blue bipartite graph with parts of infinite size and hence is infinitely linked in blue. Similarly, there is a set $Y \subseteq B_1 \cup B_2$ which is co-finite in $B_1 \cup B_2$ and infinitely linked in red or infinitely linked in blue.

Case 2. Suppose without loss of generality that R_1 is finite. It is easy to verify that $X = B_1 \cup B_2$ is infinitely linked in blue and $Y := R_2$ is infinitely linked in red. □

The following corollary might be useful to make progress on Question 1.4.6.

Corollary 4.2.18. *Suppose the edges of $K_{\mathbb{N}}$ are 3-coloured and that there is some colour α such that all but finitely many vertices have infinite degree in colour α. Then, there is a monochromatic copy of P_∞ with upper density at least $1/2$.*

Proof. Let G_α be the spanning subgraph induced by all edges of colour α. By deleting vertices if necessary we may assume that every vertex in G_α has infinite degree. We separate two cases.

Case 1 (G_α is infinitely connected)**.** In this case G_α contains a spanning path by Lemma 4.2.20 below.

Case 2 (G_α is not infinitely connected)**.** In this case there is a finite set S so that $G_\alpha[\mathbb{N} \setminus S]$ is disconnected. Let V_1 be one component and $V_2 := \mathbb{N} \setminus (S \cup V_1)$. Then V_1 and V_2 induce a 2-edge-coloured bipartite graph and both V_1 and V_2 are infinite since every vertex has infinite degree in G_α. Thus the result follows from Theorem 1.4.8. □

4.2.6 Trees

In this subsection, we will deduce Theorem 1.4.10 from the following four lemmas.

Lemma 4.2.19. *For any 2-edge-colouring of $K_{\mathbb{N}}$, there are sets R and S such that*

(i) $R \cup S$ is cofinite,

(ii) if R is infinite, then it is infinitely connected in red, and

(iii) if S is infinite, then it is infinitely connected in one of the colours.

Lemma 4.2.20. *Let T be an infinite tree which contains an infinite path. If G is an infinitely connected graph, then G contains a spanning copy of T.*

Lemma 4.2.21. *Let G be a 2-edge-coloured $K_{\mathbb{N}}$. There exists a set $A \subseteq \mathbb{N}$, a vertex $v \in A$, and a colour c such that every vertex in $H := G_c[A]$ has infinite degree and $\overline{\mathrm{d}}(N_H(v)) \geq 1/2$.*

Lemma 4.2.22. *Let T be an infinite tree with at least one vertex of infinite degree. If G is a graph in which every vertex has infinite degree, then for all $v \in V(G)$, G contains a copy of T covering $N_G(v)$.*

It is now easy to prove Theorem 1.4.10.

Proof of Theorem 1.4.10. Suppose the edges of $K_{\mathbb{N}}$ are coloured with two colours. If T does not have an infinite path, it must have one vertex of infinite degree and therefore the theorem follows immediately from Lemmas 4.2.21 and 4.2.22. So suppose T has an infinite path. By Lemma 4.2.19, there is an infinite set A with $\overline{\mathrm{d}}(A) \geq 1/2$ and a colour c, so that the induced subgraph on A is infinitely connected in c. By Lemma 4.2.20, there is a monochromatic copy of T spanning A and we are done. □

It remains to prove the four lemmas. We will use basic properties of ordinals, see [70, Chapter 6] for an introduction.

Proof of Lemma 4.2.19. Fix a 2-edge-colouring of $K_{\mathbb{N}}$. We define a sequence of sets R_α, S_α, for all ordinals α, as follows. Let $S_0 = \mathbb{N}$. For each α, we define R_α to be the set of vertices in S_α whose blue neighbourhood has finite intersection with S_α, and we set $S_{\alpha+1} = S_\alpha \setminus R_\alpha$. If λ is a limit ordinal, then we define S_λ to be the intersection of the sets S_α, for $\alpha < \lambda$.

Note that the sets R_α are pairwise disjoint, and hence there is some countable ordinal γ such that $R_\alpha = \emptyset$ for all $\alpha \geq \gamma$. Let γ^* be the minimal ordinal such that R_{γ^*} is finite; it follows then that $R_\beta = \emptyset$ for all $\beta > \gamma^*$. Set

$$R = \bigcup \{R_\alpha : \alpha < \gamma^*\}.$$

(Note that γ^* may be 0, in which case $R = \emptyset$.)

Suppose that R is infinite. Then $\gamma^* > 0$ and R_α is infinite for all $\alpha < \gamma^*$. Let $u, v \in R$ with $u \in R_\alpha$ and $v \in R_\beta$ for some $\alpha \leq \beta < \gamma^*$. It follows that the red neighbourhoods of both u and v are cofinite in R_β. Since R_β is infinite, this implies that there is a red path of length 2 connecting u and v, even after removing a finite set of vertices. Hence R is infinitely connected in red.

Set $S = S_{\gamma^*+1}$. Then $R \cup S = \mathbb{N} \setminus R_{\gamma^*}$, so $R \cup S$ is cofinite. Moreover, since $R_{\gamma^*+1} = \emptyset$, it follows that for every $v \in S$, the blue neighbourhood of v has infinite intersection with S. Now suppose that S is not infinitely connected in blue. Then there is a finite set $F \subseteq S$ and a partition $S \setminus F = X \cup Y$ such that X and Y are both nonempty, and every edge between X and Y is red. Note that X and Y must both be infinite, since if $x_0 \in X$ and $y_0 \in Y$ then $X \cup F$ and $Y \cup F$ must contain the blue neighbourhoods of x_0 and y_0 (both of which are infinite) respectively. But then the red graph restricted to $X \cup Y = S \setminus F$ is infinitely connected. □

Proof of Lemma 4.2.20. Let $t_1 \in V(T)$ and $v_1 \in V(G)$ (we think of t_1 as being the root of the tree and v_1 as the embedding of the root in G). We will build an embedding f of T into G recursively, in finite pieces, at each stage ensuring that we add the first vertices of $V(T) \setminus \operatorname{dom} f$ and $V(G) \setminus \operatorname{rg} f$ into the domain and range of f respectively. Initially, let $f(t_1) = v_1$ and $v_{last} = v_1$. Since T has an infinite path, it has an infinite path starting at t_1. Denote by P_T such a path. We will proceed as described by the following Algorithm 1.

Algorithm 1

1: **while** True **do**
2: Let $v_{next} := \min(V(G) \setminus \operatorname{rg} f)$.
3: Let $Q_{next} \subseteq G$ be a finite path from v_{last} to v_{next} which is internally disjoint from $\operatorname{rg} f$.
4: Let $T_{next} \subseteq V(T)$ be the next $|V(Q_{next})| - 1$ vertices of $P_T \setminus \operatorname{dom} f$.
5: Extend f by embedding $V(T_{next}) \setminus \{v_{last}\}$ into $V(Q_{next})$.
6: Update $v_{last} := v_{next}$.
7: Let $X \subseteq \operatorname{dom} f$ be the set of $t \in \operatorname{dom} f$ for which $S_t := N_T(t) \setminus \operatorname{dom} f \neq \emptyset$.
8: **for** $t \in S$ **do**
9: Embed $\min(S)$ into an arbitrary vertex in $N_G(f(t)) \setminus \operatorname{rg} f$.

First, note that we can always follow lines 3 and 9 of Algorithm 1 since G is infinitely connected and in particular every vertex has infinite degree. Let $f : V(T) \to V(G)$ be the function produced by Algorithm 1. We need to prove that f

is well-defined, surjective and an embedding of T.

Clearly, every vertex on the path P_T will be eventually embedded. Furthermore, since we always embed the smallest not yet embedded neighbour of every previously embedded $t \in V(T)$ in line 9, every other vertex will be embedded eventually as well. Therefore, f is well defined. Since in line 2 we choose the smallest not yet covered vertex v_{next} and later embed a vertex into v_{next}, f is surjective. Finally, by construction of f, it defines an embedding of T into G (whenever a new vertex $t \in T$ is embedded, its parent t' is already embedded and we make sure that $f(t)$ is adjacent to $f(t')$). □

Proof of Lemma 4.2.21. Fix a 2-edge-colouring of $K_{\mathbb{N}}$. We claim that there are sets $R_\alpha, B_\alpha, S_\alpha$ for all ordinals α with the following properties.

(i) There is a unique ordinal γ^* such that $R_\alpha \cup B_\alpha$ is infinite for all $\alpha < \gamma^*$, finite for $\alpha = \gamma^*$ and empty for all $\alpha > \gamma^*$. We denote $R = \bigcup_{\alpha<\gamma^*} R_\alpha$ and $B = \bigcup_{\alpha<\gamma^*} B_\alpha$.

(ii) $S_\alpha = S_{\alpha'}$ for all ordinals $\alpha, \alpha' > \gamma^*$. We denote $S = S_{\gamma^*+1}$.

(iii) R, B, S are pairwise disjoint and $R \cup B \cup S$ is cofinite.

(iv) If $v \in R_\gamma$ for some ordinal γ, then v has finitely many blue neighbours in $S \cup \bigcup_{\alpha\geq\gamma} R_\alpha$.

(v) If $v \in B_\gamma$ for some ordinal γ, then v has finitely many red neighbours in $S \cup \bigcup_{\alpha\geq\gamma} B_\alpha$.

(vi) Every $v \in S$ has infinitely many neighbours of both colours in S.

Indeed, we can proceed similarly as in Lemma 4.2.19: Let $S_0 = \mathbb{N}$. For each α, we define R_α to be the set of vertices in S_α whose blue neighbourhood has finite intersection with S_α, and B_α to be the set of vertices in $S_\alpha \setminus R_\alpha$ whose red neighbourhood has finite intersection with $S_\alpha \setminus R_\alpha$. We then set $S_{\alpha+1} = S_\alpha \setminus (R_\alpha \cup B_\alpha)$. Observe that, if $R_\alpha \cup B_\alpha$ is finite, then $R_{\alpha+1} \cup B_{\alpha+1}$ is empty. Hence, the rest of the construction is analogous.

If $R \cup B$ is empty, then let $A = S$ and choose an arbitrary vertex $v \in S$. Since A is cofinite in $\mathbb{N}$, either the blue or the red neighbourhood of v in A has upper density at least $1/2$. Since every vertex in A has infinite degree in both colours, we are done.

If $R \cup B$ is non-empty it must be infinite (by the way R and B are defined). Since $R \cup B \cup S = (R \cup S) \cup (B \cup S)$ is cofinite, we may assume without loss of generality that R is non-empty and $\overline{\mathrm{d}}(R \cup S) \geq 1/2$. Let $A = R \cup S$ and let $v \in R_0$ be arbitrary (if R is non-empty, then R_0 must be infinite). Clearly every vertex in A has infinite red degree in A and since v has only finitely many blue neighbours in A, we are done. □

Proof of Lemma 4.2.22. Let $t_1 \in V(T)$ be a vertex of infinite degree and let $v_1 = v$ from the statement of the theorem (again we think of t_1 as being the root of the tree and v_1 as the embedding of the root in G). We will build an embedding f of T into G recursively, in finite pieces, at each stage adding one more child of every previously embedded $t \in T$ (unless all children have been embedded already). The embedding strategy is very similar to that in the proof of Lemma 4.2.20. Initially, let $f(t_1) = v_1$. We will use the following Algorithm 2.

Algorithm 2

1: **while** True **do**
2:     **for** $t \in \operatorname{dom} f$ **do**
3:         **if** $S := N_T(t) \setminus \operatorname{dom} f$ is non-empty **then**
4:             Embed $\min(S)$ into $\min(N_G(f(t)) \setminus \operatorname{rg} f)$.

First, note that we can always follow line 4 of Algorithm 1 since every vertex in G has infinite degree. Let $f : V(T) \to V(G)$ be the function produced by Algorithm 2. We need to prove that f is well-defined, an embedding of T and that $N_G(v) \subseteq \operatorname{dom} f$.

Since we always embed the smallest not yet embedded neighbour of every previously embedded $t \in V(T)$ in line 4, every other vertex will be embedded eventually as well. Therefore, f is well defined. Furthermore, by construction of f, it defines an embedding of T into G (whenever a new vertex $t \in T$ is embedded, its parent t' is already embedded and we make sure that $f(t)$ is adjacent to $f(t')$). Finally note that we are infinitely often in line 4 when $t = t_1$ since $N_T(t_1)$ is infinite.

Since we always choose the smallest available vertex in $N_G(v) \setminus \operatorname{rg} f$, it follows that $N_G(v) \subseteq \operatorname{rg} f$. □

4.2.7 Graphs of Bounded Chromatic Number

In this section we will prove Theorem 1.4.11.

Proof of Theorem 1.4.11. (i). Let $r \geq 2$ and let G be a one-way 2-locally finite graph. Suppose the edges of $K_\mathbb{N}$ are coloured with r colours and let $\mathscr{U}$ be a non-trivial ultrafilter on $\mathbb{N}$. Let $c_\mathscr{U}$ be the vertex-colouring induced by $\mathscr{U}$ and for all $i \in [r]$, let A_i be the set of vertices receiving colour i. Without loss of generality, suppose that $\overline{\mathrm{d}}(A_1) \geq 1/r$. If $A_i \notin \mathscr{U}$, then $A_1^c \in \mathscr{U}$ and therefore every finite set $S \subseteq A_1$ has infinitely many common neighbours of colour 1 in A_1^c. Hence, using Proposition 4.2.15 with $U_1 = A_1$ and $U_2 = A_1^c$, we can find a monochromatic (in colour 1) copy of G containing A_1. If $A_1 \in \mathscr{U}$, then every finite set $S \subseteq A_1$ has infinitely many common neighbours of colour 1 in A_1. Hence, since every one-way k-locally finite graph is 0-ruled, we can use 4.2.16 to find a monochromatic copy of G (in colour 1) covering A_1. Since $\overline{\mathrm{d}}(A_1) \geq 1/r$, this finishes the proof.

(ii). Let $k \geq 2$ and let G be a one-way k-locally finite graph. Suppose the edges of $K_\mathbb{N}$ are coloured with 2 colours and let $\mathscr{U}_1$ be a non-trivial ultrafilter on $\mathbb{N}$. Let $c_{\mathscr{U}_1}$ be the vertex-colouring induced by $\mathscr{U}_1$ and for all $i \in [2]$, let $A_{1,i}$ be the set of vertices receiving colour i. Choose $i_1 \in [2]$ so that $A_{1,i_1} \in \mathscr{U}_1$ and let $i_1' = 3 - i_1$. Now let $\mathscr{U}_2$ be a non-trivial ultrafilter on $V_2 = A_{1,i_1'}$ and let $c_{\mathscr{U}_2}$ be the vertex-colouring of V_2 induced by $\mathscr{U}_2$ (unless V_2 is finite, in this case we will stop the iteration). For all $i \in [2]$, let $A_{2,i}$ be the set of vertices receiving colour i. Choose i_2 so that $A_{2,i_2} \in \mathscr{U}_2$ and let $i_2' = 3 - i_2$. Let $V_3 := A_{2,i_2'}$ and continue in this manner until the point at which there exist $t \in \mathbb{N}$ so that V_t is finite or there are $j \in [2]$, and a set $I \subseteq [t]$ with $|I| = k - 1$ and $A_{i,j} \in \mathscr{U}_i$ for all $i \in I$. Note that, by the pigeonhole principle, $t \leq 2k - 3$ and suppose without loss of generality that $j = 1$. Set $V_{t+1} := V_t \setminus A_{t,1}$. One of the sets $A_{1,i_1}, A_{2,i_2}, \dots, A_{t,i_t}, V_{t+1}$ has upper density at least $1/(t+1) \geq 1/(2k-2)$. If, say, $\overline{\mathrm{d}}(A_{\ell,i_\ell}) \geq 1/(2k-2)$ for some $\ell \in [t]$, then applying Proposition 4.2.16 with colour i_ℓ gives a monochromatic copy of G covering A_{ℓ,i_ℓ}. Otherwise, we have $\overline{\mathrm{d}}(V_{t+1}) \geq 1/(2k-2)$ (and in particular we did not stop because V_t is finite). Set $U_1 = V_{t+1}$, and let $\{U_2, \ldots, U_k\} = \{A_{i,1} : i \in I\}$ in reverse order (that is U_2 corresponds to $\max(I)$ and so on). Then, using properties of our ultrafilters, for every $s \in [k-1]$, every finite set $S \subseteq U_1 \cup \ldots U_s$ has infinitely many common neighbours of colour 2 in U_{s+1}. Hence, applying Proposition 4.2.15 with colour 2 gives a monochromatic copy of G covering V_{t+1}.

(iii). We will use the following notation. Given $i_1, i_2 \in \mathbb{N}_0$, and $L_1 \in \mathbb{N}^{i_1}$ and $L_2 \in \mathbb{N}^{i_2}$, we write $L_1 \prec L_2$ if L_1 is an initial segment of L_2. Furthermore, given $L = (j_1, \ldots, j_i) \in \mathbb{N}^i$ for some $i \in \mathbb{N}$, we define $L^- := (j_1, \ldots, j_{i-1})$.

The process is very similar to (ii) but more technical. Suppose the edges of $K_\mathbb{N}$ are coloured with r colours and let $q = (k-2)r + 1$. We will define sets A_L for $L \in \bigcup_{i=0}^{q} [r-1]^i$ and colourings $\chi_1 : \{A_L : L \in \bigcup_{i=0}^{q-1} [r-1]^i\} \to [r]$ and $\chi_2 : \bigcup_{i=1}^{q} [r-1]^i \to [r]$ with the following properties.

(a) The sets A_L, $L \in \bigcup_{i=0}^{q} [r-1]^i$, are pairwise disjoint and their union is cofinite.

(b) For every $L \in \bigcup_{i=1}^{q} [r-1]^i$, A_L is empty or every finite set $S \subseteq A_L$ has infinitely many common neighbours of colour $\chi_1(A_L)$ in A_L.

(c) For every $L \in \bigcup_{i=1}^{q} [r-1]^i$, A_L is empty or every finite set $S \subseteq \bigcup_{L \prec L'} A_{L'}$ has infinitely many common neighbours of colour $\chi_2(L)$ in A_{L^-}.

We will construct these sets and colourings recursively. In the process, we will also construct sets B_L and ultrafilters $\mathscr{U}_L$ on B_L for every $L \in \bigcup_{i=0}^{q} [r-1]^i$.

Let $B_{()} = \mathbb{N}$ and let $\mathscr{U}_{()}$ be a non-trivial ultrafilter on $B_{()}$, where $()$ denotes the empty sequence. Let $c_{\mathscr{U}_{()}}$ be the vertex-colouring induced by $\mathscr{U}_{()}$. Let c be the colour so that $A_{()}$, the set of vertices of colour c, is in $\mathscr{U}_{()}$ and let $\chi_1(A_{()}) = c$. Let $[r] \setminus \{c\} = \{j_1, \ldots, j_{r-1}\}$ and, for $i \in [r-1]$, let $B_{(i)}$ be the set of vertices receiving colour j_i and let let $\chi_2((i)) = j_i$.

In the next step, we proceed as follows for every $i_0 \in [r-1]$. If $B_{(i_0)}$ is finite, let $A_{(i_0)} = B_{(i_0,i)} = \emptyset$ for every $i \in [r-1]$. Otherwise, let $\mathscr{U}_{(i_0)}$ be a non-trivial ultrafilter on $B_{(i_0)}$ and let $c_{\mathscr{U}_{(i_0)}}$ be the vertex-colouring induced by $\mathscr{U}_{(i_0)}$. Let c be the colour so that $A_{(i_0)}$, the set of vertices of colour c, is in $\mathscr{U}_{(i_0)}$ and let $\chi_1(A_{(i_0)}) = c$. Let $[r] \setminus \{c\} = \{j_1, \ldots, j_{r-1}\}$ and, for $i \in [r-1]$, let $B_{(i_0,i)}$ be the set of vertices receiving colour j_i and let let $\chi_2((i_0, i)) = j_i$.

We proceed like this until we defined the sets B_L for every $L \in [r-1]^q$ and let $A_L := B_L$ for all $L \in [r-1]^q$. It is easy to see from the ultrafilter properties that the above properties hold.

Therefore, for every $L \in \bigcup_{i=0}^{q-1} [r-1]^i$, A_L is empty or can be covered by a monochromatic copy of G by Proposition 4.2.16 (since every one-way k-locally finite graph is 0-ruled). Furthermore, for every $L \in [r-1]^q$ for which A_L is non-empty, we find $k-1$ sets $L_1 \prec \ldots \prec L_{k-1} \prec L$ of the same colour w.r.t. χ_2 by pigeonhole principle. Therefore, applying Proposition 4.2.15 to $U_k := A_{L_1^-}, \ldots, U_2 := A_{L_{k-1}^-}, U_1 := A_L$, we find a monochromatic copy of G covering A_L. Since, there are $C := \sum_{i=0}^{q} (r-1)^i$ sets A_L, one of them has upper density at least $1/C$. □

The above proof immediately shows that for every one-way k-locally finite graph G and every r-edge-coloured $K_{\mathbb{N}}$, there is a collection of at most $f(r,k)$ monochromatic copies of G covering a cofinite subset of $\mathbb{N}$, where $f(r,k)$ is as in the statement of Theorem 1.4.13. In order to obtain Theorem 1.4.13, we need to guarantee that these copies can be chosen to be disjoint. To do so, instead of applying Propositions 4.2.15 and 4.2.16, we will embed the graphs simultaneously doing one step of the embedding algorithms of Propositions 4.2.15 and 4.2.16 at a time always making sure not to repeat vertices (which is possible since we have infinitely many choices in every step but only finitely many embedded vertices). Otherwise, the proof is exactly the same and therefore we will omit it.

4.2.8 Graphs of Bounded Ruling Number

In this section, we will prove Theorems 1.4.15 and 1.4.19.

Proof of Theorem 1.4.15. Let G be a finitely ruled graph and suppose the edges of $K_{\mathbb{N}}$ are coloured with r colours. Let $\mathscr{U}$ be a positive ultrafilter on $\mathbb{N}$ and, for $i \in [r]$, denote by A_i be the set of vertices of colour i in the induced vertex-colouring by $\mathscr{U}$. Suppose without loss of generality that $A_1 \in \mathscr{U}$. Since G is finitely ruled, there is a finite set S such that $G[S^c]$ does not have finite dominating sets and in particular $G[S^c]$ is 0-ruled.

We will now construct the embedding $f : V(G) \to \mathbb{N}$. First embed S into an arbitrary monochromatic clique in A_1 of colour 1 and size $|S|$ (such a clique can be found easily since $A_1 \in \mathscr{U}$ guarantees that every finite set $F \subseteq A_1$ has infinitely many common neighbours in colour 1). Let A_1' be the common neighbourhood in colour 1 of $f(S)$ restricted to A_1 and note that $A_1' \in \mathscr{U}$ and hence satisfies the assumptions of Proposition 4.2.16. Therefore, $G[S^c]$ can be surjectively embedded into A_1', and we can extend f to an embedding of G. Since $\mathscr{U}$ is a positive ultrafilter, $A_1' \subseteq f(V(G))$ has positive upper density, and we are done. □

Proof of Theorem 1.4.19. Colour the edges of $K_{\mathbb{N}}$ using the Rado colouring. Suppose now that V is the vertex set of a monochromatic copy of G, say in colour $i \in [2]$. Then for each $N \in \mathbb{N}$, we have

$$V \subseteq^* \bigcap_{n=1}^{N} \bigcup_{v \in F_n} N_i(v),$$

where we write $A \subseteq^* B$ if $A \setminus B$ is finite. It follows from Observation 4.2.4 that $d\left(\bigcup_{v\in F_n} N_i(v)\right) = 1 - d\left(\bigcap_{v\in F_n} N_{3-i}(v)\right) = 1 - 2^{-|F_n|}$ for every $n \in \mathbb{N}$. Therefore, we have

$$\overline{\mathrm{d}}\left(\bigcap_{n=1}^{N} \bigcup_{v \in F_n} N_i(v)\right) \leq \prod_{n=1}^{N}(1 - 2^{-|F_n|})$$

and hence

$$\overline{\mathrm{d}}(V) \leq \prod_{n=1}^{\infty}(1 - 2^{-|F_n|}) \leq \prod_{n=1}^{\infty} \exp\left(-2^{-|F_n|}\right).$$

It is well-known that an infinite product $\prod_{n=1}^{\infty} \alpha_n$, with $\alpha_n \in (0, 1)$, converges to 0 if and only if

$$\sum_{n=1}^{\infty} \log(\alpha_n) = -\infty.$$

Since $|F_n| \leq \log_2(n)$ for all sufficiently large n, we have

$$\log\left(\exp\left(-2^{-|F_n|}\right)\right) \leq -1/n.$$

By the limit comparison test and the divergence of the harmonic series, it follows that $d(V) = 0$. □

5 Robust Triangle Tilings

5.1 Overview

In this chapter, we are going to prove the following result.

Theorem 1.5.9 (Allen et al. [3]). *There is a constant $C > 0$ such that for all $n \in \mathbb{N}$ divisible by 3, all $p \geq C(\log n)^{1/3}n^{-2/3}$ and all n-vertex graphs G with $\delta(G) \geq 2n/3$, G_p has a triangle tiling w.h.p.*

In fact we shall prove the following similar theorem with a super-regular host-graph, which might be of independent interest.

