EXTREMAL PROBLEMS IN PSEUDO-RANDOM GRAPHS AND ASYMPTOTIC ENUMERATION

BY

WOJCIECH SAMOTIJ

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor Douglas West, Chair
Assistant Professor József Balogh, Director of Research
Professor Alexandr Kostochka
Assistant Professor Sujith Vijay

Abstract

This dissertation tackles several questions in extremal graph theory and the theory of random graphs. It consists of three more or less independent parts that all fit into one bigger picture – the meta-problem of describing the structure and properties of large random and pseudo-random graphs.

Given a positive constant c, we call an n-vertex graph G *c-Ramsey* if G does not contain a clique or an independent set of size greater than $c \log n$. Since all of the known examples of Ramsey graphs come from various constructions employing randomness, several researchers have conjectured that all Ramsey graphs possess certain pseudo-random properties. We study one such question – a conjecture of Erdős, Faudree and Sós regarding the orders and sizes of induced subgraphs of Ramsey graphs. Although we do not fully resolve this conjecture, the main theorem in the first part of this dissertation, joint work with Noga Alon, József Balogh, and Alexandr Kostochka, significantly improves the previous state-of-the-art result of Alon and Kostochka.

For a positive integer n and a real number $p \in [0, 1]$, one defines the *Erdős-Rényi random graph* $G(n, p)$ to be the probability distribution on the set of all graphs on the vertex set $\{1, \ldots, n\}$ such that the probability that a particular pair $\{i, j\}$ of vertices is an edge in $G(n, p)$ is p, independently of all other pairs. In the second part of this dissertation, we study the behavior of the random graph $G(n, p)$ with respect to the property of containing large trees with bounded maximum degree. Our first main theorem, joint work with József Balogh, Béla Csaba, and Martin Pei, gives a sufficient condition on p to imply that with probability tending to 1 as n tends to infinity, $G(n, p)$ contains all almost spanning trees with bounded maximum degree, improving a previous result of Alon, Krivelevich, and Sudakov. In the second main theorem of this part, joint work with József Balogh and Béla Csaba, we show that $G(n, p)$ almost surely contains all almost spanning trees with bounded maximum degree even after an adversary removes asymptotically half of the edges in $G(n, p)$.

Given an arbitrary graph H, we say that a graph G is H-free if G does not contain H as a subgraph. Edrős, Frankl, and Rödl generalized a famous theorem of Erdős and Stone by proving that for every non-bipartite H, the number of labeled H-free graphs on a fixed n-vertex set, $f_n(H)$, satisfies $\log_2 f_n(H) \leq (1 + o(1)) \operatorname{ex}(n, H)$. The case when H is bipartite has proved to be much harder. For all such H, apart from the cycles of length 4 and 6, it is not even known whether $\log_2 f_n(H) \leq C \operatorname{ex}(n, H)$. The main result of the last part of this thesis, joint work with József Balogh, proves such a bound for 'almost all' complete bipartite graphs. This result and the methods used to prove it have many interesting applications, some of which we study in the last chapter.

Acknowledgements

I owe my deepest gratitude to Józsi Balogh, the greatest advisor one could ever ask for. Thank you for taking me under your wings; thank you for all the time, encouragement and continuous support in every aspect of life that I have been getting from you since my day one here, at Urbana-Champaign. It has been an honor and pleasure to learn from you and work with you. Without you, I would have never accomplished so much in so little time.

I would also like to thank the many people who have taught me mathematics: my high school teachers (especially Marian Bąk and Wiesław Suchocki), my university teachers (especially Paweł Głowacki, Andrzej Hulanicki, Andrzej Kisielewicz, Ludomir Newelski, and Tadeusz Pezda) and my graduate school teachers (especially Józsi Balogh, Kevin Ford, and Douglas West). For their generous assistance with writing letters, helping with various applications, giving wise advice, and generally making my student life much easier, I wish to thank Noga Alon, Józsi Balogh, Béla Bollobás, Zoltan Füredi, Andrzej Kisielewicz, Sasha Kostochka, Rob Morris, Karen Mortensen, Miki Simonovits, Benny Sudakov, Tibor Szabó, Sujith Vijay, and Douglas West.

I am indebted to many colleagues for providing a stimulating and enjoyable studying and working environment. I am particularly grateful to Jane Butterfield, David Conlon, Béla Csaba, Paweł Gawrychowski, Simon Griffiths, Hiep Han, Artur Jeż, Łukasz Jeż, Paweł Kawa, Choongbum Lee, John Lenz, Mariusz Mirek, Rob Morris, Yury Person and Noah Prince.

I wish to thank all my friends for their companionship, emotional support, entertainment and caring they provided. I am especially grateful to my two best friends in Urbana-Champaign – Evan Erickson and Marianna Strzelecka.

Lastly, and most importantly, I wish to thank my family, and most of them all – my parents, for their ever-present love, support and understanding.

Contents

Chapter 1

Introduction, preliminaries and tools

> Extremal graph theory, in its strictest sense, is a branch of graph theory developed and loved by Hungarians.
>
> *Extremal graph theory*
> Béla Bollobás

1.1 Introduction

Extremal graph theory is a well-established branch of graph theory that studies graphs that are *extremal*, i.e., maximal or minimal with respect to a particular graph parameter, among all graphs which satisfy a certain property. A typical result in extremal graph theory, which many mathematicians consider as one of its cornerstones, is the celebrated theorem of Turán [73], which determines the maximum number of edges in an n-vertex graph that does not contain a subgraph isomorphic to the complete graph on a vertex set of prescribed size. Another typical example is the theorem of Dirac [24], which states that a graph on $n \geq 3$ vertices is Hamiltonian if each of its vertices has degree greater than or equal to $n/2$.

A *random graph*, in the broadest sense, is a probability distribution on a certain family of graphs. The theory of random graphs, an area lying at the intersection of graph theory and probability theory founded half a century ago by Erdős and Rényi [31, 32], studies the properties of a 'typical' random graph, i.e., a graph drawn from that probability distribution. Perhaps the simplest and at the same time the most natural model of generating a random graph G is the random process that for a fixed integer n, considers each pair of numbers $\{i, j\}$, where $1 \leq i < j \leq n$, and independently of all other pairs defines $\{i, j\}$ to be an edge of G with probability $1/2$.

It often happens that with probability tending to 1 as $n \to \infty$ a random graph on n vertices possesses some graph property $\mathcal{P}$; in other words, we might say that a 'typical' graph from a certain family has $\mathcal{P}$. Given such a property and a fixed graph G, if it happens that G has $\mathcal{P}$, we oftentimes say that G is *random-like*, *quasi-random*, or *pseudo-random*. Even though the notion of quasi-randomness, which first appeared in the work of Chung, Graham, and Wilson [21], Rödl [66], and Thomason [72], is quite vague and informal, nevertheless it has proved very useful in many areas of graph theory.

Ever since Erdős [25] used a probabilistic argument to prove that for all natural numbers g and k with $g \geq 3$ and $k \geq 3$, there exist graphs with girth g and chromatic number k, the two areas of extremal graph theory and the theory of random graphs have walked hand in hand and greatly influenced each other. This dissertation tackles several questions that sprouted from the interplay between them.

1.2 Basic definitions

A *graph* is an abstract representation of a set of objects together with a binary relation on that set. Formally, a graph G (more precisely, *undirected simple finite graph*) is an ordered pair composed of an arbitrary finite set $V(G)$ of *vertices* and a set $E(G)$ of unordered pairs of elements of $V(G)$ called its *edges*. An *isomorphism* from a graph G to a graph H is a bijection $f : V(G) \to V(H)$ that for all distinct pairs $v, w \in V(G)$ satisfies $\{v, w\} \in E(G)$ if and only if $\{f(v), f(w)\} \in E(H)$. Sometimes we will identify graphs that are isomorphic; if we want to differentiate between a graph G and its isomorphism class, we will explicitly say that G is a *labeled graph*.

The number of vertices of a graph G, also referred to as the *order* of G, will be denoted by $v(G)$; the number of edges in a graph G, sometimes called the *size* of G, is usually denoted by $e(G)$. The *edge density* (or simply, *density*) of a graph G is the quantity $a(G)$ defined by $a(G) = e(G)/\binom{v(G)}{2}$; clearly, the edge density of any graph is a real number in the interval $[0, 1]$.

Fix a graph G. We say that vertices $v, w \in V(G)$ are *adjacent* (or that v and w are *neighbors*) if the pair $\{v, w\}$ (abbreviated vw) belongs to the set $E(G)$. A vertex $v \in V(G)$ and an edge $e \in E(G)$ are said to be *incident* if v is an *endpoint* of e, which means $v \in e$. For $v \in V(G)$, the *neighborhood* of v, denoted by $N_G(v)$, is the set of all *neighbors* of v; the *degree* of v in G, written $\deg_G(v)$, is the number of neighbors v has in G, i.e., $\deg_G(v) = |N_G(v)|$. The *minimum degree* of G is $\delta(G)$, i.e., $\delta(G) = \min\{\deg_G(v) : v \in V(G)\}$; the *maximum degree* of G is $\Delta(G)$, i.e., $\Delta(G) = \max\{\deg_G(v) : v \in V(G)\}$. Finally, for a set $A \subseteq V(G)$, we denote the set of common neighbors of all vertices in A by $N^*_G(A)$, i.e., $N^*_G(A) = \bigcap\{N_G(v) : v \in A\}$.

A graph H is said to be a *subgraph* of G, written as $H \subseteq G$, if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$; whenever H is a subgraph of G, we also say that G is a *supergraph* of H. We say that H is an *induced subgraph* of G, denoted $H \leq G$, if H is the maximal subgraph of G with the vertex set $V(H)$, i.e., $E(H)$ consists of all the edges of G whose both endpoints lie in the set $V(H)$. Since every induced subgraph of G is uniquely determined by its vertex set, given a set $A \subseteq V(G)$, we denote the induced subgraph of G with vertex set A by $G[A]$; the notation $G - A$ abbreviates $G[V(G) - A]$. A subgraph $H \subseteq G$ is *spanning* if $V(H) = V(G)$.

Given a set $A \subseteq V(G)$, the number of edges within A, i.e., the number of edges in the graph $G[A]$, is denoted by $e_G(A)$. If $e_G(A) = 0$, then A is an *independent set*; if $e_G(A) = \binom{|A|}{2}$, then A is a *clique*. We use $\deg_G(v, A)$ to denote the number of neighbors that a particular vertex $v \in V(G)$ has in A. Given two disjoint sets $A, B \subseteq V(G)$, the number of edges with one endpoint in A and one endpoint in B is $e_G(A, B)$.

A graph G is *k-colorable* (or *k-partite*) if its vertex set can be partitioned into at most k independent sets, called *partite sets*. The *chromatic number* of G, denoted $\chi(G)$, is the smallest integer k such that G is k-colorable. We say that a graph G is *bipartite* if $\chi(G) = 2$.

The n-vertex *complete graph*, denoted K_n, is the unique n-vertex graph with $\binom{n}{2}$ edges. The *complete bipartite* graph with partite sets of size s and t, denoted $K_{s,t}$, is the unique graph with $s+t$ vertices whose vertex set can be partitioned into two partite sets of sizes s and t that has st edges. More generally, the *complete k-partite* graph with partite sets of sizes $s_1, \ldots, s_k$, denoted $K_{s_1,\ldots,s_k}$, is the unique k-partite graph with partite sets of sizes $s_1, \ldots, s_k$ and $\sum_{i<j} s_i s_j$ edges.

A *path* P in G is a list $v_0, \ldots, v_k$ of pairwise distinct vertices of G such that for each $i \in \{0, \ldots, k-1\}$, the pair $\{v_i, v_{i+1}\}$ is an edge in G; the vertices v_0 and v_k are the *endpoints* of P. We usually identify P with the subgraph of G consisting of the vertex set $\{v_0, \dots, v_k\}$ and the edge set $\{v_i v_{i+1} : 0 \le i < k\}$. A *cycle* C in G is a list $v_1, \dots, v_k$ of pairwise distinct vertices of G such that $v_1, \dots, v_k$ is a path in G and $v_k v_1 \in E(G)$; we usually identify C with the subgraph of G consisting of the vertex set $\{v_1, \dots, v_k\}$ and the edge set $\{v_i v_{i+1} : 1 \le i < k\} \cup \{v_1 v_k\}$. The *length* of a path P is the number of edges in P. The *length* of a cycle C is the number of edges (or vertices) in C; the cycle of length k is denoted by C_k. A *Hamilton cycle* in G is a cycle which is a spanning subgraph of G; a graph that contains a Hamilton cycle is called *Hamiltonian*.

Two vertices $v, w \in V(G)$ are *connected* if G contains a path with endpoints v and w. A graph G is *connected* if every two vertices of G are connected. A *component* of G is a maximal subgraph of G that is connected. Note that the vertex and edge sets of the components of a graph partition its vertex and edge sets, respectively. A *tree* is a connected graph that contains no cycles. A *spanning tree* of a graph G is a spanning subgraph of G that is a tree; it is easy to see that a graph is connected if and only if it contains a spanning tree.

For graphs H and G, an *embedding* of H into G is an injective function $f : V(H) \to V(G)$ such that if $e \in E(H)$, then $f(e) \in E(G)$. Clearly, H is a subgraph of G if and only if there exists an embedding of H into G.

A *hypergraph* is a generalization of a graph in which every edge can contain an arbitrary number of vertices. Formally, a (finite) hypergraph $\mathcal{H}$ is an ordered pair composed of an arbitrary finite set $V(\mathcal{H})$ called its *vertices* and a set $E(\mathcal{H})$ of nonempty subsets of $V(\mathcal{H})$ called its *hyperedges* (*edges*). A hypergraph $\mathcal{H}$ is called *k-uniform* if each of its edges has size exactly k; in particular, graphs are simply 2-uniform hypergraphs. Similarly as for graphs, we denote the number of edges of a hypergraph $\mathcal{H}$ by $e(\mathcal{H})$. Also, given a set $A \subseteq \mathcal{H}$, the *subhypergraph induced by A*, denoted $\mathcal{H}[A]$, is the hypergraph with vertex set A consisting only of those edges of $\mathcal{H}$ which are fully contained in A. The notation $\mathcal{H} - A$ abbreviates $\mathcal{H}[V(\mathcal{H}) - A]$.

For a subset $W \subseteq V(\mathcal{H})$, the *degree* of W in $\mathcal{H}$ is the number $\deg_{\mathcal{H}}(W)$ of edges of $\mathcal{H}$ that the set W is contained in; thus $\deg_{\mathcal{H}}(W) = \left|\{D \in E(\mathcal{H}) : W \subseteq D\}\right|$. Given a vertex $w \in V(\mathcal{H})$, we write $\deg_{\mathcal{H}}(w)$ for $\deg_{\mathcal{H}}(\{w\})$. For $\ell \ge 1$, the *maximum ℓ-degree* of $\mathcal{H}$, written $\Delta_\ell(\mathcal{H})$, is defined by

$$\Delta_\ell(\mathcal{H}) = \max\{\deg_{\mathcal{H}}(W) : W \subseteq V(\mathcal{H}) \text{ and } |W| = \ell\};$$

the *maximum degree* of $\mathcal{H}$, written $\Delta(\mathcal{H})$, is $\Delta_1(\mathcal{H})$. Note that if $\ell \ge 1$, then

$$\Delta(\mathcal{H}) \le |V(\mathcal{H})|^{\ell-1} \cdot \Delta_\ell(\mathcal{H}). \tag{1.1}$$

Finally, $\sigma(\mathcal{H})$ will denote the minimum size of a set of vertices of $\mathcal{H}$ that intersects more than half of the edges of $\mathcal{H}$, i.e.,

$$\sigma(\mathcal{H}) = \min\{|S| : S \subseteq V(H) \text{ and } e(\mathcal{H} - S) < e(\mathcal{H})/2\}.$$

Since no vertex belongs to more than $\Delta(\mathcal{H})$ edges of $\mathcal{H}$, clearly

$$\sigma(\mathcal{H}) \ge \frac{e(\mathcal{H})}{2\Delta(\mathcal{H})}. \tag{1.2}$$

The *power set* of a set X, denoted $\mathcal{P}(X)$, is the set of all subsets of X. For a positive integer

k, the set of all functions from A to the set $\{0,\ldots,k-1\}$ will be denoted by k^A. For a positive integer k, the term *k-set* (or *k-subset*) stands for k-element set (or k-element subset). Finally, "log" will always denote the natural logarithm.

Sometimes, for the sake of brevity, we will omit certain subscripts. For example, when a graph G is clear from the context, we may abbreviate $e_G(A)$ by $e(A)$, $\deg_G(v)$ by $\deg(v)$, etc.

Last but not least, we will extensively use the standard *asymptotic notation*. Let f and g be two non-negative real-valued functions defined on the set $\mathbb{N}$ of integers. Moreover, assume that $g(n) > 0$ for all $n \in \mathbb{N}$. We write

- $f = O(g)$ if there exists a $C > 0$ and an $n_0 \in \mathbb{N}$ such that $f(n) \le C \cdot g(n)$ for all $n \ge n_0$,
- $f = \Omega(g)$ if there exists a $C > 0$ and an $n_0 \in \mathbb{N}$ such that $f(n) \ge C \cdot g(n)$ for all $n \ge n_0$,
- $f = o(g)$ if the ratio $\frac{f(n)}{g(n)}$ tends to 0 as n tends to infinity,
- $f = \omega(g)$ if the ratio $\frac{f(n)}{g(n)}$ tends to infinity as n tends to infinity.

1.3 Szemerédi's Regularity Lemma

Szemerédi's proof [71] of the conjecture of Erdős and Turán [37] on the existence of arbitrarily long arithmetic progressions in dense subsets of integers is arguably one of the greatest achievements of modern combinatorics. An auxilliary lemma used in that proof, which has become known as *Szemerédi's Regularity Lemma*, has turned out to be an extremely powerful and widely applicable tool in graph theory. In very rough terms, the lemma says that every graph can be partitioned into a bounded number of pseudo-random bipartite subgraphs and a relatively small leftover. The notion of pseudo-randomness that appears in Szemerédi's Regularity Lemma is defined in terms of the uniformity of the distribution of edges and has, perhaps somewhat unexpectedly, striking implications. Due to its central role in modern extremal graph theory, there have been numerous modifications and generalizations of the Regularity Lemma. Our presentation follows the one from [54], which is especially well suited for sparse graphs.

Fix positive constants η and p, with $p \le 1$. We say that an n-vertex graph G is *η-uniform with density p*, or simply *(η, p)-uniform*, if, for all $A, B \subseteq V(G)$ with $A \cap B = \emptyset$ and $|A|, |B| \ge \eta n$, we have

$$(1-\eta)p|A||B| \le e_G(A,B) \le (1+\eta)p|A||B| \tag{1.3}$$

and

$$(1-\eta)p\binom{|A|}{2} \le e_G(A) \le (1+\eta)p\binom{|A|}{2}. \tag{1.4}$$

Furthermore, we say that G is *η-upper-uniform with density p*, or simply *(η, p)-upper-uniform*, if only the second inequalities in (1.3) and (1.4) hold for all A and B as above.

Let G be a graph, and let p be a positive constant. For any two disjoint subsets $A, B \subseteq V(G)$, let us define the *p-density* of the pair (A, B) in G to be the quantity

$$d_{G,p}(A,B) = \frac{e_G(A,B)}{p|A||B|}.$$

Now suppose that $\varepsilon > 0$ and A, B are as above. We say that the pair (A, B) is *(ε, p)-regular*, or simply *ε-regular* when $p = 1$, if for all $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq \varepsilon|A|$ and $|B'| \geq \varepsilon|B|$, we have

$$|d_{G,p}(A', B') - d_{G,p}(A, B)| \leq \varepsilon.$$

Finally, we say that a partition $(V_0, \ldots, V_k)$ of $V(G)$ is *(ε, p)-regular* if $|V_0| \leq \varepsilon|V(G)|$, and $|V_i| = |V_j|$ for all $i, j \in \{1, \ldots, k\}$, and at least $(1 - \varepsilon)\binom{k}{2}$ pairs (V_i, V_j) with $1 \leq i < j \leq k$ are (ε, p)-regular. We may now state a version of Szemerédi's Regularity Lemma for (η, p)-upper-uniform graphs.

Lemma 1.1. *For any $\varepsilon > 0$ and $k_0 \geq 1$, there are constants $\eta = \eta(\varepsilon, k_0) > 0$ and $K_0 = K_0(\varepsilon, k_0) \geq k_0$ such that any η-upper-uniform graph G with density $0 < p \leq 1$ and at least k_0 vertices admits an (ε, p)-regular partition $(V_0, \ldots, V_k)$ with $k_0 \leq k \leq K_0$.*

1.4 Bounding of Large Deviations

Many arguments in probabilistic combinatorics rely on our ability to simultaneously bound the tail probabilities of a large number of random variables. Often, the 'polynomial' bounds that we can obtain using Markov's and Chebyshev's inequalities are far too weak for our purposes. Fortunately, whenever we are dealing with large families of independent random variables, which are intrinsic to many objects studied by probabilistic combinatorics, the use of moment-generating functions yields much stronger, 'exponential' estimates. The so-called *Chernoff bounds*, named after their inventor [19], give exponentially decreasing bounds on tail probabilities of sums of independent two-valued random variables. Our presentation follows that in [8, Appendix A].