Theorem 5.1.1. *For every $d > 0$ there exist constants $\varepsilon > 0$ and $C > 0$ such that the following is true for every $n \in \mathbb{N}$ and $p \geq C(\log n)^{1/3}n^{-2/3}$. If Γ is a tripartite graph with parts V^1, V^2, V^3 of size n so that Γ is (ε, d^+, d)-super-regular, then Γ_p contains a triangle tiling w.h.p.*

We will deduce Theorem 1.5.9 from Theorem 5.1.1 in Section 5.5. The proof is quite technical and long, but uses standard methods in extremal graph theory. The proof of Theorem 5.1.1 is the main challenge and will span several sections. We will describe the general set-up and proof outline here.

5.1.1 Set-up

Constants: There are a few constants which we shall use repeatedly throughout this chapter. We will define all of them here for convenience and they remain fixed throughout the proof. They will obey the following hierarchy

$$0 < \varepsilon, 1/C \ll \varepsilon' \ll \eta \ll \eta' \ll \alpha \ll d < 1. \qquad (5.1.1)$$

By this we mean that given some d as in the statement of Theorem 5.1.1, we can choose the remaining constants from right to left such that if we choose each constant sufficiently small (in terms of some function of the already chosen constants), then all the required inequalities between the constants in the proof will hold. In more technical proofs, we sometimes require the use of other constants and we will always explicitly state their relation to the constants above. The complete list of constants we will use obeys the following hierarchy

$$0 < \varepsilon, 1/C \ll \varepsilon' \ll \eta \ll \eta' \ll \beta'' \ll \gamma \ll \beta' \ll \beta \ll \alpha \ll d < 1. \qquad (5.1.2)$$

The host-graph Γ: By Lemma 2.2.4, it suffices to prove Theorem 5.1.1 for $(\varepsilon, d, d-\varepsilon)$-super regular 3-cylinders (that is, all pairs have density exactly d). Throughout the rest of this chapter, we will fix some $n \in \mathbb{N}$, disjoint vertex-sets V^1, V^2 and $V^3$of size n, and a tripartite $(\varepsilon, d, d-\varepsilon)$-super-regular graph Γ with parts V^1, V^2, V^3. We will assume that n is sufficiently large in terms of all fixed constants whenever necessary.

(Very) high probability: We further fix some probability $p \geq C(\log n)^{1/3}n^{-2/3}$ throughout the chapter. As ε and C feature at the bottom of our constant hierarchy (5.1.1), we can always push ε as small and C as large as we want, as long as they are independent of n. Therefore, for brevity, we will say 'w.h.p. ...' as shorthand for 'there exist $C > 0$ and $\varepsilon > 0$ such that w.h.p. ...' throughout the rest of this chapter. Note that if we have two events E_1, E_2 which hold w.h.p., then by taking a union bound, it follows that $E_1 \cap E_2$ holds w.h.p. However, we can only iterate this constantly many times (i.e. $f(\varepsilon', \eta, \eta', \alpha, d)$ many times for some function f) and we sometimes need to take larger union bounds. We will also say that an event holds *with very high probability* (w.v.h.p.) if for any $K > 0$, there exists a large enough $C > 0$ and a small enough $\varepsilon > 0$ such that the probability that the event holds in Γ_p is at least $1 - n^{-K}$. This allows us to take union bounds over polynomially many events. E.g., given events $E_1, \ldots, E_{n^3}$ which hold w.v.h.p., $E_1 \cap \ldots \cap E_{n^3}$ holds w.v.h.p. as well.

Special vertex-sets and induced subgraphs: Let $\mathcal{V} := \{\emptyset\} \cup V^1 \cup (V^1 \times V^2) \cup (V^1 \times V^2 \times V^3)$. That is, an element $\underline{u} \in \mathcal{V}$ is a vector of some *length* $0 \leq \ell(\underline{u}) \leq 3$ such that for each $i \leq \ell(\underline{u})$, we have that $\underline{u}$ contains exactly one vertex from V^i.

Now given $\underline{u} = (u_1, \ldots, u_\ell) \in \mathcal{V}$ for some $\ell \in [3]_0$, let

$$V^i_{\underline{u}} := \begin{cases} V^i \setminus \{u_i\} & \text{for } 1 \leq i \leq \ell \\ V^i & \text{for } \ell + 1 \leq i \leq 3, \end{cases} \tag{5.1.3}$$

denote the vertex-sets after removing the vertices which feature in $\underline{u}$. We will also at times want to consider the subsets which leave out not only the vertices in $\underline{u}$ but also some extra vertices which lie in $V^{\ell+1}$, say u and v. We define $V^i_{\underline{u},u,v} := V^i_{\underline{u}}$ for all $i \neq \ell + 1$ and $V^{\ell+1}_{\underline{u},u,v} := V^{\ell+1} \setminus \{u, v\}$.

Given some $\underline{u} \in \mathcal{V}$, we consider the graph induced after removing the vertices of $\underline{u}$, that is

$$\Gamma_{\underline{u}} := \Gamma\left[\bigcup\nolimits_{i \in [3]} V^i_{\underline{u}}\right].$$

Similarly, given $\underline{u}$ of length ℓ and a pair of vertices u and v in $V^{\ell+1}$, $\Gamma_{\underline{u},u,v}$ denotes the graph induced on the vertex-sets $V^i_{\underline{u},u,v}$ as above.

Triangles: For a graph G, $\mathrm{Tr}(G)$ denotes the set of triangles in G, for a vertex $v \in V(G)$, $\mathrm{Tr}_v(G)$ denotes the set of pairs in G which form a triangle with v.

Embeddings of partial triangle tilings: We will be concerned with embedding partial triangle tilings in a given host graph. For $t \in [n]$, we therefore define H_t to be the graph on vertex-set $[t] \times \{3\}$, whose edges consist of the edges $\{\{(s,i),(s,j)\} : s \in [t], i \neq j \in [3]\}$. Thus H_t simply consists of t labelled vertex disjoint triangles.

Given a graph G on $V^1 \cup V^2 \cup V^3$ (usually $G = \Gamma_p$), we define $\Phi^t(G)$ to be the collection of labelled embeddings of H_t into G, which map $[t] \times \{i\}$ to a subset of V^i for $i \in [3]$. We will be interested in embeddings that fix certain vertices to be isolated.

Given $\underline{u} = (u_1, \ldots, u_\ell) \in \mathcal{V}$ of length $\ell \leq 3$ and $t \in [n-1]$, we define $\Phi^t_{\underline{u}}(G) \subseteq \Phi^t(G)$ to be those $\phi \in \Phi^t(G)$ for which $\phi((s,i)) \neq u_i$ for all $i \in [\ell]$ and $s \in [t]$ (that is, we fix the ℓ vertices in $\underline{u}$ to be isolated in the embedding of H_t). Note that if $\underline{u} = \emptyset$, then $\Phi^t_{\underline{u}}(G) = \Phi^t(G)$.

5.1.2 Proof Outline

We will deduce Theorem 5.1.1 from the following two propositions. The first proposition counts 'almost-tilings'.

Proposition 5.1.2. *W.v.h.p., we have*

$$\left|\Phi^t(\Gamma_p)\right| \geq (1-\eta)^t (pd)^{3t} (n!_t)^3 \,. \tag{5.1.4}$$

for all $t \leq (1-\eta)n$*.*

In order to go beyond Proposition 5.1.2, we need different techniques. The second proposition allows us to extend incomplete triangle tilings by embedding further triangles one by one.

Proposition 5.1.3. *W.v.h.p., the following is true for every integer* $(1-\eta)n \leq t < n$ *and every* $\tau = \tau(n) \geq (1-\eta'/2)^n$*. If*

$$\left|\Phi^t(\Gamma_p)\right| \geq \tau (pd)^{3t} (n!_t)^3$$

then

$$\left|\Phi^{t+1}(\Gamma_p)\right| \geq \alpha\tau (pd)^{3(t+1)} (n!_{(t+1)})^3.$$

The proof of Proposition 5.1.3 is the main difficulty of this chapter and is technically involved. In order to count embeddings of partial triangle tilings in $\Phi^{t+1}(\Gamma_p)$, we will first count how many candidates there are in $\Phi^t(\Gamma_p)$ (by fixing certain vertices to be isolated). If there are many triples of vertices which are isolated in many embeddings then we can argue that many of these triples actually host triangles and thus extend to embeddings in $\Phi^{t+1}(\Gamma_p)$. When considering embeddings which leave certain vertices isolated we also do this in steps, growing our set of isolated vertices one vertex at a time. The key step is captured by the following lemma.

Lemma 5.1.4 (Local distribution Lemma)**.** *W.v.h.p., the following holds for all* $\ell \in [3]$*, all integers* $(1-\eta)n \leq t < n$ *and all* $\underline{u} = (u_1 \ldots, u_{\ell-1}) \in \mathcal{V}$*. If* $\tau = \tau(n) \geq (1-\eta')^n$ *and*

$$\left|\Phi^t_{\underline{u}}(\Gamma_p)\right| \geq \tau (pd)^{3t} ((n-1)!_t)^{\ell-1} (n!_t)^{4-\ell},$$

then for all but at most αn *vertices* $u_\ell \in V^\ell$

$$\left|\Phi^t_{\underline{u}'}(\Gamma_p)\right| \geq \left(\frac{d}{10}\right)^2 \tau(pd)^{3t}((n-1)!_t)^\ell (n!_t)^{3-\ell}, \tag{5.1.5}$$

where $\underline{u}' = (\underline{u}, u_\ell) \in \mathcal{V}$.

In Section 5.2, we will provide some results on triangle counts in Γ_p, which will be useful in the remaining proofs. In Section 5.3, we will prove Proposition 5.1.3 (using Lemma 5.1.4 as a blackbox for now) and Proposition 5.1.2, and deduce Theorem 5.1.1. In Section 5.4, we will prove Lemma 5.1.4 with a delicate argument using entropy.

5.2 Counting Triangles in Γ_p

The purpose of this section is to prove that certain properties of Γ_p hold with high probability. These properties regard triangle counts in Γ_p and their proofs use the properties of regular tuples given in Section 2.2 and the probabilistic tools outlined in Section 2.4.

Lemma 5.2.1. *W.v.h.p., we have*

$$\left|\mathrm{Tr}(\Gamma_p[X_1 \cup X_2 \cup X_3])\right| = (pd)^3|X_1||X_2||X_3| \pm 1000\varepsilon(pd)^3 n^3 \tag{5.2.1}$$

for all $X_1 \subseteq V^1$, $X_2 \subseteq V^2$ *and* $X_3 \subseteq V^3$.

We will frequently replace 1000ε by ε' in the above equation, which we can do since $\varepsilon \ll \varepsilon'$.

Proof. We first show that the lower bound holds w.v.h.p. using Janson's inequality.

Claim 5.1. *W.v.h.p., we have*

$$\left|\mathrm{Tr}(\Gamma_p[X_1 \cup X_2 \cup X_3])\right| \geq (pd)^3|X_1||X_2||X_3| - 100\varepsilon(pd)^3 n^3 \tag{5.2.2}$$

for all $X_1 \subseteq V^1$, $X_2 \subseteq V^2$ *and* $X_3 \subseteq V^3$.

Proof. Fix $X_1 \subseteq V^1$, $X_2 \subseteq V^2$ and $X_3 \subseteq V^3$ and let $Y := \mathrm{Tr}(\Gamma[X_1 \cup X_2 \cup X_3])$. We may assume that

$$|X_1||X_2||X_3| \geq 100\varepsilon n^3 \tag{5.2.3}$$

for otherwise (5.2.2) is trivially true. In particular, we have $|X_i| \geq \varepsilon n$ for all $i \in [3]$ and thus Lemma 2.2.7 implies

$$|Y| \geq d^3|X_1||X_2||X_3| - 10\varepsilon n^3. \tag{5.2.4}$$

Consider now the random variable

$$X := |\,\mathrm{Tr}(\Gamma_p[X_1 \cup X_2 \cup X_3])| = \sum_{A \in \mathcal{S}} I_A,$$

where $\mathcal{S} = \{A_T : T \in Y\}$ and A_T is the event that the triangle $T \in Y$ is present in Γ_p. Let

$$\lambda := \mathbb{E}\left[X\right] = p^3 \cdot |Y| \geq (pd)^3|X_1||X_2||X_3| - 10\varepsilon p^3 n^3. \tag{5.2.5}$$

which in combination with (5.2.3) implies

$$\lambda \geq \varepsilon' p^3 n^3. \tag{5.2.6}$$

Furthermore, we define

$$\bar{\Delta} := \sum_{T,T' \in Y:\ T \cap T' \neq \emptyset} \mathbb{E}\left[I_{A_T} I_{A_{T'}}\right] \leq p^5 \cdot |Y| \cdot 3n + p^3 \cdot |Y| \tag{5.2.7}$$

$$= \lambda(3np^2 + 1). \tag{5.2.8}$$

where the inequality in (5.2.7) follows from the fact that there are at most $|Y| \cdot 3n$ pairs of triangles intersecting in exactly one edge, no pairs intersecting in exactly two edges and $|Y|$ pairs intersecting in three edges. Hence Janson's inequality

(Lemma 2.4.2) implies

$$\mathbb{P}[X \le (1-\varepsilon)\lambda] \le \exp\left(-\frac{\varepsilon^2\lambda^2}{2\bar{\Delta}}\right) \tag{5.2.9}$$

$$\le \exp\left(-\varepsilon^3 p^3 n^3 \cdot 2\lambda/\bar{\Delta}\right) \tag{5.2.10}$$

$$\le \exp\left(-\varepsilon^3 p^3 n^3 \cdot 1/(3np^2)\right) + \exp\left(-\varepsilon^3 p^3 n^3\right) \tag{5.2.11}$$

$$\le \exp(-10n) \tag{5.2.12}$$

for all large enough n. Here, we used (5.2.6) in (5.2.10), and (5.2.7) in (5.2.11) (more precisely, we used that (5.2.7) implies that $\bar{\Delta} \le 6\lambda np^2$ or $\bar{\Delta} \le 2\lambda$).

By (5.2.5), we have $(1-\varepsilon)\lambda \ge (pd)^3|X_1||X_2||X_3| - 100\varepsilon(pd)^3n^3$. Hence, taking a union bound over all choices of $X_1 \subseteq V^1, X_2 \subseteq V^2, X_3 \subseteq V^3$, we deduce that, (5.2.2) holds w.v.h.p. for all $X_1 \subseteq V^1, X_2 \subseteq V^2, X_3 \subseteq V^3$. ■

We now show that the upper bound holds w.v.h.p. in the case when $X_i = V^i$ for all $i \in [3]$.

Claim 5.2. *W.v.h.p., we have*

$$\left|\mathrm{Tr}(\Gamma_p)\right| \le (pd)^3n^3 + 100\varepsilon(pd)^3n^3. \tag{5.2.13}$$

Proof. Let $Y = \mathrm{Tr}(\Gamma)$ and let $X = |\mathrm{Tr}(\Gamma_p)| = \sum_{T\in Y} I_T$. By Lemma 2.2.7, we have

$$|Y| = d^3n^3 \pm 10\varepsilon n^3. \tag{5.2.14}$$

It follows that

$$\lambda := \mathbb{E}[X] = (pd)^3n^3 \pm 10\varepsilon p^3n^3. \tag{5.2.15}$$

Using notations from the Kim-Vu inequality (Lemma 2.4.3), we have $E_1 = np^2$, $E_2 = p$ and $E_3 = 1$. Hence $E' = \max\{1, np^2\} \le \lambda^{1/2}$ and $E = \lambda$. Let $\mu = \lambda^{1/16}$ and let $c = c(3)$ be the constant from Lemma 2.4.3. Then, for large enough n,

$$c \cdot (EE')^{1/2}\mu^3 \le c\lambda^{3/4} \cdot \lambda^{3/16} \le \varepsilon\lambda.$$

Hence, we have

$$\mathbb{P}[X \geq (1+\varepsilon)\lambda] \leq 10cn^4 \cdot e^{-\mu} \leq e^{-n^{1/16}}$$

for all large enough n. Here, the last inequality follows from (5.2.15) which implies $\lambda \geq n \log n$. This finishes the proof of the claim. ■

We now conclude the proof of the lemma. W.v.h.p., both claims above hold simultaneously. Suppose now both claims hold and fix $X_1 \subseteq V^1, X_2 \subseteq V^2, X_3 \subseteq V^3$. Let $\mathcal{U} = \big(\{X_1, V^1 \setminus X_1\} \times \{X_2, V^2 \setminus X_2\} \times \{X_3, V^3 \setminus X_3\}\big) \setminus \{(X_1, X_2, X_3)\}$ and observe that

$$\begin{aligned}\big|\mathrm{Tr}(\Gamma_p[X_1 \cup X_2 \cup X_3])\big| &= \big|\mathrm{Tr}(\Gamma_p)\big| - \sum_{(U_1,U_2,U_3)\in\mathcal{U}} \big|\mathrm{Tr}(\Gamma_p[U_1 \cup U_2 \cup U_3])\big| \\ &\leq (pd)^3|X_1||X_2||X_3| + 800\varepsilon(pd)^3n^3.\end{aligned}$$

Here we used (5.2.13) to bound $\big|\mathrm{Tr}(\Gamma_p)\big|$ and (5.2.2) to bound each term of the form $\big|\mathrm{Tr}(\Gamma_p[U_1 \cup U_2 \cup U_3])\big|$. This completes the proof. □

Corollary 5.2.2. *W.v.h.p., we have*

$$\big|\mathrm{Tr}_v(\Gamma_p)\big| = (1 \pm \varepsilon')(pd)^3n^2$$

for all but at most $\varepsilon' n$ vertices $v \in V(\Gamma)$.

Proof. Let $\tilde{\varepsilon} = 1000\varepsilon$ and let $G \subseteq \Gamma$ be any graph with

$$|\mathrm{Tr}(G[X_1 \cup X_2 \cup X_3])| = (pd)^3|X_1||X_2||X_3| \pm \tilde{\varepsilon}(pd)^3n^3 \tag{5.2.16}$$

for all $X_1 \subseteq V^1$, $X_2 \subseteq V^2$ and $X_3 \subseteq V^3$. Since (by Lemma 5.2.1) this is satisfied by Γ_p w.v.h.p., it suffices to show that G satisfies the conclusion of Corollary 5.2.2. For $i \in [3]$, let X_i be the set of vertices $v \in V^i$ with $|\mathrm{Tr}_v(G)| \leq (1-\varepsilon')(pd)^3n^2$, and let let Y_i be the set of vertices $v \in V^i$ with $|\mathrm{Tr}_v(G)| \geq (1+\varepsilon')(pd)^3n^2$. We claim that $|X_1| \leq \varepsilon'/10 \cdot n$. Indeed, assuming the contrary, we have

$$\big|\mathrm{Tr}(G[X_1 \cup V^2 \cup V^3])\big| < (pd)^3|X_1||V^2||V^3| - \tilde{\varepsilon}(pd)^3n^3,$$

contradicting (5.2.16). Using a similar argument, we can bound the sizes of X_2 and X_3, and Y_1, Y_2 and Y_3, completing the proof. □

Sometimes, we need an upper bound on $|\operatorname{Tr}_v(\Gamma_p)|$ which works for all $v \in V(\Gamma)$. For this we simply upper bound this quantity by the number of triangles in $G(3n, p)$ containing a specific vertex using a result of Spencer [113].

Lemma 5.2.3. *W.v.h.p., we have*

$$\left|\operatorname{Tr}_v(\Gamma_p)\right| \leq 10p^3n^2$$

for all vertices $v \in V(\Gamma)$.

In the remainder of this section we prove some more technical properties of Γ_p which will come in useful in the proof of Lemma 5.1.4 (more specifically, in the proof of Lemma 5.4.1).

Definition 5.2.4. For $\ell \in [3]$, $\underline{u} = (u_1, \ldots, u_{\ell-1}) \in \mathcal{V}$, $u, v \in V^\ell$ and $J = \{j_1, j_2\} := [3] \setminus \{\ell\}$, we define the following sets.

(i) $F_{\underline{u}}(u) := \{(y_1, y_2) \in V_{\underline{u}}^{j_1} \times V_{\underline{u}}^{j_2} : uy_1, uy_2 \in E(\Gamma),\ y_1y_2 \in E(\Gamma_p)\}$.

(ii) $F_{\underline{u}}(u, v) := \{(y_1, y_2) \in V_{\underline{u}}^{j_1} \times V_{\underline{u}}^{j_2} : uy_1, uy_2, vy_1, vy_2 \in E(\Gamma),\ y_1y_2 \in E(\Gamma_p)\}$.

(iii) $S_{\underline{u}}(u) := \{(y_1, y_2) \in V_{\underline{u}}^{j_1} \times V_{\underline{u}}^{j_2} : uy_1, uy_2, y_1y_2 \in E(\Gamma_p)\}$.

We will show that $F_{\underline{u}}(u, v)$ is large for most pairs of vertices u, v.

Lemma 5.2.5. *W.v.h.p., the following holds for every* $\ell \in [3]$*, every choice of* $\underline{u} = (u_1, \ldots, u_{\ell-1}) \in \mathcal{V}$ *and every* $u \in V^\ell$*. We have*

$$\left|F_{\underline{u}}(u, v)\right| \geq d^5pn^2/4$$

for all but at most $\varepsilon' n$ *vertices* $v \in V^\ell$.

Proof. Fix some $\ell \in [3]$, $\underline{u} = (u_1, \ldots, u_{\ell-1}) \in \mathcal{V}$ and $u \in V^\ell$. We first use regularity to show that there are many edges in the deterministic graph.

Claim 5.3. *We have*

$$\left|E_\Gamma(N_{\Gamma_{\underline{\boldsymbol{t}}}}(u,v))\right| \geq d^5n^2/2 \tag{5.2.17}$$

for all but at most $2\varepsilon n$ vertices $v \in V^\ell$.

Proof. Assume w.l.o.g. that $\ell = 3$. For $i \in [2]$, let $X_i = N_\Gamma(u; V^i)$ and for $v \in V^3 \setminus \{u\}$, let $Y_i(v) = N_\Gamma(u,v;V^i) \subseteq X_i$. Since Γ is $(\varepsilon, d, d-\varepsilon)$-super-regular, we have $|X_i| \geq (d-\varepsilon)n$ for both $i \in [2]$. For $i \in [2]$, let $R_i \subseteq V^3$ bet the set of vertices $v \in V^3$ for which $|Y_i(v)| < (d-\varepsilon)^2 n$ and let $R = R_1 \cup R_2$. It follows from the ε-regularity of (V_i, V_3), that $|R_i| \leq \varepsilon n$ for both $i \in [2]$ and hence $|R| \leq 2\varepsilon n$. Furthermore, for every $v \in V^3 \setminus R$, it follows from the ε-regularity of the pair (V^2, V^3) that $|E_\Gamma(Y_1(v) \cup Y_2(v))| \geq (d-\varepsilon)^5 n^2$. This completes the proof. ■

Observe now that each edge in $E_\Gamma(N_{\Gamma_{\underline{\boldsymbol{t}}}}(u,v))$ is present independently in Γ_p and hence it follows from Chernoff's inequality that, w.v.h.p., we have

$$\left|F_{\underline{\boldsymbol{t}}}(u,v)\right| \geq d^5pn^2/4$$

for all but at most $2\varepsilon n$ vertices $v \in V^3$. Taking a union bound over all $\ell \in [3]$, $\underline{u} = (u_1, \ldots, u_{\ell-1}) \in \mathcal{V}$ and $u \in V^\ell$ completes the proof. □

An application of Janson's inequality shows that, w.v.h.p., $S_{\underline{\boldsymbol{t}}}(u)$ is large for all choices of $(\underline{u}, u) \in \mathcal{V}$. We will require a slightly stronger and more technical statement than this, showing that $S_{\underline{\boldsymbol{t}}}(u)$ has large intersection with a given, reasonably large set $F \subseteq F_{\underline{\boldsymbol{t}}}(u)$.

Lemma 5.2.6. *Let $\ell \in [3]$, $\underline{u} = (u_1, \ldots, u_{\ell-1}) \in \mathcal{V}$ and $u \in V^\ell$. Let $\{j_1, j_2\} = [3] \setminus \{\ell\}$. Let $F \subseteq F_{\underline{\boldsymbol{t}}}(u)$ be a possibly random set which is independent from $\{e \in E(\Gamma_p) : e \cap \{u\} \neq \emptyset\}$*[1] *and assume $|F| \geq \alpha pn^2$. Then, w.v.h.p., we have*

$$\left|F \cap S_{\underline{\boldsymbol{t}}}(u)\right| \geq p^2|F|/2.$$

Proof. Fix $\ell \in [3]$, $t \in [n]$, $\underline{u} = (u_1, \ldots, u_{\ell-1}) \in \mathcal{V}$ and $u, v \in V^\ell$. Let $G_u \subseteq \Gamma_p$ be the graph on $V(\Gamma)$ consisting of all edges in $E(\Gamma_p)$ which are incident to u and let

[1] In other words, F is completely determined by revealing all edges which are not incident to u.

$G_u^c = \Gamma_p \setminus G_u$. Reveal G_u^c and assume that $\Delta(G_u^c) \leq 4pn$ (a simple application of Chernoff's bound (Lemma 2.4.1) shows that this holds w.v.h.p.).

Let $\mathcal{F} = \{\{ux_1, ux_2\} : x_1x_2 \in F\}$. We will now use Janson's inequality to show that many pairs of edges in $\mathcal{F}$ are present in G_u (we can treat $\mathcal{F}$ as a fixed set here since G_u^c is revealed).