Theorem 1.2. *Let $p \in [0, 1]$, and suppose that $X_1, \ldots, X_n$ are mutually independent random variables with $P(X_i = 1) = p$ and $P(X_i = 0) = 1 - p$ for all i. Let $S_n = X_1 + \ldots + X_n$, and note that $\mathbb{E}[S_n] = pn$. Then for every $a > 0$,*

$$P\big(S_n - pn > a\big) \leq \exp\left(-\frac{a^2}{2pn} + \frac{a^3}{2(pn)^2}\right)$$

and

$$P\big(S_n - pn < -a\big) \leq \exp\left(-\frac{a^2}{2pn}\right).$$

Furthermore, regardless of p,

$$P\big(|S_n - pn| > a\big) \leq 2\exp\left(-\frac{2a^2}{n}\right).$$

A discrete-time *martingale* is a sequence $X_1, X_2, \ldots$ of absolutely integrable random variables that satisfies for all n,

$$\mathbb{E}\big[X_{n+1}|X_1, \ldots, X_n\big] = X_n.$$

A discrete-time martingale can be thought of as a generalization of a sequence of growing sums of independent random variables. It is a very useful tool in situations where we do not have independence. In such settings, when the Chernoff bounds are inapplicable, very strong tail probability bounds can be still obtained from Azuma-Hoeffding inequality. The following result is proved in [8, Chapter 7].

Theorem 1.3. *Let $X_0, X_1, \ldots$ be a martingale such that for all n, we have $|X_{n+1} - X_n| \leq 1$. Let $\mu = \mathbb{E}[X_0]$. For all $n \in \mathbb{N}$ and positive λ,*

$$P\big(|X_n - \mu| > \lambda\sqrt{n}\big) \leq 2e^{-\lambda^2/2}.$$

1.5 Resilience

A *graph property* is a class of graphs closed under isomorphism. A graph property $\mathcal{P}$ is monotone increasing (decreasing) if, whenever a graph G is in $\mathcal{P}$, every graph H that is a supergraph (subgraph) of G with the same vertex set also belongs to $\mathcal{P}$. Let $\mathcal{P}$ be a graph property, and let G be an arbitrary graph from $\mathcal{P}$. The *resilience* of G with respect to the property $\mathcal{P}$ measures how much one has to change G in order to destroy $\mathcal{P}$. Although the notion of resilience, also called *fault tolerance*, has been present in the literature for several years (see, e.g., [3]), only recently it was given a more systematic treatment by Sudakov and Vu [69], who define it as follows.

Definition 1.4. Let $\mathcal{P}$ be a monotone increasing (decreasing) graph property and let G be an arbitrary graph.

1. The *global resilience* of G with respect to $\mathcal{P}$ is the minimum number r such that by deleting (adding) at most $r \cdot e(G)$ edges from (to) G, one can obtain a graph not in $\mathcal{P}$.

2. The *local resilience* of G with respect to $\mathcal{P}$ is the minimum number r such that by deleting (adding) at most $r \cdot \deg_G(v)$ edges at each vertex v of G, one can obtain a graph not in $\mathcal{P}$.

Using the resilience terminology, one can restate many classic results in graph theory. For example, the famous theorem of Turán [73] determines the global resilience of the complete graph on n vertices with respect to the property of containing a k-vertex clique. The theorem of Dirac [24] states that the local resilience of the n-vertex complete graph with respect to the property of containing a Hamilton cycle is $1/2$. In this respect, the notion of resilience has proved very useful and initiated a series of generalizations of classic theorems to the more general setting of random and pseudo-random graphs, see [17, 23, 40, 59, 69].

Chapter 2

Properties of Ramsey graphs

2.1 Introduction

Given a graph G, we call a subset W of its vertices *homogeneous* if W is a clique or an independent set. A classic result in Ramsey theory [35, 65] states that every n-vertex graph contains a homogeneous set of size at least logarithmic in n. On the other hand, a simple application of the probabilistic method proves that in almost every n-vertex graph, the size of the largest homogeneous set is at most logarithmic in n. Moreover, the only known examples of graphs containing no super-logarithmic homogeneous sets come from various constructions based on random graphs with edge density bounded away from 0 and 1. Despite numerous efforts and a promised prize of \$100 for the discoverer [20], no one has yet found any explicit constructions of such graphs. Therefore it is natural to ask whether all graphs with this property look 'random' in some sense.

For a graph G, let $\hom(G)$ denote the size of a largest homogeneous set of vertices in G. Given a positive constant c, we say that an n-vertex graph G is *c-Ramsey* if $\hom(G) \leq c \log n$. The first definite result in the study of c-Ramsey graphs is due to Erdős and Szemerédi [36], who showed that the edge density of a c-Ramsey graph is bounded away from 0 and 1. Soon thereafter, Erdős and Hajnal [30] proved that for a fixed integer k, almost all c-Ramsey graphs are *k-universal*, i.e., they contain every graph on k vertices as an induced subgraph. This was improved by Prömel and Rödl [64], who proved that in fact c-Ramsey graphs are $d \log n$-universal, where the constant d depends only on c. Since by definition a c-Ramsey graph does not contain a clique of super-logarithmic size, this clearly is asymptotically best possible. A similar result was obtained by Shelah [68], who proved that every c-Ramsey graph contains 2^{dn} non-isomorphic induced subgraphs, where again d is some constant depending only on c. This settled a conjecture of Erdős and Rényi.

We tackle a similar problem, first posed by Erdős, Faudree, and Sós (see [26, 27]), who stated the following conjecture.

Conjecture 2.1. *For every positive constant c, there is a positive constant b (depending on c) such that if G is a c-Ramsey graph on n vertices, then the number of distinct pairs $(v(H), e(H))$ realized by induced subgraphs H of G is at least $bn^{5/2}$.*

At the time the conjecture was stated, its authors knew how to prove an $\Omega(n^{3/2})$ lower bound.

The material presented in this chapter is joint work with Noga Alon, József Balogh and Alexandr Kostochka. It was originally published under the title *Sizes of induced subgraphs of Ramsey graphs* in Combinatorics, Probability and Computing (see [2]).

Recently, it was improved to $\Omega(n^2)$ by Alon and Kostochka [5]. We further improve this bound to $\Omega(n^{1+\sqrt{30}/4-\varepsilon})$; note that $1+\sqrt{30}/4 \approx 2.3693$.

Theorem 2.2 ([2])**.** *For all positive constants c and ε, there is a positive constant b (depending on c and ε) such that if G is a c-Ramsey graph on n vertices, then the number of distinct pairs $(v(H), e(H))$ realized by induced subgraphs H of G is at least $bn^{1+\frac{\sqrt{30}}{4}-\varepsilon}$*

In fact, as will become clear in the proof of Theorem 2.2, we prove a slightly stronger statement. Namely, we show that for $\Theta(n)$ different values of k, there are k-vertex induced subgraphs with $\Omega(n^{\frac{\sqrt{30}}{4}-\varepsilon})$ different sizes.

The remainder of this chapter is organized as follows. In Section 2.2, we introduce some notation and cite several results that will be used repeatedly throughout the chapter; at the end of the section, we formulate Theorem 2.12 (a rather technical statement), from which we will quite easily derive (in Section 2.3) the main result, Theorem 2.2. In Section 2.4, we prove three technical lemmas that will later be used in the proof of Theorem 2.12 in Section 2.5.

2.2 Counting induced subgraphs

We start this section with two important definitions. For every integer k with $0 \le k \le v(G)$, we let

$$\psi(k, G) = \max\{e(H) - e(H') : H, H' \le G \text{ with } v(H) = v(H') = k\}, \text{ and}$$
$$\phi(k, G) = |\{e(H) : H \le G \text{ with } v(H) = k\}|.$$

Note that the number of distinct pairs $(v(H), e(H))$ as H ranges over all induced subgraphs of G, as in Conjecture 2.1, can now be computed by

$$\big|\{(v(H), e(H)) : H \le G\}\big| = \sum_{k=0}^{v(G)} \phi(k, G).$$

Erdős, Goldberg, Pach, and Spencer [29] derived the following lower bound on $\psi(k, G)$ for graphs with edge density bounded away from 0 and 1; see also [5, 16].

Theorem 2.3. *For any positive ε with $\varepsilon < 1/2$ and k and n satisfying $5/\varepsilon < k < n/2$, and for any graph G on n vertices with density satisfying $\varepsilon < a(G) < 1-\varepsilon$, we have $\psi(k, G) \ge 10^{-4}k^{3/2}\varepsilon^{1/2}$.*

Suppose that each vertex $v \in V(G)$ is given a nonnegative weight $\omega(v)$. For a subgraph G' of G, let its *weight* be defined as $\omega(G') = e(G') + \sum_{v \in V(G')} \omega(v)$. Generalizing the above definitions to weighted graphs, we introduce a new parameter

$$\phi_\omega(k, G) = |\{\omega(G') : G' \le G \text{ with } v(G') = k\}|.$$

Also, for a vertex v, let $\deg^\omega(v, W) = \deg(v, W) + \omega(v)$ and similarly, $\deg^\omega_H(u) = \deg_H(u) + \omega(u)$. We will refer to these values as *weighted degrees.*

In the sequel we will repeatedly use the following results of Alon and Kostochka [5]. Although Theorem 2.4 does not appear there in the form in which it is stated below, it can be inferred from the proof of the main result of [5] (see the concluding remarks in [5]).

Theorem 2.4. *For every positive ε with $\varepsilon < 1/2$, there is an integer n_0 (depending on ε) such that the following holds. Let G be a graph with $\varepsilon < a(G) < 1 - \varepsilon$. If $n = v(G) \geq n_0$, and $k \leq \frac{\varepsilon n}{3}$ and every vertex $v \in V(G)$ is given a weight $\omega(v)$ in the interval $[0, x \cdot \psi(k, G)/k]$, where $x \geq 1$, then*

$$\phi_\omega(k, G) \geq 10^{-8}\frac{k}{x}.$$

Moreover, one can find $10^{-8}k/x$ distinct sizes of induced k-vertex subgraphs of G such that the difference between consecutive weights is at least mx, where $m = 500\psi(k, G)/k$.

Definition 2.5. Let G be a graph on n vertices. For $0 \leq k \leq n$, define

$$m(k, G) = 500\frac{\psi(k, G)}{k}.$$

For a k-element subset W of $V(G)$, say that a vertex v in G is *W-typical* if

$$\left|\deg(v, W) - a(G)(k-1)\right| \leq m(k, G) + 1.$$

Theorem 2.6 ([5])**.** *Let G be a graph on n vertices, and let W be a k-element subset of $V(G)$, where $20 < k \leq n/3$. All but at most $|W|/5$ vertices inside W are W-typical, and all but at most $|W|/5$ vertices outside W are W-typical.*

It is also good to keep in mind the following simple observation.

Observation 2.7. *Let G be a graph, and let W be a k-element subset of $V(G)$. If ω is a nonnegative weight function on $V(G)$, then every W-typical vertex v satisfies*

$$\left|\deg^\omega(v, W) - a(G)(k-1)\right| \leq m + \omega(v) + 1.$$

In particular, if the weights are in the interval $[0, x \cdot \psi(k, G)/k]$ for some $x \geq 1$, then all typical vertices satisfy

$$\left|\deg^\omega(v, W) - a(G)(k-1)|\right| \leq 2mx.$$

The two main definitions we are about to state are motivated by the following result of Erdős and Szemerédi [36].

Theorem 2.8. *For every positive constant c, there is some positive ε (depending on c) such that if G is an n-vertex c-Ramsey graph, then $\varepsilon < a(G) < 1 - \varepsilon$.*

Assume that G is a graph on n vertices and H is an induced subgraph of G with n^δ vertices. It is clear that $\hom(H) \leq \hom(G)$. Therefore, if G is c-Ramsey, then H is c/δ-Ramsey, so in particular $a(H)$ is bounded away from 0 and 1. Below we define a similar property for an arbitrary graph.

Definition 2.9. For $0 < \varepsilon < 1/2$ and $0 < \delta \leq 1$, let $\mathcal{D}(\epsilon, \delta)$ denote the family of graphs G such that all induced subgraphs $H \leq G$ with $v(H) \geq v(G)^\delta$ have density $a(H)$ in the interval $(\varepsilon, 1-\varepsilon)$.

We immediately derive the following corollary of Theorem 2.8.

Corollary 2.10. *Let c and δ be positive constants. There are constants $\varepsilon \in (0, 1/2)$ and n_0 (depending on c and δ) such that every n-vertex c-Ramsey graph with $n \geq n_0$ belongs to $\mathcal{D}(\epsilon, \delta)$.*

Keeping in mind the statement of Corollary 2.10, from now on we can focus our attention on graphs in the classes $\mathcal{D}(\epsilon, \delta)$. Our aim will be to show that large enough graphs in $\mathcal{D}(\epsilon, \delta)$ have many induced subgraphs that differ either by number of vertices or weight, for a reasonably chosen weight function. The following definition should make this a little more precise.

Definition 2.11. For $0 < \varepsilon < 1/2$ and $0 < \delta \leq 1$, let $\mathcal{P}(\epsilon, \delta)$ be the set of pairs (α, β) such that for some positive constants C, D, F, and n_0 the following holds:
If $n \geq n_0$ and G is an n-vertex graph in $\mathcal{D}(\epsilon, \delta)$, then

$$\phi_\omega(k, G) \geq C \frac{k}{x^F \log^D n} \min \left\{ k^\alpha, \left(\frac{\psi(k,G)}{k} \right)^\beta \right\} \tag{2.1}$$

for all $k \in [\frac{\epsilon n}{100}, \frac{\epsilon n}{3}]$ and nonnegative weight functions ω bounded by $x \cdot \psi(k, G)/k$, where $x \geq 1$.

We will be working only with graphs whose edge density is bounded away from 0 and 1. For all such G, Theorem 2.3 guarantees that $(\psi(k,G)/k)^\beta = \Omega(k^{\beta/2})$ when $k \in [\frac{\epsilon n}{100}, \frac{\epsilon n}{3}]$. Therefore, if $\beta > 2\alpha$, the minimum in (2.1) is k^α, and it can change only by a constant multiplicative factor when we decrease β to 2α. Since we do not care about the constants, but only the order of magnitude of $\phi_\omega(k, G)$, we can always assume that $\beta \leq 2\alpha$ whenever $(\alpha, \beta) \in \mathcal{P}(\epsilon, \delta)$.

Also, since $\psi(k,G) \leq k(k-1)/2$, trivially $(\psi(k,G)/k)^\beta \leq k^\beta$. Therefore, when $\alpha > \beta$, the minimum in (2.1) is $(\psi(k,G)/k)^\beta$, and it will not change when we decrease α to β. Therefore we can also assume that $\alpha \leq \beta$ when $(\alpha, \beta) \in \mathcal{P}(\epsilon, \delta)$.

Finally, we state the main theorem, from which the main result, Theorem 2.2, will be derived as a simple corollary.

Theorem 2.12 ([2])**.** *If* $(\alpha, \beta) \in \mathcal{P}(\epsilon, \delta)$*, then* $\left(\frac{\beta+2}{\beta+5}, \frac{\alpha+1}{2}\right) \in \mathcal{P}(\epsilon, \delta/10)$.

We postpone the proof of Theorem 2.12 to Section 2.5. Instead we now show how it implies the main result, Theorem 2.2.

2.3 Proof of Theorem 2.2

First note that, by Theorem 2.4, when $0 < \varepsilon' < 1/2$, the pair $(0, 0)$ is in $\mathcal{P}(\varepsilon', 1)$. Define

$$i(\alpha, \beta) = \left(\frac{\beta+2}{\beta+5}, \frac{\alpha+1}{2} \right),$$

and note that

$$i^2(\alpha, \beta) = \left(\frac{\alpha+5}{\alpha+11}, \frac{2\beta+7}{2\beta+10} \right).$$

Now it is easy to see that both coordinates of the sequence $i^{2n}(0,0)$ are increasing and bounded, and hence the sequence converges to a pair (α, β) satisfying

$$\alpha = \frac{\alpha+5}{\alpha+11} \quad \text{and} \quad \beta = \frac{2\beta+7}{2\beta+10},$$

namely,

$$(\alpha, \beta) = (\sqrt{30} - 5, \frac{\sqrt{30}}{2} - 2) \approx (0.4772, 0.7386).$$

By iteratively applying Theorem 2.12, we infer that for every positive constant ε, there is some δ such that $(\alpha-\varepsilon, \beta-\varepsilon) \in \mathcal{P}(\epsilon', \delta)$ for every $\varepsilon' > 0$. Let G be a c-Ramsey graph with $v(G)$ sufficiently large, and let $n = v(G)$. By Corollary 2.10, $G \in \mathcal{D}(\epsilon', \delta)$ for some positive ε'. Set $x = 1$ and let $\omega(v) = 0$ for all $v \in V(G)$. By the definition of $\mathcal{P}(\epsilon', \delta)$, for all $k \in [\frac{\epsilon' n}{100}, \frac{\epsilon' n}{3}]$,

$$\begin{aligned}\phi(k, G) &\geq C \frac{k}{\log^D n} \min \left\{ k^{\alpha-\varepsilon}, \left(\frac{\psi(k, G)}{k} \right)^{\beta-\varepsilon} \right\} \geq \Omega \left(\frac{k \cdot \min \left\{ k^{\alpha-\varepsilon}, k^{(\beta-\varepsilon)/2} \right\}}{\log^D n} \right) \\ &\geq \Omega \left(\min \left\{ k^{1+\alpha-2\varepsilon}, k^{1+\beta/2-\varepsilon} \right\} \right) \geq \Omega \left(k^{\frac{\sqrt{30}}{4}-\varepsilon} \right),\end{aligned}$$

where the first inequality follows from Theorem 2.3, the second inequality holds because $k = \Theta(n)$ and hence $\log^D n = o(k^\varepsilon)$, and the last one is due to $\beta/2 < \alpha$. Hence the number of distinct pairs $(v(H), e(H))$ can be bounded below as follows:

$$\sum_{k=0}^{n} \phi(k, G) \geq \sum_{k \in [\frac{\epsilon' n}{100}, \frac{\epsilon' n}{3}]} \phi(k, G) \geq \Omega(n^{1+\frac{\sqrt{30}}{4}-\varepsilon}).$$

2.4 Technical lemmas

Theorem 2.13 ([2]). *Let M_j denote the family of all $\prod_{i=0}^{j-1}(n-i)$ ordered subsets of $[n]$ of cardinality j. Let $f : M_j \to \mathbb{R}$ be a real function. Suppose that $|F(A) - f(B)| \leq 1$ for every $A = \{a_1, \ldots, a_j\} \in M_j$ and $B = \{b_1, \ldots, b_j\} \in M_j$ such that the number of indices i for which $a_i \neq b_i$ is at most 2. Let μ denote the expected value of $f(T)$, where T is chosen randomly and uniformly in M_j. Then, for every $\lambda > 0$,*

$$P\big(|f(T) - \mu| \geq \lambda \sqrt{j}\big) \leq 2e^{-\lambda^2/2}.$$

Proof. We apply the method from [4, Lemma 2.2]. Define a martingale $X_0, X_1, \ldots, X_j$ on the members T of M_j, where $X_i(T)$ is the expected value of $f(T')$ as T' ranges over all ordered subsets T of size j whose first i entries agree with the first i entries of T. Thus $X_0 = \mu$ is a constant and $X_j(T) = f(T)$. We claim that if two ordered sets A and B agree on their first i elements and differ in position $i+1$, then $|X_{i+1}(A) - X_{i+1}(B)| \leq 1$. For every $T \in M_j$, let $\pi(T)$ be the set obtained from T by swapping a_{i+1} and b_{i+1}. Clearly, π is a bijection between the ordered sets $T \in M_j$ that agree with A on their first $i+1$ elements and those that agree with B on their first $i+1$ elements, such that the symmetric difference between T and $\pi(T)$ is at most 2 for all T. Thus the two averages $X_{i+1}(A)$ and $X_{i+1}(B)$ differ by at most 1. This easily implies that $|X_{i+1}(T) - X_i(T)| \leq 1$ for all i, as $X_i(T)$ is the average of numbers of the form $X_{i+1}(T')$ any pair of which, by the hypothesis, differ by at most 1. The result now follows from Theorem 1.3. □

From the above it is easy to get the following.

Lemma 2.14 ([2]). *Let s be a fixed integer. Let G be a graph on n vertices, and take $N_1, \ldots, N_{n^s} \subseteq V(G)$ with $n_i = |N_i|$. There is an ordering $(v_1, \ldots, v_n)$ of the vertices in $V(G)$ such that, if $S_j = \{v_1, \ldots, v_j\}$ for $1 \leq j \leq n$, then*

1. *for each i and j, the size of $S_j \cap N_i$ differs from its expectation $\frac{j}{n} \cdot n_i$ by at most $2j^{1/2}\sqrt{2(s+1)\log(2n)}$,*

2. for each j, the number of edges in $G[S_j]$ differs from its expectation $a(G)\cdot\binom{j}{2}$ by at most $2j^{3/2}\sqrt{2\log(2n)}$.

Proof. Take a random ordering of the vertices of $V(G)$. For every fixed j, the set S_j is a uniform random j-subset of $V(G)$. Fix $i\in[n^s]$. By applying Theorem 2.13 to the function f defined by $f(T)=|T\cap N_i|/2$, we obtain

$$P\left(\left||S_j\cap N_i|-\frac{j}{n}\cdot n_i\right|>2j^{1/2}\sqrt{2(s+1)\log(2n)}\right)\le\frac{1}{(2n)^{s+1}}.$$

It follows that the probability that our ordering does not satisfy 1 is at most $n^s\cdot n\cdot 1/(2n)^{s+1}$, which is at most $1/2^{s+1}$.

Similarly, applying Theorem 2.13 when $f(T)$ if defined to be $e(G[T])/(2|T|)$ yields

$$P\left(\left|e(G[S_j])-a(G)\cdot\binom{j}{2}\right|>2j^{3/2}\sqrt{\log(2n)}\right)\le\frac{1}{2n}.$$

Therefore, the probability that our ordering does not satisfy 2 is at most $1/2$. Hence the probability that a random ordering of $V(G)$ satisfies both 1 and 2 is greater than zero. □

Finally, we need a folklore lemma, whose proof we present for the sake of completeness. For a set $X\subseteq\mathbb{R}$ that is a union of finitely many disjoint intervals, let $l(X)$ denote the sum of the lengths of these intervals.

Lemma 2.15 ([2])**.** *Given n bounded open intervals $I_1,\dots,I_n$, there exists a set $J\subseteq[n]$ such that the intervals indexed by J are pairwise disjoint and*

$$\sum_{j\in J}l(I_j)=l\left(\bigcup_{j\in J}I_j\right)\ge\frac{1}{2}l\left(\bigcup_{j=1}^{n}I_j\right).$$

Proof. First, we delete all 'redundant' intervals, i.e. every time some $I_i\subseteq\bigcup_{j\neq i}I_j$, we remove I_i. Since the union of all intervals does not change after any such deletion, without loss of generality we can assume that the family $I_1,\dots,I_n$ contains no redundant intervals. We may also assume that the left ends of our intervals form a nondecreasing sequence. It easily follows that also the right ends form a non-decreasing sequence, or otherwise $I_{i+1}\subseteq I_i$ for some i.

Now observe that whenever $j>i+1$, the interval I_j lies to the right of I_i (so in particular they are disjoint), or otherwise $I_{i+1}\subseteq I_i\cup I_j$. Hence all the intervals with even indices are pairwise disjoint, and similarly, all the intervals with odd indices are pairwise disjoint. Clearly, one of those families has to cover at least half of the total length. □

2.5 Proof of Theorem 2.12

Fix some $\delta,\varepsilon>0$ and any pair $(\alpha,\beta)\in\mathcal{P}(\epsilon,\delta)$. Let $\gamma=(\beta+2)/(\beta+5)$ and C,D,F be as in Definition 2.11. Recall that by the remark following Definition 2.11, we may assume that $\beta/2\le\alpha\le\beta$. Furthermore, choose $G\in\mathcal{D}(\epsilon,\delta/10)$ with n vertices, fix $k\in[\frac{\epsilon n}{100},\frac{\epsilon n}{3}]$, and fix $x\ge 1$. For each $v\in V(G)$, let $\omega(v)$ be its weight, satisfying $0\le\omega(v)\le x\cdot\psi(k,G)/k$. To simplify the notation we let $m = m(k, G)$. Throughout the proof we will assume that n is sufficiently large. We will also omit all ceiling and floor signs, as they are not crucial. Finally, in order to avoid tedious constant computations, $C', D', F', C'', D'', F'' \dots$ will denote some constants that depend only on δ, ε, α, and β, and not on k, n, ω, or G. In order to limit the number of different symbols in the proof, these constants will often be 'recycled'. We hope that this does not cause too much confusion. Similarly, $C_1, C_2, \dots$ will denote some constants depending only on δ, ε, α, and β, but their values will remain fixed throughout the entire proof.

Theorem 2.4 guarantees the existence of a sequence $H_1, \dots, H_{10^{-8}k/x}$ of k-vertex induced subgraphs of G such that $\omega(H_{i+1}) - \omega(H_i) \geq mx$ for every $1 \leq i < 10^{-8}k/x$.

Before we start, let us outline our general strategy. First, for each i in the above range, we will find an interval I_i centered at $\omega(H_i)$ that contains some N_i different weights of k-vertex induced subgraphs of G. Then, using Lemma 2.15, we will find a large subfamily of pairwise disjoint intervals (thus making sure that they all contain different weights) and add up the corresponding values of N_i. The sum we obtain will surely be a lower bound on the number of distinct weights of k-vertex induced subgraphs of G. In order to prove the promised lower bound, we will make sure that for all i, the ratio of N_i and the length $l(I_i)$ of the interval I_i will satisfy

$$\frac{N_i}{l(I_i)} \geq \frac{C'}{mx^{F'} \log^{D'} n} \min\left\{k^{\gamma}, m^{\frac{1+\alpha}{2}}\right\}. \tag{2.2}$$

We also ensure that the total length of this family of disjoint intervals will be of order $\Omega(k \cdot m)$.

Fix i. By Theorem 2.6, at least $0.8k$ vertices in $V(H_i)$ are $V(H_i)$-typical. Hence, we can find either a list $u_1, \dots, u_{0.5k^{\gamma}}$ of typical vertices with different values under $\deg^{\omega}_{H_i}$ or a set $B_i \subseteq V(H_i)$ of typical vertices with the same value under $\deg^{\omega}_{H_i}$, say d_i', of size at least $k^{1-\gamma}$. Similarly, there are either typical vertices $v_1, \dots, v_{0.5k^{\gamma}} \in V(G) - V(H_i)$ with different weighted degrees $\deg^{\omega}(v_j, V(H_i))$ or a set $A_i \subseteq V(G) - V(H_i)$ of typical vertices with the same value of $\deg^{\omega}(-, V(H_i))$, say d_i, of size at least $k^{1-\gamma}$.

Assume first that we have found a list $u_1, \dots, u_{0.5k^{\gamma}} \in V(H_i)$ of typical vertices with different weighted degrees $\deg^{\omega}_{H_i}$. Let v be an arbitrary $V(H_i)$-typical vertex from $V(G) - V(H_i)$. Either at least $0.25k^{\gamma}$ vertices in the list (u_j) are adjacent to v, or at least $0.25k^{\gamma}$ vertices in that list are non-neighbors of v. Without loss of generality we may assume that the former holds and $u_1, \dots, u_{0.25k^{\gamma}} \in N_G(v)$. Consider graphs $H_{i,j} = G[V(H_i) \cup \{v\} - \{u_j\}]$. For all j in $\{1, \dots, 0.25k^{\gamma}\}$, the weights $\omega(H_{i,j})$, which satisfy

$$\omega(H_{i,j}) = \omega(H_i) + \deg^{\omega}(v, H_i) - 1 - \deg^{\omega}_{H_i}(u_j),$$

are all distinct. Moreover, since both u_j and v are $V(H_i)$-typical, by our assumption on ω and Observation 2.7,

$$|\omega(H_{i,j}) - \omega(H_i)| \leq |\deg^{\omega}(v, H_i) - \deg^{\omega}_{H_i}(u_j)| + 1 < 5mx.$$

Hence if we set $I_i = \omega(H_i) + (-5mx, 5mx)$, then all the weights $\omega(H_{i,j})$ will belong to I_i. Therefore $N_i \geq C'k^{\gamma}$, and so (2.2) is satisfied.

We deal with the case when we can find a list $v_1, \dots, v_{0.5k^\gamma}$ in a similar fashion. Fix a typical vertex $u \in V(H_i)$. Similarly as above, without loss of generality we may assume that $v_1, \dots, v_{0.25k^\gamma}$ are adjacent to u. The $0.25k^\gamma$ graphs obtained from $V(H_i)$ by exchanging u with a vertex from the list $v_1, \dots, v_{0.25k^\gamma}$ have distinct weights from the interval $I_i = \omega(H_i) + (-5mx, 5mx)$. Hence for the remainder of the proof we assume that there are sets A_i, B_i and numbers d_i, d_i', as described above.

Let $t = k^{2(1-\gamma)/3}$. Fix an arbitrary subset B_i' of B_i of size t. We can find a t-element subset $A_i' \subseteq A_i$ such that the numbers of neighbors in B_i' of every pair of vertices of A_i' differ by at most $\sqrt{t}$. It is possible since

$$|A_i| \geq t^{3/2} = t \cdot \frac{|B_i'|}{\sqrt{t}}.$$

Let d_i^* be the edge density between A_i' and B_i', that is

$$d_i^* = \sum_{v \in A_i'} \frac{\deg(v, B_i')}{t^2}.$$

By the choice of A_i', we have $|\deg(v, B_i') - td_i^*| \leq \sqrt{t}$ for all $v \in A_i'$. Applying Lemma 2.14 to the graph $G[B_i']$ and the neighborhoods in the set B_i' of vertices from A_i', one gets the following statement.

Claim 2.16. *It is possible to order B_i' as $b_1, \dots, b_t\}$ so that, for all $1 \leq z \leq t$,*

1. $|e(S_z) - a_i'\binom{z}{2}| \leq 2z^{3/2}\sqrt{2\log k}$, *and*
2. $|\deg(v, S_z) - zd_i^*| \leq 2z^{1/2}\sqrt{5\log k}$ *for all* $v \in A_i'$,

where $S_z = \{b_1, \dots, b_z\}$ and $a_i' = a(G[B_i'])$.

Proof. Let $A_i' = \{v_1, \dots, v_t\}$. Define N_j to be the set of neighbors of v_j in the set B_i'. By the remark preceding the statement of this Claim, we have

$$n_j = |N_j| = \deg(v_j, B_i') \in \left[td_i^* - \sqrt{t}, td_i^* + \sqrt{t}\right]. \tag{2.3}$$

Lemma 2.14 applied to the graph $G[B_i']$ and the sets $N_1, \dots, N_t$ yields the desired ordering $b_1, \dots, b_t$. To see that 1 holds, it suffices to note that $t \ll k$, so $\log(2t) \leq \log k$. For 1, Lemma 2.14 guarantees that

$$\left|d(v_j, S_z) - \frac{z}{t}n_j\right| \leq 2z^{1/2}\sqrt{4\log(2t)}; \tag{2.4}$$

combining (2.3) with (2.4) gives the desired bound. □

What we would like to do now is to obtain many k-vertex induced subgraphs of G with different weights by exchanging the set of vertices $S_z \subseteq B_i'$ with many subsets of A_i', possibly for many values of z.

To get started and see how this idea works in practice, let T_z be some set of z vertices from A_i', and let $H_i'(z) = G[V(H_i) \cup T_z - S_z]$. We compute the weight of this graph.

$$
\begin{aligned}
\omega(H_i'(z)) &= \omega(H_i) - \sum_{j=1}^{z}(\omega(b_j) + \deg_{H_i}(b_j)) + e(S_z) + \sum_{v\in T_z}(\omega(v) + \deg(v, H_i - S_z)) + e(T_z) \\
&= \omega(H_i) - d_i'z + e(S_z) + \sum_{v\in T_z}(d_i - \deg(v, S_z)) + e(T_z) \\
&= \omega(H_i) + \Delta_i z + e(S_z) + e(T_z) - \sum_{v\in T_z}\deg(v, S_z), \qquad (2.5)
\end{aligned}
$$

where $\Delta_i = d_i - d_i' \in [-4mx, 4mx]$. If all the degrees $\deg(v, S_z)$, where v ranges over $A_i' \supseteq T_z$, were equal, then for fixed i and z the weight of $H_i'(z)$ would depend only on $e(T_z)$. Even though it does not have to be the case (our last claim only guarantees that $\deg(v, S_z)$ are all 'close' to $d_i^* z$), this will not be a big issue for us, since, as we will later see, by assigning carefully chosen weights to vertices in A_i', we can compensate for the possibly uneven distribution of the degrees.

Fix z with $z \geq n^{1/10}$. Let $d_i^{\max}(z) = \max_{v\in A_i'}\deg(v, S_z)$ and for each $v \in A_i'$ set $\omega'(v) = d_i^{\max}(z) - \deg(v, S_z)$. If we again let T_z be some z-subset of A_i', then

$$
\omega'(T_z) = e(T_z) + \sum_{v\in T_z}\omega'(v) = e(T_z) - \sum_{v\in T_z} d(v, S_z) + d_i^{\max}(z)\cdot z. \qquad (2.6)
$$

Hence if we combine (2.5) with (2.6), we can express the weight $\omega(H_i'(z))$ of $H_i'(z)$ as

$$
\omega(H_i'(z)) = \omega(H_i) + \Delta_i z + e(S_z) - d_i^{\max}(z)\cdot z + \omega'(T_z),
$$

where only the last term depends on the choice of T_z as a particular z-subset of A_i'.

Claim 2.17. *There are positive constants C_1 and D_1 such that if $n^{1/10} \leq z \leq t' = \varepsilon t/3$, then*

$$
\phi_{\omega'}(z, G[A_i']) \geq \frac{C_1 z}{\log^{D_1} n}\min\left\{z^\alpha, \left(\frac{\psi(z, A_i')}{z}\right)^\beta\right\}. \qquad (2.7)
$$

Proof. Let A_i'' be any $(3z/\varepsilon)$-element subset of A_i' that contains some z vertices spanning the most edges among all z-vertex subsets of A_i' and some z vertices spanning the least edges among all z-vertex subsets of A_i'. By construction, $\psi(z, A_i'') = \psi(z, A_i')$. Since we assumed that z is large enough, $G \in \mathcal{D}(\epsilon, \delta/10)$ implies that $G[A_i''] \in \mathcal{D}(\epsilon, \delta)$. By our assumption that $(\alpha, \beta) \in \mathcal{P}(\epsilon, \delta)$, we have

$$
\phi_{\omega'}(z, G[A_i'']) \geq \frac{C' z}{\log^D z \cdot \log^F n}\min\left\{z^\alpha, \left(\frac{\psi(z, A_i'')}{z}\right)^\beta\right\},
$$

since from Claim 2.16 (ii) and Theorem 2.3 it follows that (provided n is large enough) $0 \leq \omega' \leq \log n \cdot \psi(z, A_i'')/z$. Finally, (2.7) follows because $A_i'' \subseteq A_i'$, and therefore $\phi_{\omega'}(z, G[A_i']) \geq \phi_{\omega'}(z, G[A_i''])$. □

We rewrite formula (2.5) as

$$
\omega(H_i'(z)) = \omega(H_i) + \Delta_i z + a_i'\binom{z}{2} + (e(S_z) - a_i'\binom{z}{2}) + e(T_z) - \sum_{v\in T_z}(\deg(v, S_z) - d_i^* z) - d_i^* z^2.
$$

Now let $a_i = a(G[A'_i])$, and recall that $z \ge n^{1/10}$. Since:

- $|e(S_z) - a'_i\binom{z}{2}| \le 2z^{3/2}\sqrt{2\log k}$ by Claim 2.16 (i),
- $|\deg(v, S_z) - d^*_i z| \le 2z^{1/2}\sqrt{5\log k}$ by Claim 2.16 (ii),
- $|e(T_z) - a_i \cdot \binom{z}{2}| \le \psi(z, A'_i)$ by the definition of ψ, and
- $\psi(z, A'_i) \ge 10^{-4}\varepsilon^{1/2}z^{3/2}$ by Theorem 2.3 and the assumptions on G and z,

we conclude that $\omega(H'_i(z)) - \omega(H_i)$ lands in the interval $I_i(z)$ obtained by scaling and shifting the interval $(-\psi(z, A'_i), \psi(z, A'_i))$. That is,

$$I_i(z) = \Delta_i z + (a'_i + a_i)\binom{z}{2} - d^*_i z^2 + C_2 \log k \cdot \big(-\psi(z, A'_i), \psi(z, A'_i)\big), \tag{2.8}$$

where C_2 is some constant depending only on ε. In particular, the following is true.

Claim 2.18. *We can find k-vertex induced subgraphs of G with at least $\phi_{\omega'}(z, G[A'_i])$ different weights in the interval $\omega(H_i) + I_i(z)$.*

In the remainder of the proof we will carefully bound the number of different weights in all these intervals. Recall that the number z of vertices we want to exchange satisfies $n^{1/10} \le z \le t' = \frac{\varepsilon t}{3}$. For the sake of brevity, let $|I_i(z)|$ denote $\max\{|\inf I_i(z)|, |\sup I_i(z)|\}$, and

$$c(I_i(z)) = \Delta_i z + (a'_i + a_i)\binom{z}{2} - d^*_i z^2 \tag{2.9}$$

will denote the center of the interval $I_i(z)$.

Before we proceed with the counting, first we prove a technical lemma.

Claim 2.19. *For $z \ge n^{1/10}$, the centers $c(I_i(z+1))$ and $c(I_i(z))$ satisfy*

$$|c(I_i(z+1)) - c(I_i(z)) - \Delta_i| < 5z.$$

In particular,

$$|\Delta_i| - 5z < |c(I_i(z+1)) - c(I_i(z))| < |\Delta_i| + 5z.$$

Proof. Let $\delta = c(I_i(z+1)) - c(I_i(z))$. Looking at the definition of $c(I_i(z))$ in (2.9), it is easy to see that δ can be computed as follows:

$$\delta = \Delta_i + (a'_i + a_i) \cdot \left[\binom{z+1}{2} - \binom{z}{2}\right] + d^*_i(z^2 - (z+1)^2) = \Delta_i + (a'_i + a_i)z - d^*_i(2z+1).$$

Hence,

$$\big|\delta - \Delta_i\big| \le (a'_i + a_i)z + d^*_i(2z+1) \le 4z + 1,$$

where the last inequality holds because $a_i, a'_i, d^*_i \in [0, 1]$. □

We will now analyze the function $z \mapsto |I_i(z)|$ and split the proof into several cases. First, recall that $m = 500\psi(k, G)/k$ and z is in the range $n^{1/10} \le z \le t' = \varepsilon t/3$.

Case 1. $\max_z |I_i(z)| < 3mx$.

In particular, $I_i(t') \subseteq (-3mx, 3mx)$. We set $I_i = \omega(H_i) + (-3mx, 3mx)$ and note that by Claims 2.17 and 2.18,

$$\begin{aligned} N_i \geq \phi_{\omega'}(t', G[A_i']) &\geq \frac{C_1 t'}{\log^{D_1} n} \min\left\{(t')^\alpha, \left(\frac{\psi(t', A_i')}{t'}\right)^\beta\right\} \\ &\geq \frac{C'}{\log^{D'} n} \min\left\{k^{\frac{2}{3}(1-\gamma)(1+\alpha)}, k^{\frac{2}{3}(1-\gamma)(1+\frac{\beta}{2})}\right\} = \frac{C'}{\log^{D'} n} k^\gamma, \end{aligned} \tag{2.10}$$

since $\alpha \geq \beta/2$ and $\frac{2}{3}(1-\gamma)(1+\frac{\beta}{2}) = \gamma$. Finally, note that $l(I_i) = 6mx$, and therefore inequality (2.2) is satisfied. This completes the proof in Case 1.

Case 2. $\max_z |I_i(z)| \geq 3mx$.

Let z_0 be the minimal z such that $|I_i(z)| \geq 3mx$. First we show that $|I_i(z_0)|$ is in fact not much larger than $3mx$. To make this precise, we prove the following claim.

Claim 2.20. *If $z_0 > n^{1/10}$, then there is a constant C_3 depending only on ε such that*

$$|I_i(z_0)| < C_3 mx.$$

Proof. Minimality of z_0 implies that $|c(I_i(z_0 - 1))|$ and $C_2 \log k \cdot \psi(z_0 - 1, A_i')$ are at most $3mx$. By Claim 2.19,

$$|c(I_i(z_0))| \leq 3mx + |\Delta_i| + 5z_0 < C' mx,$$

where the second inequality holds since, by Theorem 2.3, we have $3mx \geq \psi(z_0 - 1, A_i') = \Omega\left(z_0^{3/2}\right)$ and, by Observation 2.7, we have $|\Delta_i| \leq 4mx$ (recall that we work only with typical vertices). Finally, note that $\psi(z_0, A_i')$ differs from $\psi(z_0 - 1, A_i')$ by at most z_0. Therefore,

$$|I_i(z_0)| = |c(I_i(z_0))| + C_2 \log k \cdot \psi(z_0, A_i') < C' mx + C'' \log k \cdot \psi(z_0 - 1, A_i') < C_3 mx.$$

□

From (2.8) it easily follows that

$$|I_i(z)| \leq |\Delta_i| z + \left|(a_i' + a_i)\binom{z}{2} - d_i^* z^2\right| + C_2 \log k \cdot \psi(z, A_i'), \tag{2.11}$$

and therefore we can split the proof into further subcases, depending on which of the three terms on the right-hand side of (2.11) is the 'dominant' term.