Claim 5.4. *W.v.h.p. (in G_u), we have*

$$\left|F \cap S_{\underline{tt}}(u)\right| \geq p^2|F|/2.$$

Proof. Note that $|\mathcal{F}| = |F| \geq \alpha pn^2$. Let

$$X = \left|F \cap S_{\underline{tt}}(u)\right| = \sum_{A \in \mathcal{F}} I_A$$

and note that we only need to reveal G_u in order to determine I_F. Let

$$\begin{aligned} \lambda := \mathbb{E}[X] &= p^2 |\mathcal{F}| && (5.2.18)\\ &\geq \alpha p^3 n^2 && (5.2.19)\\ &\geq C^2 \log n. && (5.2.20) \end{aligned}$$

Furthermore, let

$$\begin{aligned} \bar{\Delta} :=& \sum_{(A,A') \in \mathcal{F}^2:\ A \cap A' \neq \emptyset} \mathbb{E}[I_A I_{A'}] && (5.2.21)\\ &\leq 8p^4|\mathcal{F}|n + p^2|\mathcal{F}| && (5.2.22)\\ &= \lambda(1 + 8p^2 n). && (5.2.23) \end{aligned}$$

Here, (5.2.22) follows from the fact that there are at most $|\mathcal{F}| \cdot \Delta(G_u^c) = |\mathcal{F}| \cdot 8pn$ pairs $(A, A') \in \mathcal{F}^2$ intersecting in exactly one edge, and $\mathcal{F}$ pairs intersecting in two edges. Hence Janson's inequality (Lemma 2.4.2) implies

$$\mathbb{P}\left[X \leq 1/2 \cdot \lambda\right] \leq \exp\left(-\frac{\lambda^2}{8\bar{\Delta}}\right) \tag{5.2.24}$$

$$\leq \exp\left(-\lambda/(8(1+8pn^2))\right) \tag{5.2.25}$$

$$\leq \exp\left(-\lambda/16\right) + \exp\left(-\lambda/(128pn^2)\right) \tag{5.2.26}$$

$$\leq n^{-C} + e^{-n^{1/3}} \tag{5.2.27}$$

for all large enough n. Here, we used (5.2.23) in (5.2.25), the fact that $1+8pn^2 \leq 2$ or $1+8pn^2 \leq 16pn^2$ in (5.2.26), and (5.2.19) and (5.2.20) in (5.2.27). This completes the proof of the claim. ■

□

5.3 Embedding (Partial) Triangle Tilings

5.3.1 Counting Almost Triangle Tilings

Here we prove Proposition 5.1.2.

Proof of Proposition 5.1.2. By Lemma 5.2.1, we have w.v.h.p.

$$\left|\mathrm{Tr}(\Gamma_p[X_1 \cup X_2 \cup X_3])\right| = (pd)^3|X_1||X_2||X_3| \pm \varepsilon'(pd)^3 n^3 \tag{5.3.1}$$

for all $X_1 \subseteq V^1$, $X_2 \subseteq V^2$ and $X_3 \subseteq V^3$. We will show by induction on t that if Γ_p satisfies (5.3.1), then it satisfies

$$\left|\Phi^t(\Gamma_p)\right| \geq (1-\eta)^t (pd)^{3t} \left(n!_t\right)^3 \tag{5.3.2}$$

for all integers $t \leq (1-\eta)n$, as claimed. (5.3.2) is trivial for $t = 0$. Suppose now (5.3.2) holds for some integer $0 \leq t \leq (1-\eta)n$. Fix some $\phi \in \Phi^t(\Gamma_p)$ and let $X_i \subseteq V^i$, $i \in [3]$, be the sets of vertices which are not in $\phi(H^t)$. Note that $|X_i| = n-t$ for all $i \in [3]$. Now the number of triangles which extend ϕ to an embedding in

$\Phi^{t+1}(\Gamma_p)$ is precisely $\left|\mathrm{Tr}(\Gamma_p[X_1 \cup X_2 \cup X_3])\right|$ and by (5.3.1), we have

$$\begin{aligned}\left|\mathrm{Tr}(\Gamma_p[X_1 \cup X_2 \cup X_3])\right| &\geq (pd)^3|X_1||X_2||X_3| - \varepsilon'(pd)^3n^3\\ &\geq (pd)^3(n-t)^3 - \varepsilon'/\eta^3 \cdot (pd)^3(n-t)^3\\ &\geq (1-\eta)(pd)^3(n-t)^3.\end{aligned}$$

It follows from the induction hypothesis that

$$\begin{aligned}\left|\Phi^{t+1}(\Gamma_p)\right| &\geq \left|\Phi^t(\Gamma_p)\right| \cdot (1-\eta)(pd)^3(n-t)^3\\ &\geq (1-\eta)^{t+1}(pd)^{3(t+1)}\left(n!_{t+1}\right)^3,\end{aligned}$$

finishing the proof. □

5.3.2 Extending Almost Triangle Tilings

In this subsection, we will prove Proposition 5.1.3 using the local distribution lemma (Lemma 5.1.4) as a blackbox for now. The main difficulty of this proof lies in the following lemma. Here, given a graph $G \subseteq \Gamma$, a vertex $v \in V^1$ and some $t \in \mathbb{N}_0$, we denote by $\Phi^t_v(G) \subseteq \Phi^t(G)$ the set of embeddings $\phi \in \Phi^t(G)$ for which $\phi((1,1)) = v$.

Lemma 5.3.1. *W.v.h.p., the following is true for every $\tau \geq (1-\eta')^n$ and every integer $(1-\eta)n \leq t < n$. For all $v \in V^1$, if*

$$\left|\Phi^t_v(\Gamma_p)\right| \geq \tau(pd)^{3t}(n-1)!_t(n!_t)^2,$$

then

$$\left|\Phi^{t+1}_v(\Gamma_p)\right| \geq \sqrt{\alpha}\tau(pd)^{3(t+1)}(n-1)!_t(n!_{t+1})^2.$$

We first show how Proposition 5.1.3 follows from this.

Proof of Proposition 5.1.3. Let $(1-\eta)n \leq t < n$ and choose τ_0 so that

$$\left|\Phi^t(\Gamma_p)\right| = \tau_0(pd)^{3t}(n!_t)^3.$$

If $\tau > \tau_0$, there is nothing to prove, hence it is sufficient to prove the proposition for all $(1-\eta'/2)^n \leq \tau \leq \tau_0$. Let $U_1 \subseteq V^1$ be the set of vertices $u_1 \in V^1$ for which

$$\left|\Phi^t_{u_\mathrm{T}}(\Gamma_p)\right| \geq d^2\tau_0/100 \cdot (pd)^{3t}(n-1)!_t(n!_t)^2.$$

It follows from Lemma 5.1.4 (with $\ell = 1$) that, w.v.h.p., we have $|U_1| \geq n/2$. It further follows from Lemma 5.3.1 that, w.v.h.p., we have

$$\left|\Phi^{t+1}_{u_1}(\Gamma_p)\right| \geq \sqrt{\alpha}d^2\tau_0/100 \cdot (pd)^{3(t+1)}(n-1)!_t(n!_{t+1})^2.$$

for every $u_1 \in U_1$. Therefore, w.v.h.p., we have

$$\begin{aligned}
\left|\Phi^{t+1}(G)\right| &\geq \sum_{u_1 \in U_1} \left|\Phi^{t+1}_{u_1}(\Gamma_p)\right| \\
&\geq n \cdot \sqrt{\alpha}d^2\tau_0/200 \cdot (pd)^{3(t+1)}(n-1)!_t(n!_{t+1})^2 \\
&\geq \alpha\tau_0 \cdot (pd)^{3(t+1)}(n)!_t(n!_{t+1})^2 \\
&\geq \alpha\tau \cdot (pd)^{3(t+1)}(n)!_t(n!_{t+1})^2
\end{aligned}$$

for every $\tau \leq \tau_0$. Taking a union bound over all $(1-\eta)n \leq t < n$ finishes the proof. □

It remains to prove Lemma 5.3.1.

Proof of Lemma 5.3.1. Let $K \geq 100$ and let $\tilde{\alpha} = \alpha^{1/5}$. Let $G = \Gamma_p$ and fix some $v_1 \in V^1$. Let $(1-\alpha)n \leq t < n$ and choose τ_0 so that

$$\left|\Phi^t_{v_\mathrm{T}}(\Gamma_p)\right| = \tau_0(pd)^{3t}(n-1)!_t(n!_t)^2.$$

If $\tau > \tau_0$, there is nothing to prove, hence it is sufficient to prove the proposition for all $(1-\eta')^n \leq \tau \leq \tau_0$.

Given some $i \in [3]$, we call a sequence of vertices $\underline{u} = (u_1, \ldots, u_i) \in V^1 \times \ldots \times V^i$ *good* if

$$\left|\Phi^t_{\underline{u}}(G)\right| \geq \tilde{\alpha}^{i-1}\tau_0(pd)^{3t}((n-1)!_t)^i(n!_t)^{3-i}.$$

Note that (v_1) is good by definition. Let $X_2(v_1) \subseteq N_\Gamma(v_1; V^2)$ be the set of vertices $u_2 \in N_\Gamma(v_1; V^2)$ such that (v_1, u_2) is good and $\deg_\Gamma(v_1, u_2) \geq d^2n/2$.

Step 1. *With probability at least* $1 - n^{-8K}$, *we have* $|X_2(v_1)| \geq dn/2$.

Given $v_2 \in V^2$, let $X_3(v_1, v_2) \subseteq N_\Gamma(v_1, v_2; V^3)$ be the set of vertices $u_3 \in N_\Gamma(v_1, v_2; V^3)$ such that $\{v_1, v_2, u_3\}$ is good. Furthermore, let $Y_3(v_1, v_2) \subseteq X_3(v_1, v_2)$ be the set of those u_3 such that $v_2u_3 \in E(G)$.

Step 2. *With probability at least* $1 - n^{-8K}$ *the following is true for every* $v_2 \in V^2$. *If* $v_2 \in X_2(v_1)$, *then we have* $|Y_3(v_1, v_2)| \geq pd^2n/8$.

Let now $Z' = Z'(v_1) = \big\{(u_2, u_3) \in V^2 \times V^3 : u_2 \in X_2(v_1), u_3 \in Y_3(v_1, u_2)\big\}$ and $Z = Z(v_1) = \{(u_2, u_3) \in Z' : \{v_1, u_2, u_3\}$ is a triangle in $G\}$. We will use Steps 1 and 2 to deduce the following.

Step 3. *With probability* $1 - n^{-4K}$, *we have* $|Z'| \geq pd^3n^2/16$ *and* $\Delta(G) \leq 4pn$.

Step 4. *With probability at least* $1 - n^{-4K}$, *the following is true in* G. *If* $|Z'| \geq pd^3n^2/16$ *and* $\Delta(G) \leq 4pn$, *then we have* $|Z| \geq (pd)^3n^2/32$.

Before we prove the claims in Steps 1 to 4, let us deduce the lemma. Combining Steps 3 and 4, we have $|Z| \geq (pd)^3n^2/32$ with probability at least $1 - n^{-2K}$. Furthermore, for all $(u_2, u_3) \in Z$, the vector (v_1, u_2, u_3) is good and hence $\left|\Phi^t_{(v_1, u_2, u_3)}(G)\right| \geq \tilde{\alpha}^2\tau_0(pd)^{3t}((n-1)!_t)^3$. Therefore,

$$
\begin{aligned}
\left|\Phi^{t+1}_{v_1}(G)\right| &\geq \sum_{(u_2,u_3)\in Z} \left|\Phi^t_{(v_1, u_2, u_3)}(G)\right| \\
&\geq (pd)^3n^2/32 \cdot \tilde{\alpha}^2\tau_0(pd)^{3t}((n-1)!_t)^3 \\
&\geq \sqrt{\alpha}\tau(pd)^{3(t+1)}(n-1)!_t(n!_{t+1})^2
\end{aligned}
$$

for all $\tau \leq \tau_0$. Taking a union bound over all $v_1 \in V^1$ and all $(1-\alpha)n \leq t < n$ shows that the claimed result holds with probability at least n^{-K}, finishing the proof. It remains to prove Steps 1 to 4.

Proof of Step 1. For $i = 2, 3$, let $A_i := N(v_1; V^i)$. Furthermore, let $A_2' \subseteq V^2$ be the set of vertices $u_2 \in V^2$ for which (v_1, u_2) is good and let $A_2'' \subseteq V^2$ be the set of vertices

$u_2 \in V^2$ for which $\deg(v_1, u_2; V^3) \geq d^2n/2$. Note that $X_2(v_1) = A_2 \cap A_2' \cap A_2''$. Since (V^1, V^i) is $(\varepsilon, d, d-\varepsilon)$-super-regular, we have $|A_i| \geq (d-\varepsilon)n$ for $i = 2, 3$. Since (V^2, V^3) is ε-regular, we have $|A_2''| \geq (1-\varepsilon)n$. Finally, it follows from Lemma 5.1.4, that $|A_2'| \geq (1-\alpha)n$ with probability at least $1 - n^{-8K}$ and hence

$$|X_2(v_1)| = \left|A_2 \cap A_2' \cap A_2''\right| \geq dn/2,$$

as claimed. ∎

Proof of Step 2. Fix some $v_2 \in V^2$. We may assume that $v_2 \in X_2(v_1)$, otherwise we are done. In particular, we have $v_1v_2 \in E(\Gamma)$ and $\deg_\Gamma(v_1, v_2; V^3) \geq d^2n/2$. Let $X_3 = X_3(v_1, v_2)$ and $Y_3 = Y_3(v_1, v_2) \subseteq X_3$. It follows from Lemma 5.1.4 that

$$\mathbb{P}\left[|X_3| < d^2n/4\right] \leq n^{-10K}. \tag{5.3.3}$$

For $u_3 \in V^3$, let Y_{u_3} be the indicator random variable which is one if and only if $v_2u_3 \in E(G)$. Then $|Y_3| = \sum_{u_3 \in X_3} Y_{u_3}$. Note that, in order to determine X_3, we do not need to reveal edges adjacent to v_2. Therefore, X_3 is independent from $\{Y_{u_3} : u_3 \in V^3\}$[2] and we have by Chernoff's bound (Lemma 2.4.1)

$$\mathbb{P}\left[Y_3 \leq p|X_3|/2\right] \leq e^{-p|X_3|/8}. \tag{5.3.4}$$

Furthermore, if $|X_3| \geq d^2n/4$, we have

$$p|X_3| \geq pd^2n/4 \geq 100K \log n \tag{5.3.5}$$

for all large enough n. Hence, taking a union bound over (5.3.3) and (5.3.4), we have

$$|Y_3| \geq p|X_3|/2 \geq pd^2n/8$$

with probability at least $1 - n^{-10K} - e^{p|X|/8} \geq 1 - 2n^{-10K}$. Taking another union bound over all $v_2 \in V^2$ completes the proof. ∎

[2]One can see this as a two-step process: first we reveal edges not incident to v_2 to determine X_3, and then we consider X_3 as a fixed deterministic set.

Proof of Step 3. By Steps 1 and 2

$$|Z'| = \sum_{u_2 \in X_2(v_1)} |Y_3(v_1, u_2)| \geq dn/2 \cdot pd^2 n/8 = pd^3 n^2/16$$

holds with probability at least $1 - 2n^{-8K}$. Furthermore, by Chernoff's bound (Lemma 2.4.1), we have $\Delta(G) \leq 4pn$ with probability at least $1 - n^{-8K}$. Hence the claimed result follows from another union bound. ■

Proof of Step 4. Let $G_1 = G[V(G) \setminus \{v_1\}]$ and let $G_2 = G \setminus G_1$ (i.e. the graph with all edges incident to v_1). Note that $G = G_1 \cup G_2$. Call an edge u_2u_3 with $u_2 \in V^2$ and $u_3 \in V^3$ nice if $v_1u_2, v_1u_3 \in E(G)$. Observe that in order to determine Z', we only need to reveal G_1 and in order to determine the events $\{u_2u_3 \text{ is nice}\}$, we only need to reveal G_2. Therefore, we can treat Z' as a fixed set in the following application of Janson's inequality (Lemma 2.4.2). Furthermore, we may assume that $|Z'| \geq pd^3n^2/16$ and $\Delta(G) \leq 4pn$, otherwise there is nothing to prove. In order to apply Janson's inequality, define

$$\lambda = \mathbb{E}\left[|Z|\right] = p^2|Z'| \geq (pd)^3 n^2/16 \geq C^2 \cdot \log n$$

and

$$\Delta := \sum_{e,f \in Z' \colon e \cap f \neq \emptyset} \mathbb{P}\left[e, f \text{ are both nice}\right] \leq p^2|Z'| + p^3|Z'|8pn$$
$$= p^2|Z'|(1 + 8p^2 n) \leq 2\lambda.$$

Here, we used in the first inequality that there are $p^2|Z'|$ pairs (e, f) sharing two vertices and at most $|Z'|8pn$ pairs (e, f) sharing exactly one vertex (using $\Delta(G) \leq 4pn$). Thus, by Lemma 2.4.2, we have

$$\mathbb{P}\left[|Z| \leq \lambda/2\right] \leq e^{-\lambda^2/8\Delta} \leq e^{-\lambda/16} \leq n^{-C^2/16} \leq n^{-4K}, \tag{5.3.6}$$

finishing the proof. ■

□

5.3.3 Completing Triangle Tilings

With Proposition 5.1.2 and Proposition 5.1.3 in hand the proof of Theorem 5.1.1 follows easily.

Proof of Theorem 5.1.1. First note that, w.v.h.p., both the conclusions of Propositions 5.1.2 and 5.1.3 hold simultaneously. We will now assume they hold and show that this implies

$$\left|\Phi^t(\Gamma_p)\right| \geq (1-\eta)^n \alpha^{t-(1-\eta)n} (pd)^{3t} \left(n!_t\right)^3 \tag{5.3.7}$$

for all $(1-\eta)n \leq t \leq n$. Indeed, for $t = (1-\eta)n$, (5.3.7) readily follows from (the assumed conclusion of) Proposition 5.1.2. Assume now (5.3.7) holds for some $(1-\eta)n \leq t < n$. Since $\eta \ll \eta' \ll \alpha$, we have

$$\begin{aligned}
\tau := (1-\eta)^n \alpha^{t-(1-\eta)n} &\geq (1-\eta)^n \alpha^{\eta n} \\
&\geq (1-\eta)^n e^{-\log(1/\alpha)\eta n} \\
&\geq (1-\eta)^n (1-\log(1/\alpha)\eta)^n \\
&\geq (1-\eta')^n.
\end{aligned}$$

It follows from (the assumed conclusion of) Proposition 5.1.3 that (5.3.7) holds for $t+1$. In particular, we have

$$\left|\Phi^n(\Gamma_p)\right| \geq (1-\eta)^n \alpha^{\eta n} (pd)^{3t} \left(n!\right)^3 \geq 1,$$

completing the proof. □

5.4 Proof of the Local Distribution Lemma

5.4.1 A Simplification

Given some some t, ℓ and $\underline{u} = (u_1, \ldots, u_{\ell-1})$ as in the statement of Lemma 5.1.4, we aim to prove a lower bound on the size of $\Phi^t_{(\underline{u}, u_\ell)}$ for almost all of the $u_\ell \in V^\ell$. Given that $\Phi^t_{\underline{u}}$ is large, a simple averaging argument shows that (5.1.5) is true 'on average' (i.e. if we take the average of $\left|\Phi^t_{(\underline{u}, u_\ell)}(\Gamma_p)\right|$ over all $u_\ell \in V^\ell$). The challenge comes in proving that (5.1.5) holds for almost all choices of u_ℓ. In order to do this, we compare the difference in the sizes of $\Phi^t_{(\underline{u}, u_\ell)}$ for different choices of $u_\ell \in V^\ell$. The key step is given in the following lemma.

Lemma 5.4.1. *W.v.h.p., the following holds for all* $\ell \in [3]$, $(1-\eta)n \leq t < n$, $\underline{u} = (u_1 \dots, u_{\ell-1}) \in \mathcal{V}$ *and* $u \in V^\ell$. *If*

$$\left|\Phi^t_{(\underline{u}, u)}(\Gamma_p)\right| \geq (1-\eta')^n (pd)^{3t}((n-1)!_t)^\ell (n!_t)^{3-\ell},$$

then

$$\left|\Phi^t_{(\underline{u}, v)}(\Gamma_p)\right| \geq \left(\frac{d}{10}\right)^2 \cdot \left|\Phi^t_{(\underline{u}, u)}(\Gamma_p)\right|.$$

for at least $(1-\alpha)n$ *vertices* $v \in V^\ell$.

Indeed, with Lemma 5.4.1 in hand, Lemma 5.1.4 follows easily.

Proof of Lemma 5.1.4. Assume that the conclusion of Lemma 5.4.1 holds in Γ_p. Fix $\ell \in [3]$, $(1-\eta)n \leq t < n$, $\underline{u} = (u_1 \dots, u_{\ell-1}) \in \mathcal{V}$ and $\tau = \tau(n) \geq (1-\eta')^n$. We may assume that

$$\left|\Phi^t_{\underline{u}}(G)\right| \geq \tau (pd)^{3t}((n-1)!_t)^{\ell-1}(n!_t)^{4-\ell},$$

otherwise there is nothing to prove. Now, for each $\phi \in \Phi^t_{\underline{u}}(G)$, we have $\phi \in \Phi^t_{(\underline{u}, u_\ell)}(G)$ for exactly $n-t$ choices of $u_\ell \in V^\ell$. Therefore, we have that

$$\begin{aligned}\sum_{u \in V^\ell} \left|\Phi^t_{(\underline{u}, u)}(G)\right| &= (n-t)\left|\Phi^t_{\underline{u}}(G)\right| \\ &\geq \tau (pd)^{3t}((n-1)!_t)^{\ell-1}(n!_t)^{4-\ell}(n-t) \\ &= \tau (pd)^{3t}((n-1)!_t)^{\ell-1}(n!_{t+1})(n!_t)^{3-\ell}.\end{aligned}$$

Therefore, by averaging, there must be some $u^* \in V^\ell$ such that

$$\begin{aligned}\left|\Phi^t_{\cancel{(\underline{u}, u^*)}}(G)\right| &\geq \frac{1}{n}\left(\tau(pd)^{3t}((n-1)!_t)^{\ell-1}(n!_{t+1})(n!_t)^{3-\ell}\right) \\ &= \tau(pd)^{3t}((n-1)!_t)^{\ell}(n!_t)^{3-\ell}.\end{aligned}$$

The result now follows from applying the conclusion of Lemma 5.4.1 with u^* playing the role of u. □

5.4.2 Entropy lemma

Let us fix constants $\beta, \beta', \beta'', \gamma > 0$ satisfying $\eta' \ll \beta'' \ll \gamma \ll \beta' \ll \beta \ll \alpha$ for this section. We start with a few definitions. Given some $\ell \in [3]$, $t \in [n]$ and some $\phi \in \Phi^t(\Gamma)$, we define $I^\ell(\phi) \subseteq V^\ell$ to be the vertices in V^ℓ which are isolated in the embedded subgraph $\phi(H_t)$ (that is, $I^\ell(\phi) := V^\ell \setminus \operatorname{rg}\phi$). If ℓ is clear from context, we will drop the superscript. If we are further given some $v \in V^\ell$, we define

$$\phi_v = \begin{cases} \emptyset & \text{if } v \in I(\phi), \\ \left(N^j_{\phi(H_t)}(v) : j \in J\right) & \text{if } v \notin I(\phi), \end{cases}$$

where $J = [3] \setminus \{\ell\}$. So ϕ_v either returns an empty set, indicating that the vertex v is isolated in $\phi(H_t)$, or it returns the pair of vertices which are contained in the triangle containing v in $\phi(H_t)$. We also define the function

$$Y_v(\phi) = \mathbb{1}[\{\phi_v \neq \emptyset\}] = \begin{cases} 1 & \text{if } \phi_v \neq \emptyset, \\ 0 & \text{if } \phi_v = \emptyset, \end{cases}$$

which returns 1 if $v \notin I(\phi)$ and 0 otherwise. If the input ϕ of Y_v is clear from context then we simply denote $Y_v(\phi)$ by Y_v. Note that the set $\{\phi_v : v \in V^\ell\}$ completely determines the subgraph $\phi(H_t)$.

For a fixed $v \in V^\ell \setminus \{u\}$, we will be interested in the distribution of ϕ_v if ϕ is chosen randomly among a set of embeddings we wish to extend. In order to analyse this, we use entropy. See Section 2.5 for the definition and basic properties. We remark that there will be two independent stages of randomness in the argument.

First, there is the random subgraph $\Gamma_p \subseteq \Gamma$, and second, there will be a randomly chosen $\phi^* \in \Phi^t(\Gamma_p)$. In particular, the values of the entropy function $h(\phi^*), h(\phi^*_v)$ are random variables themselves. However, once we reveal a particular instance $G = \Gamma_p$, these values are deterministic. We proceed with the following important definition.

Definition 5.4.2. Let $\hat{h} = \hat{h}(n) := \log\left((pd)^3 \cdot n^2\right)$.

The relevance of this definition comes in the form of the following observation.

Observation 5.4.3. *W.v.h.p., the following is true for all $\ell \in [3]$, $(1-\eta)n \le t < n$, $\underline{u} = (u_1 \dots, u_{\ell-1}) \in \mathcal{V}$ and $u \in V^\ell$. Let $\underline{u}' := (\underline{u}, u)$ and let ϕ^* be chosen uniformly from $\Phi^t_{\underline{u}'}(\Gamma_p)$. Then, $h(\phi^*_v | Y_v = 1) \le \hat{h} + \varepsilon'$ for all but at most $\varepsilon' n$ vertices $v \in V^\ell$.*

Proof. Let $G = \Gamma_p$ be revealed. By Corollary 5.2.2, we have w.v.h.p.

$$|\mathrm{Tr}_v(G)| = (1 \pm \varepsilon')(pd)^3 n^2$$

and in partcular $\log|S_{\underline{u}}(v)| \le \hat{h} + \varepsilon'$ for all but at most $\varepsilon' n$ vertices $v \in V^\ell$. Therefore, by Lemma 2.5.1, we have $h(\phi^* | Y_v = 1) \le \hat{h} + \varepsilon'$. □

The purpose of this section is to prove the following lemma, which provides a partial converse to the above observation, showing that for almost all vertices $v \in V^\ell$, $\hat{h}$ is a good approximation for the entropy $h(\phi^*_v | Y_v = 1)$.