Case 2a. $C_2 \log k \cdot \psi(z_0, A_i') \geq mx$**.**

First note that $z_0 = \Omega\left(\sqrt{\frac{mx}{\log k}}\right)$, simply because $\psi(z,-) \leq \binom{z}{2}$. Claim 2.20 allows us to set $I_i = \omega(H_i) + (-C_3mx, C_3mx) \supseteq \omega(H_i) + I_i(z_0)$. Finally by Claims 2.17 and 2.18,

$$\begin{aligned} N_i \geq \phi_{\omega'}(z_0, G[A_i']) &\geq \frac{C_1 z_0}{\log^{D_1} n} \min\left\{(z_0)^\alpha, \left(\frac{\psi(z_0, A_i')}{z_0}\right)^\beta\right\} \\ &= \frac{C_1}{\log^{D_1} n} \min\left\{z_0^{(1+\alpha)}, \psi(z_0, A_i')^\beta z_0^{1-\beta}\right\} \\ &\geq \frac{C'}{\log^{D'} n} \min\left\{(mx)^{\frac{1+\alpha}{2}}, (mx)^\beta (mx)^{\frac{1-\beta}{2}}\right\} = \frac{C'}{\log^{D'} n}(mx)^{\frac{1+\alpha}{2}}, \end{aligned} \tag{2.12}$$

since $\alpha \leq \beta$. The length of I_i is $l(I_i) = 2C_3mx$, hence the inequality (2.2) is satisfied. This completes the proof in Case 2a.

Case 2b. $z_0 \geq \sqrt{mx/3}$ **(takes care of** $\left|(a_i' + a_i)\binom{z_0}{2} - d_i^* z_0^2\right| \geq mx$**).**

Note that $|\Delta_i| < \sqrt{mx \log k}$ or else the center of $I_i(z_1)$, where $z_1 = \sqrt{mx}/\log^{1/4} k < z_0$ would be at distance at least

$$|\Delta_i| z_1 - |(a_i' - a_i(z_1))\binom{z_1}{2} - d_i^* z_1^2| \geq mx \log^{1/4} k - O\left(\frac{mx}{\sqrt{\log k}}\right) > 3mx$$

from 0, contradicting the minimality of z_0. From (2.11) and the above bound on $|\Delta_i|$ it follows that $|I_i(\sqrt{mx}/\log k)| \leq C' \frac{mx}{\sqrt{\log k}} \leq 0.1mx$. In the sequel we will combine this simple observation with the following claim.

Claim. For $\sqrt{mx}/\log k \leq z < z_0$, the intervals $I_i(z)$ and $I_i(z+1)$ intersect.

By Claim 2.19, the distance δ between the centers of these two intervals is at most $|\Delta_i| + 5z$, and $\psi(z, A_i') \geq C' z^{3/2} \gg (mx)^{1/2+1/6} \gg |\Delta_i|$. Now we are done, since $l(I_i(z)) = 2C_2 \log k \cdot \psi(z, A_i')$.

The upper bound on $|I_i(\sqrt{mx}/\log k)|$, together with this Claim show that the family $\{I_i(z) : \sqrt{mx}/\log k \leq z \leq z_0\}$ covers an interval of length at least $2.9mx$ (either $[0.1mx, 3m]$ or $[-3m, -0.1mx]$). Also, Claim 2.20 shows that $I_i(z_0)$ (and by the choice of z_0 also all the other intervals $I_i(z)$ with $\sqrt{mx}/\log k \leq z \leq z_0$) is entirely contained in $-\omega(H_i) + I_i = (-C_3mx, C_3mx)$.

Again, by Claims 2.17 and 2.18, in each of the intervals $\omega(H_i) + I_i(z)$, where $\sqrt{mx}/\log k \leq z \leq z_0$, we can find at least

$$\phi_{\omega'}(z, G[A_i']) \geq \frac{C_1}{\log^{D_1} n} \min\left\{z^{1+\alpha}, \psi(z, A_i')^\beta z^{1-\beta}\right\} \geq \frac{C'}{\log^{D'} n} \min\left\{m^{\frac{1+\alpha}{2}}, l(I_i(z))^\beta m^{\frac{1-\beta}{2}}\right\}$$

weights. Lemma 2.15 ensures we can find a collection of disjoint $I_i(z)$'s, indexed by $z \in Z$, of total length at least $1.45m$. Hence

$$\begin{aligned}
N_i &\geq \sum_{z \in Z} \frac{C'}{\log^{D'} n} \min\left\{m^{\frac{1+\alpha}{2}}, l(I_i(z))^{\beta} m^{\frac{1-\beta}{2}}\right\} \geq \frac{C'}{\log^{D'} n} \min\left\{m^{\frac{1+\alpha}{2}}, \sum_{z \in Z} l(I_i(z))^{\beta} m^{\frac{1-\beta}{2}}\right\} \\
&\geq \frac{C'}{\log^{D'} n} \min\left\{m^{\frac{1+\alpha}{2}}, \left(\sum_{z \in Z} l(I_i(z))\right)^{\beta} m^{\frac{1-\beta}{2}}\right\} \geq \frac{C''}{\log^{D''} n} \min\left\{m^{\frac{1+\alpha}{2}}, m^{\beta} m^{\frac{1-\beta}{2}}\right\} \\
&= \frac{C''}{\log^{D''} n} m^{\frac{1+\alpha}{2}},
\end{aligned} \tag{2.13}$$

where the third inequality holds because $0 \leq \beta \leq 1$ and therefore $y \mapsto y^{\beta}$ is concave. Once again, $l(I_i) = 2C_3 mx$ and therefore inequality (2.2) is satisfied. This completes the proof in Case 2b.

Case 2c. $|\Delta_i| z_0 \geq mx$ **and** $\psi(z_0, A_i') \geq |\Delta_i|$**.**

We may assume that neither of the previous subcases holds, so $C_2 \log k \cdot \psi(z_0, A_i') < mx$ and $z_0 < \sqrt{mx/3}$, which implies $|\Delta_i| > \sqrt{3mx}$. If $|\Delta_i| z_0 > 8mx$, then by Claim 2.19,

$$|c(I_i(z_0 - 1))| > 8mx - |\Delta_i| - 5z \geq 8mx - 4mx - mx = 3mx,$$

contradicting the minimality of z_0. Hence $|\Delta_i| z_0 \leq 8mx$. Moreover, z_0 is not too small either, since

$$z_0^2 > \psi(z_0, A_i') \geq |\Delta_i| > \sqrt{3mx} = \sqrt{1500x \cdot \psi(k, G)/k} \geq C' k^{1/4}. \tag{2.14}$$

Before we proceed, we need the following claim.

Claim. For all z such that $z_0/30 \leq z \leq z_0$,

$$10^{-3} \psi(z_0, A_i') \leq \psi(z, A_i') \leq 48 \psi(z_0, A_i'). \tag{2.15}$$

The first inequality follows from a simple averaging argument (see [5, Observation 4]), which implies that

$$\psi(z, A_i') \geq \psi(z_0, A_i') \cdot \binom{z}{2} / \binom{z_0}{2}.$$

Finally, Lemma 6 in [5] yields that for every n-vertex graph G and all $0 < s < k < n/3$, we have $\psi(s, G) \leq 48 \psi(k, G)$. This implies the second inequality.

This Claim implies that $0 \leq \omega' \leq C' z^{1/2} \sqrt{\log k} \leq \log n \cdot \psi(z, A_i')/z$ for all z in the range $[z_0/30, z_0]$. Moreover, since (2.14) implies that $z_0/30 \geq n^{1/10}$, Claims 2.17 and 2.18 imply that in each interval $\omega(H_i) + I_i(z)$ we find at least

$$\begin{aligned}
\phi_{\omega'}(z, G[A_i']) &\geq \frac{C_1 z}{\log^{D_1} n} \min\left\{z^{\alpha}, \left(\frac{\psi(z, A_i')}{z}\right)^{\beta}\right\} = \frac{C_1}{\log^{D_1} n} \min\left\{z^{1+\alpha}, \psi(z, A_i')^{\beta} z^{1-\beta}\right\} \\
&\geq \frac{C'}{\log^{D'} n} \min\left\{l(I_i(z))^{\frac{1+\alpha}{2}}, l(I_i(z))^{\beta} l(I_i(z))^{\frac{1-\beta}{2}}\right\} = \frac{C'}{\log^{D''}} l(I_i(z))^{\frac{1+\alpha}{2}}
\end{aligned}$$

k-vertex induced subgraphs with different weights. Now, since $|\Delta_i|z_0 \le 8mx$,

$$|c(I_i(z_0/30))| \le |\Delta_i|\frac{z_0}{30} + 3\left(\frac{z_0}{30}\right)^2 \le \frac{4mx}{15} + \frac{mx}{900} \le \frac{mx}{2}.$$

Moreover, by Claim 2.19 and (2.15),

$$|c(I_i(z+1)) - c(I_i(z))| < |\Delta_i| + 5z \le 2\psi(z_0, A_i') \le C_2 \log k \cdot \psi(z, A_i') = l(I_i(z))/2,$$

so the intervals $I_i(z)$ and $I_i(z+1)$ intersect and hence the family $\{I_i'(z) : z_0/30 \le z \le z_0\}$ will cover an interval of length at least $2.5mx$. Lemma 2.15 ensures we can find a collection of disjoint $I_i(z)$'s, indexed by $z \in Z$, of total length not smaller than $1.25m$. By (2.11),

$$|I_i(z_0)| \le |\Delta_i|z_0 + z^2 + C_2 \log k \cdot \psi(z_0, A_i') < 8mx + mx + mx,$$

and therefore all intervals $I_i(z)$ in question are entirely contained in the interval $(-10mx, 10mx)$. Hence, if we set $I_i = \omega(H_i) + (-10mx, 10mx)$, we will have

$$N_i \ge \frac{C'}{\log^{D'} n} \sum_{z\in Z} l(I_i(z))^{\frac{1+\alpha}{2}} \ge \frac{C'}{\log^{D'} n} \left(\sum_{z\in Z} l(I_i(z))\right)^{\frac{1+\alpha}{2}} \ge \frac{C'}{\log^{D'} n} m^{\frac{1+\alpha}{2}}, \qquad (2.16)$$

where the second inequality follows by concavity of the function $y \mapsto y^{\frac{1+\alpha}{2}}$ (recall that $0 \le \alpha \le 1$). Finally, note that inequality (2.2) is satisfied. This completes the proof in Case 2c.

Case 2d. $|\Delta_i|z_0 \ge mx$ **and** $\psi(z_0, A_i') < |\Delta_i|$**.**

Recall that $t' = \varepsilon t/3$. This time we have to let z be a little larger. Define

$$z_1 = \min\left\{t', \min\{z : \psi(z, A_i') \ge |\Delta_i|\}\right\}.$$

There are two distinct cases to consider, depending on which value in the above minimum is smaller.

Case 2d-A. $z_1 = t'$ and $\psi(z, A_i') < |\Delta_i|$ for all $z \le z_1$.

First note that $z_1 \ll \psi(z_1, A_i') < |\Delta_i|$, so

$$c(I_i(z_1)) \ge |\Delta_i|z_1 - z_1^2 \ge 0.5|\Delta_i|z_1,$$

and

$$c(I_i(z_1/30)) \le |\Delta_i|\frac{z_1}{30} + \left(\frac{z_1}{30}\right)^2 \le 0.1|\Delta_i|z_1.$$

Claim. There are at least $C'z_1/\log k$ pairwise disjoint intervals among $\{I_i(z) : z_1/30 \le z \le z_1\}$.

Since $|\Delta_i|$ is larger than both z and $l(I_i(z))$, intuitively it is clear that whenever $z_2 - z_1$ is big enough, $I_i(z_1)$ and $I_i(z_2)$ are disjoint. Formally, by Claim 2.19,

$$|c(I_i(z_2)) - c(I_i(z_1))| \ge (z_2 - z_1) \cdot |\Delta_i| - 5z_2(z_2 - z_1) = (z_2 - z_1) \cdot (|\Delta_i| - 5z_2) \ge \frac{(z_2 - z_1)|\Delta_i|}{2},$$

and therefore whenever

$$z_2 - z_1 \geq 4C_2 \log k \geq \frac{l(I_i(z_1)) + l(I_i(z_2))}{|\Delta_i|},$$

the intervals $I_i(z_1)$ and $I_i(z_2)$ are disjoint.

For each z,

$$|I_i(z)| \leq |\Delta_i|z + z^2 + C_2 \log k \cdot \psi(z, A_i') < 2|\Delta_i|z,$$

so $(-2|\Delta_i|t', 2|\Delta_i|t')$ contains all the intervals $I_i(z)$ for $z_1/30 \leq z \leq z_1$. Finally, set $I_i = \omega(H_i) + (-2|\Delta_i|t', 2|\Delta_i|t')$. Since $z_1/30 \geq n^{1/10}$,

$$\begin{aligned} N_i &\geq \frac{C'}{\log k} t' \cdot \frac{C_1 t'}{\log^{D_1} n} \min\left\{ (t')^\alpha, \left(\frac{\psi(t', A_i')}{t'} \right)^\beta \right\} \\ &\geq t' \cdot \frac{C''}{\log^{D'} n} \min\left\{ k^{\frac{2}{3}(1-\gamma)(1+\alpha)}, k^{\frac{2}{3}(1-\gamma)(1+\frac{\beta}{2})} \right\} = t' \cdot \frac{C''}{\log^{D'} n} k^\gamma. \end{aligned} \tag{2.17}$$

Recall that we are exchanging only $V(H_i)$-typical vertices and therefore $|\Delta_i| \leq 4mx$. Hence $l(I_i) \leq 16mx \cdot t'$ and therefore inequality (2.2) is satisfied. That completes the proof in Case 2d-A.

Case 2d-B. $\psi(z_1, A_i') \geq |\Delta_i|$.

We can simply rewrite the proof of Case 2c here, replacing z_0 with z_1. The only change is that $I_i = \omega(H_i) + (-C'Mx, C'Mx)$, where $M = |\Delta_i| z_1$ and $|c(I_i'(\frac{z_1}{30}))| \leq 0.5Mx$, and in (2.16), $m^{\frac{1+\alpha}{2}}$ will be replaced by $M^{\frac{1+\alpha}{2}}$. Hence we consider Case 2d-B resolved.

To finish the proof, note that each time (see (2.10), (2.12), (2.13), (2.16), (2.17)) we were able to construct at least N_i graphs with different weights in the interval I_i such that the aforementioned inequality (2.2) holds:

$$\frac{N_i}{l(I_i)} \geq \frac{C'}{mx^{F'} \log^{D'} n} \min\left\{ k^\gamma, m^{\frac{1+\alpha}{2}} \right\}. \tag{2.2}$$

Moreover, for each i, the interval I_i, which is centered at $\omega(H_i)$, has length at least mx. Therefore these intervals cover the (disjoint!) family $\{\omega(H_i) + [-0.5mx, 0.5mx] \colon 1 \leq i \leq 10^{-8}k/x\}$. Hence

$$l\left(\bigcup_{i=1}^{10^{-8}k/x} I_i \right) \geq 0.5 \cdot 10^{-8} k \cdot m.$$

By Lemma 2.15, we can find a set J such that the intervals indexed by J are pairwise disjoint and the sum of their lengths is $C'km$. Hence the total number of different weights satisfies

$$\begin{aligned} \phi_\omega(k, G) &\geq C'km \cdot \min \frac{N_i}{l(I_i)} \geq \frac{C'k}{x^{F'} \log^{D'} n} \min\left\{ k^\gamma, m^{\frac{1+\alpha}{2}} \right\} \\ &= \frac{C''k}{x^{F'} \log^{D'} n} \min\left\{ k^{\frac{\beta+2}{\beta+5}}, \left(\frac{\psi(k,G)}{k} \right)^{\frac{1+\alpha}{2}} \right\} \end{aligned}$$

for some absolute constants C'', D', F'. This completes the proof.

Chapter 3

Large trees in random graphs

3.1 Introduction

3.1.1 The Erdős-Rényi random graph

Ever since its introduction in a paper of Gilbert [48] from 1959, the Erdős-Rényi random graph model has been one of the main objects of study of probabilistic combinatorics. It got this name in honor of the authors of one of the most important and influential papers on random graphs [32]. Given a positive integer n and a real number $p \in [0,1]$, the *Erdős-Rényi random graph* $G(n,p)$ is a random variable taking values in the set of all labeled graphs on the vertex set $[n]$. Most commonly, one describes the probability distribution of $G(n,p)$ by saying that each pair of elements of $[n]$ forms an edge in $G(n,p)$ independently with probability p. Usually, the questions considered in this model have the following generic form. Let $\mathcal{P}$ be some graph property, by which we mean a family of graphs that is closed under graph isomorphisms. Given a function $p : \mathbb{N} \to [0,1]$, determine if $G(n,p)$ satisfies $\mathcal{P}$ *asymptotically almost surely* (a.a.s. for short), i.e., if the probability that $G(n,p)$ satisfies $\mathcal{P}$ tends to 1 as n tends to infinity. The core meta-problem in the area of random graphs is the study of the *evolution* of $G(n,p)$, i.e., analyzing how $G(n,p)$ behaves with respect to certain graph properties as p traverses the interval $[0,1]$. This task is inherently connected with the problem of determining so-called *threshold functions*, i.e., optimal conditions on the edge probability function p that guarantee almost sure appearance (or almost sure non-appearance) of certain subgraphs in $G(n,p)$. Apart from substructures of fixed (independent of n) size, large trees seem to be one of the most natural classes of graphs, for which we can hope to find such thresholds.

3.1.2 Tree embeddings

A very well known folklore result on tree embedding states that every graph with minimum degree at least k contains all trees with at most k edges. The threshold for minimum degree is best possible, as illustrated by an arbitrarily large disjoint union of $(k+1)$-vertex complete graphs. A natural question arises – what additional assumptions on a graph can force it to contain certain trees? Extending a

The material presented in this chapter is joint work with József Balogh, Béla Csaba and Martin Pei. Part of it was originally published under the title *Large bounded degree trees in expanding graphs* in The Electronic Journal of Combinatorics, see [11]; part of it was accepted to Random Structures & Algorithms under the title *Local resilience of almost spanning trees in random graphs*, see [12].

path-embedding result of Pósa [63], Friedman and Pippenger [38] proved that all graphs satisfying certain expansion properties contain all small trees with bounded maximum degree.

Theorem 3.1 ([38])**.** *Let m and D be positive integers and let H be a non-empty graph. If every $X \subseteq V(H)$ with $|X| \leq 2m$ satisfies $|N_H(X)| \geq (D+1)|X|$, then H contains every tree with m vertices and maximum degree at most D.*

An apparent limitation of Theorem 3.1 is that it can be helpful in finding only relatively small trees. Namely, in a graph of order n, the size of the largest tree whose existence is guaranteed by Theorem 3.1 is only about $n/(2D+2)$, where D is the maximum degree of the tree. Building on the ideas developed by Friedman and Pippenger [38], Haxell [49] managed to overcome this problem by requiring different types of expansion for sets of different sizes. Haxell's very general result allows one to conclude that certain graphs contain trees that are almost spanning. Due to its somewhat technical nature, we defer its statement to Section 3.2.1,

It turns out that the following simple corollary of Haxell's result has a few very interesting and yet quite straightforward consequences.

Theorem 3.2 ([11])**.** *Let D, m, L, and M be positive integers such that $0 \leq L \leq 2Dm$. If H is a non-empty graph satisfying the following two conditions:*

1. *For every $X \subseteq V(H)$ with $0 < |X| \leq m$, $|N_H(X)| \geq D|X| + 1$,*
2. *For every $X \subseteq V(H)$ with $m < |X| \leq 2m$, $|N_H(X)| \geq D|X| + M$,*

then H contains every tree T with $M + L$ vertices and maximum degree at most D, provided that T has at least L leaves.

3.1.3 Almost spanning trees in random graphs

Even though the problem of existence of large trees in random graphs has been long studied, until a few years ago most of the results had focused on finding a very specific tree − a long path. Erdős conjectured that the random graph $G(n, c/n)$ almost surely contains a path of length at least $(1 - \varepsilon(c))n$, where $\varepsilon(c) \in (0, 1)$ if $c > 1$, and $\lim_{c\to\infty} \varepsilon(c) = 0$. This was first confirmed by Ajtai, Komlós, and Szemerédi [1]; then Frieze [39] proved that $\varepsilon(c) = (1+o(1))ce^{-c}$. Alon, Krivelevich, and Sudakov [6] extended this result from paths to all almost spanning trees with bounded maximum degree. They showed that edge probability p of order at least $(D^3 \log D \log^2(1/\varepsilon))/\varepsilon \cdot 1/n$ is enough to guarantee, a.a.s., the appearance of all trees with at most $(1 - \varepsilon)n$ vertices and maximum degree D in the random graph $G(n, p)$. Combining Theorem 3.2 with a few easy estimates on the expansion properties of the random graph $G(n, p)$, we obtain the following improvement of the above mentioned result of Alon et al.

Theorem 3.3 ([11])**.** *Let D be an integer with $D \geq 2$ and let $\varepsilon \in (0, 1/2)$. If*

$$c > \max\left\{1000D\log(20D), \frac{30D}{\varepsilon}\log\frac{4e}{\varepsilon}\right\},$$

then the random graph $G(n, c/n)$ a.a.s. contains every tree with maximum degree D and at most $(1 - \varepsilon)n$ vertices.

Recently, using Theorem 3.5 and a refinement of the piece-by-piece embedding method of Alon et al. [6], Pei [62] obtained an improvement of their result that is slightly weaker than Theorem 3.3. We would also like to remark that in [6] it is suggested that the statement of Theorem 3.3 is still not optimal – and the condition on the constant c could be weakened to '$c > c_0$', for some c_0 of order $D \log(1/\varepsilon)$.

3.1.4 Local resilience of tree-universality

For a family of graphs $\mathcal{H}$, we say that a graph G is *$\mathcal{H}$-universal* if G contains every member of $\mathcal{H}$ as a subgaph. The main result of Section 3.1.3, Theorem 3.3, states that for every D and ε, the random graph $G(n,p)$ is asymptotically almost surely $\mathcal{T}_{n,D,\varepsilon}$-universal, where $\mathcal{T}_{n,D,\varepsilon}$ denotes the family of all trees with maximum degree at most D and order at most $(1-\varepsilon)n$. A natural question arises – how resilient is $G(n,p)$ with respect to tree-universality? The following theorem gives an answer to that question. For the definition of local resilience, we refer the reader to Section 1.5.

Theorem 3.4 ([12])**.** *Let α and γ be positive constants, and assume that $D \geq 2$. There exists a constant C_0 (depending on α, γ, and D) such that for all p satisfying $p(n) \geq C_0/n$, the local resilience of $G(n,p)$ with respect to the property of containing all trees of order at most $(1-\alpha)n$ and maximum degree at most D is a.a.s. greater than $(1/2-\gamma)$.*

Note that the constant $(1/2-\gamma)$ in the statement of Theorem 3.4 is best possible, provided that $\alpha < 1/2$. Dellamonica, Kohayakawa, Marciniszyn, and Steger [23] proved that for every positive constant γ, if pn is a large enough constant, one can almost surely find an approximately even bipartition of the vertex set of the random graph $G(n,p)$ such that each vertex v has at most $(1/2+\gamma)\deg(v)$ neighbors in the other partite set. It follows that a.a.s. by deleting at most a $(1/2+\gamma)$-fraction of the edges incident to each vertex, one can turn the random graph $G(n,p)$ into a graph whose largest connected component has about $n/2$ vertices, and hence does not contain any tree with $(1-\alpha)n$ vertices.

A simple argument proves that the constant $(1/2-\gamma)$ is sharp if $D \geq 3$ and $\alpha < (D-2)/(2D-2)$. Recall that one can make an arbitrary graph bipartite by removing at most half the edges at each vertex. It is easy to check that in the random graph $G(n,p)$, where pn is a large enough constant, a.a.s. every bipartite graph obtained in such way has partite sets of approximately even size. Since for all sufficiently large n, there are trees with n vertices and maximum degree D, whose color classes have sizes differing by a factor arbitrarily close to $D-1$, it follows that a.a.s. after deleting at most half the edges at each vertex of $G(n,p)$, the remaining graph cannot contain all trees with maximum degree D of size greater than $D/(2D-2) \cdot n$.

Recall that if $p(n) \leq C/n$, then there is a positive constant α (depending on C) such that almost surely the size of the largest connected component of $G(n,p)$ does not exceed $(1-\alpha)n$. Moreover, if $D(n) \to \infty$ as $n \to \infty$, then a.a.s. $G(n,p)$ contains $o(n/D)$ vertices with degree at least D, and hence it cannot contain all trees with maximum degree D. Therefore, Theorem 3.4 is in a sense sharp.

3.2 Tools

3.2.1 Embedding trees in expanding graphs

As we briefly mentioned in Section 3.1.2, every graph having certain expansion properties must also contain all large trees with bounded degree. The following quantitative version of this statement, which we also already mentioned in Section 3.1.2, was proved by Haxell [49].

Theorem 3.5 ([49])**.** *Let T be a tree with t edges and maximum degree d. Let $\emptyset = T_0 \subseteq T_1 \subseteq \cdots \subseteq T_\ell \subseteq T$ be a sequence of subtrees of T such that T can be obtained by attaching new leaves to T_ℓ. Let $d = d_1 \geq \cdots \geq d_\ell$ be a sequence of integers such that for each i with $1 \leq i \leq \ell$ and each $v \in V(T)$ we have*

$$\delta_T(v) - \delta_{T_{i-1}}(v) \leq d_i,$$

where $\delta_S(v)$ denotes the degree of v in the subtree S (if $v \notin V(S)$, then we let $\delta_S(v) = 0$). Let $t_i = |E(T_i)|$. Suppose that $m \geq 1$ is an integer and H is a graph satisfying the following conditions

1. *For every subset $X \subseteq V(H)$ with $0 < |X| \leq m$, $|N_H(X)| \geq d|X| + 1$.*
2. *For every subset $X \subseteq V(H)$ with $m < |X| \leq 2m$, and for each $i \in \{1, \ldots, \ell\}$, $|N_H(X)| \geq d_i|X| + t_i + 1$.*
3. *For every subset $X \subseteq V(H)$ with $|X| = 2m + 1$, $|N_H(X)| \geq t + 1$.*

Then H contains T as a subgraph. Moreover, for any vertex x_0 of T_1 and any $y \in V(H)$, there exists an embedding f of T into H such that $f(x_0) = y$.

In the remainder of this chapter, we will not be requiring Theorem 3.5 in its full generality, but rather we will use its simplified version, Theorem 3.2. First, let us define the following class of expanding graphs.

Definition 3.6. Let α, ε, and b satisfy $b \geq 2$, $0 < \alpha < 1$, and $0 < \varepsilon < 1/b$. We say that an n-vertex graph G is an (ε, b, α)-expander if it possesses the following two properties.

1. Every subset $X \subseteq V(G)$ of size at most εn satisfies $|N_G(X)| \geq b|X|$.
2. Every subset $X \subseteq V(G)$ of size at least εn satisfies $|N_G(X)| \geq (1 - \alpha)n$.

As an immediate consequence of Theorem 3.2, we derive the following sufficient conditions on the expansion parameters ε, b, and α that guarantee that all (ε, b, α)-expanders contain every almost spanning tree with bounded maximum degree.

Corollary 3.7 ([11])**.** *Let D be an integer with $D \geq 2$ and let $\varepsilon \in (0, 1)$. Suppose that α and ε_0 are positive reals satisfying $2D\varepsilon_0 + \alpha \leq \varepsilon$. Then every n-vertex $(\varepsilon_0, D + 1, \alpha)$-expander contains all trees of order $(1 - \varepsilon)n$ and maximum degree D.*

Proof. Let G be an n-vertex $(\varepsilon_0, D + 1, \alpha)$-expander. It is straightforward to check that G satisfies assumptions of Theorem 3.2 with $m = \varepsilon_0 n$, $M = (1 - 2D\varepsilon_0 - \alpha)n$, and $L = 0$. Hence, G contains every tree with maximum degree D and order M. □

It is not hard to see that even the general Theorem 3.5 is not suitable when the target graph is bipartite, as it does not allow to embed trees of order greater than the number of vertices in the smaller partite set. Luckily, a natural and straightforward modification of the argument of Friedman and Pippenger [38] allows one to prove the following theorem. We defer its proof to the end of this section.

Theorem 3.8 ([12])**.** *Let D, m_1, M_1, m_2 and M_2 be positive integers. Assume that H is a non-empty bipartite graph with color classes V_1 and V_2, satisfying the following conditions.*

1. *For every $X \subseteq V_i$ with $0 < |X| \le m_i$, $|N_H(X)| \ge D|X| + 1$ for $i \in \{1, 2\}$.*
2. *For every $X \subseteq V_i$ with $m_i < |X| \le 2m_i$, $|N_H(X)| \ge D|X| + M_{3-i}$ for $i \in \{1, 2\}$.*

Furthermore, let T be a tree with maximum degree at most D and color classes of sizes M_1 and M_2, respectively, and let v be an arbitrary vertex of T belonging to the first color class. Then every mapping of v to a vertex in V_1 extends to an embedding of T in H.

With the aim of deriving a short statement along the lines of Corollary 3.7, let us define the following class of expanding biparite graphs.

Definition 3.9. Let $b \ge 2$ and let H be a bipartite graph with color classes V_1 and V_2, where $|V_1| \le |V_2|$. Let q be a positive integer with $q < |V_1|$. We will say that H is a *bipartite* (q, b)-*expander* if it possesses the following properties.

1. Every subset $X \subseteq V_i$ of size at most q satisfies $|N_H(X)| \ge b|X|$ for $i \in \{1, 2\}$.
2. Every subset $X \subseteq V_i$ of size at least q satisfies $|N_H(X)| \ge |V_{3-i}| - q$ for $i \in \{1, 2\}$.

Corollary 3.10 ([12])**.** *Let $D \ge 2$ and let H be a bipartite graph with color classes V_1 and V_2, where $|V_1| \le |V_2|$. Suppose that H is a bipartite $(q, D+1)$-expander for some $0 < q < |V_1|/(2D+1)$. Then H contains all trees with maximum degree at most D and color classes of sizes at most $|V_1| - (2D+1)q$ and $|V_2| - (2D+1)q$, respectively. Furthermore, any such tree can be embedded even if we require that a particular vertex of the tree is mapped to a particular vertex of H, as long as this mapping respects the color classes.*

Proof. It is straightforward to check that H satisfies the assumptions of Theorem 3.8 with $m_i = q \ge 1$ and $M_i = |V_i| - (2D+1)q$ for $i \in \{1, 2\}$. □

We close this section with the proof of Theorem 3.8.

Proof of Theorem 3.8. A tree T will be called (M_1, M_2, D)-*small*, or simply *small*, if $\Delta(T) \le D$ and the two color classes of T, the sets $U_1, U_2 \subseteq V(T)$, satisfy $|U_1| \le M_1$ and $|U_2| \le M_2$. Using induction on the size of T, we will prove that H contains all small trees. First, we need a few definitions. Let f be an embedding of some small tree T into our expanding graph H. The *liability* $B_f(x)$ of a vertex $x \in V(H)$ with respect to f is defined by

$$B_f(x) = \begin{cases} D - \deg_T(v), & x = f(v) \text{ for some } v \in V(T), \\ D, & x \notin f(V(T)). \end{cases}$$

We define the *assets* $\mathcal{A}_f(X)$ of a set $X \subseteq V(H)$ to be the set of neighbors of X that are not used in the embedding, i.e., $\mathcal{A}_f(X) = N_H(X) - f(V(T))$. For every set $X \subseteq V(H)$, define $A_f(X) = |\mathcal{A}_f(X)|$ and $B_f(X) = \sum_{x \in X} B_f(x)$. The quantity $A_f(X) - B_f(X)$ will be called the *balance* of the set X, denoted $C_f(X)$ .Finally, an embedding f of a small tree T into our graph H will be called *good* if it maps U_1 into V_1 and U_2 into V_2, and moreover, every set $X \subseteq V_1$ of size at most $2m_1$ and every set $X \subseteq V_2$ of size at most $2m_2$ have non-negative balance with respect to f.

In order to prove the existence of a good embedding of an arbitrary small tree into our graph H, it clearly suffices to show that the class of good embeddings satisfies the following two properties.

Property 1. Every embedding of a single-vertex tree into V_1 is good.

Property 2. If T is a small tree and S is its subtree obtained by deleting a leaf and the edge incident to it, then any good embedding of S in H can be extended to a good embedding of T in H.

To prove Property 1, suppose that T is a tree consisting of a single vertex. Let f be an arbitrary embedding of T in V_1. We show that f is good. Fix an arbitrary $i \in \{1, 2\}$ and suppose that $X \subseteq V_i$ and $|X| \leq 2m_i$. We can easily assume that $X \neq \emptyset$, since $C_f(\emptyset) = 0$. We have

$$A_f(X) = |N_H(X) - f(V(T))| \geq |N_H(X)| - 1 \geq D|X| + 1 - 1 = B_f(X).$$

Thus every set $X \subseteq V_i$ of size at most $2m_i$ has non-negative balance, and so Property 1 is satisfied.

To prove Property 2, assume that f is a good embedding of a tree S, obtained from a small tree T by removing a leaf v. Without loss of generality we may assume that $v \in U_2$, as the proof for the case $v \in U_1$ is identical. For the sake of brevity, let $U_2' = U_2 - \{v\}$.

Claim 3.11. *If for some $X \subseteq V_1$ with $|X| \leq 2m_1$, $C_f(X) = 0$, then $|X| \leq m_1$.*

Proof of Claim 3.11. Assume that some $X \subseteq V_1$ satisfies $C_f(X) = 0$, but $m_1 < |X| \leq 2m_1$. Since T is a small tree, $|U_2'| = |U_2| - 1 \leq M_2 - 1$, and thus

$$\begin{aligned} A_f(X) &= |N_H(X) - f(V(S))| = |N_H(X) - f(U_2')| \geq |N_H(X)| - |f(U_2')| \\ &\geq D|X| + M_2 - (M_2 - 1) = D|X| + 1 \geq B_f(X) + 1, \end{aligned}$$

where the second equality follows from the fact that $N_H(X) \subseteq V_2$ and f maps U_1 to V_1. This contradicts the assumption that $A_f(X) = B_f(X)$. □

Claim 3.12. *If some $X, Y \subseteq V_1$, with $|X| \leq |Y| \leq 2m_1$, satisfy $C_f(X) = C_f(Y) = 0$, then also $C_f(X \cup Y) = 0$ and $|X \cup Y| \leq m_1$.*

Proof of Claim 3.12. Since B_f is a measure on $V(H)$, clearly

$$B_f(X \cup Y) + B_f(X \cap Y) = B_f(X) + B_f(Y).$$

Moreover, since $\mathcal{A}_f(X \cup Y) = \mathcal{A}_f(X) \cup \mathcal{A}_f(Y)$ and $\mathcal{A}_f(X \cap Y) \subseteq \mathcal{A}_f(X) \cap \mathcal{A}_f(Y)$, we have

$$A_f(X \cup Y) + A_f(X \cap Y) \leq |\mathcal{A}_f(X) \cup \mathcal{A}_f(Y)| + |\mathcal{A}_f(X) \cap \mathcal{A}_f(Y)| = A_f(X) + A_f(Y).$$

It follows that $C_f(X \cup Y) \leq C_f(X) + C_f(Y) - C_f(X \cap Y) = -C_f(X \cap Y)$. Since $|X \cap Y| \leq |Y| \leq 2m_1$ and f is good, $C_f(X \cap Y) \geq 0$, and hence $C_f(X \cup Y) \leq 0$. By Claim 3.11, $|X|, |Y| \leq m_1$, and so $|X \cup Y| \leq 2m_1$. Hence, also $C_f(X \cup Y) \geq 0$, and by Claim 3.11, $|X \cup Y| \leq m_1$. □

Corollary 3.13. *Suppose $X_1, \ldots, X_k \subseteq V_1$ are sets of size at most $2m_1$ having zero balance. Then $C_f(X_1 \cup \cdots \cup X_k) = 0$ and $|X_1 \cup \cdots \cup X_k| \leq m_1$.*

Let w be the only neighbor of v in T. Since $v \in U_2$, clearly $w \in U_1$. Recall that f was a good embedding of S in H and hence $f(w) \in V_1$. Let $Y = \mathcal{A}_f(f(w))$ and note that $Y \subseteq V_2$. We can extend f to an embedding of T by mapping v to any vertex in Y. Suppose that for no $y \in Y$, the extension f_y, defined by

$$f_y(x) = \begin{cases} y & \text{if } x = v, \\ f(x) & \text{if } x \neq v, \end{cases}$$

is good. Since clearly f_y maps U_1 to V_1 and U_2 to V_2, this means that for every $y \in Y$, there is an $i(y) \in \{1, 2\}$ and a set $X_y \subseteq V_{i(y)}$ of size at most $2m_{i(y)}$ with $C_{f_y}(X_y) < 0$. Clearly, for all $y \in Y$, $i(y) = 1$, since for every $X \subseteq V_2$, $\mathcal{A}_f(X) = \mathcal{A}_{f_y}(X)$ and $B_{f_y}(X) \leq B_f(X)$. Moreover, for each $y \in Y$, we must have $y \in \mathcal{A}_f(X_y)$, $f(w) \notin X_y$ and $C_f(X_y) = 0$, or otherwise $C_{f_y}(X_y) \geq 0$. Let $X^* = \bigcup_{y \in Y} X_y$. By Corollary 3.13, $C_f(X^*) = 0$ and $|X^*| \leq m_1$. Moreover, if we let $X' = X^* \cup \{f(w)\} \subseteq V_1$, then

$$\mathcal{A}_f(X') = \mathcal{A}_f(X^*) \cup \mathcal{A}_f(f(w)) = \mathcal{A}_f(X^*) \cup Y = \mathcal{A}_f(X^*),$$

since $\mathcal{A}_f(X^*) = \cup_{y \in Y} \mathcal{A}_f(X_y)$ and for all $y \in Y$, $y \in \mathcal{A}_f(X_y)$. Also, since $f(w) \notin X^*$ and $\deg_S(w) = \deg_T(w) - 1 \leq D - 1$,

$$B_f(X') = B_f(X^*) + B_f(f(w)) \geq B_f(X^*) + 1.$$

This implies that $C_f(X') < 0$, which is a clear contradiction, since $|X'| \leq m_1 + 1 \leq 2m_1$ and f was good. Hence, f can be extended to a good embedding of T, and so Property 2 holds. □

3.2.2 Uniformity and expansion in random graphs

In this section, we collect a few straightforward facts about uniformity and expansion properties of the random graph $G(n, p)$.