Lemma 5.4.4 (Entropy lemma)**.** *W.v.h.p., the following is true for all $\ell \in [3]$, $(1-\eta)n \le t < n$, $\underline{u} = (u_1 \dots, u_{\ell-1}) \in \mathcal{V}$ and $u \in V^\ell$. Let $\underline{u}' := (\underline{u}, u)$ and let ϕ^* be chosen uniformly from $\Phi^t_{\underline{u}'}(\Gamma_p)$. Then, if*

$$\left|\Phi^t_{\underline{u}'}(\Gamma_p)\right| \ge (1-\eta')^n (pd)^{3t} ((n-1)!_t)^\ell (n!_t)^{3-\ell},$$

we have $h(\phi^*_v | Y_v = 1) \ge \hat{h} - \beta'$ for all but at most $\beta' n$ vertices $v \in V^\ell$.

In the remainder of this section, we will prove Lemma 5.4.4. Recall that $V(\Gamma) = V(\Gamma_p) = V^1 \cup V^2 \cup V^3$. As above, for an embedding $\phi \in \Phi(\Gamma)$ and some $\ell \in [3]$, we denote by $I(\phi) = I^\ell(\phi)$ the vertices in V^ℓ which are not contained in the subgraph

$\phi(H_t)$. In the proof, we will describe ϕ by revealing the status of ϕ_v one by one for each $v \in V^\ell$ according to some linear order σ of V^ℓ for some $\ell \in [3]$. In order to do so, we need to make a couple of definitions. Given some $\ell \in [3]$, an ordering σ of V^ℓ, some $j \in [3] \setminus \{\ell\}$, some $v \in V^\ell$ and some $\underline{u} \in \mathcal{V}$, we define

$$A_v^j(\phi, \sigma, \underline{u}) := \left\{ a \in V_{\underline{u}}^j : a \notin \bigcup_{w \in V^\ell :\, w <_\sigma v} N_{\phi(H_t)}(w) \right\}$$

and $A_v(\phi, \sigma, \underline{u}) := \bigcup_{j \in J} A_v^j(\phi, \sigma, \underline{u})$, wehre $J := [3] \setminus \{\ell\}$. We think of these vertices as being *'alive'* at the point we reveal ϕ_v. By 'alive', we mean that it is still possible that ϕ_v reveals that $a \in A_v^j(\phi, \sigma, \underline{u})$ is in a triangle with v. All other vertices $a \in V^j \setminus A_v^j(\phi, \sigma, \underline{u})$ are already embedded in triangles with vertices $w \in V^\ell$ which come before v in the ordering σ.

Triangles with alive vertices

In this subsection, we will prove that most vertices $v \in V^\ell$ are in the expected number of triangles with the other two vertices still being 'alive'. This will be useful in the proof of Lemma 5.4.4.

Lemma 5.4.5. *W.v.h.p., the following is true for all $\ell \in [3]$, $t \in [n-1]$, $\underline{u} = (u_1 \dots, u_{\ell-1}) \in \mathcal{V}$, $u_\ell \in V^\ell$, $\phi \in \Phi^t_{(\underline{u}, u_\ell)}(\Gamma_p)$ and every ordering σ of V^ℓ. There are at most $\varepsilon' n$ vertices $v \in V^\ell$ for which*

$$\left|\mathrm{Tr}_v(\Gamma_p[A_v(\phi, \sigma, \underline{u})])\right| > (pd)^3 \prod_{j \in J} \left|A_v^j(\phi, \sigma, \underline{u})\right| + \varepsilon'(pd)^3 n^2, \tag{5.4.1}$$

where, as above, $J = [3] \setminus \{\ell\}$.

Proof. Let $G \subseteq \Gamma$ be any graph satisfying

$$|\mathrm{Tr}(G[X_1 \cup X_2 \cup X_3])| \le (pd)^3 |X_1||X_2||X_3| + 1000\varepsilon(pd)^3 n^3 \tag{5.4.2}$$

for all $X_1 \subseteq V^1$, $X_2 \subseteq V^2$, $X_3 \subseteq V^3$ and note that Γ_p satisfies (5.4.2) w.v.h.p. by Lemma 5.2.1. We will show that G already satisfies the conclusion of Lemma 5.4.5.

Let $\ell \in [3]$, $t \in [n-1]$, $\underline{u} = (u_1 \dots, u_{\ell-1}) \in \mathcal{V}$, $u_\ell \in V^\ell$, $\phi \in \Phi^t_{\cancel{(\underline{u}, u_\ell)}}(\Gamma_p)$ and let σ be an ordering of V^ℓ. Enumerate $V^\ell = \{v^\ell_1, \dots, v^\ell_n\}$ according to the ordering σ, that is in such a way that $v^\ell_1 <_\sigma \dots <_\sigma v^\ell_n$. Define $U \subseteq V^\ell$ to be the set of vertices satisfying (5.4.1). We will show that $|U| < \varepsilon' n$. We split V^ℓ into intervals as follows. Let $\varepsilon'' = \varepsilon'/4$ and for $k = 1, \dots, 1/\varepsilon''$, let

$$W_k = \{v^\ell_i : 1 + (k-1) \cdot \varepsilon'' n \le i < 1 + k \cdot \varepsilon'' n\}$$

and $U_k := U \cap W_k$. Fix some $k \in [4/\varepsilon']$ and let $i_k := 1 + (k-1) \cdot \varepsilon'' n$ and $w_k := v^\ell_{i_k}$ (that is, w_k is the first vertex in W_k). Let $X_\ell = U_k$ and $X_j = A^j_{w_k}(\phi, \sigma, \underline{u})$ for $j \in J = [3] \setminus \{\ell\}$. It follows that, for any $u \in U_k$,

$$\begin{aligned}
\left|\mathrm{Tr}_u(G[\cup_{i\in[3]} X_i])\right| &\ge \left|\mathrm{Tr}_u(G[A_u(\phi, \sigma, \underline{u})])\right| && (5.4.3)\\
&\ge (pd)^3 \prod_{j\in J} \left|A^j_u(\phi, \sigma, \underline{u})\right| + \varepsilon'(pd)^3 n^2. && (5.4.4)\\
&\ge (pd)^3 \prod_{j\in J} \left(|X_j| - \varepsilon'' n\right) + \varepsilon'(pd)^3 n^2 && (5.4.5)\\
&\ge (pd)^3 \prod_{j\in J} |X_j| + \varepsilon'(pd)^3 n^2/2. && (5.4.6)
\end{aligned}$$

Here, (5.4.3) follows from the fact that $u >_\sigma w_k$ and thus $A_u(\phi, \sigma, \underline{u}) \subseteq A_{w_k}(\phi, \sigma, \underline{u})$ for every $u \in U_k$. Furthermore, (5.4.4) follows from $u \in U$ and (5.4.5) from the fact that $|A^j_u(\phi, \sigma, \underline{u})| \ge |A^j_{w_k}(\phi, \sigma, \underline{u})| - \varepsilon'' n$ for all $u \in U_k$ since u and w_k are close in the ordering σ. By summing over all $u \in U_k$, it follows that

$$|\mathrm{Tr}(G[X_1 \cup X_2 \cup X_3])| \ge (pd)^3 \, |X_1| \, |X_2| \, |X_3| + \varepsilon'(pd)^3 |X_\ell| n^2/2.$$

Combining this with (5.4.2) gives $|U_k| = |X_\ell| \le \frac{2000\varepsilon}{\varepsilon'} n < \frac{\varepsilon'^2}{4} n$. It follows that $|U| = \sum_{k=1}^{1/\varepsilon''} |U_k| < \varepsilon' n$, as claimed. □

Proof of entropy lemma

Here, we will prove Lemma 5.4.4. The proof is quite technical and long, so we will break it up in smaller claims along the way.

Proof of Lemma 5.4.4. Assume $G \subseteq \Gamma$ is a subgraph of Γ with $V(G) = V(\Gamma)$ which satisfies the following properties for all $\ell \in [3]$, $(1-\eta)n \leq t < n$, $\underline{u} = (u_1 \ldots, u_{\ell-1}) \in \mathcal{V}$, $u \in V^\ell$, $\phi \in \Phi^t_{(\underline{u},u)}(G)$ and every ordering σ of V^ℓ.

(P.1) For all vertices $v \in V(G)$, we have.

$$|\mathrm{Tr}_v(G)| \leq 10p^3n^2.$$

(P.2) There are at most $\varepsilon' n$ vertices $v \in V^\ell$ for which

$$\left|\mathrm{Tr}_v(G[A_v(\phi, \sigma, \underline{u})])\right| > (pd)^3 \prod_{j \in [3] \setminus \{\ell\}} \left|A_v^j(\phi, \sigma, \underline{u})\right| + \varepsilon'(pd)^3 n^2.$$

By Lemmas 5.2.3 and 5.4.5, G satisfies those properties w.v.h.p. and therefore it suffices to show that G satisfies the conclusion of Lemma 5.4.4.

Fix now $\ell \in [3]$, $(1-\eta)n \leq t < n$, $\underline{u} = (u_1 \ldots, u_{\ell-1}) \in \mathcal{V}$ and $u \in V^\ell$. Let $\underline{u}' := (\underline{u}, u)$ and let $\Phi := \Phi^t_{\underline{u}'}(G)$. Let ϕ^* be chosen uniformly from Φ. We may assume that

$$|\Phi| \geq (1-\eta')^n (pd)^{3t} ((n-1)!_t)^\ell (n!_t)^{3-\ell}, \tag{5.4.7}$$

otherwise there is nothing to prove. In particular, by Lemma 2.5.1, we have

$$\begin{aligned} h(\phi^*) &\geq n\log(1-\eta') + 3t\log(pd) + 3\log(n!_t) - 2\log(n) && (5.4.8)\\ &\geq 3t\log(pd) + 3\log(n!_t) - \beta'' n, && (5.4.9)\end{aligned}$$

where we used $\eta' \ll \beta''$ and that n is large enough in the last step.

Assume for a contradiction that there are at least $\beta' n$ vertices $v \in V^\ell$ such that $h(\phi^*_v | Y_v = 1) < \hat{h} - \beta'$ and let $U \subseteq V^\ell$ be a set of these exceptional vertices of size $|U| = \gamma n$. We will derive an upper bound on $h(\phi^*)$ which contradicts (5.4.9). Recall that $I(\phi) = I^\ell(\phi) \subseteq V^\ell$ is the set of vertices which are isolated in $\phi(H_t)$. We begin as follows

$$
\begin{aligned}
h(\phi^*) &= h\left(\phi^*, \{\phi^*_v\}_{v\in V^\ell}, I(\phi^*)\right) && (5.4.10)\\
&= h\left(\{\phi^*_v\}_{v\in V^\ell}, I(\phi^*)\right) + h\left(\phi^*|\{\phi^*_v\}_{v\in V^\ell}, I(\phi^*)\right) && (5.4.11)\\
&\le h\left(\{\phi^*_v\}_{v\in V^\ell}, I(\phi^*)\right) + \log(t!) && (5.4.12)\\
&= h\left(\{\phi^*_v\}_{v\in V^\ell}|I(\phi^*)\right) + h\left(I(\phi^*)\right) + \log(t!) && (5.4.13)\\
&\le h\left(\{\phi^*_v\}_{v\in V^\ell}|I(\phi^*)\right) + \log(t!) + \log\left(\binom{n}{t}\right) && (5.4.14)\\
&= h\left(\{\phi^*_v\}_{v\in V^\ell}|I(\phi^*)\right) + \log(n!_t). && (5.4.15)
\end{aligned}
$$

Here, we used Lemma 2.5.5 in (5.4.10) and the chain rule (Lemma 2.5.4) in (5.4.11) and (5.4.13). In (5.4.12), we used Lemma 2.5.6 and that the set $\{\phi_v\}_{v\in V^\ell}$ completely determines the subgraph $\phi(H_t)$ and there are $t!$ embeddings $\phi \in \Phi$ which map to the same subgraph $\phi(H_t)$, namely one for each choice of ordering of the triangles. Finally, in (5.4.14) we used Lemma 2.5.1.

Now, in order to estimate this sum further, we fix some ordering σ of V^ℓ in which the vertices in U come first, that is $w <_\sigma w'$ for all $w \in U$ and $w' \in V^\ell \setminus U$. We then reveal vertices in that order and apply the conditional chain rule (Lemma 2.5.8). That is,

$$
\begin{aligned}
h\left(\{\phi^*_v\}_{v\in V^\ell}|I(\phi^*)\right) &= \sum_{v\in V^\ell} h\left(\phi^*_v|\{\phi^*_u : u <_\sigma v\}, I(\phi^*)\right) && (5.4.16)\\
&\le \sum_{v\in U} h\left(\phi^*_v|I(\phi^*)\right) + \sum_{v\in V^\ell\setminus U} h\left(\phi^*_v|\{\phi^*_u : u <_\sigma v\}, I(\phi^*)\right), && (5.4.17)
\end{aligned}
$$

where we applied Lemma 2.5.3 in the second step. We treat the vertices in U separately to those in $V^\ell \setminus U$. To ease notation, we make the following definition. For $\phi \in \Phi$, and $v \in V^\ell$, we let $t_v(\phi)$ denote the number of vertices $u \in V^\ell$ such that $u <_\sigma v$ and $u \notin I(\phi)$. Let us first address the vertices in U.

Claim 5.5. *For all $v \in U$, we have that*

$$h(\phi_v^*|I(\phi^*)) \le \frac{1}{|\Phi|}\sum_{\phi\in\Phi}\mathbb{1}[v \notin I(\phi)]\left(\log\left((pd)^3(n-t_v(\phi))^2\right)-\frac{\beta'}{2}\right) \tag{5.4.18}$$

Proof. To ease notation as before, we let $Y_v = Y_v(\phi^*)$. Now, for each $v \in U$, we have

$$\begin{aligned} h(\phi_v^*|I(\phi^*)) &\le h(\phi_v^*|Y_v) && (5.4.19)\\ &= \mathbb{P}\left[Y_v = 1\right] h(\phi_v^*|\{Y_v = 1\}) + \mathbb{P}\left[Y_v = 0\right] h(\phi_v^*|\{Y_v = 0\}) && (5.4.20)\\ &\le \mathbb{P}\left[Y_v = 1\right] (\hat{h} - \beta') && (5.4.21)\\ &= \frac{1}{|\Phi|}\sum_{\phi\in\Phi}\mathbb{1}\left[v \notin I(\phi)\right](\hat{h}-\beta'). && (5.4.22) \end{aligned}$$

Here we used Lemma 2.5.3 in (5.4.19), and the definition of conditional entropy (2.5.1) in (5.4.20), and the definition of U in (5.4.21). Furthermore, we have $t_v(\phi) \le \gamma n$ for all $v \in U$ and $\phi \in \Phi$ since U comes at the beginning of the ordering σ. Therefore,

$$\begin{aligned} \log\left((pd)^3(n-t_v(\phi))^2\right) &\ge \log\left((pd)^3(1-\gamma)^2n^2\right)\\ &= \hat{h} + 2\log(1-\gamma)\\ &\ge \hat{h} - 4\gamma\\ &\ge \hat{h} - \beta'/2. \end{aligned}$$

This completes the proof of the claim. ∎

We will now deal with the vertices outside U. Given $v \in V^\ell$ and $\phi \in \Phi$, we write

$$h'(v,\phi) := h\left(\phi_v^*|\{I(\phi^*) = I(\phi), \{\phi_u^* = \phi_u\}_{u<_\sigma v}\}\right).$$

Claim 5.6. *The following is true for all $\phi \in \Phi$.*

(i) For all $v \in V^\ell$, we have

$$h'(v,\phi) \le \log\left((pd)^3(n-t_v(\phi))^2\right) + \log\left(10/d^3\right) + \log\left(\frac{n^2}{(n-t_v(\phi))^2}\right).$$

(ii) There exists a set $B(\phi) \subseteq V^\ell$ with $|B(\phi)| \le \beta'' n$, such that for all $v \in V^\ell \setminus B(\phi)$, we have

$$h'(v,\phi) \le \log\left((pd)^3(n-t_v(\phi))^2\right) + \beta''.$$

Proof. The first inequality follows from **(P.1)** and Lemma 2.5.6. Indeed, for all $v \in V^\ell$, we have

$$\begin{aligned}
h'(v,\phi) &\le \log\left(|\mathrm{Tr}_v(G)|\right) \\
&\le \log(10p^3n^2) \\
&= \log\left((pd)^3(n-t_v(\phi))^2\right) + \log\left(10/d^3\right) + \log\left(\frac{n^2}{(n-t_v(\phi))^2}\right).
\end{aligned}$$

For the second inequality, we will use **(P.2)** in combination with Lemma 2.5.6. We have that for all but at most $\varepsilon' n$ vertices,

$$\begin{aligned}
h'(v,\phi) &\le \log\left(\left|\mathrm{Tr}_v(G[A_v(\phi,\sigma,\underline{u})])\right|\right) && (5.4.23)\\
&\le \log\left((pd)^3 \prod_{j\in J}\left|A_v^j(\phi,\sigma,\underline{u})\right| + \varepsilon'(pd)^3n^2\right) && (5.4.24)\\
&\le \log\left((pd)^3(n-t_v(\phi))^2 + \varepsilon'(pd)^3n^2\right). && (5.4.25)
\end{aligned}$$

Observe that $t_v(\phi) \le (1-\beta''/2)n$ for all but at most $\beta'' n/2$ vertices $v \in V^\ell$. In particular, we have

$$(n-t_v(\phi))^2 \ge (\beta'' n)^2/4 \ge (\beta'')^2/(4\varepsilon')\cdot\varepsilon' n^2 \ge 1/\beta''\cdot\varepsilon' n^2$$

for all but at most $\beta'' n/2$ vertices $v \in V^\ell$. (We used that $\varepsilon' \ll \beta''$ here). Plugging this back into (5.4.25), we get

$$\begin{aligned} h'(v,\phi) &\le \log\left((1+\beta'')\cdot(pd)^3(n-t_v(\phi))^2\right) \\ &\le \beta'' + \log\left((pd)^3(n-t_v(\phi))^2\right) \end{aligned}$$

for all but at most $\beta'' n$ vertices $v \in V^\ell$. ■

We will now use Claims 5.5 and 5.6 to finish the proof. Indeed, it follows from Claim 5.5 that

$$\sum_{v\in U} h(\phi_v^*|I(\phi^*)) \le \frac{1}{|\Phi|}\sum_{\phi\in\Phi}\sum_{v\in U} \mathbb{1}[v\notin I(\phi)]\left(\log\left((pd)^3(n-t_v(\phi))^2\right) - \frac{\beta'}{2}\right). \tag{5.4.26}$$

Furthermore, using Claim 5.6 and the definition of conditional entropy, (2.5.2), we have

$$\sum_{v\in V^\ell\setminus U} h\left(\phi_v^*|\{\phi_u^* : u <_\sigma v\}, I(\phi^*)\right) = \sum_{v\in V^\ell\setminus U}\frac{1}{|\Phi|}\sum_{\phi\in\Phi}\mathbb{1}[v\notin I(\phi)]h'(v,\phi) \tag{5.4.27}$$

$$\le \frac{1}{|\Phi|}\sum_{\phi\in\Phi}\left(\beta'' n + N_1(\phi) + \sum_{v\in V^\ell\setminus U}\mathbb{1}[v\notin I(\phi)]\log\left((pd)^3(n-t_v(\phi))^2\right)\right), \tag{5.4.28}$$

where

$$N_1(\phi) = \sum_{v\in B(\phi)} \mathbb{1}[v\notin I(\phi)]\left(\log\left(10/d^3\right) + 2\log\left(\frac{n}{n-t_v(\phi)}\right)\right).$$

Let now

$$\begin{aligned} M(\phi) &:= \sum_{v\in V^\ell}\mathbb{1}[v\notin I(\phi)]\log\left((pd)^3(n-t_v(\phi))^2\right), \text{ and} \\ N_2(\phi) &:= \sum_{v\in U}\mathbb{1}[v\notin I(\phi)]\cdot\beta'/2. \end{aligned}$$

Then, combining (5.4.17), (5.4.26) and (5.4.28), we get

$$h\left(\{\phi_v^*\}_{v\in V^\ell}|I(\phi^*)\right) \le \frac{1}{|\Phi|}\sum_{\phi\in\Phi}\left(M(\phi)+N_1(\phi)+\beta''n-N_2(\phi)\right). \tag{5.4.29}$$

We will bound each of these terms one by one.

Claim 5.7. *For all $\phi \in \Phi$, we have*

$$\begin{aligned}M(\phi) &\le 3t\log(pd)+2\log(n!_t),\\ N_1(\phi) &\le \sqrt{\beta''}n \quad \textit{and}\\ N_2(\phi) &\ge \gamma^2\cdot n.\end{aligned}$$

Before we prove the claim, let us finish the main proof. Combining the claim with (5.4.29), we get (using $\beta'' \ll \gamma$)

$$\begin{aligned}h\left(\{\phi_v^*\}_{v\in V^\ell}|I(\phi^*)\right) &\le 3t\log(pd)+2\log(n!_t)+(\beta''+\sqrt{\beta''}-\gamma^2)n\\ &\le 3t\log(pd)+2\log(n!_t)-2\beta''n.\end{aligned}$$

Plugging this back into (5.4.15), we get

$$h(\phi^*) \le 3t\log(pd)+3\log(n!_t)-2\beta''n,$$

contradicting (5.4.9). Hence it remains to prove Claim 5.7.

Proof. Let $\phi \in \Phi$ and observe that $\{t_v(\phi) : v \in V^\ell \setminus I(\phi)\} = [t-1]_0$. Thus

$$\begin{aligned}M(\phi) &= \sum_{v\in V^\ell\setminus I(\phi)}\log\left((pd)^3(n-t_v(\phi))^2\right)\\ &= \sum_{k=0}^{t-1}\log\left((pd)^3(n-k)^2\right)\\ &= 3t\log(pd)+2\log(n!_t).\end{aligned}$$

Let $B' = B(\phi) \setminus I(\phi)$ and observe that $|B'| \le |B(\phi)| \le \beta''n$. Let $K = \{t_v(\phi) : v \in B'\}$. Enumerate $K = \{k_1, \ldots, k_{|B'|}\}$ so that $k_1 > \ldots > k_{|B'|}$ and observe that

$k_i \leq n - i$ for all $i \in [|B'|]$ and hence

$$
\begin{aligned}
N_1(\phi) &= \sum_{v \in B'} \mathbb{1}[v \notin I(\phi)] \left(\log\left(10/d^3\right) + 2\log\left(\frac{n}{(n - t_v(\phi))}\right)\right) \\
&\leq \beta'' n \log(10/d^3) + \sum_{k=1}^{\beta'' n} 2\log(n/k) \\
&\leq \beta'' n \log(10/d^3) + 2\beta'' n \log(n) - 2\log((\beta'' n)!) \\
&\leq \beta'' n \log(10/d^3) + 2\beta'' n \left(\log(n) - \log(\beta'' n/e)\right) \\
&\leq \sqrt{\beta''} \cdot n,
\end{aligned}
$$

where we used $(\gamma n)! \geq (\gamma n/e)^{\gamma n}$ in the second to last line. Finally, let $U' = U \setminus I(\phi)$ and observe that, since $\eta \ll \gamma$, we have $|U'| \geq \gamma n/2$. Therefore,

$$N_2(\phi) = \sum_{v \in U'} \beta'/2 \geq \gamma^2 n,$$

as claimed. ■

□

5.4.3 Counting via Comparison

In this subsection, we will prove Lemma 5.4.1. The idea is, given $u, v \in V^3$ (say), to define a appropriate weight function ζ on pairs of vertices $(y_1, y_2) \in V^1 \times V^2$ which encodes how many partial triangle tilings there are which leave u, v, y_1 and y_2 isolated. We then apply Lemmas 5.4.4 and 2.5.9 to show that this weight function is approximately constant and use this to bound the number of triangles tilings which leave u isolated but not v and vice-versa. The reader might want to recall Definition 5.2.4 before the proof.

Proof of Lemma 5.4.1. Let $\ell \in [3]$, $(1 - \alpha)n \leq t < n$, $\underline{u} = (u_1 \ldots, u_{\ell-1}) \in \mathcal{V}$ and $u \in V^\ell$. We write $J := [3] \setminus \{\ell\}$ and choose $j_1, j_2 \in [3]$ so that $J = \{j_1, j_2\}$.

Let $G = \Gamma_p$ and reveal all edges which are not incident to u. Let ϕ be chosen uniformly from $\Phi^t_{\cancel{(\underline{u}, u)}}(G)$. Define

$$Z_1 := \{v \in V^\ell : h(\phi_v | Y_v = 1) \geq \hat{h} - \beta\}$$

and note that, by Lemma 5.4.4, we have $|Z_1| \geq (1-\beta)n$ w.v.h.p. Furthermore, define

$$\begin{aligned} Z_2 &:= \left\{v \in V^\ell : |\mathrm{Tr}_v(G)| = (1 \pm \varepsilon')(pd)^3 n^2\right\}, \\ Z_3 &:= \{v \in V^\ell : \left|F_{\underline{\cancel{u}}}(u, v)\right| \geq d^5 p n^2/4\}. \end{aligned}$$

It follows from Lemmas 5.4.4 and 5.2.5 and Corollary 5.2.2 that $Z := Z_1 \cap Z_2 \cap Z_3$ satisfies $|Z| \geq (1 - \beta - 2\varepsilon')n \geq (1-2\beta)n$ w.v.h.p. Let us fix some vertex $v \in Z$. We will show that v satisfies the conclusion of Lemma 5.4.1. Define

$$\begin{aligned} \Phi_{\cancel{uv}} &:= \Phi^t_{\cancel{(\underline{u}, u)}}(G) \cap \Phi^t_{\cancel{(\underline{u}, v)}}(G), \\ \Phi_{v\cancel{u}} &:= \Phi^t_{\cancel{(\underline{u}, u)}}(G) \setminus \Phi_{\cancel{uv}} \text{ and} \\ \Phi_{u\cancel{v}} &:= \Phi^t_{\cancel{(\underline{u}, v)}}(G) \setminus \Phi_{\cancel{uv}}. \end{aligned}$$

In words, $\Phi_{\cancel{uv}}$ consists of those embeddings which leave both u and v isolated whilst embeddings in $\Phi_{v\cancel{u}}$ leave u isolated but have v contained in a triangle, and similarly for $\Phi_{u\cancel{v}}$. Clearly, we have

$$\begin{aligned} \left|\Phi^t_{\cancel{(\underline{u}, u)}}(G)\right| &= |\Phi_{\cancel{uv}}| + |\Phi_{v\cancel{u}}|, \text{ and} \\ \left|\Phi^t_{\cancel{(\underline{u}, v)}}(G)\right| &= |\Phi_{\cancel{uv}}| + |\Phi_{u\cancel{v}}|. \end{aligned}$$

If $|\Phi_{\cancel{uv}}| \geq (d/10)^2 \left|\Phi^t_{\cancel{(\underline{u}, u)}}(G)\right|$, we are done and so we may assume that

$$|\Phi_{v\cancel{u}}| \geq (1 - (d/10)^2) \left|\Phi^t_{\cancel{(\underline{u}, u)}}(G)\right| \geq \frac{1}{2} \left|\Phi^t_{\cancel{(\underline{u}, u)}}(G)\right|. \tag{5.4.30}$$

In what remains, we will compare the sizes of $\Phi_{v\cancel{u}}$ and $\Phi_{u\cancel{v}}$. For $(w_1, w_2) \in V^{j_1}_{\underline{\cancel{u}}} \times V^{j_2}_{\underline{\cancel{u}}}$, define $\zeta(w_1, w_2)$ to be t times the number of labelled embeddings of H_{t-1} into $G_{\underline{u}, \cancel{u}, \cancel{v}}$

in which both w_1 and w_2 are isolated vertices. That is

$$\zeta(w_1, w_2) := t \cdot \left| \Phi^{t-1}_{\cancel{(w_1, w_2)}} \left(G_{\cancel{\underline{u}, u, v}} \right) \right| . \tag{5.4.31}$$

Note that ζ is independent from all edges incident to u or v. Observe that

$$|\Phi_{v\cancel{u}}| = \sum_{(y_1,y_2) \in S_{\underline{\cancel{u}}}(v)} \zeta(y_1, y_2), \text{ and}$$
$$|\Phi_{\cancel{v}u}| = \sum_{(y_1,y_2) \in S_{\underline{\cancel{u}}}(u)} \zeta(y_1, y_2).$$

Since $\phi_v | \{Y_v = 1\}$ is a random variable taking values in $S_{\underline{\cancel{u}}}(v)$ and $v \in Z_1$, we can apply Lemma 2.5.9 to get some set $W^* \subseteq S_{\underline{\cancel{u}}}(v)$ with the following properties:

(i) $\displaystyle\sum_{(w_1,w_2) \in W^*} \zeta(w_1, w_2) \geq (1 - \beta) \, |\Phi_{v\cancel{u}}|$;

(ii) There exists some value $\bar{\zeta}$ such that for each $(w_1, w_2) \in W^*$, we have that

$$\zeta(w_1, w_2) = (1 \pm \beta)\bar{\zeta};$$

(iii) We have $(1 - \beta)|S_{\underline{\cancel{u}}}(v)| \leq |W^*| \leq |S_{\underline{\cancel{u}}}(v)|$.

Therefore we can estimate the size of $\Phi_{v\cancel{u}}$ using (i) to (iii) in that order.

$$|\Phi_{v\cancel{u}}| \leq \left(\frac{1}{1 - \beta} \right) \sum_{(w_1,w_2) \in W^*} \zeta(w_1, w_2) \tag{5.4.32}$$
$$\leq \left(\frac{1 + \beta}{1 - \beta} \right) |W^*| \bar{\zeta} \tag{5.4.33}$$
$$\leq \left(\frac{1 + \beta}{1 - \beta} \right) |S_{\underline{\cancel{u}}}(v)| \bar{\zeta} \tag{5.4.34}$$
$$\leq 2\bar{\zeta}(pd)^3 n^2. \tag{5.4.35}$$

In the last inequality, we used that $|S_{\underline{\cancel{u}}}(v)| = (1 \pm 2\varepsilon')(pd)^3 n^2$ since $v \in Z_2$.

We are now going to lower bound $\Phi_{\cancel{v}u}$ in a similar way. However, so far we only know that ζ is 'well-behaved' on $S_{\underline{\cancel{u}}}(v)$ but nothing about $S_{\underline{\cancel{u}}}(u)$. Using Lemma 5.2.6, we can infer though that ζ is 'well-behaved' on a large part of $F_{\underline{u}}(u,v)$. To do so, partition $F_{\underline{u}}(u,v)$ into two sets F_S, F_L of equal size so that $\zeta(y_1,y_2) \le \zeta(z_1,z_2)$ for all $(y_1,y_2) \in F_S$ and $(z_1,z_2) \in F_L$. Since ζ is independent from all edges adjacent to u or v, so are F_S and F_L. Furthermore, since $v \in Z_3$, we have $|F_S| = |F_L| \ge d^5pn^2/4$. Hence, we can use Lemma 5.2.6 with F_S and F_L. We will assume for the rest of the proof that its conclusion holds in G, both with input $\ell, \underline{u}, u, F_S$ and with input $\ell, \underline{u}, v, F_L$.

Claim 5.8. *We have* $\zeta(y_1,y_2) \ge (1-\beta)\bar{\zeta}$ *for all* $(y_1,y_2) \in F_L$.

Proof. The conclusion of Lemma 5.2.6 (with input u and F_S) implies that

$$\left|S_{\underline{u}}(v) \cap F_S\right| \ge p^2|F_S|/2 \ge d^5p^3n^2/8.$$

Furthermore, it follows from (iii) and $v \in Z_2(u)$, that

$$\left|S_{\underline{u}}(v) \setminus W^*\right| \le \beta\left|S_{\underline{u}}(v)\right| \le 2\beta(pd)^3n^2.$$

Hence, as $\beta \ll d$, we can conclude that $W^* \cap F_S^t \ne \emptyset$ and so

$$(1-\beta)\bar{\zeta} \le \min_{(y_1,y_2)\in W^*} \zeta(y_1,y_2) \le \max_{(y_1,y_2)\in F_S} \zeta(y_1,y_2) \le \min_{(y_1,y_2)\in F_L} \zeta(y_1,y_2),$$

using (ii) in the first inequality. ■

We now use the conclusion of Lemma 5.2.6 again (with input v and F_L) to lower bound the size of Φ_{uv} as follows.

$$\begin{aligned}
|\Phi_{uv}| &= \sum_{(y_1,y_2)\in S_{\underline{u}}(u)} \zeta(y_1,y_2) && (5.4.36)\\
&\ge \sum_{(y_1,y_2)\in S_{\underline{u}}(u)\cap F_L} \zeta(y_1,y_2) && (5.4.37)\\
&\ge (1-\beta)\bar{\zeta}\left|S_{\underline{u}}(u) \cap F_L\right| && (5.4.38)\\
&\ge \bar{\zeta}d^5p^3n^2/20, && (5.4.39)
\end{aligned}$$

where (5.4.38) follows from Claim 5.8. Putting (5.4.30), (5.4.35) and (5.4.39) together, we get that

$$\begin{aligned}\left|\Phi^t_{(\underline{\cancel{u}},\cancel{v})}(G)\right| &\geq |\Phi_{u\cancel{v}}| \\ &\geq \frac{\bar{\zeta} d^5 p^3 n^2}{20} \\ &\geq \frac{d^2}{40}\,|\Phi_{v\cancel{u}}| \\ &\geq \frac{d^2}{80}\left|\Phi^t_{(\underline{\cancel{u}},\cancel{u})}(G)\right|,\end{aligned}$$

as required. □

5.5 Reducing to Regular Triples

Here we show how Theorem 1.5.9 follows from Theorem 5.1.1. The proof mainly relies on fairly standard though somewhat involved extremal graph theory. We will collect a number of lemmas in the following subsection and then proceed to the main proof.

5.5.1 Preparation

Fractional Stability for the Hajnal–Szemerédi Theorem

Recall the Hajnal-Szemerédi Theorem (Theorem 2.2.17), which states that any graph with maximum degree Δ has an equitable colouring with $\Delta + 1$ colours (that is, a colouring where the colour classes differ in size by at most one). Applying this to the complement of G, which has maximum degree $n - 1 - \delta(G)$, we find a collection of $n - \delta(G)$ vertex-disjoint cliques in G whose sizes differ by at most one and which cover G. We will make use of the following corollary.

Theorem 5.5.1 (Hajnal–Szemerédi [66]). *Let $n, k \geq 2$ be integers and let $0 \leq x < 1$. Suppose that G is an n-vertex graph with $\delta(G) \geq (\frac{k-1}{k} - x)n$. Then G contains $(1 - (k-1)kx)\lfloor \frac{n}{k} \rfloor$ vertex-disjoint cliques of size k.*

This statement is often used in extremal graph theory, and in particular the case $x = 0$, which gives the best possible minimum degree condition for containing a spanning K_k-tiling. We require the following result, a stability statement for a fractional version of the Hajnal-Szemerédi theorem, which is of independent interest.

Theorem 5.5.2. *For every integer $k \geq 2$ and every $\delta > 0$, there is some $\gamma > 0$ such that the following is true for all $n \in \mathbb{N}$. Let G be an n-vertex graph with $\delta(G) \geq \left(\frac{k-1}{k} - \gamma\right)n$ and $\alpha(G) < \left(\frac{1}{k} - \delta\right)n$. Let $\lambda : V(G) \to \mathbb{N}$ be a weight function with $\lambda(u) = (1 \pm \gamma)\frac{1}{n}\sum_{v \in V(G)} \lambda(v)$ for all $u \in V(G)$. Then there is a weight function $\omega : K_k(G) \to \mathbb{R}_{\geq 0}$ such that $\sum_{K \in K_k(G,u)} \omega(K) = \lambda(u)$ for all $u \in V(G)$, where $K_k(G,u)$ denotes the set of unlabelled copies of K_k in G containing u.*

Proof. First, we claim that it suffices to prove the theorem for the function λ which gives weight 1 to every vertex. Indeed, let k, δ, G and λ as in the statement and let $\gamma \leq \delta$ be small enough to apply the result with weight function $\lambda \equiv 1$ and input $\delta/2$. We define an auxiliary graph H by blowing-up every $v \in V(G)$ to an independent set of size $\lambda(v)$ (that is, every edge is replaced by a complete bipartite graph). Then, with $N := v(H) = \sum_{v \in V(G)} \lambda(v)$, we have $\delta(H) \geq (\frac{k}{k-1} - 2\gamma)N$ and $\alpha(H) \leq (1/k - \delta/2)N$. Hence, we can apply the theorem to H (with $\lambda \equiv 1$ and $\delta/2$) and obtain a weight function $\omega_H : K_k(H) \to \mathbb{R}_{\geq 0}$ with the desired properties. We define $\omega : K_k(G) \to \mathbb{R}_{\geq 0}$ by $\omega(K) = \sum_{K' \in K_k(H[K])} \omega_H(K')$, where $H[K]$ is the subgraph of H induced by the blown-up vertices of K.

Hence, we may assume that $\lambda(v) = 1$ for all $v \in V(G)$. We will prove this case of the theorem for $\gamma = \frac{\delta}{8^k (k!)^2}$. Observe that the existence of the claimed weight function is the same as saying that the value of the following packing linear program is $\frac{n}{k}$. We ask for non-negative real weights on the elements of $K_k(G)$ with maximum sum, subject to the condition that the total weight on copies of K_k at any given vertex is at most 1. The dual of this is the covering linear program in which we place nonnegative weights on the vertices of G, with minimum sum, subject to the constraint that the total weight on the vertices of each element of $K_k(G)$ is at least 1. The duality theorem for linear programs (see [111, Corollary 7.1g]) implies that these two linear programs have the same value. So it is enough to show that the latter linear program has value $\frac{n}{k}$, which we do inductively. In other words, we want to prove the following claim by induction on k. We define $z_2 = 3$ and inductively $z_k = 8k^2 z_{k-1}$ for $k \geq 3$.

Claim 5.9. *Given any $k \geq 2$ and $\gamma > 0$, suppose that G is an n-vertex graph with minimum degree at least $\left(\frac{k-1}{k} - \gamma\right)n$ and no independent set of size $\left(\frac{1}{k} - z_k\gamma\right)n$. Suppose $c : V(G) \to [0,1]$ is any weight function such that for each $Q \in K_k(G)$ we have $\sum_{v \in Q} c(v) \geq 1$. Then $\sum_{v \in V(G)} c(v) \geq \frac{n}{k}$.*

Proof. It is convenient to let the vertices of G be $v_1, \ldots, v_n$ in order of decreasing weight, i.e. $c(v_i) \geq c(v_j)$ if $i \leq j$. If $\sum_{i \in [n]} c(v_i) \geq \frac{n}{k}$ there is nothing to prove, so we can assume the sum is less than $\frac{n}{k}$. If $c(v_n) > 0$, then we can define a new weight function by $c'(v_i) := \frac{1}{k} + \alpha\big(c(v_i) - \frac{1}{k}\big)$ for all $i \in [n]$, where $\alpha > 1$ is chosen so that $c'(v_n) = 0$. We have

$$\begin{aligned}\sum_{i \in [n]} c'(v_i) &= n/k + \alpha \sum_{i \in [n]} (c(v_i) - 1/k) \\ &= \sum_{i \in [n]} c(v_i) + (\alpha - 1)\left(\sum\nolimits_{i \in [n]} c(v_i) - n/k\right) < \sum_{i \in [n]} c(v_i).\end{aligned}$$

However, for every $Q \in K_k(G)$,

$$\sum_{v \in Q} c'(v) = \sum_{v \in Q} \left(\tfrac{1}{k} + \alpha\big(c(v_i) - \tfrac{1}{k}\big)\right) = 1 + \alpha\left(\sum\nolimits_{v \in Q} c(v) - 1\right) \geq 1.$$

Therefore, c' also satisfies the condition of Claim 5.9 and we can assume $c(v_n) = 0$.

We are now in a position to prove the base case $k = 2$. Since v_n has at least $\left(\frac{1}{2} - \gamma\right)n$ neighbours, and $c(v_n) = 0$, we see that for each i such that $v_i v_n \in E(G)$, we have $c(v_i) = 1$. In particular, $c(v_i) = 1$ for each $i \leq \left(\frac{1}{2} - \gamma\right)n$. Furthermore, the vertices $\{v_i : i \geq \frac{n}{2} + 2\gamma n\}$ do not form an independent set, so there is an edge within this set; at least one endpoint of this edge has weight at least $\frac{1}{2}$, and in particular each vertex v_i with $\frac{n}{2} - \gamma n < i < \frac{n}{2} + 2\gamma n$ has weight at least $\frac{1}{2}$. Summing, we obtain weight at least $\frac{n}{2}$ as desired.

Next, we prove the induction step; let $k \geq 3$. We build a copy of K_k containing v_n as follows: we take $u_1 = v_n$, and then for each $2 \leq i \leq k-2$ in succession, we take u_i to be the common neighbour of $u_1, \ldots, u_{i-1}$ with smallest weight. From the minimum degree condition, when we choose u_i there are at least $n - (i-1)\big(\frac{1}{k} + \gamma\big)$ common neighbours to choose from; in particular, we always choose a vertex v_j with $j \geq \frac{3n}{k} - (k-3)\gamma n$, and the common neighbourhood of all $k-2$ vertices we choose has size at least $\frac{2n}{k} - (k-2)\gamma n$. Now consider the last $\big(\frac{1}{k} - k(k-1)\gamma\big)n$ of these common neighbours. Since $z_k \geq k(k-1)$, they do not form an independent set, so contain an edge $u_{k-1}u_k$. Since $\sum_{i=1}^{k} c(u_i) \geq 1$, and $c(u_1) = 0$, one of these vertices has weight at least $\frac{1}{k-1}$. In particular, $c(v_i) \geq \frac{1}{k-1}$ whenever $i \leq \big(\frac{1}{k} + (k-1)^2\gamma\big)n$.

Now let $c^* := c(v_{n/k-(k-1)\gamma n})$, and let G' denote the subgraph of G induced by vertices v_i with $i \geq \big(\frac{1}{k} + (k-1)^2\gamma\big)n$. If $c^* = 1$ then we have

$$\sum_{i\in[n]} c(v_i) \geq \tfrac{n}{k} - (k-1)\gamma n + \tfrac{1}{k-1} \cdot k(k-1)\gamma n > \tfrac{n}{k}$$

and we are done; so we can assume $c^* < 1$. If Q is any copy of K_{k-1} in G', then Q has a common neighbourhood in G of size at least $\frac{n}{k} - (k-1)\gamma n$, and so in particular Q extends to a copy of K_k in G by adding a vertex whose weight is at most c^*. Thus the function $c'(u) := \frac{1}{1-c^*}c(u)$ on $V(G')$ is a weight function on $V(G')$ taking values in $[0,1]$ and such that $\sum_{u\in Q} c(u) \geq 1$ for each $Q \in K_{k-1}(G')$. Furthermore every vertex in G' has at most $\frac{n}{k} + \gamma n$ non-neighbours in G, at most all of which are in G', so the minimum degree of G' is at least $\frac{(k-2)n}{k} - ((k-1)^2+1)\gamma n$. Since $v(G') = \frac{(k-1)n}{k} - (k-1)^2\gamma n$, we have $\delta(G') \geq \frac{k-2}{k-1}v(G') - \gamma' v(G')$ where $\gamma' := 2k^2\gamma$. Furthermore G' has no independent set of size

$$\tfrac{1}{k}n - z_k\gamma n = \tfrac{1}{k}n - 4z_{k-1}\gamma' n \leq \tfrac{1}{k-1}v(G') - z_{k-1}\gamma' v(G')\,.$$

We are therefore in a position to apply the induction hypothesis (that is, Claim 5.9 for $k-1$) to G', with γ' replacing γ. We conclude that

$$\sum_{u\in V(G')} c'(u) \geq \tfrac{1}{k-1}v(G') \geq \frac{(1 - 1/k - (k-1)^2\gamma)n}{k-1} = (\tfrac{1}{k} - (k-1)\gamma)n$$

and so

$$\sum_{i\in[n]} c(v_i) \geq c^*\big(\tfrac{1}{k} - (k-1)\gamma\big)n + \tfrac{1}{k-1}\cdot k(k-1)\gamma n + (1-c^*)\cdot(\tfrac{1}{k} - (k-1)\gamma)n$$
$$= \big(1/k - (k-1)\gamma\big)n + k\gamma n > \tfrac{n}{k},$$

as desired. ■

This completes the proof by LP-duality. □

Note that we obtain from this proof a little more: the unique optimal cover is the uniform cover (since after assuming $c(v_n) < \frac{1}{k}$ we eventually conclude the total weight is strictly bigger than $\frac{n}{k}$). However we will not need this fact. We will also need only the $k = 2$ and $k = 3$ cases, but for future use give the general result. We will further show that we can find an integer-valued weight-function if λ is sufficiently large.

Lemma 5.5.3. *For every integer $k \geq 2$ and every $\delta > 0$, there is some $\gamma > 0$ such that the following is true for all $n \in \mathbb{N}$. Let G be a connected[3] n-vertex graph with $\delta(G) \geq \big(\frac{k-1}{k} - \gamma\big)n$ and $\alpha(G) < \big(\frac{1}{k} - \delta\big)n$. Let $\lambda : V(G) \to \mathbb{N}$ be a weight function such that $\lambda(u) = (1 \pm \delta)\frac{1}{n}\sum_{v\in V(G)} \lambda(v)$, $\lambda(u) \geq n^{2k}$ for all $u \in V(G)$ and k divides $\sum_{v\in V(G)} \lambda(v)$. Then there is a weight function $\omega : K_k(G) \to \mathbb{N}_0$ such that $\sum_{K\in K_k(G,u)} \omega(K) = \lambda(u)$ for all $u \in V(G)$, where $K_k(G,u)$ denotes the set of unlabelled copies of K_k in G containing u.*

Proof. We will construct ω in three steps. Define $\lambda' : V(G) \to \mathbb{N}$ by $\lambda'(u) = \lambda(u) - |K_k(G,u)|n^k \geq 0$. By Theorem 5.5.2, there is some weight function $\omega' : K_k(G) \to \mathbb{R}_{\geq 0}$ such that $\sum_{K\in K_k(G,u)} \omega'(K) = \lambda'(u)$ for all $u \in V(G)$. We define $\omega'' : V(G) \to \mathbb{N}_0$ such that, for each $K \in K_k(G)$,

(i) $\omega''(K) \in \big\{\lfloor \omega'(K) + kn^k \rfloor, \lceil \omega'(K) + kn^k \rceil\big\}$, and

(ii) $k\sum_{K\in K_k(G)} \omega''(K) = \sum_{v\in V(G)} \lambda(v)$.

[3]Note that, if $k \geq 3$, this is already implied by the minimum degree condition.

Note that this is possible since by construction the unrounded sum satisfies (ii) and since k divides $\sum_{v\in V(G)} \lambda(v)$. Furthermore, for each $u \in V(G)$, we have $\sum_{K\in K_k(G,u)} \omega''(K) = \lambda(u) \pm n^{k-1}$ (since the unrounded sum would be exactly correct and $|K_k(G,u)| \le n^{k-1}$).

Finally, we obtain ω from ω'' via the following iterative process. As long as possible, we identify pairs $u, v \in V(G)$ such that $\sum_{K\in K_k(G,u)} \omega''(K) > \lambda(u)$ and $\sum_{K\in K_k(G,v)} \omega''(K) < \lambda(v)$. If $k \ge 3$, we claim that there is a clique of size $k-1$ in the common neighbourhood of u and v. Indeed, since $\delta(G) \ge \left(\frac{k-1}{k} - \gamma\right)n$, we can iteratively find a clique with vertices $u_2, \dots, u_{k-2}$ in the common neighbourhood of u and v and the common neighbourhood of $u, v, u_2, \dots, u_{k-2}$ has size at least $(\frac{1}{k} - (k-1)\gamma))n > (\frac{1}{k} - \delta)n$. In particular, there is an edge $u_{k-1}u_k$ in there, completing the clique. Let $K_u = \{u, u_2, \dots, u_k\}$ and $K_v = \{v, u_2, \dots, u_k\}$, and decrease the weight of K_u by 1 and increase the weight of K_v by 1. If $k = 2$, we do the following: Since $\alpha(G) < n/2$, G is not bipartite and hence contains an odd cycle. Since G is connected, this implies that there is a trail (that is a path with possibly repeated vertices but no repeated edges) from u to v of even length (even number of edges). We decrease the weight of the edge at u and then alternate increasing and decreasing the weight of the edges along the path. Note that in both cases the total weight at u decreases by 1 and the total weight at v increases by 1, and the total weight at any other vertex remains unchanged.

Note that $\sum_{v\in V(G)} \left|\lambda(v) - \sum_{K\in K_k(v,G)} \omega(K)\right|$ decreases by 2 in every step. So this process finishes after at most n^k steps. Clearly, at this time, we have $\sum_{K\in K_k(v,G)} \omega(K) = \lambda(v)$ for all $v \in V(G)$ and $\omega(K) \ge \omega''(K) - n^k \ge 0$ for all $K \in K_k(G)$, completing the proof. □

Probabilistic lemmas

We need a few probabilistic lemmas which allow us to find a collection of vertex-disjoint triangles in G_p with specified properties. In this section we write $V(\mathcal{T})$ for the set of vertices covered by a collection of triangles $\mathcal{T}$. The first lemma allows us to find a large collection of triangles in G_p if G contains many triangles.

Lemma 5.5.4. *For all $\mu > 0$ there exists $C > 0$ such that the following holds for all large enough $n \in \mathbb{N}$. Let $k > 0$, $p \ge Cn^{-2/3}$ and let G be an n-vertex graph.*

(i) *Assume that for every set $X \subseteq V(G)$ with $|X| \geq 3k$, $G[X]$ contains at least μn^3 triangles. Then, w.h.p., G_p contains a collection of at least $\frac{n}{3} - k$ vertex-disjoint triangles.*

(ii) *Assume that $V(G) = V_1 \cup V_2 \cup V_3$ is a partition into sets of size at least n_0 so that for every $X_i \subseteq V_i$ with $|X_i| \geq k$ for all $i \in [3]$, $G[X_1, X_2, X_3]$ contains at least μn^3 triangles. Then, w.h.p., G_p contains a collection of at least $\frac{n_0}{3} - k$ vertex-disjoint triangles.*

Proof. Let $\mu > 0$ and set $C = 50\mu^{-2}$. Let p, k, G be given as in the statement. We will deduce the lemma from the following claim.