Lemma 3.14. *Let η be a positive constant. If $pn > \frac{8}{\eta^4(1-\eta)}$, then a.a.s. the random graph $G(n,p)$ is (η, p)-uniform.*

Proof. Let $G = G(n, p)$. The probability that a fixed pair of sets A and B violates the η-uniformity condition (1.3) is

$$P\big(e_G(A,B) - p|A||B| > \eta p|A||B|\big) + P\big(e_G(A,B) - p|A||B| < -\eta p|A||B|\big).$$

By Theorem 1.2, this is at most

$$\exp\left(-\frac{(\eta p|A||B|)^2}{2p|A||B|}\right) + \exp\left(-\frac{(\eta p|A||B|)^2}{2p|A||B|} + \frac{(\eta p|A||B|)^3}{2(p|A||B|)^2}\right) \leq 2\exp\left(-\frac{\eta^4(1-\eta)pn^2}{2}\right),$$

where the inequality holds since we assumed that $|A|, |B| \geq \eta n$. Similarly, the probability that a fixed set A of size at least ηn violates the η-uniformity condition (1.4) is at most

$$\exp\left(-\frac{(\eta p\binom{|A|}{2})^2}{2p\binom{|A|}{2}}\right) + \exp\left(-\frac{(\eta p\binom{|A|}{2})^2}{2p\binom{|A|}{2}} + \frac{(\eta p\binom{|A|}{2})^3}{2(p\binom{|A|}{2})^2}\right) \leq 2\exp\left(-\frac{\eta^4(1-\eta)pn^2}{8}\right).$$

By our assumption on p and the union bound, the probability that G is not η-uniform is at most $2^n \cdot 2^n \cdot (2e^{-4n}) + 2^n \cdot (2e^{-n}) = o(1)$. □

Lemma 3.15. *Let β and γ be positive constants satisfying $\beta \leq \gamma \leq 1/2$. If $c \geq \frac{3}{\beta}\log\frac{e}{\gamma}$, then a.a.s. the random graph $G(n, c/n)$ does not contain two disjoint sets B, C of sizes at least βn and γn, respectively, such that $e(B, C) = 0$.*

Proof. If $G(n, c/n)$ contains two sets B and C as in the statement of this lemma, clearly we can also find two disjoint sets B' and C' of sizes exactly βn and γn, with $e(B', C') = 0$. The probability that such a pair exists is at most

$$\binom{n}{\beta n}\binom{n}{\gamma n} \cdot \left(1 - \frac{c}{n}\right)^{\beta\gamma n^2} \leq \binom{n}{\gamma n}^2 \cdot e^{-c\beta\gamma n} \leq \left(\frac{en}{\gamma n}\right)^{2\gamma n} \cdot \left(\frac{e}{\gamma}\right)^{-3\gamma n} = o(1).$$

□

Lemma 3.16. *Let β and γ be positive constants satisfying $\beta \leq \gamma \leq 1/2$. If $c \geq \frac{6\gamma}{\beta}\log\frac{e}{\gamma}$, then a.a.s. $G(n, c/n)$ does not contain a pair of disjoint sets B and C of sizes at least βn and at least $(1-\gamma)n$, with $e(B, C) = 0$.*

Proof. As in the proof of Lemma 3.15, we only need to show that a.a.s. there is no such pair with sizes exactly βn and $(1-\gamma)n$. The probability that such a pair exists is at most

$$\binom{n}{\beta n}\binom{n}{(1-\gamma)n}\left(1 - \frac{c}{n}\right)^{\beta(1-\gamma)n^2} \leq \binom{n}{\gamma n}^2 \cdot e^{-c\beta n/2} \leq \left(\frac{en}{\gamma n}\right)^{2\gamma n} \cdot \left(\frac{e}{\gamma}\right)^{-3\gamma n} = o(1).$$

□

Lemma 3.17. *If $k \geq 2$ and $c \geq 10k\log_2 k$, then almost surely every subset A of at most $n/(ek)$ vertices in the random graph $G(n, c/n)$ spans less than $c|A|/k$ edges.*

Proof. Certainly, if a subset A of size a violates the assertion, then $a \geq c/k$. The probability that there is a bad subset A of size a, with $c/k \leq a \leq n/ek$, is at most

$$\begin{aligned}\binom{n}{a}\binom{a^2/2}{ac/k} \cdot \left(\frac{c}{n}\right)^{ca/k} &\leq \left(\frac{en}{a}\right)^a \cdot \left(\frac{ea^2}{2} \cdot \frac{k}{ac}\right)^{ca/k} \cdot \left(\frac{c}{n}\right)^{ca/k} \\ &= \left(\frac{en}{a} \cdot \left(\frac{eka}{2n}\right)^{c/k}\right)^a \leq \left(\frac{(ek/2)^{c/k+1}}{(n/a)^{c/k-1}}\right)^a.\end{aligned} \tag{3.1}$$

If $\sqrt{n} \leq a \leq n/ek$, then

$$\frac{(ek/2)^{c/k+1}}{(n/a)^{c/k-1}} \leq \left(\frac{1}{2}\right)^{c/k} \cdot (ek)^2 \leq k^{-10} \cdot (ek)^2 \leq \frac{1}{2},$$

and consequently (3.1) is bounded by $2^{-\sqrt{n}}$. In the case $c/k \le a \le \sqrt{n}$, the right-hand side of (3.1) can be further estimated as follows

$$\left(\frac{(ek/2)^{c/k+1}}{(n/a)^{c/k-1}}\right)^{a} \le \left(\frac{(ek/2)^{11}}{(\sqrt{n})^{9}}\right)^{10} = o(n^{-1}).$$

Summing these estimates over all values of a yields the desired result. □

Lemma 3.18. *Let $\rho \in (0, 1/2)$. If $c > 64 \log \frac{e}{\rho}$, then a.a.s. the random graph $G(n, c/n)$ contains an induced subgraph G' with at least $(1-\rho)n$ vertices and minimum degree at least $c/4$.*

Proof. Let $G = G(n, c/n)$. While G contains a vertex with degree less than $c/4$, delete this vertex. Denote the remaining induced subgraph of G by G'. If G' has at least $(1-\rho)n$ vertices, we have found the subgraph we were looking for. It suffices to show that the probability of G' having less than $(1-\rho)n$ vertices is small. First observe that if we were forced to delete more than ρn vertices, then the original graph G contained a set A of size ρn such that $e_A = e(A, V(G) - A) < \rho cn/4$. Note that $\mathbb{E}[e_A] = \rho(1-\rho)cn \ge \rho cn/2$. By Theorem 1.2, the probability of this event in our random graph satisfies

$$P\big(e_A < c\rho n/4\big) \le P\big(e_A - \mathbb{E}[e_A] < -\rho cn/4\big) \le e^{-\frac{(\rho cn/4)^2}{2\rho cn}} = e^{-\rho cn/32}.$$

Hence the probability that such a set A exists in our graph G is bounded by

$$\binom{n}{\rho n} \cdot e^{-\rho cn/32} \le \left(\frac{en}{\rho n}\right)^{\rho n} \cdot \left(\frac{e}{\rho}\right)^{-2\rho n} = o(1).$$

□

Theorem 3.19 ([11])**.** *Let α, ε, ρ, and b satisfy $b \ge 2$, $0 < \rho \le \varepsilon \le \alpha < 1/2$, and $\varepsilon < 1/(2b+4)$. If*

$$c > \max\left\{500b\log(12b), \frac{6}{\varepsilon}\log\frac{2e}{\alpha}, 64\log\frac{e}{\rho}\right\},$$

then a.a.s. the random graph $G(n, c/n)$ contains an induced subgraph G' of order at least $(1-\rho)n$ that is an (ε, b, α)-expander.

Proof. By Lemma 3.18, almost surely $G(n, c/n)$ contains an induced subgraph G' of order n', with $n' \ge (1-\rho)n$ and $\delta(G') \ge c/4$. Conditioning on that event, we will show that G' is almost surely an (ε, b, α)-expander.

Suppose that G' fails to possess property 1 from Definition 3.6. Then there is a set $X \subseteq V(G')$ of size t, with $t \le \varepsilon n'$ and $|N_{G'}(X)| \le bt$. Let $A = X \cup N_{G'}(X)$. Clearly, $|A| \le (b+1)t$. We consider three cases, depending on the order of t.

Case 1. $t \le \frac{n}{8e(b+1)^2}$.
Let $k = 8(b+1)$. Since edges incident to vertices in X are contained in A, $e(A) \ge \delta(G')|X|/2 \ge ct/8 \ge c|A|/k$. By our assumptions, $|A| \le n/(ek)$ and $c > 10k\log_2 k$. By Lemma 3.17, such a non-expanding set X a.a.s. does not exist.

Case 2. $\frac{n}{8e(b+1)^2} \le t \le \frac{n}{20e(b+1)}$.

Since G' is an induced subgraph, in G there are no edges between X and $Y = V(G') - A$. By our assumptions on t and ε, the latter set has at least

$$n' - |A| \geq (1-\rho)n - (b+1)t \geq n - n/(b+1) - (b+1)t \geq n - (8e+1)(b+1)t$$

vertices. Let $\beta = t/n$ and $\gamma = (8e+1)(b+1)\beta$. By our assumption on t, we have that $\beta \geq \frac{1}{8e(b+1)^2}$ and consequently $e/\gamma < 12b$. Moreover, note that $6\gamma/\beta < 500b$. It follows that $c > 6\frac{\gamma}{\beta}\log\frac{e}{\gamma}$ and, by Lemma 3.16, such non-expanding set X a.a.s. does not exist.

Case 3. $\frac{n}{20e(b+1)} \leq t \leq \varepsilon n'$.
Again, in G there are no edges between X and $V(G') - A$. By our assumptions on t and ε, the latter set has at least

$$n' - |A| \geq (1-(b+1)\varepsilon)n' \geq (1-(b+1)\varepsilon)(1-\varepsilon)n \geq (1-(b+2)\varepsilon)\, n \geq \frac{n}{2}$$

vertices. Let $\beta = \frac{1}{20e(b+1)}$ and $\gamma = 1/2$. Clearly $c > (3/\beta)\log(e/\gamma)$. By Lemma 3.15, such non-expanding set X a.a.s. does not exist.

Hence, a.a.s. the graph G' satisfies property 1 from definition 3.6. Finally, suppose that G' fails to possess the other property. Then there is a set X of size exactly $\varepsilon n'$ with $|N_{G'}(X)| \leq (1-\alpha)n'$. It follows that in G there are no edges between X and $V(G') - X - N_G(X)$. Clearly, the latter set contains at least $\alpha n/2$ vertices. Let $\beta = \varepsilon/2$ and $\gamma = \alpha/2$. Since $c > (3/\beta)\log\frac{e}{\gamma}$, by Lemma 3.15, this a.a.s. does not happen. □

3.3 Proof of Theorem 3.3

Since most of the work was done in Section 3.2, the proof of Theorem 3.3 now fits into a single paragraph.

Proof of Theorem 3.3. Let $\varepsilon_0 = \frac{\varepsilon}{4d+2}$. By Theorem 3.19, with $\alpha = \varepsilon/2$, $b = D+1$, $\rho = \varepsilon_0$ and $\varepsilon = \varepsilon_0$, $G(n, c/n)$ almost surely contains a subgraph G' of order at least $(1-\varepsilon_0)n$ that is an $(\varepsilon_0, D+1, \varepsilon/2)$-expander. By Corollary 3.7, G' contains every tree with maximum degree D and order at most M, where

$$M = (1 - 2D\varepsilon_0 - \varepsilon/2)|V(G')| \geq (1-(4D+1)\varepsilon_0)\cdot(1-\varepsilon_0)n \geq (1-\varepsilon)n.$$

□

3.4 Proof of Theorem 3.4

Theorem 3.4 is a direct consequence of Lemma 3.14 and the following more general Theorem 3.20.

Theorem 3.20 ([12])**.** *Let α and γ be positive constants, and assume that $D \geq 2$. There exist η_0 and n_0 (both depending on α, γ, and D) such that the following holds. Let G be an n-vertex (η, p)-uniform graph, with $p > 0$, $\eta < \eta_0$ and $n \geq n_0$. Let G' be a subgraph of G such that $\deg_{G'}(v) \geq (1/2+\gamma)\deg_G(v)$ for each vertex v. Then G' contains all trees with at most $(1-\alpha)n$ vertices and maximum degree at most D.*

3.4.1 Setup

We start by defining some constants. Let

$$\delta = \frac{\alpha}{16D^2}, \quad \varepsilon = \min\left\{\frac{\alpha}{64D^3}, \frac{\alpha\gamma\delta}{96}, \frac{\gamma^2}{36}\right\}, \text{ and } \eta_0 = \min\left\{\frac{\varepsilon}{2}, \frac{1}{2K_0}\right\},$$

where K_0 is given by Lemma 1.1 with $k_0 = \lceil 1/\varepsilon \rceil$. Let G be an (η, p)-uniform n-vertex graph, with $p > 0$, $\eta < \eta_0$ and n larger than some constant n_0 depending on α, γ, and D.[1] We let our adversary remove edges from G, so that no more than $(1/2 - \gamma)\deg_G(v)$ edges incident to every vertex $v \in V(G)$ are deleted. Denote the leftover graph by G'. Clearly, G' is η-upper-uniform with density p. Finally, let T be a tree with at most $(1 - \alpha)n$ vertices and maximum degree at most D; without loss of generality we may also assume that T has at least $n/2$ vertices. We will show that $T \subseteq G'$.

3.4.2 Proof outline

In Section 3.4.3, we apply Szemerédi's Regularity Lemma to G', and show that the cluster graph, whose edges are regular pairs with density bounded away from zero, contains an almost spanning subgraph H'' with minimum degree slightly larger than $|V(H'')|/2$. Such large minimum degree guarantees the existence in H'' of an almost perfect matching M. The tree T will be embedded into G'' – the subgraph of G' induced by the union of the clusters in H''; moreover, most edges of T will be mapped to edges inside the dense, regular pairs in G'' that appear in M.

In Section 3.4.4, we partition the tree T into a bounded number of small subtrees in such a way that none of these subtrees is adjacent to more than D^3 others and every subtree contains all the children of its root.

In Section 3.4.5, the vertex set of G'' is partitioned into linear-sized subsets, which are then assigned to subtrees from our partition of T and the edges of T joining those subtrees. Each subtree S is assinged two subsets of the opposite ends of some edge in M, one for each color class of S; both subsets are slightly larger than the color class of S they are assigned to. An edge e joining two subtrees S and S' is assigned a small subset (a 'connecting' set) of a cluster that is adjacent (in H'') to the two clusters that were assigned to the color classes of S and S' that contain the endpoints of e. In Section 3.4.6, we trim all these subsets so that the pair assigned to every subtree is a bipartite expander, and every 'connecting' set has many neighbors in both sets it 'connects'.

Finally, in Section 3.4.7, we embed T in G'' in a top-down fashion. The subtree containing the root of T is embedded into the pair of sets assigned to it. For every other subtree, its root is mapped to an appropriate 'connecting' set and the remainder of that subtree is embedded into its pair of sets that, as we arranged before, induces a bipartite expander in G''.

3.4.3 Preparing G'

Since G' is (η, p)-upper-uniform and n is large, we may apply Szemerédi's regularity lemma (Lemma 1.1) with ε as above and $k_0 = \lceil 1/\varepsilon \rceil$. Let $(V_0, \ldots, V_k)$ be the resulting (ε, p)-regular partition of $V(G')$,

[1] Although we do not give a particular value of n_0, the existence of such a constant will become clear from the proof. The lower bound on n_0 comes mainly from the fact that we apply Szemerédi's regularity lemma to G; additional requirements on the largeness of n_0 are discussed in a footnote at the end of Section 3.4.5.

and recall that $k \le K_0$ by the definition of K_0. Define an auxiliary graph H' on the vertex set $\{V_1, \dots, V_k\}$ as follows. For all i and j with $1 \le i < j \le k$, the pair $\{V_i, V_j\}$ will be an edge in H' if and only if the p-density of the pair (V_j, V_j) in G' is at least $\gamma/6$.

Claim 3.21. *The minimum degree in H' is at least $(1/2 + 2\gamma/3)k$.*

Proof. Fix some $i \in \{1, \dots, k\}$ and let $V = V(G')$. Since G was η-uniform with density p, $|V_i| \ge n/2k > \eta n$ and $1/k \le \varepsilon$, then

$$\begin{aligned} e_G(V_i, V - V_0 - V_i) &\ge (1-\eta)p|V_i|(n - |V_0| - |V_i|) \ge (1-\eta)p|V_i|(n - \varepsilon n - n/k) \\ &\ge (1 - \eta - \varepsilon - 1/k)pn|V_i| \ge (1 - 3\varepsilon)pn|V_i|, \end{aligned}$$

and

$$\begin{aligned} \sum_{v \in V_i} \deg_G(v) = e_G(V_i, V - V_i) + 2e_G(V_i) &\le (1+\eta)p\left[|V_i|(n - |V_i|) + 2\binom{|V_i|}{2}\right] \\ &\le (1+\eta)p|V_i|(n - |V_i| + |V_i| - 1) \le (1+\varepsilon)pn|V_i|. \end{aligned}$$

Since our adversary deleted at most $(1/2 - \gamma)\deg_G(v)$ edges at every vertex v, the number of edges of G' that leave the set V_i can be bounded as follows:

$$\begin{aligned} e_{G'}(V_i, V - V_0 - V_i) &\ge e_G(V_i, V - V_0 - V_i) - (1/2 - \gamma) \cdot \sum_{v \in V_i} \deg_G(v) \\ &\ge (1/2 + \gamma - 7\varepsilon/2)pn|V_i|. \end{aligned}$$

Recall that $i \in \{1, \dots, k\}$ is fixed. The total number of edges in all pairs (V_i, V_j) with $j \in \{1, \dots, k\} - \{i\}$ whose density is smaller than $\gamma/6$ is at most $\gamma/6 \cdot pn|V_i|$. Moreover, the η-upper-uniformity of G' implies that $e_{G'}(V_i, V_j) \le (1+\eta)p|V_i||V_j| \le (1+\varepsilon)p(n/k)|V_i|$ for all $j \ne i$. Therefore,

$$\delta(H') \ge (1/2 + 5\gamma/6 - 7\varepsilon/2)(1+\varepsilon)^{-1}k \ge (1/2 + 5\gamma/6 - 5\varepsilon)k \ge (1/2 + 2\gamma/3)k.$$

□

Now, delete from H' all edges that correspond to pairs (V_i, V_j) that are not (ε, p)-regular in G' and let H'' be the subgraph of H' induced by the set of vertices whose degree in H' after that deletion exceeds $(1/2 + \gamma/2)k$.

Claim 3.22. *The graph H'' has at least $(1 - \alpha/8)k$ vertices and $\delta(H'') \ge (1/2 + \gamma/3)k$.*

Proof. Since H' contains at most $\varepsilon\binom{k}{2}$ edges corresponding to non-(ε, p)-regular pairs, their deletion lowers the degree sum of H' by no more than εk^2. Since $\delta(H') \ge (1/2 + 2\gamma/3)k$, the degree of at most $(6\varepsilon/\gamma)k$ vertices will fall below the $(1/2 + \gamma/2)k$ threshold after the deletion. Recall that $\varepsilon \le \min\{\gamma^2/36, \alpha\gamma/96\}$, and thus H'' will have at least $(1 - \alpha/8)k$ vertices, and its minimum degree will satisfy $\delta(H'') \ge (1/2 + \gamma/2)k - (6\varepsilon/\gamma)k \ge (1/2 + \gamma/3)k$. □

Let $k' = |V(H'')|$ and let $m' = \lfloor k'/2 \rfloor$. Since $\delta(H'') > k'/2$, H'' contains a matching of size m'. Fix any such matching M and denote its edges by $\{A_1, B_1\}, \dots \{A_{m'}, B_{m'}\}$. Finally, let G''

denote the subgraph of G' induced by the union of all vertices of H'' (which are clusters in G'). Let $n' = |V(G'')|$ and note that $n' \geq (1 - \alpha/8)(1 - \varepsilon)n \geq (1 - \alpha/4)n$.

3.4.4 Partitioning the tree

Every partition of the vertex set of a tree into connected subsets gives rise to a natural tree structure on the set of parts. Namely, we make two parts adjacent if and only if the subtrees they induce in the original tree are joined by an edge. Let us call this tree the *cluster tree* of our partition. The following general lemma will be crucial in the remainder of the proof.

Lemma 3.23. *Let t and D be positive integers with $D \geq 2$. Let T be a rooted tree with t vertices and maximum degree at most D. If $\beta \geq 1/t$, then there exists a partition of $V(T)$ into at most $4/\beta$ rooted subtrees of size at most $D^2\beta t$ each such that the maximum degree of the corresponding cluster tree does not exceed D^3 and all children of the root of each subtree belong to the same subtree (the subtree containing that root).*

The proof of Lemma 3.23 will make use of the following simple statement, whose proof is a straightforward modification of the proof of Proposition 4.2 from [6].

Proposition 3.24. *Let s and D be positive integers with $D \geq 2$. Let T be a tree with maximum degree at most D and S be a subset of $V(T)$ containing at least $s+1$ vertices. Then there exists an edge $e \in E(T)$ such that at least one of the two trees obtained from T by deleting e contains at least s and at most $(D-1)(s-1)+1$ vertices from S.*

Proof of Lemma 3.23. We will construct the required partition in three stages. The first stage will guarantee that the subtrees in our partition are not too large, i.e., they contain no more than $D\beta t$ vertices each. In the second stage we will refine the partition to reduce the maximum degree in the cluster tree to at most D^2. In the third stage we will merge some subtrees to guarantee that each root has children only in its own subtree, and we will do it in such a way that neither the upper bound on the sizes of the subtrees nor the maximum degree of the cluster tree grow more than by a factor of D, and hence in the end they are bounded by $D^2\beta t$ and D^3, respectively.

Stage 1. Start with the trivial partition of $V(T)$ into a single set. We will keep refining it until all parts are small enough, making sure that at all times at most one of the parts is larger than $D\beta t$ and at most one of the parts is smaller than βt. Clearly, our initial partition has that property. Suppose that our partition contains a subtree T' with more than $D\beta t$ vertices. Proposition 3.24 guarantees that T' contains an edge that splits it into two trees, one of which has at least βt and at most $(D-1)(\lceil \beta t \rceil - 1) + 1 \leq D\beta t$ vertices. We refine our partition by replacing T' with these two trees. Finally, we iterate this procedure until all parts have at most $D\beta t$ vertices and all but at most one has at least βt vertices. Denote that partition by Π. Clearly, the number of parts is at most $1/\beta + 1$.