Claim 5.10. *The following holds w.h.p. for all $X \subseteq V(G)$. If $G[X]$ contains at least $t \geq \mu n^3$ copies of K_3, then the number of triangles in $G_p[X]$ is at least $\frac{1}{2}p^3 t$.*

Proof. This is a straightforward application of Janson's inequality and the union bound. Note that the total number of choices of X is at most 2^n. Fix one such choice. The expected number of triangles in $G_p[X]$ is $p^3 t \geq \mu p^3 n^3$, and we have $\bar{\Delta} \leq \max(p^5 n^4, p^3 n^3)$. Hence Janson's inequality tells us that the probability of having less than $\frac{1}{2}p^3 t$ triangles is at most

$$\exp\Big(-\frac{\mu^2 p^6 n^6}{8\max(p^5 n^4, p^3 n^3)}\Big) \leq \exp\Big(-\frac{\mu^2}{8}\min(pn^2, p^3 n^3)\Big) \leq \exp\big(-\frac{C\mu^2}{8}n\big)$$

and by choice of C and the union bound the lemma statement follows. ■

We only prove (i) as (ii) is similar. Suppose that $\mathcal{T}$ is a maximal collection of vertex-disjoint triangles with $|\mathcal{T}| < \frac{n}{3} - k$. Then $X := V(G) \setminus V(\mathcal{T})$ has size at least $3k$ but $G_p[X]$ does not contain a triangle. Thus, the claimed result follows from the above claim. □

The next lemma allows us to find triangles which cover a given small set of vertices, using edges in specified places.

Lemma 5.5.5. *For any $0 < \mu < 1/100$ there exists $C > 0$ such that the following holds for every $n \in \mathbb{N}$ and $p \geq Cn^{-2/3}(\log n)^{1/3}$. Let G be an n-vertex graph, and let $v_1, \ldots, v_\ell \in V(G)$ be distinct vertices with $\ell \leq \mu^2 n$. For each $i \in [\ell]$,* *let $E_i \subseteq E(G)$ be a set of edges with $|E_i| \geq \mu n^2$ such that $v_i e$ is a triangle of G for each $e \in E_i$. Let $A_1, \ldots, A_t \subseteq V(G) \setminus \{v_1, \ldots, v_\ell\}$ be disjoint sets for some $t \in \mathbb{N}_0$. Then, w.h.p., there is a set $\mathcal{T} = \{T_1, \ldots, T_\ell\}$ of vertex-disjoint triangles in G_p such that for each $i \in [\ell]$ the triangle T_i consists of v_i joined to an edge of E_i and $|A_i \cap V(\mathcal{T})| \leq 12\mu|A_i| + 1$ for all $i \in [t]$.*

Proof. Given $0 < \mu < 1/100$, we set $C = 1000\mu^{-1}$. Given n, we can assume $p = Cn^{-2/3}(\log n)^{1/3}$, since the probability of any given collection of triangles of G appearing in G_p is monotone increasing in p.

We use a careful step-by-step revealing argument and choose $T_1, \ldots, T_\ell$ one at a time. Given $k \in [t]$ and $i \in [\ell]$, say that A_k is *full* at time i if $|A_k \cap V(\{T_1, \ldots, T_{i-1}\})| \geq 12\mu|A_k|$. Let X_i be the set of vertices in some set A_k which is full at time i. For each step $i \in [\ell]$ in succession, we will reveal certain edges of G_p and then choose a triangle T_i among the edges revealed. Specifically, we reveal first the edges S_i of G_p at v_i which do not go to $v_1, \ldots, v_\ell$, X_i or a vertex of $T_1, \ldots, T_{i-1}$. We then reveal all edges of E_i surviving in G_p which form a triangle using two edges of S_i and which were not previously revealed. From these edges we pick any triangle T_i, and move on to the next i.

Observe that by definition we do not reveal any edge of G_p twice; and if we successfully choose a triangle at each step we indeed obtain the desired collection of vertex-disjoint triangles. To begin with, we argue that when we come to v_i, most edges of E_i are still available to choose from. Note that we cannot use an edge of E_i which is adjacent to any v_j or T_j; there are at most $3\mu^2 n$ such vertices, which are adjacent to at most $3\mu^2 n^2$ edges of E_i. We also cannot use an edge adjacent to X_i; we have $|X_i| \leq \frac{3\ell}{12\mu} \leq \frac{\mu}{4}n$ and hence there are at most $\frac{\mu}{4}n^2$ edges adjacent to X_i. We also cannot use an edge of E_i which was previously revealed. When we reveal edges at some v_j, we reveal by Chernoff's inequality w.v.h.p. at most $2pn = 2Cn^{1/3}(\log n)^{1/3}$ edges, and hence we reveal at most $4C^2 n^{2/3} \log^{2/3} n$ edges of E_i in this step. Since there are at most $\mu^2 n$ steps, in total we reveal less than $n^{7/4}$ edges. Putting this together, w.v.h.p., for each i there remain at least $\frac{1}{2}\mu n^2$ edges of E_i which we could use to make triangles with v_i. Let F_i denote the set of edges of E_i which remain usable at the beginning of step i.

When we reveal edges at v_i, for each edge of F_i by definition we keep the edges from v_i to the endpoints of F_i with probability p^2, and so the expected number of edges of F_i whose ends are both adjacent to v_i is $p^2|F_i| \geq \frac{1}{2}p^2\mu n^2$. Applying Janson's inequality, we have $\bar{\Delta} \leq p^3 n^3$, which is tiny compared to the square of the expectation, so w.v.h.p. at least $\frac{1}{4}p^2\mu n^2$ edges of F_i are revealed to be neighbours of v_i. We now reveal which of these edges survive in G_p; by Chernoff's inequality and by choice of C, with probability at least $1 - n^{-2}$, at least $\frac{1}{8}p^3\mu n^2$ of these edges survive in G_p, and in particular T_i exists.

Taking a union bound, the probability of failure at any step is $o(1)$. □

The last lemma allows us to find a reasonably large collection of vertex-disjoint triangles using a possibly sparse set of edges each of which extends to many triangles; we will use this to deal with nearly independent sets which have size larger than $\frac{1}{3}n$. Here we denote by $\deg_G(e, X)$ the size of the common neighbourhood of an edge e inside a set X and, given a set of edges E, we will sometimes think of E as the graph $H_E := (\bigcup E, E)$ and use notation like te $\delta(E) := \delta(H_E)$ or $\deg_E(v) := \deg_{H_E}(v)$.

Lemma 5.5.6. *For any $0 < \mu < \frac{1}{1000}$ there exists $C > 0$ such that the following holds for all $n, \delta, \delta_1, \delta_2 \in \mathbb{N}$, every n-vertex graph G and every $p \geq Cn^{-2/3}(\log n)^{1/3}$.*

(i) Let $X_1, X_2, X_3 \subseteq V(G)$ be disjoint sets of size at least $n/10$, and let $E \subseteq E(G[X_1])$ be a set of edges such that $\deg_E(v) \geq \delta$ for all $v \in X_1$ and $\deg_G(e, X_i) \geq \mu n$ for all $e \in E$ and $i = 2, 3$. Let n_2, n_3 be non-negative integers with $n_2 + n_3 \leq \min(\delta, \mu^5 n)$. Then, w.h.p., there is a set $T_1, \ldots, T_{n_2+n_3}$ of vertex-disjoint triangles in G_p, n_i of which consist of an edge $e \in E$ together with a vertex of X_i for each $i = 2, 3$.

(ii) Let $X_1, X_2 \subseteq V(G)$ be disjoint sets of size at least $n/10$. Let $E_i \subseteq E[G[X_i]]$ be sets of edges such that $\deg_{E_i}(v) \geq \delta_i$ for all $v \in X_i$ and $\deg(e, X_{3-i}) \geq \mu n$ for all $e \in E_i$ and $i \in [2]$. Let $n_i \leq \min(\delta_i, \mu^5 n)$ be non-negative integers for each $i \in [2]$. Then, w.h.p., there is a set $T_1, \ldots, T_{n_1+n_2}$ of vertex-disjoint triangles in G_p, n_i of which consist of an edge $e \in E_i$ together with a vertex of X_{3-i} for each $i \in [2]$.

Observe that, unlike other lemmas in this section, both cases of this lemma are very tight and we can't even guarantee more vertex-disjoint triangles in the underlying graph G. If the edges E have small maximum degree however, the situation is somewhat easier and we will make use of this in the proof of Lemma 5.5.6:

Lemma 5.5.7. *For all $\mu > 0$ there exists $C > 0$ such that the following holds for all $n \in \mathbb{N}$, every n-vertex graph G and every $p \geq Cn^{-2/3}(\log n)^{1/3}$. Suppose that E is a subset of $E(G)$ with $\Delta(E) \leq \mu n$ and $\mu n \leq |E| \leq \mu^2 n^2$. Suppose in addition that for each edge $e \in E$ there is a given set X_e of size $|X_e| \geq \mu n$ consisting of vertices $v \in V(G) \setminus \bigcup E$ with $ve \in \mathrm{Tr}(G)$. Then, w.h.p., there is a set $T_1, \ldots, T_\ell$ of vertex-disjoint triangles in G_p, where each T_i consists of an edge $e \in E$ together with a vertex of X_e, such that $\ell \geq \frac{|E|}{10\mu n}$.*

Proof. Let $0 < 1/C \ll \mu$. We may assume that $p = Cn^{-2/3}(\log n)^{1/3}$ and that n is large enough for the following arguments. We will deduce the lemma from the following claim.

Claim 5.11. *W.h.p. the following is true for all $X \subseteq V(G)$ with $|X| \leq \frac{|E|}{\mu n}$. If $|E \setminus \binom{X}{2}| \geq |E|/2$ and $|X_e \setminus X| \geq \mu n/2$ for all $e \in E$, then there is a triangle in $G_p[V(G) \setminus X]$ consisting of an edge $e \in E$ together with a vertex of X_e.*

Proof. This is a straightforward application of Janson's inequality and the union bound. Note that the total number of choices of X is at most $n^{|E|/(\mu n)}$. Fix one such choice. Let Y denote the number of triangles in $G_p[V(G) \setminus X]$ and note that $\lambda := \mathbb{E}[Y] \geq \frac{p^3 \mu |E| n}{4} \geq C^2 \log(n) \frac{|E|}{n}$. Furthermore, we have $\bar{\Delta} \leq \max(p^5 |E| n^2, \lambda) \leq \max(p^2 n \frac{4\lambda}{\mu}, \lambda) \leq \lambda$. Hence, by Janson's inequality (see Lemma 2.4.2), the probability of having less than $\lambda/2$ triangles is at most

$$\exp\Big(-\frac{\lambda^2}{8\bar{\Delta}}\Big) \leq n^{-C|E|/n}.$$

The claim now follows by taking a union bound and noting $C \gg 1/\mu$. ■

Assume now the high probability event in the claim occurs and let $T_1, \ldots, T_\ell$ be a maximal collection of triangles as guaranteed by the statement. Suppose for contradiction that $\ell < \frac{|E|}{10\mu n}$ and let X be the set of vertices covered by $T_1, \ldots, T_\ell$. We have $|E \setminus \binom{X}{2}| \geq |E| - |X|\mu n \geq |E|/2$ and $|X_e \setminus X| \geq \mu n - \frac{3|E|}{10\mu n} \geq \mu n/2$ for all $e \in E$, and hence there is a triangle in $G_p[V(G) \setminus X]$, a contradiction. □

We are now ready to prove Lemma 5.5.6. We will only prove (i) as both proofs are very similar.

Proof. Let $0 < 1/C \ll \mu$ and let $\mu' = \mu/2$. We may assume that $\delta \leq \mu^5 n$ and that n is large enough for the following arguments. Let G_1, G_2, G_3 be (independent copies) of $G_{p/3}$. Observe that $G_1 \cup G_2 \cup G_3$ is distributed like $G_{p'}$ for some $p' \leq p$ and therefore it suffices to show that $G_1 \cup G_2 \cup G_3$ satisfies the desired properties w.h.p. We will reveal G_1, G_2 and G_3 at different stages and make use of their independence.

Let $A := \{v \in X_1 : \deg_E(v, X_1) \geq \mu' n\}$, and let $A' := X_1 \setminus A$ and $E' = \{e \in E : e \subseteq A\}$. If $|A| \geq \delta$, let $\mathcal{T}_1 = \mathcal{T}_2 = \emptyset$ and move to the next stage. Otherwise, in a first and second round of probability (G_1 and G_2), we will find for each $i \in [2]$ collections $\mathcal{T}_i$ of at least $n'_{i+1} := \min(10(\delta - |A|), n_{i+1})$ triangles in G_i with two endpoints in A' and one in X_{i+1}, so that all triangles in $\mathcal{T}_1 \cup \mathcal{T}_2$ are vertex disjoint. Indeed, since $\Delta(E') \leq \mu' n$ and $\deg(e, X_2) \geq \mu n \geq \mu' n$ for all $e \in E'$, we can apply Lemma 5.5.7 to get a collection of such triangles of size at least

$$\frac{|E'|}{10\mu' n} \geq \frac{|A'|(\delta - |A|)}{10\mu' n} \geq \frac{(\frac{n}{10} - \delta)(\delta - |A|)}{10\mu' n} \geq 10(\delta - |A|)$$

and we can pick a subset $\mathcal{T}_1$ of the desired size. We can repeat this process to get the desired collection $\mathcal{T}_2$, making sure these copies are vertex-disjoint (since we removed at most $3n_1 \leq 3\mu^5 n$ vertices, all inequalities still hold). Let $n''_i = n_i - n'_i$ and pick disjoint subsets $A_i \subseteq A \setminus (V(\mathcal{T}_1 \cup \mathcal{T}_2))$ of size n''_i for each $i = 2, 3$. Since $n''_2 + n''_3 \leq \mu^5 n$, it follows from Lemma 5.5.5, that w.h.p. there is a collection $\mathcal{T}_3$ of $n''_2 + n''_3$ vertex-disjoint triangles in $G_3 - V(\mathcal{T}_1 \cup \mathcal{T}_2)$ consisting of n''_i triangles with two vertices in X_1 and one in X_i for each $i = 2, 3$. □

5.5.2 Reduction

We are now in a position to prove Theorem 1.5.9, assuming Theorem 5.1.1. Before giving the details, let us briefly sketch the approach. Given G with $n \in 3\mathbb{N}$ vertices and minimum degree at least $\frac{2}{3}n$, we separate three cases.

First, there is no set S of about $\frac{n}{3}$ vertices such that $G[S]$ has small maximum degree. In this case, we apply the Regularity Lemma and observe that the (ε, d)-

reduced graph R has no large independent set. By the Hajnal–Szemerédi Theorem, we find a large collection $\mathcal{T}^*$ of vertex disjoint triangles in R, and make the corresponding pairs of clusters super-regular by removing a few vertices to obtain a subgraph T of G. If T were spanning in G, and the clusters were balanced, we would be done by Theorem 5.1.1. In order to fix this, we need to remove a few more triangles covering the vertices outside T (which we do using Lemma 5.5.5) and then further triangles to balance the clusters of T (using Lemma 5.5.4). For the latter we use the $k = 3$ case of Theorem 5.5.2 to find a fractional triangle tiling which tells us where to remove triangles: this is the point where we use the fact that G has no large sparse set. We obtain the following lemma (whose proof we defer to later).

Lemma 5.5.8 (No large sparse set)**.** *For each $\mu > 0$ there are constants $C, d > 0$ such that the following holds. Let $n \in 3\mathbb{N}$, $p \geq Cn^{-2/3}(\log n)^{1/3}$ and suppose G is an n-vertex graph with $\delta(G) \geq (\frac{2}{3} - \frac{d}{2})n$ such that there is no $S \subseteq V(G)$ of size at least $\big(\frac{1}{3} - 2\mu\big)n$ with $\Delta\big(G[S]\big) \leq 2dn$. Then w.h.p. G_p contains a triangle tiling.*

Second, there is a set S of about $\frac{n}{3}$ vertices such that $G[S]$ has maximum degree at most $2dn$, but there is no second such set in $G - S$. The idea here is that we will remove a few triangles from G in order to obtain a subgraph of G which can be partitioned into sets X_1, X_2 of size $|X_2| = 2|X_1| \approx \frac{2n}{3}$, such that all vertices of X_1 are adjacent to almost all vertices of X_2 and vice versa (here Lemma 5.5.6 will be very useful). Note that, with this degree condition, X_2 can be very close to the union of two cliques of size about $\frac{1}{3}n$; this leads to a 'parity case' in which we have to be very careful, which is something of a complication. If we can arrange for the correct parities however, it will be easy to split X_1 into two sets, each of which induces a super-regular triple with one of the 'near-cliques' and apply our Theorem 5.1.1. If we are not in the parity case, we will apply the Regularity Lemma to X_2 and find an almost-spanning matching $\mathcal{M}^*$ in the reduced graph R. We proceed similarly as in the previous case, making these pairs super-regular, removing 'atypical' vertices and then balance the pairs. Here, we make sure that every triangle we remove has two vertices in X_2 and one in X_1 to keep the right balance between the two parts. Finally we can partition X_1 into smaller sets and form balanced super-regular triples with the edges of $\mathcal{M}^*$ in order to apply our Theorem 5.1.1. We obtain the following lemma (whose proof we defer to later).

Lemma 5.5.9 (One large sparse set)**.** *For each $\mu > 0$ there are $C, \tau, d > 0$ such that the following holds for all $n \in 3\mathbb{N}$ and $p \geq Cn^{-2/3}(\log n)^{1/3}$. Suppose G is an n-vertex graph with $\delta(G) \geq \frac{2}{3}n$, and suppose S is a subset of $V(G)$ with $|S| \geq \left(\frac{1}{3} - \tau\right)n$ and $\Delta(G[S]) \leq \tau n$. Suppose that there is no $S' \subseteq V(G) \setminus S$ of size at least $\left(\frac{1}{3} - 2\mu\right)n$ with $\Delta\big(G[S]\big) \leq 2dn$. Then w.h.p. G_p contains a triangle tiling.*

Third, there are two vertex-disjoint sets S_1, S_2 each of size about $\frac{n}{3}$ in G, each of which has small maximum degree. In this case G must be very close to a balanced complete tripartite graph. We start by partitioning $V(G)$ into sets X_1, X_2 and X_3 of size around $n/3$, so that (X_1, X_2, X_3) is a $(\varepsilon, d^+, \delta)$-super-regular triple, where d is close to 1, but δ can be quite small (we need $\delta \gg \varepsilon$ in order to apply Theorem 5.1.1). We remove some carefully chosen vertex-disjoint triangles in order to balance the X_i and to remove some 'atypical' vertices. This leaves us with a balanced (ε, d^+, d)-super-regular triple for some d close to 1, and Theorem 5.1.1 finds the required triangle tiling, giving the following.

Lemma 5.5.10 (Two large sparse sets)**.** *There are $C, \tau > 0$ such that the following holds for all $n \in 3\mathbb{N}$ and $p \geq Cn^{-2/3}(\log n)^{1/3}$. Suppose G is an n-vertex graph with $\delta(G) \geq \frac{2}{3}n$, and suppose S_1 and S_2 are disjoint subsets of $V(G)$ with $|S_i| \geq \left(\frac{1}{3} - \tau\right)n$ and $\Delta(G[S_i]) \leq \tau n$ for each i. Then w.h.p. G_p contains a triangle tiling.*

Before we give proofs of these three lemmas, we show how they imply Theorem 1.5.9.

Proof of Theorem 1.5.9. Let $0 < 1/C \ll d_1 \ll \mu_1 \ll \tau_2, d_2 \ll \mu_2 \ll \tau_3 \ll 1$, let $n \in 3\mathbb{N}$ and let $p \geq Cn^{-2/3}(\log n)^{1/3}$. Suppose that G is an n-vertex graph with $\delta(G) \geq \frac{2n}{3}$.

If G contains no subset of size at least $\left(\frac{1}{3} - 2\mu_1\right)n$ vertices with maximum degree at most $2d_1 n$, then by Lemma 5.5.8, G_p contains a triangle tiling w.h.p. We may therefore suppose G contains a subset S_1 of vertices of size at least $\left(\frac{1}{3} - 2\mu_1\right)n \geq \left(\frac{1}{3} - \tau_2\right)n$ with maximum degree at most $2d_1 n \leq \tau_2 n$. If there is no $S_2 \subseteq V(G) \setminus S_1$ of size at least $\left(\frac{1}{3} - 2\mu_2\right)n$ with maximum degree at most $2d_2 n$, then by Lemma 5.5.9, G_p contains a triangle tiling w.h.p. We can therefore suppose that G contains a subset S_2 disjoint from S_1 of size at least $(\frac{1}{3} - 2\mu_2)n \geq (\frac{1}{3} - \tau_3)n$ with maximum degree at most $2d_2n \leq \tau_3 n$. So by Lemma 5.5.10, G_p contains a triangle tiling w.h.p. □

We now give the proofs of the three lemmas.

Proof of Lemma 5.5.8. Suppose that $0 < 1/m_0 \ll \varepsilon \ll d, 1/C \ll \mu \ll 1$. Let $M \geq m_0$ be returned by Theorem 2.2.11 with input m_0, ε, d and $\gamma = \frac{2}{3}$. Assume that $n \gg M$. Let p and G as in the statement. Let G_1, G_2, G_3 be (independent copies) of $G_{p/3}$; we will show that $G_1 \cup G_2 \cup G_3$ satisfies the desired properties w.h.p.

We apply Theorem 2.2.11 to G, and obtain an (ε, d)-reduced graph R on r vertices with $m_0 \leq r \leq M$ and minimum degree at least $(\frac{2}{3} - \frac{d}{2} - d - \varepsilon)r \geq (\frac{2}{3} - 2d)r$. We will see R as a graph on $[r]$ in the obvious way.

Claim 5.12. *We have* $\alpha(R) < (\frac{1}{3} - \mu)r$.

Proof. Suppose for contradiction that R contains an independent set I of size $(\frac{1}{3} - \mu)r$. Then in $\bigcup_{i\in I} V_i$ there must exist at least $\frac{1}{2}\mu n$ vertices each of whose degree into $\bigcup_{i\in I} V_i$ exceeds $2dn$, otherwise removing these we would have a set S whose existence is forbidden in the lemma statement. By averaging, there is some $i \in I$ such that $\frac{1}{2}\mu|V_i|$ of these vertices V_i' are in V_i. Now vertices of V_i can have at most $|V_i|$ neighbours in V_i, and at most εrn neighbours in sets V_j such that $j \in I$ and (V_i, V_j) is not ε-regular. So the vertices of V_i' all have at least $\frac{3}{2}dn$ neighbours in total in sets V_j such that $j \in I$, $j \neq i$ and (V_i, V_j) is ε-regular. By averaging, there is one of these sets V_j such that the density between V_i' and V_j exceeds $\frac{3}{2}d$. But (V_i, V_j) is ε-regular and has density less than d; this is a contradiction. ■

We apply the Hajnal-Szemerédi Theorem (Theorem 5.5.1) to R, which gives us a collection $\mathcal{T}^*$ of vertex-disjoint triangles in R covering at least $(1 - 12d)r$ vertices. We denote by $T^* := V(\mathcal{T}^*)$ the set of indices in triangles of $\mathcal{T}^*$. By Lemma 2.2.3, there are $V_i' \subseteq V_i$ for each $i \in T^*$ such that $|V_i'| = \lceil(1-4\varepsilon)|V_i|\rceil$ and, for every triangle $ijk \in \mathcal{T}^*$, the triple (V_i', V_j', V_k') is $(2\varepsilon, (d - 4\varepsilon)^+)$-super-regular. Let $T = \bigcup_{i\in T^*} V_i'$ be the set of vertices in G which are in a cluster V_i' corresponding to a triangle of $\mathcal{T}^*$. Let $X = V(G) \setminus T$. Observe that $|X| \leq 4\varepsilon n + \varepsilon n + 12dn \leq 16dn$. Let $W \subseteq T$ be a set such that *(i)* $\left|W \cap V_i'\right| = (\frac{1}{2} \pm \frac{1}{4})\frac{n}{r}$ for each $i \in T^*$,

(ii) $\deg_G(v, W) \geq \frac{3}{5}|W|$ for each $v \in V(G)$, and

(iii) $\deg_G(v, V_i' \cap W) = (\frac{1}{2} \pm \frac{1}{4})\deg_G(v, V_i')$ for each $i \in T^*$ and $v \in V(G)$ with $\deg_G(v, V_i') \geq \varepsilon|V_i'|$.

(Such a set W can be found by choosing each vertex of T independently with probability $\frac{1}{2}$ and applying Chernoff's inequality and a union bound.)

We will start by covering X. We will not touch vertices outside of W in order to maintain super-regularity properties.

Claim 5.13. *W.h.p. in G_1, there is a set of vertex-disjoint triangles $\mathcal{T}_1 \subseteq \mathrm{Tr}(G_1[W \cup X])$ so that $X \subseteq V(\mathcal{T}_1)$ and $\left|V(\mathcal{T}_1) \cap V_i'\right| \leq 20\sqrt{d}|V_i'|$ for all $i \in T^*$.*

Proof. Let $\tilde{\mu} := 4\sqrt{d}$ and enumerate $X = \{v_1, \ldots, v_\ell\}$ and note that $\ell \leq \tilde{\mu}^2 n$. For each $i \in [\ell]$, let $E_i := \{e \in E(G[W]) : ev \in \mathrm{Tr}(G)\}$. Note that, since $\deg(v, W) \geq \frac{3}{5}|W|$ for all $v \in V(G)$, we have $|E_i| \geq \tilde{\mu}n^2$ for all $i \in [\ell]$. Finally, let $A_i = V_i'$ for each $i \in T^*$. The claim now follows readily from Lemma 5.5.5. ■

Let now $V_i'' = V_i' \setminus V(\mathcal{T}_1)$ for each $i \in T^*$. We would like to apply Theorem 5.1.1 to the super-regular triples (V_i'', V_j'', V_k'') for each $ijk \in \mathcal{T}^*$. However, these triples are not necessarily balanced, but we can correct this.