Stage 2. Let T_Π be the cluster tree corresponding to the partition Π and let

$$E(\Pi) = \sum_{V \in \Pi} \max\{0, \deg_{T_\Pi}(V) - D^2\}.$$

Clearly,

$$0 \leq E(\Pi) \leq \sum_{V \in \Pi} \deg_{T_\Pi}(V) = 2(r-1) \leq 2/\beta.$$

Suppose that the partition Π does not satisfy our maximum degree requirement, i.e., $\Delta(T_\Pi) > D^2$. Then there must be some $V \in \Pi$ whose degree in T_Π is larger than D^2. Let S be the set of all neighbors of V in T outside of V. Clearly, $|S| = \deg_{T_\Pi}(V)$. Finally, let T' be the subtree of T induced by $V \cup S$. By Proposition 3.24, T' contains an edge e whose deletion splits T' into two trees, one of which contains at least D and at most $(D-1)^2 + 1 \leq D^2 - 1$ vertices from S, see Figure 3.1. Note that none of the endpoints of e lies in S, or otherwise the two trees would contain 1 and $|S| - 1$ vertices from S respectively, and this is impossible since $1 < D$ and $|S| - 1 > D^2 - 1$. Hence, e partitions V into two connected subsets V' and V''. Let Π' be the partition obtained from Π by replacing V with V' and V''. Note that

$$\deg_{T_{\Pi'}}(V') + \deg_{T_{\Pi'}}(V'') = \deg_{T_\Pi}(V) + 2,$$

and by the choice of e, either $\deg_{T_{\Pi'}}(V')$ or $\deg_{T_{\Pi'}}(V'')$ is at least $D+1$ and at most D^2. Hence, $E(\Pi') < E(\Pi)$. It follows that by refining our initial partition Π at most $2/\beta$ times, each time increasing the number of parts by one, we will arrive at a partition Π^* with $E(\Pi^*) = 0$. Clearly, the maximum degree of the cluster tree T_{Π^*} is at most D^2, and the number of parts is not greater than $1/\beta + 1 + 2/\beta$, which is at most $4/\beta$.

Stage 3. Root the cluster tree T_{Π^*} at the subset containing the root of the original tree. Order the subsets in Π^* in such a way that all descendants (in the cluster tree) of every subset come later in the ordering (e.g., by performing a breadth-first search on T_{Π^*}). Now, iterate the following procedure until the 'children of the root only in its own tree' condition is satisfied. If there is a subtree whose root has children in other subtrees, merge the first (with respect to our order) such tree with the subtrees containing children of its root and remove all the merged pieces from the ordered list. Clearly, after the procedure terminates, the new partition satisfies the required condition. As a consequence of our 'parent-first' ordering, each tree can be merged only once and hence both the upper bound on the sizes of the parts and the maximum degree of the cluster tree can increase at most D times. Finally, the number of parts in the partition can only become smaller. □

Recall that the number n' of vertices in the graph G'', which is an induced subgraph of G', satisfies $n' \geq (1 - \alpha/4)n$. Root our tree T at an arbitrary vertex. Since T has fewer than n' and more than $n/2$ vertices, by Lemma 3.23, where we let $\beta = \delta/k'$, there is a partition Π of $V(T)$ into connected subsets $S_1, \ldots, S_\tau$ such that

$$\tau \leq 4k'/\delta, \quad \text{and} \quad \max_{1 \leq j \leq \tau} |S_j| \leq D^2 \delta \cdot n'/k', \tag{3.2}$$

and the maximum degree of the cluster tree T_Π does not exceed D^3. Moreover, assume that the subtrees S_j are ordered in such a way that all descendants (in the cluster tree) of every subtree come later in the ordering. Since each S_j induces a connected bipartite subgraph of T (a subtree of T), it can be uniquely decomposed into two independent sets $S_{j,1}$ and $S_{j,2}$ – the color classes in the unique proper 2-coloring of the tree T restricted to S_j. Let $\mathcal{S}$ be the collection of all these color classes, i.e.,

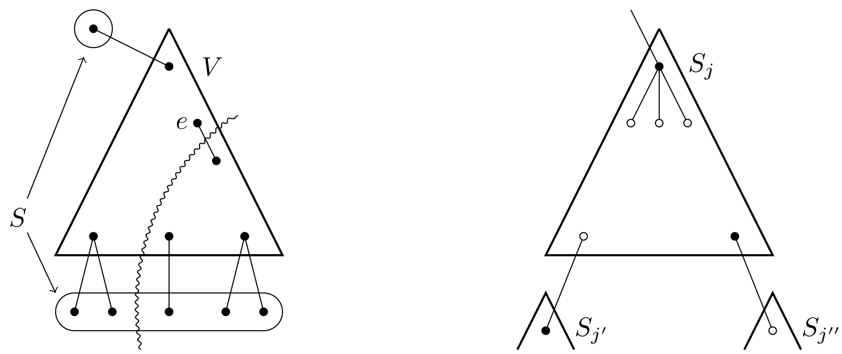

Figure 3.1: On the left, the edge e splits $T[V \cup S]$ into two trees, partitioning the set S. On the right, a typical subtree S_j; vertices in color classes 1 and 2 are drawn as white and black circles, respectively. Note that all children of the root of S_j belong to S_j. Since $S_{j'}$ and $S_{j''}$ are below S_j in T_Π, we have $j', j'' > j$. Finally, observe that $\{S_{j,1}, S_{j',2}\}, \{S_{j,2}, S_{j'',1}\} \in \mathcal{E}$.

$\mathcal{S} = \{S_{j,l} : 1 \le j \le \tau,\ l \in \{1,2\}\}$. Also, let $\mathcal{E}$ be the set of all pairs $\{S_{j,l}, S_{j',l'}\}$ such that $j \neq j'$ and in T there is an edge joining $S_{j,l}$ and $S_{j',l'}$. Finally, note that the graph obtained from the graph $(\mathcal{S}, \mathcal{E})$ by identifying all pairs $\{S_{j,1}, S_{j,2}\}$ is the cluster tree T_Π. It follows that $|\mathcal{E}| = \tau - 1$. For better visualization, the reader is encouraged to consult Figure 3.1.

3.4.5 Planning out the embedding

Recall that M is a maximum matching in the cluster graph H'', and $V(M) = \{A_1, B_1, \dots, A_{m'}, B_{m'}\}$. Our plan is to embed (most of) the tree T into regular pairs forming M. We will do it piece-by-piece, according to the partition Π. Before we start the actual embedding, we need to lay out a plan in order to make sure that we will never run out of vacant vertices or edges. We start by assigning to each edge in the matching a collection of subtrees of T that we plan to embed in the (ε, p)-regular pair in G'' that is represented by this edge.

Lemma 3.25. *There is an assignment $\varphi : \mathcal{S} \to V(M)$ with the following two properties.*

1. *For each j, there is an i such that the sets $S_{j,1}$ and $S_{j,2}$ are assigned to two different clusters in the pair $\{A_i, B_i\}$.*

2. *Let $X \in V(M)$ and let $\mathcal{S}(X)$ be the family of sets that φ assigns to X. Define the usage $U(X)$ of the cluster X by*
$$U(X) = \sum_{S \in \mathcal{S}(X)} \left(|S| + 4D^3\varepsilon \cdot \frac{n'}{k'} \right).$$
Then $U(X) \le (1 - \alpha/4) \cdot \frac{n'}{k'}$ for all $X \in V(M)$.

Proof. We can easily construct such a map φ using the following greedy procedure. Start with an empty map φ_0. Assume that $1 \le j \le \tau$ and we have already defined φ_{j-1}. For a cluster $X \in V(M)$, define the usage of X at step $j-1$ as
$$U_{j-1}(X) = \sum_{S \in \mathcal{S}_{j-1}(X)} \left(|S| + 4D^3\varepsilon \cdot \frac{n'}{k'} \right),$$

where $\mathcal{S}_{j-1}(X)$ is the family of sets from $\mathcal{S}$ that φ_{j-1} assigns to X. We claim that there exists an $i \in \{1, \ldots, m'\}$ such that

$$\max\{U_{j-1}(A_i), U_{j-1}(B_i)\} \leq (1 - \alpha/4 - 4D^3\varepsilon - D^2\delta) \cdot \frac{n'}{k'}. \tag{3.3}$$

We postpone the verification of this claim till the end of the proof of Lemma 3.25. Let $i(j)$ to be the smallest i satisfying (3.3). Let φ_j be an extension of φ_{j-1} that maps the pair $\{S_{j,1}, S_{j,2}\}$ to $\{A_{i(j)}, B_{i(j)}\}$ in such a way that the smaller of the sets $S_{j,1}, S_{j,2}$ is mapped to the cluster with larger usage U_{j-1} (we break ties arbitrarily). Finally, put $\varphi = \varphi_\tau$.

Note that the way we construct φ guarantees that condition 1 is satisfied. Moreover, since by (3.2), for all j and l, we have $|S_{j,l}| \leq |S_j| \leq D^2\delta \cdot n'/k'$, it follows that $U_j(X) - U_{j-1}(X) \leq (D^2\delta + 4D^3\varepsilon) \cdot n'/k'$ for every $X \in \{A_{i(j)}, B_{i(j)}\}$ (and clearly $U_j(X) = U_{j-1}(X)$ if $X \notin \{A_{i(j)}, B_{i(j)}\}$), and hence the choice of $i(j)$ at each step guarantees that φ will satisfy condition 2 as well. The only thing we still have to check is that for all j, the index $i(j)$ is well-defined, i.e., inequality (3.3) is satisfied for some i.

Observe that our strategy of balancing the usage of each pair $\{A_i, B_i\}$ guarantees that for all i and j,

$$\big|U_j(A_i) - U_j(B_i)\big| \leq \max_{j' \leq j} \big||S_{j',1}| - |S_{j',2}|\big| \leq \max_{j' \leq j} |S_{j'}| \leq D^2\delta \cdot \frac{n'}{k'}.$$

Hence, if for some j, inequality (3.3) is not satisfied for all i, then

$$U_j(X) \geq (1 - \alpha/4 - 4D^3\varepsilon - 2D^2\delta) \cdot \frac{n'}{k'}$$

for all $X \in V(M)$. Recall that $m' = \lfloor k'/2 \rfloor$. It follows that

$$\begin{aligned} \sum_{j' \leq j} |S_{j'}| &= \sum_{X \in V(M)} \left(U_j(X) - 4D^3\varepsilon \cdot \frac{n'}{k'} \right) \geq (1 - \alpha/4 - 8D^3\varepsilon - 2D^2\delta) \cdot \frac{2m'}{k'} \cdot n' \\ &\geq (1 - \alpha/4 - 8D^3\varepsilon - 2D^2\delta - 1/k') \cdot n' \geq (1 - \alpha/2) \cdot n' \geq (1 - 3\alpha/4) \cdot n. \end{aligned}$$

This would be a clear contradiction, since $(1 - \alpha) \cdot n \geq |V(T)| = \sum_j |S_j|$. □

Let φ be a map satisfying both conditions in Lemma 3.25. Our next step will be planning out connections between all the subtrees in our partition Π, whose locations in the graph G' have already been determined by φ.

Lemma 3.26. *There is an assignment $\psi : \mathcal{E} \to V(H'')$ with the following two properties.*

1. *For all $e \in \mathcal{E}$, the following holds. Suppose that $e = \{S_{j,l}, S_{j',l'}\}$, where $j < j'$. Then $\psi(e)$ is a common neighbor in H'' of the clusters $\varphi(S_{j,l})$ and $\varphi(S_{j',3-l'})$.*

2. *Every cluster is assigned to at most $6/(\gamma\delta)$ edges in $\mathcal{E}$, i.e., $|\psi^{-1}(X)| \leq 6/(\gamma\delta)$ for all $X \in V(H'')$.*

Proof. We construct such a map greedily, starting from the empty map and extending it one-by-one to the whole set $\mathcal{E}$. Let $e = \{S_{j,l}, S_{j',l'}\} \in \mathcal{E}$, where $j < j'$. Since $\delta(H'') \geq (1/2 + \gamma/3)k \geq (1/2 + \gamma/3)k'$, the clusters $\varphi(S_{j,l})$ and $\varphi(S_{j',3-l'})$ have at least $2\gamma k'/3$ common neighbors. One of them has been used fewer than $|\mathcal{E}|/(2\gamma k'/3) \leq \tau/(2\gamma k'/3) \leq 6/(\gamma\delta)$ times, where the second inequality follows from (3.2). We let $\psi(e)$ be an arbitrary cluster with that property. □

Now that we have laid out a general plan for the embedding, it is time to assign to each $S_{j,l}$ a particular subset of $V(G'')$, where we will map $S_{j,l}$. We start by choosing in each cluster $X \in V(H'')$ an arbitrary subset $C(X)$ of size $\alpha/8 \cdot n'/k'$. Let e_X be the number of edges in $\mathcal{E}$ that are assigned to X. We partition $C(X)$ into e_X subsets of equal sizes and label those subsets with elements of $\psi^{-1}(X)$ such that each $e \in \psi^{-1}(X)$ gets its own set $\psi'(e)$ of size at least $\alpha/(4e_X) \cdot n'/k'$, which is at least $\alpha\gamma\delta/48 \cdot n'/k'$.

Next, fix a cluster $X \in V(M)$. For each $S \in \mathcal{S}(X)$, we choose a subset $\varphi'(S)$ of $X - C(X)$ with size $|S| + 4D^3\varepsilon \cdot n'/k'$ such that all these sets are disjoint. By the choice of φ, which satisfies condition 2 in Lemma 3.25, this is possible. We do this for all clusters in $V(M)$[2].

Finally, note that for all j, $\varphi'(S_{j,1})$ and $\varphi'(S_{j,2})$ are subsets of opposite clusters in an (ε, p)-regular pair in G'' with p-density at least $\gamma/6$ and for each $e = \{S_{j,l}, S_{j',l'}\}$, where $j < j'$, also $\{\varphi'(S_{j,l}), \psi'(e)\}$ and $\{\varphi'(S_{j',3-l'}), \psi'(e)\}$ are pairs of subsets of opposite classes in an (ε, p)-regular pair in G'', whose p-density is at least $\gamma/6$.

3.4.6 Cleaning up G''

Recall that we have ordered the subtrees in our partition in such a way that all descendants of a tree S in the cluster tree come later in the ordering, i.e., if $S_{j'}$ is a descendant of S_j, then $j' > j$. Our goal in the cleaning-up stage is the following.

Goal. Construct functions $\varphi'' : \mathcal{S} \to \mathcal{P}(V(G''))$ and $\psi'' : \mathcal{E} \to \mathcal{P}(V(G''))$ with the following properties.

1. For all $S \in \mathcal{S}$, $\varphi''(S) \subseteq \varphi'(S)$ and $|\varphi''(S)| \geq |S| + (2D+1)\varepsilon \cdot n'/k'$.
2. For all j, the graph $(\varphi''(S_{j,1}), \varphi''(S_{j,2}))$ is a bipartite $(\varepsilon \cdot n'/k', 2D+2)$-expander.
3. For all $e \in \mathcal{E}$, the following holds. Suppose that $e = \{S_{j_1,l_1}, S_{j_2,l_2}\}$, where $j_1 < j_2$. Then:
   (a) $\psi''(e) \subseteq \psi'(e)$ and $|\psi''(e)| \geq \varepsilon \cdot n'/k'$,
   (b) each vertex in $\varphi''(S_{j_1,l_1})$ has a neighbor in $\psi''(e)$,
   (c) each vertex in $\psi''(e)$ has at least $D+1$ neighbors in $\varphi''(S_{j_2,3-l_2})$.

In the process of achieving our goal, we will extensively use the following two technical lemmas.

Lemma 3.27. *Let (A, B) be an (ε, p)-regular pair, whose p-density is larger than ε. Suppose that $|A| = |B| = n'/k'$, and $A' \subseteq A$ and $B' \subseteq B$ are sets of size at least $(4D+6)\varepsilon \cdot n'/k'$. Then there are subsets $A'' \subseteq A'$ and $B'' \subseteq B'$ satisfying the following two conditions.*

1. *$|A' - A''| \leq \varepsilon \cdot n'/k'$ and $|B' - B''| \leq \varepsilon \cdot n'/k'$.*
2. *The subgraph (A'', B'') is a bipartite $(\varepsilon \cdot n'/k', 2D+2)$-expander.*

[2] For the sake of clarity of the presentation we tacitly assumed that the numbers $\alpha/(4e_X) \cdot n'/k'$ and $4D^3\varepsilon \cdot n'/k'$ were integers. This is clearly not true in general, but since we assume that n' is large, we can utilize the remaining $\alpha/8 \cdot n'/k'$ unused vertices in each cluster to account for all rounding errors, as the number of sets $\varphi'(S)$ and $\psi'(e)$ is independent of n.

Proof. We will greedily construct such subsets A'' and B''. Before we start, we would like to remark that all neighborhoods are computed in the subgraph (A', B'), and not the graph (A, B) itself. First, let $X = \emptyset$ and $Y = \emptyset$. We will iterate the following procedure. If there is a set $X' \subseteq A' - X$, with $|X'| \leq \varepsilon \cdot n'/k'$, such that $|N(X') - Y| < (2D+2)|X'|$, then set $X = X \cup X'$. Similarly, if there is a set $Y' \subseteq B' - Y$, with $|Y'| \leq \varepsilon \cdot n'/k'$ such that $|N(Y') - X| < (2D+2)|Y'|$, then set $Y = Y \cup Y'$.

First we show that at all times $|X| \leq \varepsilon \cdot n'/k'$ and $|N(X) - Y| \leq (2D+2)|X|$, and similarly, $|Y| \leq \varepsilon \cdot n'/k'$ and $|N(Y) - X| \leq (2D+2)|Y|$. Certainly, this is true at the beginning of the procedure, since then $|X| = |Y| = |N(X) - Y| = |N(Y) - X| = 0$. Suppose that all four inequalities hold at the beginning of some iteration. Assume that the procedure finds an $X' \subseteq A' - X$ with $|X'| \leq \varepsilon \cdot n'/k'$ and $|N(X') - Y| < (2D+2)|X'|$. Then

$$\begin{aligned} |N(X \cup X') - Y| &= |(N(X) - Y) \cup (N(X') - Y)| \leq |N(X) - Y| + |N(X') - Y| \\ &\leq (2D+2)|X| + (2D+2)|X'| = (2D+2)|X \cup X'|. \end{aligned}$$

Note that in (A, B) there are no edges between $X \cup X'$ and $B' - N(X \cup X')$. Also, since $|X \cup X'| = |X| + |X'| \leq 2\varepsilon \cdot n'/k'$, then

$$|B' - N(X \cup X')| \geq |B'| - |N(X \cup X') - Y| - |Y| \geq |B'| - (4D+5)\varepsilon \cdot n'/k' \geq \varepsilon \cdot n'/k'.$$

Since (A, B) was (ε, p)-regular with p-density larger than ε, and $|A| = |B| = n'/k'$, it must be that $|X \cup X'| < \varepsilon \cdot n'/k'$. A symmetric argument proves the other two inequalities.

Now, put $A'' = A' - X$ and $B'' = B' - Y$. We have already proved that condition 1 holds for this choice of A'' and B''. As for the other condition, the definition of X and Y guarantees that all small subsets of A'' and B'' expand at least $2D + 2$ times. It suffices to prove that also large sets expand well enough. Suppose that there is an $X' \subseteq A''$ with $|X'| \geq \varepsilon \cdot n'/k'$ such that $|N(X') \cap B''| < |B''| - \varepsilon \cdot n'/k'$. There are no edges in (A, B) between the sets X' and $B'' - N(X')$, but this is impossible, since (A, B) is (ε, p)-regular with p density larger than ε, and both sets are larger than $\varepsilon \cdot n'/k'$. $\square$

Lemma 3.28. *Let $b \geq 1$ and let (A, B) be an (ε, p)-regular pair, whose p-density is larger than ε. Suppose that $|A| = |B| = n'/k'$, and $A' \subseteq A$ and $B' \subseteq B$ are sets of size at least $2\varepsilon \cdot n'/k'$ and $b\varepsilon \cdot n'/k'$ respectively. Then there is a subset $A'' \subseteq A'$ such that $|A' - A''| \leq \varepsilon \cdot n'/k'$ and every vertex in A'' has at least b neighbors in B'.*

Proof. Let $X \subseteq A'$ be the set of all vertices in A that have fewer than b neighbors in B and put $A'' = A' - X$. If $|X| \leq \varepsilon \cdot n'/k'$, then there is nothing left to prove. Otherwise, let X' be an arbitrary subset of X of size $\varepsilon \cdot n'/k'$. Clearly, there are no edges in (A, B) between X' and $B' - N(X')$. This is impossible, since (A, B) is (ε, p)-regular with p-density larger than ε, $|X'| \geq \varepsilon \cdot n'/k'$ and by the definition of X, we have $|B' - N(X')| \geq |B'| - (b-1)|X'| \geq \varepsilon \cdot n'/k'$. $\square$

An immediate consequence of Lemma 3.28 is the following Corollary.