Claim 5.14. *W.h.p. in G_2, there is a set of vertex-disjoint triangles $\mathcal{T}_2 \subseteq \mathrm{Tr}(G_2[W \setminus V(\mathcal{T}_1)])$ so that $|V_i'' \setminus V(\mathcal{T}_2)| = \lfloor 0.9\frac{n}{r} \rfloor$ for all $i \in T^*$.*

Proof. Let $R' = R[T^*]$ and let $\lambda : T^* \to \mathbb{N}$ be given by $\lambda(i) = |V_i''| - \lfloor 0.9\frac{n}{r} \rfloor$. Note that $(0.1 - 30\sqrt{d})\frac{n}{r} \leq \lambda(i) \leq \lceil 0.1\frac{n}{r} \rceil$, and that $\sum_{i \in T^*} \lambda(i) = n - 3\,|\mathcal{T}_1| - 3\,|\mathcal{T}^*|\,\lfloor 0.9\frac{n}{r} \rfloor$ is divisible by 3. Hence, by Lemma 5.5.3, there is a weight function $\omega : \mathrm{Tr}(R') \to \mathbb{N}_0$ such that for each $i \in T^*$ we have $\sum_{t \in \mathrm{Tr}_i(R')} \omega(t) = \lambda(i)$. We claim that we can remove $\omega(ijk)$ triangles from $G_2[V_i'' \cap W, V_j'' \cap W, V_k'' \cap W]$ for each triangle ijk of R', making sure that all our choices are vertex-disjoint. Indeed, observe that the triple $(V_i'' \cap W, V_j'' \cap W, V_k'' \cap W)$ is $(10\varepsilon, (d/2)^+, d/10)$-super-regular for each $ijk \in \mathrm{Tr}(R')$ by the slicing lemma (see Lemma 2.2.1) and the choice of W. Furthermore, observe that $|V_i'' \cap W| \geq \frac{1}{5} \cdot \frac{n}{r}$. Hence, Lemma 5.5.4 implies that w.h.p. there are $\frac{1}{6} \cdot \frac{n}{r}$

vertex-disjoint triangles in $G_2[V_i'' \cap W, V_j'' \cap W, V_k'' \cap W]$ for each $ijk \in \mathrm{Tr}(R')$, so we can select the desired amount of triangles for each $t \in \mathrm{Tr}(R')$ one at a time. ■

Let now $V_i''' = V_i'' \setminus V(\mathcal{T}_2)$ for all $i \in T^*$ and observe that we have covered all vertices except for those in $\bigcup_{i \in T^*} V_i'''$. We claim that (V_i''', V_j''', V_k''') is $(8\varepsilon, (d/2)^+, d/8)$-super-regular for all $ijk \in \mathcal{T}^*$. Indeed, this follows from the slicing lemma, and from $\deg(v, V_j''') \geq \deg(v, V_j' \setminus W) \geq \frac{1}{4}\deg_G(v, V_j') \geq \frac{d}{8}|V_j'|$ and the analogous inequalities for other pairs. Finally, we apply Theorem 5.1.1 to each of these triples individually in G_3 to obtain (w.h.p.) a collection $\mathcal{T}_3$ of vertex-disjoint triangles covering exactly $\bigcup_{i \in T^*} V_i'''$. □

Next, we deal with the case when G has two large sparse sets; i.e. it looks similar to the extremal complete tripartite graph. This is the easiest case; we will not need the Regularity Lemma.

Proof of Lemma 5.5.10. Suppose that $0 < \tau, 1/C \ll 1$ and let $\rho := \tau^{1/8}$. Let $n \in 3\mathbb{N}$ be large enough for the following arguments and let $p \geq Cn^{-2/3}(\log n)^{1/3}$. Let G and sparse sets S_1 and S_2 be given as in the statement. Let $G_1, \ldots, G_4$ be independent copies of $G_{p/5}$. We will find a triangle tiling in $G_1 \cup \ldots \cup G_4$.

Claim 5.15. *There is a partition $V(G) = X_1 \cup X_2 \cup X_3$ such that*

(i) $|X_i| = (1/3 \pm \rho^6)n$ *for all* $i \in [3]$,

(ii) $\deg(v, X_j) \geq \rho n$ *for all* $i \neq j \in [3]$ *and* $v \in X_i$,

(iii) $d(X_i, X_j) \geq 1 - \rho^6$ *for all* $1 \leq i < j \leq 3$.

(iv) For each $i \in [3]$, *if* $|X_i| \geq \frac{n}{3}$, *then* $\deg(v, X_j) \geq |X_j| - 4\rho n$ *for all* $v \in X_i$ *and* $j \in [3] \setminus \{i\}$.

Proof. For $i \in [2]$, let $Z_i = \{v \in V(G) \setminus (S_1 \cup S_2) : \deg(v, S_i) \leq \rho n\}$. Let $X_i' = S_i \cup Z_i$ for $i \in [2]$ and $X_3' = \{v \in V(G) : \deg(v, S_i) \geq (\frac{1}{3} - 2\rho)n$ for each $i \in [2]\}$. Note that, since $\delta(G) \geq \frac{2}{3}n$, Z_1 and Z_2 are disjoint and hence X_1' and X_2' are disjoint as well. Furthermore, by definition, X_3' is disjoint from X_1' and X_2'. Let $Z' := V(G) \setminus (X_1' \cup X_2' \cup X_3')$ be the set of remaining vertices. Partition $Z' = Z_1' \cup Z_2' \cup Z_3'$ so that $Z_i' = \emptyset$ if $|X_i'| \geq \frac{n}{3}$ and $|X_i'| + |Z_i'| \leq \frac{n}{3}$ otherwise. Finally, let and $X_i = X'_i \cup Z'_i$ for all $i \in [3]$. Note that $V(G) = X_1 \cup X_2 \cup X_3$ is indeed a partition.

We will show now that the sets Z_1, Z_2 and Z' are small. Let $i \in [2]$. Since $|S_i| \geq \left(\frac{1}{3} - \rho^8\right)n$, each vertex of S_i has at least $\left(\frac{1}{3} - 2\rho^8\right)n$ non-neighbours in S_i, and so at most $2\rho^8 n$ non-neighbours outside S_i. Therefore, the total number of non-edges between S_i and $V(G) \setminus S_i$ is at most $\rho^8 n^2$. Since every $v \in Z_i$ has at least $n/4$ non-neighbours in S_i, this implies $|Z_i| \leq 4\rho^8 n$. Moreover, the number of non-edges between $X'_1 \cup X'_2$ and Z' is at most $2\rho^8 n^2 + (|Z_1| + |Z_2|)n \leq 10\rho^8 n^2$. Observe that every $v \in Z'$ has at least ρn non-neighbours in $X'_1 \cup X'_2$ (otherwise it would be in X'_3), and therefore $|Z'| \leq 10\rho^7 n$. In particular, (i) holds.

Furthermore, for each $v \in Z'$, we have $\deg(v, S_i) \geq \rho n$ since $v \notin Z_i$ for $i \in [2]$, and $\deg(v, X'_3) \geq \rho n$ for otherwise v would be in X'_3. Clearly, we have $\deg(v, X_j) \geq \rho n$ for all $i \in [2]$, $j \in [3] \setminus \{i\}$ and $v \in Z_i$. Therefore, (ii) holds.

Moreover, we have $\deg(v, X_i) \geq |X_i| - \rho^7 n$ for all $x \in S_1$ and $i = 2, 3$. Since $|Z_1 \cup Z'_1| \leq \rho^7 n$, this implies $d(X_1, X_i) \geq 1 - \rho^6$ for $i = 2, 3$. Similarly $d(X_2, X_3) \geq 1 - \rho^6$.

Finally, let $i, j \in [3]$ be distinct. If $|X_i| \geq n/3$, then $X_i \cap Z' = \emptyset$ by construction and therefore $\deg(v, X_j) \geq \frac{n}{3} - 2\rho n \geq |X_j| - 4\rho n$ for all $v \in X_i$. ∎

We now perform a stage of removing some vertex-disjoint triangles in order to obtain a balanced tripartite graph.

Claim 5.16. *W.h.p. in G_1, there is a set of triangles $\mathcal{T}_1 \subseteq \mathrm{Tr}(G_1)$ so that $|X_1 \setminus V(\mathcal{T}_1)| = |X_2 \setminus V(\mathcal{T}_1)| = |X_3 \setminus V(\mathcal{T}_1)| \geq (\frac{1}{3} - \rho^4)n$.*

Proof. If all three sets X_1, X_2, X_3 have size exactly $\frac{n}{3}$, we are done. Otherwise, one or two of these sets have size exceeding $\frac{n}{3}$.

Case 1. Assume first that only one set exceeds $\frac{n}{3}$ in size and, without loss of generality, this set is X_1. Let $n_2 := \frac{n}{3} - |X_3|$ and $n_3 := \frac{n}{3} - |X_2|$, and let $E = E(G[X_1])$. Observe that $\delta(E) \geq |X_1| - \frac{n}{3} = n_1 + n_2$. Furthermore, we have $\deg(e, X_i) \geq |X_i| - 10\rho n \geq \frac{n}{4}$ for both $i = 2, 3$. Therefore, by Lemma 5.5.6 (i), there is a collection $\mathcal{T}_1$ of $n_2 + n_3$ vertex-disjoint triangles in G_1 with two vertices in X_1, n_2 of which have the third vertex in X_2, and n_3 of which have the third vertex in

X_3. Clearly, we have $|X_1 \setminus T_1| = |X_2 \setminus T_1| = |X_3 \setminus T_1| = \frac{2n}{3} - |X_1| \geq (\frac{1}{3} - 5\rho^6)n$, as claimed.

Case 2. Assume now that there are two sets (say X_1 and X_2) exceeding $\frac{n}{3}$ in size. For $i \in [2]$, let $n_i := |X_i| - \frac{n}{3}$ and $E_i = E(G[X_i])$. Observe that, for $i \in [2]$, $\delta(E_i) \geq n_i$ and $\deg(e, X_{3-i}) \geq |X_{3-i}| - 10\rho n \geq \frac{n}{4}$ for all $e \in E_i$. Therefore, by Lemma 5.5.6 (ii), there is a collection $\mathcal{T}_1$ of $n_1 + n_2$ vertex disjoint triangles in G_1, n_1 of which have two vertices in X_1 and one in X_2, and n_2 of which have two vertices in X_2 and one in X_1. Clearly, we have $|X_1 \setminus T_1| = |X_2 \setminus T_1| = |X_3 \setminus T_1| = |X_3| \geq (\frac{1}{3} - \rho^6)n$, as claimed. ■

Let now $X_i' = X_i \setminus V(\mathcal{T}_1)$ and observe that $|X_1'| = |X_2'| = |X_3'|$. Define

$$Y_i' := \{v \in X_i' : \deg(v, X_j') \leq (1 - \rho/2)|X_j'| \text{ for some } j \in [3] \setminus \{i\}\}.$$

Since $d(X_i', X_j') \geq 1 - 2\rho^6$ for all $1 \leq i < j \leq 3$, we have $\left|Y_i'\right| \leq 10\rho^5 n$ for each $i \in [3]$. Furthermore, each vertex $v \in Y_1' \cup Y_2' \cup Y_3'$ is in at least $\frac{1}{4}\rho^2 n^2$ triangles of G with one vertex in each X_i'. By applying Lemma 5.5.5 (with $t = 0$), w.h.p. in G_2, we can find a set $\mathcal{T}_2 \subseteq \mathrm{Tr}(G_2)$ of vertex-disjoint triangles using one vertex from each part with $Y_1' \cup Y_2' \cup Y_3' \subseteq V(\mathcal{T}_2) \subseteq X_1' \cup X_2' \cup X_3'$ and $|V(\mathcal{T}_2)| \leq 3(|Y_1'| + |Y_2'| + |Y_3'|) \leq \rho^4 n$.

Let now $X_i'' := X_i' \setminus V(\mathcal{T}_2)$ for each $i \in [3]$ and observe that $|X_1''| = |X_2''| = |X_3''| \geq (\frac{1}{3} - 2\rho^4)n$. Furthermore, (X_1'', X_2'', X_3'') is $(\rho, (1-\rho)^+)$-super-regular. Hence, by Theorem 5.1.1, there is a triangle tiling $\mathcal{T}_3$ in G_3 covering it. □

Finally, we deal with the second case in the above sketch; we will use several of the ideas from the previous two lemmas, so we abbreviate the details.

Proof of Lemma 5.5.9. Suppose that $0 < 1/m_0 \ll \varepsilon \ll d, 1/C \ll \mu \ll 1$. Let $M \geq m_0$ be returned by Theorem 2.2.11 with input m_0, ε, d and $\gamma = \frac{2}{3}$. Let $0 < \tau \ll 1/M$ and $\rho = \tau^{1/8}$. Assume that $n \in 3\mathbb{N}$ is large enough for the following arguments. Let p, G and S be as in the statement of the lemma. Let $G_1, \ldots, G_5$ be (independent copies) of $G_{p/5}$. We will show that $G_1 \cup \ldots \cup G_5$ contains a triangle tiling w.h.p.

We begin with a claim that gives us a lot of structure. We will call a set $X \subseteq V(G)$ *ε-strongly connected* if $\overline{e}(X', X \setminus X') \leq \frac{|X|^2}{4} - \varepsilon n^2$ for all $X' \subseteq X$, where we denote by

$\overline{e}(X,Y) = |X||Y| - e(X,Y)$ the number of non-edges between X and Y.[4] Furthermore, we say that X is *ε-close to complete* if $e(G[X]) \geq (\frac{1}{2} - \varepsilon)|X|^2$.

Claim 5.17. *There is a collection $\mathcal{T}_1$ of vertex disjoint triangles in $G_1 \cup G_2$ and disjoint sets $X_1, X_2 \subseteq V(G)$ so that*

(i) $X_1 \cup X_2 = V(G) \setminus V(\mathcal{T}_1)$ *and* $|X_1| = |X_2|/2 = (\frac{1}{3} \pm \rho n)$.

(ii) $\deg(v, X_{3-i}) \geq (1 - 4\rho)|X_{3-i}|$ *for all* $i \in [2]$ *and* $v \in X_i$,

(iii) X_2 *is* $10d$*-strongly connected or there is a partition* $X_2 = X_{2,1} \cup X_{2,2}$ *so that* $|X_{2,i}| \geq \frac{n}{4}$ *is even and* $50d$*-close to complete for both* $i \in [2]$.

Proof. Let $Y_1 = \{v \in V(G) \setminus S : \deg(v,S) \leq \rho n\}$. Let $X_1' = S \cup Y_1$ and $X_2' = V(G) \setminus X_1'$. Observe that, similar to the proof of Claim 5.15, we have

(P1) $\deg(v, X_2') \geq |X_2'| - 2\rho n$ for all $v \in X_1'$ and $\deg(v, X_1') \geq \rho n$ for all $v \in X_2'$,

(P2) $|X_1'| = (\frac{1}{3} \pm \rho^6)n$ and $|X_2'| = (\frac{2}{3} \pm \rho^6)n$, and

(P3) $d(X_1', X_2') \geq 1 - \rho^6$.

Let $\sigma = 10d$ and let $X_2' = X_{2,1}' \cup X_{2,2}'$ be the partition of X_2' which maximises $\overline{e}(X_{2,1}' \cup X_{2,2}')$. Throughout this proof, we will have to distinguish between two cases: either X_2' is σ-strongly-connected (this we will call the *connected case* from now on) or $\overline{e}(X_{2,1}' \cup X_{2,2}') \geq \frac{|X_2|^2}{4} - \sigma n^2$. Although the process is very similar for both, we will handle them separately, starting with the disconnected case.

The disconnected case. We claim that

(Q1) $|X_{2,i}'| = (\frac{1}{3} \pm 2\sigma)n$ and $e(X_{2,i}') \geq \frac{1}{2}|X_{2,i}'|^2 - 2\sigma n^2$ for both $i \in [2]$, and

(Q2) $\deg(v, X_{2,i}') \geq \frac{n}{10}$ for all $i \in [2]$ and $v \in X_{2,i}'$.

Indeed, (Q1) follows from the case assumption and $\delta(G) \geq \frac{2n}{3}$, and (Q2) since $X_{2,1}', X_{2,1}'$ are chosen to maximise non-edges in between (otherwise, moving a vertex violating (Q2) to the other set increases the count).

[4] This definition might appear somewhat strange now but will assure that the reduced graph in this proof is connected.

In a first round of probability (G_1), our goal is to balance the sizes. Assume first that $|X'_1| > \frac{n}{3}$. Let $n_2 = 0$ if $X'_{2,1}$ is even and $n_2 = 1$ otherwise, and let $n_3 = |X'_1| - \frac{n}{3} - n_2$. Let $E = E(G[X'_1])$, and observe that $\delta(E) \geq n_2 + n_3$. Furthermore, we have $\deg(e, X'_{2,i}) \geq |X'_{2,i}| - 10\rho n \geq \frac{n}{4}$ for both $i \in [2]$ by (P1) and (Q1). Therefore, by Lemma 5.5.6, w.h.p. there is a collection $\mathcal{T}'_1$ of exactly $|X'_1| - \frac{n}{3}$ vertex-disjoint triangles in G_1 with two vertices in X'_1 and one vertex in X'_2. Let $X''_i = X'_i \setminus V(\mathcal{T}'_1)$ and $X''_{2,i} = X'_{2,i} \setminus V(\mathcal{T}'_1)$ for $i \in [2]$. By construction, we have $|X''_2| = 2|X''_1| = \frac{4n}{3} - 2|X'_1| \geq 2(\frac{1}{3} - \rho^5)n$ and $X''_{2,i}$ is even for both $i \in [2]$.

Assume now that $|X'_2| \geq \frac{2n}{3}$. Observe that for each $i \in [2]$ and $X \subseteq X'_{2,i}$ of size $|X| \geq \frac{1}{2}|X'_{2,i}| \geq \frac{n}{8}$, we have $\mathrm{Tr}(G[X]) \geq \frac{n^3}{1000}$ by (Q1). Thus, by Lemma 5.5.4, there are collections of $n/8$ vertex-disjoint triangles in each of $G_1[X'_{2,i}]$ w.h.p. Thus, we can pick a collection $\mathcal{T}'_1$ of exactly $\frac{n}{3} - |X'_1|$ from these, again taking either one or no triangle in $X'_{2,1}$ depending on its parity. By construction, we have $|X''_2| = 2|X''_1| = 2|X'_1| \geq 2(\frac{1}{3} - \rho^5)n$ and $X''_{2,i}$ is even for both $i \in [2]$ (where X''_i and $X''_{2,i}$ are defined as above).

In a second round of probability (G_2), we will remove 'atypical' vertices in X''_2. From this point onwards, we will only remove triangles with one vertex in X''_1 and two vertices in $X''_{2,i}$ for some $i \in [2]$, thus maintaining the right balance between X''_1 and X''_2 and the parity of $X''_{2,1}$ and $X''_{2,2}$. For $i \in [2]$, let $Y_{2,i} := \{v \in X''_{2,i} : \deg(v, X''_1) \leq |X''_1| - \frac{\rho n}{2}\}$ and for each $v \in Y_{2,i}$ let $E_v := \{u_1u_2 : u_1 \in X''_{2,i}, u_2 \in X''_2, vu_1u_2 \in \mathrm{Tr}(G)\}$. It follows from (P3) (and counting non-edges between X'_1 and X'_2) that $|Y_{2,i}| \leq 2\rho^5 n$ for both $i \in [2]$. Furthermore, this implies that $|E_v| \geq (\frac{1}{10} - 2\rho^5)n \cdot (\rho - \frac{\rho}{2})n \geq \rho^2 n$ for all $v \in Y_{2,1} \cup Y_{2,2}$. Thus, by Lemma 5.5.5, w.h.p. there is a collection $\mathcal{T}''_1$ of at most $4\rho^5 n$ vertex-disjoint triangles in $G_2[X''_1 \cup X''_2]$ of the desired form (one vertex in X''_1 and two vertices in $X''_{2,i}$ for some $i \in [2]$). Let $X_i = X''_i \setminus V(\mathcal{T}''_1)$ and $X_{2,i} = X''_{2,i} \setminus V(\mathcal{T}''_1)$ for each $i \in [2]$. Let $\mathcal{T}_1 = \mathcal{T}'_1 \cup \mathcal{T}''_1$, and $X_i = X''_i \setminus V(\mathcal{T}''_1)$ and $X_{2,i} = X''_{2,i} \setminus V(\mathcal{T}''_1)$ for each $i \in [2]$. It is easy to see that these sets have all the desired properties.

<u>The connected case.</u> This case is very similar but less technical since we do not have to worry about the sets $X'_{2,1}$ and $X'_{2,2}$. We will therefore skip some details.

In a first round of probability (G_1), our goal is to balance the sizes. The case $|X'_1| > \frac{n}{3}$ is completely analogous to the disconnected case and we find a collection of $\mathcal{T}_1'$ of exactly $|X_1'| - \frac{n}{3}$ vertex-disjoint triangles in G_1 with two vertices in X_1' and one vertex in X_2'. Let $X_i'' = X_i' \setminus V(\mathcal{T}_1')$ for $i \in [2]$. By construction, we have $|X_2''| = 2|X_1''| = \frac{4n}{3} - 2|X_1'| \geq 2(\frac{1}{3} - \rho^5)n$.

Assume now that $|X_2'| \geq \frac{2n}{3}$. Observe that for every set $Z \subseteq X_2'$ with $|Z| \leq dn$ and every $v \in X_2' \setminus Z$, we have $\deg(v, X_2' \setminus Z) \geq (\frac{1}{3} - d)n$ and thus there are at least dn edges in $N(v, X_2')$ (since there is no set $S' \subseteq X_2$ with $\Delta(G[S']) \leq 2dn$ by assumption). Thus there at least $\frac{d}{10}n^3$ triangles in $G[X_2' \setminus Z]$. It follows from Lemma 5.5.4 that w.h.p. there are at least $dn/3$ vertex-disjoint triangles in $G_1[X_2']$. Let $\mathcal{T}_1'$ be a collection of exactly $\frac{n}{3} - |X_1'|$ of these. By construction, we have $|X_2''| = 2|X_1''| = 2|X_1'| \geq 2(\frac{1}{3} - \rho^5)n$.

The process of removing atypical vertices in X_2'' is analogous to the connected case and we will omit the details. ■

From here on, the disconnected case is very simple: Let us first remove more atypical vertices of our near cliques. For $i \in [2]$, let $Z_{2,i} := \{v \in X_{2,i} : \deg(v, X_{2,i}) \leq |X_{2,i}| - \sqrt{d}n\}$. Observe that, since $X_{2,i}$ is $50d$-close to complete, we have $|Z_{2,i}| \leq 100\sqrt{d}n$. Since any two vertices in X_2 have at least $\frac{n}{4}$ common neighbours in X_1, it follows from Lemma 5.5.5 that w.h.p. (in G_3) there is a collection $\mathcal{T}_2$ of at most $200\sqrt{d}n$ vertex-disjoint triangles in $G_3[X_1 \cup X_2]$ with one vertex in X_1 and two vertices in X_2 (both of which are in the same $X_{2,i}$) covering $Z_{2,1} \cup Z_{2,2}$. Let $X_i' = X_i \setminus V(\mathcal{T}_2)$ and $X_{2,i}' = X_{2,i} \setminus V(\mathcal{T}_2)$ for each $i \in [2]$. Let $X_1' = X_{1,1}' \cup X_{2,2}'$ be a partition such that $|X_{1,i}'| = \frac{1}{2}|X_{2,i}'|$ for each $i \in [2]$ (note that here the parity of $|X_{2,i}'|$ is important). Now, $X_{1,i}' \cup X_{2,i}'$ induces a $(d^{1/3}, 0.99^+)$-super-regular triple for both $i \in [2]$ (after splitting $X_{2,i}'$ arbitrarily in two sets of equal sizes). Therefore, by Theorem 5.1.1, w.h.p. there are vertex-disjoint triangles in G_4 covering the remaining vertices.

Thus, we may assume that X_2 is $10d$-strongly connected. This case is very similar to the proof of Lemma 5.5.8. Let $n_i := |X_i|$ for both $i \in [2]$ and recall that $n_2 = 2n_1$. We apply Theorem 2.2.11 to $G[X_2]$ with input m_0, ε, d to get an ε-regular partition $X_2 = V_0 \cup V_1 \cup \ldots \cup V_r$ for some $m_0 \leq r \leq M$. Let R be the corresponding (ε, d)-reduced graph (seen as a graph on $[r]$) and observe that we have $\delta(R) \geq \left(\frac{1}{2} - 2d\right)r$ and, as in the proof of Lemma 5.5.8, we have $\alpha(R) < \left(\frac{1}{2} - \mu\right)r$. It is well-known that every graph H contains a matching of size $\min\{\delta(H), \lfloor\frac{v(H)}{2}\rfloor\}$.[5] Thus R contains a matching $\mathcal{M}^*$ of size $\left(\frac{1}{2} - 2d\right)r$; let R' be the subgraph of R induced by $M^* := V(\mathcal{M}^*)$. Note that $\delta(R') \geq (\frac{1}{2} - 6d)r$ and we claim that R' is connected. Indeed, if not, there is a set $X_R \subseteq V(R')$ such that $e(X_R, V(R') \setminus X_R) = 0$. Observe that $|X_R|, |V(R') \setminus R| \geq \delta(R') \geq (\frac{1}{2} - 6d)r$. Let now $X' := \bigcup_{i \in X_R} V_i$ and observe that $|X'| = (\frac{1}{2} \pm 20d)|X_2|$. Furthermore, we have $e(X', X_2 \setminus X') \leq (d + 4d + \varepsilon)dn^2$ and consequently $\overline{e}(X', X_2 \setminus X') \geq |X'||X_2 \setminus X'| - 6dn^2 > \frac{|X_2|^2}{4} - 10dn^2$, contradicting the fact that X_2 is $10d$-strongly connected.

By Lemma 2.2.3, there are $V_i' \subseteq V_i$ for each $i \in M^*$ such that $|V_i'| = \lceil(1-4\varepsilon)|V_i|\rceil$ and, for every triangle $ij \in \mathcal{M}^*$, the pair (V_i', V_j') is $(2\varepsilon, (d - 4\varepsilon)^+)$-super-regular. Let $Y = X_2 \setminus \bigcup_{i \in M^*} V_i'$ be the set of vertices in X_2 which are not in a cluster V_i' corresponding to an edge of $\mathcal{M}^*$. Observe that $|Y| \leq 4\varepsilon n + \varepsilon n + 4dn \leq 5dn$. Let $W \subseteq X_2 \setminus Y$ be a set such that

(i) $\left|W \cap V_i'\right| = (\frac{1}{2} \pm \frac{1}{4})\frac{n_2}{r}$ for each $i \in M^*$,

(ii) $\deg_G(v, W) \geq \frac{1}{3}|W|$ for each $v \in X_2$, and

(iii) $\deg_G(v, V_i' \cap W) = (\frac{1}{2} \pm \frac{1}{4})\deg_G(v, V_i')$ for each $i \in M^*$ and $v \in X_2$ with $\deg_G(v, V_i') \geq \varepsilon|V_i'|$.

(Such a set W can be found by choosing each vertex of $X_2 \setminus Y$ independently with probability $\frac{1}{2}$ and applying Chernoff's inequality and a union bound.)

We will start by covering Y. We will not touch vertices outside of W in order to maintain super-regularity properties.

Claim 5.18. *W.h.p. in G_3, there is a set of vertex-disjoint triangles $\mathcal{T}_2 \subseteq \mathrm{Tr}(G_1)$ each of which has two vertices in $W \cup Y \subseteq X_2$ and one in X_1, so that $Y \subseteq V(\mathcal{T}_2)$ and $\left|V(\mathcal{T}_2) \cap V_i'\right| \leq 20\sqrt{d}|V_i'|$ for all $i \in M^*$.*

The proof is very similar to the proof of Claim 5.13 and we will omit the details. Let now $X_i'' = X_i \setminus V(\mathcal{T}_2)$ for each $i \in [2]$ and let $V_i'' = V_i' \setminus V(\mathcal{T}_2)$ for each $i \in M^*$. We will now balance the sizes.

[5]If $v(H)$ is even, this is the $k = 2$ case of Theorem 5.5.1.

Claim 5.19. *W.h.p. in G_4, there is a set of vertex-disjoint triangles $\mathcal{T}_3 \subseteq \mathrm{Tr}(G_4)$ with one vertex in X_1'' and two vertices in W, so that $|V_i'' \setminus V(\mathcal{T}_3)| = \lfloor 0.9\frac{n_2}{r} \rfloor$ for all $i \in M^*$.*

Proof. Let $\lambda : M^* \to \mathbb{N}$ be given by $\lambda(i) = |V_i''| - \lfloor 0.9\frac{n}{r} \rfloor$. Note that we have $(0.1 - 30\sqrt{d})\frac{n_2}{r} \leq \lambda(i) \leq \lceil 0.1\frac{n_2}{r} \rceil$, and that $\sum_{i \in M^*} \lambda(i) = n_2 - 2|\mathcal{T}_2| - 2|\mathcal{M}^*| \lfloor 0.9\frac{n}{r} \rfloor$ is even. Hence, by applying Lemma 5.5.3 to the connected graph R', there is a weight function $\omega : E(R') \to \mathbb{N}_0$ such that for each $i \in M^*$ we have $\sum_{j \in N_{R'}(i)} \omega(ij) = \lambda(i)$. We claim that we can remove $\omega(ij)$ triangles from $G_2[X_1'', V_i'' \cap W, V_j'' \cap W]$ for each edge ij of R', making sure that all our choices are vertex-disjoint. Indeed, let $Y_1, \ldots, Y_r \subseteq X_1''$ be disjoint sets of size at least $\frac{1}{5} \cdot \lceil \frac{n_2}{r} \rceil$ and observe that $\deg(v, X_{3-1}'') \geq |X_{3-i}''| - 4\rho n$ for each $v \in X_i''$ by Claim 5.17. Since $\rho \ll 1/r \ll \varepsilon$, this implies that, for each $k \in [r]$ and $i \in M^*$, the pair (Y_k, V_i'') is $(\varepsilon, (1-\varepsilon)^+)$-super-regular. It further follows from the slicing lemma (see Lemma 2.2.1) and the choice of W that $(V_i'' \cap W, V_j'' \cap W)$ is $(10\varepsilon, (d/10)^+)$-super-regular for each $ij \in E(R')$. Hence the triple $(Y_k, V_i'' \cap W, V_j'' \cap W)$ is $(10\varepsilon, (d/10)^+)$-super-regular for each $ij \in E(R')$ and $k \in [r]$. Furthermore, we have $|V_i'' \cap W| \geq \frac{1}{5} \cdot \frac{n_2}{r}$. Hence, Lemma 5.5.4 implies that w.h.p., there are $\frac{1}{6} \cdot \frac{n_2}{r}$ vertex-disjoint triangles in $G_4[Y_k, V_i'' \cap W, V_j'' \cap W]$ for each $ij \in E(R')$ and $k \in [r]$. Thus we can select the desired amount of triangles for each $e \in E(R')$ one at a time. ■

Let now $X_i''' = X_i'' \setminus V(\mathcal{T}_3)$ for each $i \in [2]$ and $V_i''' = V_i'' \setminus V(\mathcal{T}_3)$ for all $i \in M^*$ and observe that we have covered all vertices except for those in $X_1''' \cup X_2'''$. Since $|X_1'''| = \frac{1}{2}|X_2'''|$, we can partition $X_1''' = \bigcup_{e \in \mathcal{M}^*} X_e'''$ into $|\mathcal{M}^*|$ sets of size exactly $\lfloor 0.9\frac{n_2}{r} \rfloor$. Observe that $\deg(v, X_2''') \geq |X_2'''| - 4\rho n$ for each $v \in X_1'''$ and vice versa by Claim 5.17. Since $\rho \ll 1/r \ll \varepsilon$, this implies that, for each $e \in \mathcal{M}^*$ and $i \in M^*$, the pair (X_e''', V_i''') is $(\varepsilon, (1-\varepsilon)^+)$-super-regular. Furthermore, the pair (V_i''', V_j''') is $(8\varepsilon, (d/8)^+)$-super-regular for each $ij \in \mathcal{M}^*$ by the slicing lemma (Lemma 2.2.1) and since $\deg(v, V_j''') \geq \deg(v, V_j' \setminus W) \geq \frac{1}{4}\deg_G(v, V_j') \geq \frac{d}{8}|V_j'|$ for all $v \in V_i'''$ and vice-versa. Therefore, $(X_{ij}''', V_i''', V_j''')$ is $(8\varepsilon, (d/8)^+)$-super-regular for all $ij \in \mathcal{M}^*$. Finally, we apply Theorem 5.1.1 to each of these triples individually in G_5 to obtain (w.h.p.) a collection $\mathcal{T}_4$ of vertex-disjoint triangles covering exactly $X_1''' \cup X_2'''$. □

Bibliography

[1] W. Ackermann, *Die Widerspruchsfreiheit der allgemeinen Mengenlehre*, Math. Ann. 114.1 (1937), 305–315.

[2] P. Allen, *Covering two-edge-coloured complete graphs with two disjoint monochromatic cycles*, Combin. Probab. Comput. 17.4 (2008), 471–486.

[3] P. Allen, J. Böttcher, J. Corsten, E. Davies, M. Jenssen, P. Morris, B. Roberts and J. Skokan, *Robust triangle factors*, In preparation (2020+).

[4] P. Allen, J. Böttcher, E. Davies, M. Jenssen, Y. Kohayakawa and B. Roberts, *Shamir's problem revisited*, In preparation (2020+).

[5] P. Allen, G. Brightwell and J. Skokan, *Ramsey-goodness – and otherwise*, Combinatorica 33.2 (2013), 125–160.

[6] N. Alon and R. Yuster, *Threshold functions for H-factors*, Combin. Probab. Comput. 2.2 (1993), 137–144.

[7] J. Ayel, *Sur l'existence de deux cycles supplémentaires unicolores, disjoints et de couleurs différentes dans un graphe complet bicolore*, PhD thesis, Université Joseph-Fourier - Grenoble I, 1979.

[8] F. S. Benevides and J. Skokan, *The 3-colored Ramsey number of even cycles*, J. Combin. Theory Ser. B 99.4 (2009), 690–708.

[9] S. Bessy and S. Thomassé, *Partitioning a graph into a cycle and an anticycle, a proof of Lehel's conjecture*, J. Combin. Theory Ser. B 100.2 (2010), 176–180.

[10] J. Bierbrauer and A. Gyárfás, *On (n, k)-colorings of complete graphs*, Congr. Numer. 58 (1987), 123–139.

[11] B. Bollobás, *The evolution of random graphs*, Trans. Amer. Math. Soc. 286 (1984), 257–274.

[12] B. Bollobás and A. Thomason, *Threshold functions*, Combinatorica 7.1 (1987), 35–38.

[13] J. A. Bondy and P. Erdös, *Ramsey numbers for cycles in graphs*, J. Combin. Theory Ser. B 14.1 (1973), 46–54.

[14] J. Böttcher, Y. Kohayakawa, A. Taraz and A. Würfl, *An extension of the blow-up lemma to arrangeable graphs*, SIAM J. Discrete Math. 29.2 (2015), 962–1001.

[15] J. R. Burke and W. A. Webb, *Asymptotic behavior of linear recurrences*, Fibonacci Quarterly 19.4 (1981), 318–321.

[16] S. A. Burr and P. Erdős, *On the magnitude of generalized Ramsey numbers for graphs*, Colloq. Math. Soc. János Bolyai 10 (1975), 215–240.

[17] S. A. Burr and P. Erdős, *Extremal Ramsey theory for graphs*, Util. Math. 9 (1976), An International Journal of Discrete and Combinatorial Mathematics, and Statistical Design, 247–258.

[18] S. Bustamante, J. Corsten, N. Frankl, A. Pokrovskiy and J. Skokan, *Partitioning Edge-Colored Hypergraphs into Few Monochromatic Tight Cycles*, SIAM J. Discrete Math. 34.2 (2020), 1460–1471.

[19] G. Chen and R. H. Schelp, *Graphs with linearly bounded Ramsey numbers*, J. Combin. Theory Ser. B 57.1 (1993), 138–149.

[20] H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Statistics 23 (1952), 493–507.

[21] V. Chvatál, V. Rödl, E. Szemerédi and W. T. Trotter, *The Ramsey number of a graph with bounded maximum degree*, J. Combin. Theory Ser. B 34.3 (1983), 239–243.

[22] V. Chvátal, *The tail of the hypergeometric distribution*, Discrete Math. 25.3 (1979), 285–287.

[23] D. Conlon, *Hypergraph packing and sparse bipartite Ramsey numbers*, Combin. Probab. Comput. 18.6 (2009), 913–923.

[24] D. Conlon and J. Fox, *Bounds for graph regularity and removal lemmas*, Geom. Funct. Anal. 22.5 (2012), 1191–1256.

[25] D. Conlon, J. Fox and B. Sudakov, *Ramsey numbers of sparse hypergraphs*, Random Structures & Algorithms 35.1 (2009), 1–14.

[26] D. Conlon, J. Fox and B. Sudakov, *On two problems in graph Ramsey theory*, Combinatorica 32.5 (2012), 513–535.

[27] O. Cooley, N. Fountoulakis, D. Kühn and D. Osthus, *Embeddings and Ramsey numbers of sparse κ-uniform hypergraphs*, Combinatorica 29.3 (2009), 263–297.

[28] K. Corrádi and A. Hajnal, *On the maximal number of independent circuits in a graph*, Acta Math. Hungar. 14 (1963), 423–439.

[29] J. Corsten, L. DeBiasio and P. McKenney, *Density of monochromatic infinite subgraphs*, In preparation (2020+).

[30] J. Corsten, L. DeBiasio, A. Lamaison and R. Lang, *Upper density of monochromatic infinite paths*, Advances in Combinatorics (2019).

[31] J. Corsten and W. Mendonça, *Tiling edge-coloured graphs with few monochromatic bounded-degree graphs*, In preparation (2020+).

[32] E. Davies, M. Jenssen and B. Roberts, *Multicolour Ramsey numbers of paths and even cycles*, European J. Combin. 63 (2017), 124–133.

[33] L. DeBiasio and P. McKenney, *Density of monochromatic infinite subgraphs*, Combinatorica 39.4 (2019), 847–878.

[34] G. A. Dirac, *Some theorems on abstract graphs*, Proc. Lond. Math. Soc. 3 (1952), 69–81.

[35] R. A. Duke, H. Lefmann and V. Rödl, *A fast approximation algorithm for computing the frequencies of subgraphs in a given graph*, SIAM J. Comput. 24.3 (1995), 598–620.

[36] N. Eaton, *Ramsey numbers for sparse graphs*, Discrete Math. 185.1-3 (1998), 63–75.

[37] M. Elekes, D. T. Soukup, L. Soukup and Z. Szentmiklóssy, *Decompositions of edge-colored infinite complete graphs into monochromatic paths*, Discrete Math. 340.8 (2017), 2053–2069.

[38] P. Erdős and T. Gallai, *On maximal paths and circuits of graphs*, Acta Math. Hungar. 10.3-4 (1959), 337–356.

[39] P. Erdős and F. Galvin, *Monochromatic infinite paths*, Discrete Math. 113.1 (1993), 59–70.

[40] P. Erdős and G. Szekeres, *A combinatorial problem in geometry*, Compositio Math. 2 (1935), 463–470.

[41] P. Erdős, *On the combinatorial problems which I would most like to see solved*, Combinatorica 1.1 (1981), 25–42.

[42] P. Erdős, A. Gyárfás and L. Pyber, *Vertex coverings by monochromatic cycles and trees*, J. Combin. Theory Ser. B 51.1 (1991), 90–95.

[43] P. Erdős and A. Rényi, *Asymmetric graphs*, Acta. Math. Acad. Sci. Hungar. 14.3-4 (1963), 295–315.

[44] P. Erdős and A. Rényi, *On random graphs*, Publ. Math. Debrecen 6 (1959), 290–297.

[45] P. Erdős and A. Rényi, *On random matrices*, Publ. Math. Inst. Hung. Acad. Sci 8 (1964), 455–461.

[46] P. Erdős and A. Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hung. Acad. Sci. 5 (1960), 17–60.

[47] P. Erdős and A. Rényi, *On the existence of a factor of degree one of a connected random graph*, Acta. Math. Acad. Sci. Hungar. 17 (1966), 359–368.

[48] P. Erdős and A. Rényi, *On the strength of connectedness of a random graph*, Acta. Math. Acad. Sci. Hungar. 12 (1961), 261–267.

[49] R. J. Faudree and R. H. Schelp, *All Ramsey numbers for cycles in graphs*, Discrete Math. 8.4 (1974), 313–329.

[50] R. J. Faudree and R. H. Schelp, *Path Ramsey numbers in multicolorings*, J. Combin. Theory Ser. B 19.2 (1975), 150–160.

[51] R. J. Faudree and R. H. Schelp, *Path-path Ramsey-type numbers for the complete bipartite graph*, J. Combin. Theory Ser. B 19.2 (1975), 161–173.

[52] J. Fox and B. Sudakov, *Density theorems for bipartite graphs and related Ramsey-type results*, Combinatorica 29.2 (2009), 153–196.

[53] A. Frieze and S. Janson, *Perfect matchings in random s-uniform hypergraphs*, Random Structures & Algorithms 7 (1995), 41–57.

[54] D. Galvin, *Three tutorial lectures on entropy and counting*, arXiv:1406.7872 (2014).

[55] L. Gerencsér and A. Gyárfás, *On Ramsey-type problems*, Annales Univ. Sci. Budapest., Sec. Math. 10 (1967), 167–170.

[56] E. N. Gilbert, *Random Graphs*, Ann. Math. Statist. 30.4 (1959), 1141–1144.

[57] R. L. Graham, V. Rödl and A. Ruciński, *On graphs with linear Ramsey numbers*, J. Graph Theory 35.3 (2000), 176–192.

[58] A. Grinshpun and G. N. Sárközy, *Monochromatic bounded degree subgraph partitions*, Discrete Math. 339.1 (2016), 46–53.

[59] A. Gyárfás, *Covering complete graphs by monochromatic paths*, Irregularities of partitions, Springer, 1989, 89–91.

[60] A. Gyárfás, *Vertex covers by monochromatic pieces — A survey of results and problems*, Discrete Math. 339.7 (2016), 1970–1977.

[61] A. Gyárfás and J. Lehel, *A Ramsey-type problem in directed and bipartite graphs*, Period. Math. Hungar. 3.3-4 (1973), 299–304.

[62] A. Gyárfás, M. Ruszinkó, G. N. Sárközy and E. Szemerédi, *An improved bound for the monochromatic cycle partition number*, J. Combin. Theory Ser. B 96.6 (2006), 855–873.

[63] A. Gyárfás and G. N. Sárközy, *Monochromatic loose-cycle partitions in hypergraphs*, Electron. J. Combin. 21.2 (2014), P2–36.

[64] A. Gyárfás and G. N. Sárközy, *Monochromatic path and cycle partitions in hypergraphs*, Electron. J. Combin. 20.1 (2013), P18.

[65] A. Gyárfás, M. Ruszinkó, G. N. Sárközy and E. Szemerédi, *Three-color Ramsey numbers for paths*, Combinatorica 28.4 (2008), 499–502.

[66] A. Hajnal and E. Szemerédi, *Proof of a conjecture of P. Erdős*, Combinatorial theory and its applications, vol. 2, 1970, 601–623.

[67] P. Haxell, *Partitioning complete bipartite graphs by monochromatic cycles*, J. Combin. Theory Ser. B 69.2 (1997), 210–218.

[68] A. Heckel, *Random triangles in random graphs*, arXiv:1802.08472 (2018).

[69] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. 58.301 (1963), 13–30.

[70] K. Hrbacek and T. Jech, *Introduction to set theory*, Third Edition, Crc Press, 1999.

[71] Y. Ishigami, *Linear Ramsey Numbers for Bounded-Degree Hypergraphs*, Electr. Notes Discrete Math. 29 (2007), 47–51.

[72] S. Janson, *Poisson approximation for large deviations*, Random Structures & Algorithms 1.2 (1990), 221–229.

[73] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, vol. 45, John Wiley & Sons, 2011.

[74] M. Jenssen and J. Skokan, *Exact Ramsey numbers of odd cycles via nonlinear optimisation*, arXiv:1608.05705 (2016).

[75] A. Johansson, J. Kahn and V. Vu, *Factors in random graphs*, Random Structures & Algorithms 33.1 (2008), 1–28.

[76] J. Kahn, *Asymptotics for Shamir's Problem*, arXiv:1909.06834 (2019).

[77] J. H. Kim and V. H. Vu, *Concentration of multivariate polynomials and its applications*, Combinatorica 20.3 (2000), 417–434.

[78] J. H. Kim, *Perfect matchings in random uniform hypergraphs*, Random Structures & Algorithms 23 (2003), 111–132.

[79] C. Knierim and P. Su, *Improved Bounds on the Multicolor Ramsey Numbers of Paths and Even Cycles*, Electron. J. Combin. 26.1 (2018).

[80] J. Komlós, G. N. Sárközy and E. Szemerédi, *An algorithmic version of the blow-up lemma*, Random Structures & Algorithms 12.3 (1998), 297–312.

[81] J. Komlós, G. N. Sárközy and E. Szemerédi, *Blow-up lemma*, Combinatorica 17.1 (1997), 109–123.

[82] J. Komlós and M. Simonovits, *Szemerédi's regularity lemma and its applications in graph theory*, Paul Erdős is eighty, vol. 2, Bolyai Soc. Math. Stud., Budapest, 1996, 295–352.

[83] J. Komlós and E. Szemerédi, *Limit distribution for the existence of Hamiltonian cycles in a random graph*, Discrete Math. 43.1 (1983), 55–63.

[84] A. D. Korshunov, *Solution of a problem of Erdős and Renyi on Hamiltonian cycles in nonoriented graphs*, Russian Academy of Sciences 228.3 (1976).

[85] M. Krivelevich, *Triangle factors in random graphs*, Combin. Probab. Comput. 6 (1997), 337–347.

[86] M. Krivelevich, C. Lee and B. Sudakov, *Robust Hamiltonicity of Dirac graphs*, Trans. Amer. Math. Soc. 366 (2014), 3095–3130.

[87] D. Kühn and D. Osthus, *Embedding large subgraphs into dense graphs*, Surveys in Combinatorics (2009), 137–168.

[88] D. Kühn, D. Osthus and A. Taraz, *Large planar subgraphs in dense graphs*, J. Combin. Theory Ser. B 95 (2005), 263–282.

[89] A. Lamaison, *Ramsey upper density of infinite graphs*, arXiv:2003.06329 (2020).

[90] C. Lee, *Ramsey numbers of degenerate graphs*, Annals of Mathematics 185.3 (2017), 791–829.

[91] A. Lo, N. Sanhueza-Matamala and G. Wang, *Density of monochromatic infinite paths*, Electron. J. Combin. 25.4 (2018).

[92] T. Łuczak, V. Rödl and E. Szemerédi, *Partitioning two-coloured complete graphs into two monochromatic cycles*, Combin. Probab. Comput. 7.4 (1998), 423–436.

[93] T. Łuczak, M. Simonovits and J. Skokan, *On the multi-colored Ramsey numbers of cycles*, J. Graph Theory 69.2 (2011), 169–175.

[94] W. Mantel, *Problem 28*, Wiskundige Opgaven 10.60-61 (1907), 320.

[95] A. Pokrovskiy, *Partitioning edge-coloured complete graphs into monochromatic cycles and paths*, J. Combin. Theory Ser. B 106 (2014), 70–97.

[96] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Math. 14.4 (1976), 359–364.

[97] L. Pósa, *On the circuits of finite graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl 8 (1963), 355–361.

[98] R. Rado, *Universal graphs and universal functions*, Acta Arith. 9 (1964), 331–340.

[99] R. Rado, *Monochromatic paths in graphs*, Ann. Discrete Math. 3 (1978), Advances in graph theory (Cambridge Combinatorial Conf., Trinity College, Cambridge, 1977), 191–194.

[100] F. P. Ramsey, *On a Problem of Formal Logic*, Proc. Lond. Math. Soc. 2.1 (1930), 264–286.

[101] O. Riordan, *Random cliques in random graphs*, arXiv:1802.01948 (2018).

[102] V. Rödl and M. Schacht, *Regular Partitions of Hypergraphs: Regularity Lemmas*, Combin. Probab. Comput. 16.06 (2007).

[103] V. Rosta, *On a Ramsey-type problem of J. A. Bondy and P. Erdös. I,II*, J. Combin. Theory Ser. B 15.1 (1973), 94–104.

[104] A. Ruciński, *Matching and covering the vertices of a random graph by copies of a given graph*, Discrete Math. 105.1-3 (1992), 185–197.

[105] A. Ruciński, *Open problem*, Proceedings of Random Graphs (2), Poznan, Wiley, 1989.

[106] G. N. Sárközy, *A quantitative version of the Blow-up Lemma*, arXiv:1405.7302 (2014).

[107] G. N. Sárközy, *Improved monochromatic loose cycle partitions in hypergraphs*, Discrete Math. 334 (2014), 52–62.

[108] G. N. Sárközy, *Monochromatic cycle power partitions*, Discrete Math. 340.2 (2017), 72–80.

[109] G. N. Sárközy, *On the Multi-Colored Ramsey Numbers of Paths and Even Cycles*, Electron. J. Combin. 23.3 (2016).

[110] J. Schmidt and E. Shamir, *A threshold for perfect matchings in random d-pure hypergraphs*, Discrete Math. 45.2-3 (1983), 287–295.

[111] A. Schrijver, *Theory of linear and integer programming*, John Wiley & Sons, 1998.

[112] D. T. Soukup, *Colouring problems of Erdős and Rado on infinite graphs*, PhD thesis (2015), University of Toronto (Canada).

[113] J. Spencer, *Counting extensions*, J. Combin. Theory Ser. A 55.2 (1990), 247–255.

[114] E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Hungar. 20.1-2 (1969), 89–104.

[115] E. Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes, Colloq. Internat. CNRS 260 (1978), 399–401.

[116] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27.21 (1975), 299–345.

[117] Y. Zhao, *Proof of the* $(n/2 - n/2 - n/2)$ *Conjecture for Large n*, Electron. J. Combin. 18.1 (2011).