Corollary 3.29. *Let $d \geq 1$ and let $(A, B_1), \ldots, (A, B_d)$ be (not necessarily distinct) (ε, p)-regular pairs in G''. Suppose that $|A| = |B_1| = \ldots = |B_d| = n'/k'$, $A' \subseteq A$ is a set of size at least $(d+1)\varepsilon \cdot n'/k'$ and $B'_i \subseteq B_i$ are sets of size at least $\varepsilon \cdot n'/k'$ for each $i \in \{1, \ldots, d\}$. Then there is a subset $A'' \subseteq A'$ such that $|A' - A''| \leq d\varepsilon \cdot n'/k'$ and every vertex in A'' has a neighbor in each B'_i.*

We start cleaning up by setting $\varphi'' = \varphi'$ and $\psi'' = \psi'$. Next, we will iteratively, starting with $j = t$ and each time reducing j by one, keep fixing the two functions by making sure that after we have finished step j, the requirements 1, 2 and 3b are met as long as they involve only sets $S_{j',l'}$ with $j' \geq j$ and $l' \in \{1, 2\}$ (i.e., $j_1 \geq j$ in 3), and the requirements 3a and 3c are met as long as they involve sets $\{S_{j_1,l_1}, S_{j_2,l_2}\}$ with $\max\{j_1, j_2\} \geq j$. If we manage to do that, after completing the final step ($j = 1$) our goal will be reached.

Assume that we are at a step j and our functions φ'' and ψ'' satisfy all requirements involving only sets $S_{j',l'}$ with $j' > j$ and requirements 3a and 3c, where $j_2 > j$. Let $\mathcal{D}(S_{j,1})$ and $\mathcal{D}(S_{j,2})$ be the families of the color classes of all the children (in the cluster tree T_Π) of S_j that are adjacent to the color classes $S_{j,1}$ and $S_{j,2}$, respectively. In other words, $S_{j',l'} \in \mathcal{D}(S_{j,l})$ if and only if $S_{j'}$ is a child of S_j in the cluster tree T_Π, and the edge connecting S_j and $S_{j'}$ in T has endpoints in the sets $S_{j,l}$ and $S_{j',l'}$. For each $l \in \{1, 2\}$, the following is true. Fix a set $S_{j',l'} \in \mathcal{D}(S_{j,l})$ and let $e = \{S_{j,l}, S_{j',l'}\}$. Since $S_{j',l'}$ is a descendant of $S_{j,l}$ in the cluster tree, we have $j' > j$, and therefore $|\psi''(e)| \geq \varepsilon \cdot n'/k'$. Since $|\mathcal{D}(S_{j,l})| \leq \Delta(T_\Pi) \leq D^3$, and $|\varphi'(S_{j,l})| \geq |S_{j,l}| + 4D^3\varepsilon \cdot n'/k'$, by Corollary 3.29, there is a subset $A'_l \subseteq \varphi'(S_{j,l})$ of size at least $|S_{j,l}| + 3D^3\varepsilon \cdot n'/k'$ such that every vertex in A'_l has a neighbor in every $\psi''(\{S_{j,l}, S\})$, for each $S \in \mathcal{D}(S_{j,l})$.

By Lemma 3.27, there are subsets $A''_1 \subseteq A'_1$ and $A''_2 \subseteq A'_2$ of sizes at least $|S_{j,1}| + 2D^3\varepsilon \cdot n'/k' \geq |S_{j,1}| + (2D+1)\varepsilon \cdot n'/k'$ and $|S_{j,2}| + 2D^3\varepsilon \cdot n'/k' \geq |S_{j,2}| + (2D+1)\varepsilon \cdot n'/k'$, respectively, such that the induced graph (A''_1, A''_2) is a bipartite $(\varepsilon \cdot n'/k', 2D+2)$-expander. We put $\varphi''(S_{j,l}) = A''_l$ for both l.

Finally, let l be such that the set $S_{j,l}$ contains the root of the tree $T[S_j]$. If $j \neq 1$, then $S_{j,l}$ has a unique parent $S \in \mathcal{S}$. Let $e = \{S, S_{j,l}\}$. Clearly, e belongs to $\mathcal{E}$. Since $|\psi'(e)| \geq \alpha\gamma\delta/48 \cdot n'/k' \geq 2\varepsilon \cdot n'/k'$, by Lemma 3.28 there is a subset $A'' \subseteq \psi'(e)$ of size at least $\varepsilon \cdot n'/k'$ such that every vertex in A'' has at least $D+1$ neighbors in $\varphi''(S_{j,l})$. We put $\psi''(e) = A''$. If $j = 1$, then S_j is the root of T_Π, and there is nothing to do.

3.4.7 Embedding T into G''

Finally, we are ready to embed our tree T into G''. We will do the actual embedding in a top-down fashion, starting with S_1 and extending our embedding to all other S_j one-by-one. For each j, the subtree $T[S_j]$ will be embedded into the bipartite expanding graph $(\varphi''(S_{j,1}), \varphi''(S_{j,2}))$ with the small exception that, unless $j = 1$, the root of the tree will be embedded into the appropriate 'connecting' set $\psi''(e)$, where $e \in \mathcal{E}$ represents the edge between $T[S_j]$ and its parent in the cluster tree T_Π.

We start by embedding the first subtree, S_1, into the bipartite graph H_1 induced on the pair $(\varphi''(S_{1,1}), \varphi''(S_{1,2}))$. Since $|\varphi''(S_{1,l})| \geq |S_{1,l}| + 2D^3\varepsilon \cdot n'/k'$ for each $l \in \{1, 2\}$, and H_1 is a bipartite $(\varepsilon \cdot n'/k', D+1)$-expander, Corollary 3.10 guarantees that this is possible. Suppose we have already embedded $S_1, \ldots, S_{j-1}$ into T in such a way that for every $j' < j$ and $l' \in \{1, 2\}$, all vertices in $S_{j',l'}$, except the root of $T[S_{j'}]$, are mapped into the set $\varphi''(S_{j',l'})$. Let $l \in \{1, 2\}$ be such that the root of $T[S_j]$, call it r_j, is in $S_{j,l}$. Let p_j be the parent of r_j in the tree T and let j' and l' be such that $p_j \in S_{j',l'}$. Finally, let $e := \{S_{j',l'}, S_{j,l}\}$. The way we defined φ'' and ψ'' guarantees that the image of p_j, which by the definition of Π cannot be the root of its tree and hence has not been mapped to a vertex in one of the 'connecting' sets $\psi''(e)$, is in $\varphi''(S_{j',l'})$ and has a neighbor x in the set $\psi''(e)$,

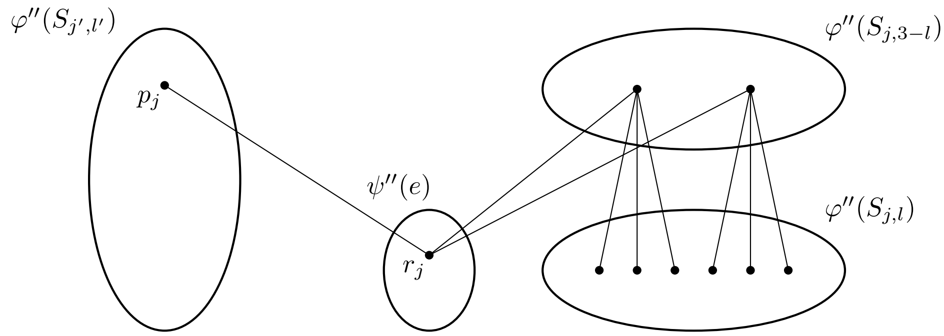

Figure 3.2: Embedding the subtree S_j into G''

and x has at least $D+1$ neighbors in $\varphi''(S_{j,3-l})$.

Claim 3.30. *The graph H_j induced on the pair $(\varphi''(S_{j,l}) \cup \{x\}, \varphi''(S_{j,3-l}))$ is a bipartite $(\varepsilon \cdot n'/k' + 1, D+1)$-expander.*

Proof. For the sake of brevity, let $A := \varphi''(S_{j,l}) \cup \{x\}$, $B := \varphi''(S_{j,3-l})$ and $q := \varepsilon \cdot n'/k'$. Recall that the graph $(A - \{x\}, B)$ is a bipartite $(q, 2D+2)$-expander. It is easy to check that H_j satisfies all conditions from Definition 3.9.

For example, let $X \subseteq A$ be a set of size at most $q+1$. If $x \notin X$, then $|N_{H_j}(X)| \geq (2D+2)\min\{|X|, q\} \geq (D+1)|X|$. If $x \in X$ but $X \neq \{x\}$, then $|N_{H_j}(X)| \geq |N_{H_j}(X - \{x\})| \geq (2D+2)(|X|-1) \geq (D+1)|X|$. Finally, if $X = \{x\}$, then $|N_{H_j}(X)| \geq (D+1)|X|$ by the choice of x.

A similar straightforward case analysis shows that the other conditions from Definition 3.9 are also satisfied. We omit the details. □

Since $|\varphi''(S_{j,l}) \cup \{x\}| \geq |S_{j,l}| + 2D^3\varepsilon \cdot n'/k'$ and $|\varphi''(S_{j,3-l})| \geq |S_{j,3-l}| + 2D^3\varepsilon \cdot n'/k'$, Corollary 3.10 says that we can embed $T[S_j]$ in H_j in such a way that r_j is mapped to x. Note that necessarily $S_{j,l} - \{r_j\}$ is mapped to $\varphi''(S_{j,l})$ and $S_{j,3-l}$ is mapped to $\varphi''(S_{j,3-l})$, see Figure 3.4.7. This completes the proof.

Chapter 4

Asymptotic counting of $K_{s,t}$-free graphs

4.1 Introduction

Let H be an arbitrary graph. We say that a graph G is *H-free* if G does not contain H as a (not necessarily induced) subgraph. Denote by $f_n(H)$ the number of labeled H-free graphs on a fixed vertex set of size n. Let $\mathrm{ex}(n, H)$ denote the *Turán number for H*, i.e., the maximum number of edges in an H-free graph on n vertices. Extending the classical theorem of Turán [73], Erdős and Stone [34] proved that the order of magnitude of $\mathrm{ex}(n, H)$ depends only on the chromatic number of H, i.e.,

$$\mathrm{ex}(n, H) = \left(1 - \frac{1}{\chi(H) - 1}\right) \frac{n^2}{2} + o(n^2). \tag{4.1}$$

Since every subgraph of an H-free graph is also H-free, it follows that $f_n(H) \geq 2^{\mathrm{ex}(n,H)}$. Erdős, Frankl and Rödl [28] proved that this crude lower bound is in fact tight whenever $\chi(H) \geq 3$, namely,

$$f_n(H) = 2^{(1+o(1))\cdot \mathrm{ex}(n,H)}. \tag{4.2}$$

The picture changes dramatically when one drops the $\chi(H) \geq 3$ assumption. This is not at all surprising, since when $\chi(H) = 2$, the right hand side of (4.1) collapses to the $o(n^2)$ error term. Nevertheless, Erdős [20] asked if (4.2) is still true if H is a bipartite graph containing a cycle. His question remains unanswered, and for most such H not even the correct order of magnitude of $\log_2 f_n(H)$ is known. The only results in this direction are due to Kleitman and Winston [53], who proved that $\log_2 f_n(C_4) \leq 2.17 \cdot \mathrm{ex}(n, C_4)$, and Kleitman and Wilson [52], who proved that $\log_2 f_n(C_6) \leq O(\mathrm{ex}(n, C_6))$. It is worth mentioning that the $2^{O(n^{5/4})}$ bound on the number of C_8-free graphs obtained by Kleitman and Wilson [52] may turn out to be asymptotically tight once the order of magnitude of the Turán number for C_8 is determined.

Below we prove the best possible result that one can expect for all complete bipartite graphs.

Definition 4.1. The binary entropy function $H: [0, 1] \to \mathbb{R}$ is defined by

$$H(x) = -x \log_2 x - (1 - x) \log_2(1 - x).$$

The material presented in this chapter is joint work with József Balogh. Part of it was accepted for publication in The SIAM Journal on Discrete Mathematics under the title *Almost all C_4-free graphs have less than $(1-\varepsilon)\,\mathrm{ex}(n, C_4)$ edges*, see [13]; part of it was submitted to Combinatorica under the title *The number of $K_{m,m}$-free graphs*, see [14]; part of it was submitted to the Journal of the London Mathematical society under the title *The number of $K_{s,t}$-free graphs*, see [15].

For every positive integer s, let

$$C_s = \sup_{x\in(0,1)} \left(x^{-1+1/s} H(x)\right) \in [1, 0.55s + 1.5).$$

Theorem 4.2 ([15])**.** *For all s and t with $2 \le s \le t$, the number of labeled $K_{s,t}$-free graphs on n vertices satisfies*

$$\log_2 f_n(K_{s,t}) \le (1 + o(1)) \frac{s(t-1)^{1/s}}{2s-1} C_s \cdot n^{2-1/s}.$$

Erdős conjectured [25] that $\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s})$ for all s and t with $2 \le s \le t$. If this conjecture is true, Theorem 4.2 would be asymptotically sharp for all pairs (s, t). So far, Erdős' conjecture has been resolved in the affirmative in the case $s \le 3$ (see [18, 42, 43]) or $t > (s-1)!$ (see [57, 7]); therefore, Theorem 4.2 is sharp for 'most' pairs of s and t.

Füredi [42] proved that for all $t \ge 2$, $\mathrm{ex}(n, K_{2,t}) = \frac{1}{2}\sqrt{t-1} \cdot n^{3/2} + O(n^{4/3})$; together with Theorem 4.2, it implies the following.

Corollary 4.3 ([15])**.** *If $t \ge 2$, then the number of $K_{2,t}$-free graphs of order n is satisfies*

$$\mathrm{ex}(n, K_{2,t}) \le \log_2 f_n(K_{2,t}) \le 2.16384 \cdot \mathrm{ex}(n, K_{2,t}).$$

Let $f_{n,m}(H)$ denote the number of H-free graphs on a fixed n-element vertex set, having exactly m edges. The methods used in the proof of Theorem 4.2 also give an upper bound on $f_{n,m}(K_{s,t})$.

Theorem 4.4 ([15])**.** *For every s and t with $2 \le s \le t$, let*

$$\mu_{s,t} = \frac{1}{s} + \frac{s-1}{s^2(t-1)(t-s+1)+s}.$$

There is an n_0 (depending on s and t) such that for all n and m with $n \ge n_0$ and $m \ge n^{2-\mu_{s,t}}(\log n)^{3t/s+2}$, the number $f_{n,m}(K_{s,t})$ of labeled $K_{s,t}$-free graphs of order n and size m satisfies

$$f_{n,m}(K_{s,t}) \le \left(\frac{3tn^{2s-1}}{m^s}\right)^m.$$

The remainder of this chapter is organized as follows. In Section 4.2, we study several implications of Theorems 4.2 and 4.4. In Section 4.3, we state and prove a few technical lemmas. We defer the proofs of Theorems 4.2, 4.4 and 4.7 to Sections 4.4, 4.5 and 4.6, respectively.

4.2 Implications of the main results

The main results of this chapter, Theorems 4.2 and 4.4 have various interesting consequences, some of which we list below. Most of the statements in this section are straightforward applications of the main results, and hence their proofs are omitted.

4.2.1 Balogh-Bollobás-Simonovits conjecture

Let H be a fixed non-bipartite graph. For every positive constant ε, almost all H-free graphs on n vertices have between $(\frac{1}{2} - \varepsilon)\,\mathrm{ex}(n, H)$ and $(\frac{1}{2} + \varepsilon)\,\mathrm{ex}(n, H)$ edges. It is not known if a similar concentration around one half still occurs when H is bipartite. Nevertheless, one would expect that the number of edges in a 'typical' H-free graph is at least bounded away from the extremal values, 0 and $\mathrm{ex}(n, H)$. Balogh, Bollobás, and Simonovits [9] formalized this intuition by stating the following conjecture.

Conjecture 4.5. *For every bipartite graph H that contains a cycle, there is a positive constant c_H such that almost all H-free graphs on n vertices have at least $c_H \cdot \mathrm{ex}(n, H)$ and at most $(1 - c_H) \cdot \mathrm{ex}(n, H)$ edges.*

So far, Conjecture 4.5 has been partially (only the lower bound) proved in the case $H = C_4$ [41] and $H = C_6$ [41, 52]. In [9], the precise structure of almost all octahedron-free ($K_{2,2,2}$-free) graphs was described. The main obstacle to extending this result to other complete multipartite graphs was the lack of results showing that the lower bound in Conjecture 4.5 holds for complete bipartite graphs other than C_4. An immediate corollary of Theorem 4.4 provides such a lower bound.

Corollary 4.6 ([15])**.** *Let s and t be integers satisfying $s \in \{2, 3\}$ and $t \geq s$, or $s > 3$ and $t > (s-1)!$. There exists a positive constant $c_{s,t}$ such that almost all $K_{s,t}$-free graphs of order n have at least $c_{s,t} \cdot \mathrm{ex}(n, K_{s,t})$ edges. Moreover, if $t \geq 2$, then we may choose $c_{2,t} = 1/12$.*

Building on the methods developed in the proof of Theorem 4.2 to obtain an upper bound on the number of one-vertex extensions of a $K_{2,t}$-free graph, one gets the following.

Theorem 4.7 ([13])**.** *There exists a positive constant ε such that for every t with $t \geq 2$, almost all $K_{2,t}$-free graphs of order n have at most $(1 - \varepsilon) \cdot \mathrm{ex}(n, K_{2,t})$ edges.*

We defer the proof of Theorem 4.7 to Section 4.6.

4.2.2 Haxell-Kohayakawa-Łuczak conjecture

Given two graphs G and H, one defines the *generalized Turán number for H in G*, denoted $\mathrm{ex}(G, H)$, by

$$\mathrm{ex}(G, H) = \max\{e(K) : H \nsubseteq K \subseteq G\}.$$

A simple averaging argument implies that for every positive integer k, an arbitrary graph G has a k-partite subgraph with at least $(1 - 1/k) \cdot e(G)$ edges. It follows that for every G and H,

$$\mathrm{ex}(G, H) \geq \left(1 - \frac{1}{\chi(H) - 1}\right) \cdot e(G).$$

It is natural to ask for which graphs G the above inequality becomes an equality. Haxell, Kohayakawa and Łuczak [50] conjectured that whenever p is large enough, so that the random graph $G(n, p)$ has many uniformly distributed copies of H, then asymptotically almost surely $\mathrm{ex}(G(n, p), H) = (1 - \frac{1}{\chi(H)-1} + o(1)) \cdot e(G(n, p))$.

Definition 4.8. Let H be a fixed graph. We define the 2-*density* of H, denoted $d_2(H)$, by

$$d_2(H) = \max\left\{\frac{|E(K)| - 1}{|V(K)| - 2} : K \subseteq H, |V(K)| \geq 3\right\}.$$

Conjecture 4.9 ([50])**.** *Let H be a fixed balanced graph and suppose that $p\colon \mathbb{N} \to [0,1]$ satisfies $p(n) = \omega\left(n^{-1/d_2(H)}\right)$. Then with probability tending to 1 as $n \to \infty$,*

$$\mathrm{ex}(G(n,p), H) = \left(1 - \frac{1}{\chi(H)-1} + o(1)\right) \cdot e(G(n,p)).$$

So far, Conjecture 4.9 has been proved for all cycles [50, 51], K_4 [55], and K_5 [47]. Some partial results are also known for larger complete graphs, see [56, 70]. Recently, Conlon and Gowers [22] and, independently, Schacht [67], have announced that they have proved a meta-theorem that implies Conjecture 4.9 in its full generality. A straightforward application of Theorem 4.4 and the first moment method gives the following relaxed version of Conjecture 4.9 when H is a complete bipartite graph.

Corollary 4.10 ([15])**.** *Let s and t be integers satisfying $2 \le s \le t$ and let $\mu_{s,t}$ be as in the statement of Theorem 4.4. If $pn^{\mu_{s,t}} = \omega\left((\log n)^{3t/s+2}\right)$, then asymptotically almost surely*

$$\mathrm{ex}(G(n,p), K_{s,t}) = o\big(e(G(n,p))\big). \tag{4.3}$$

Note that in order to prove Conjecture 4.9, one has to show that (4.3) is still true if we only assume that $pn^{-\frac{s+t-2}{st-1}} \to \infty$. Still, unless $pn^{1/s} \to \infty$, and hence $\mathrm{ex}(n, K_{s,t}) = o\left(\mathbb{E}\big[e(G(n,p))\big]\right)$, the result proved by Corollary 4.10 is non-trivial.

4.2.3 Kohayakawa-Łuczak-Rödl conjecture

Recall from Section 1.3 that a bipartite graph $B = (V_1 \cup V_2, E)$ is ε-regular if for all sufficiently large sets $V_1' \subseteq V_1$ and $V_2' \subseteq V_2$, the density $d(V_1', V_2')$ differs from the density of B by at most $\varepsilon \cdot d(V_1, V_2)$.

Definition 4.11. For a graph H, let $\mathcal{G}(H,n,m)$ be the family of graphs on the vertex set $\bigcup_{x \in V(H)} V_x$, where the sets V_x are pairwise disjoint sets of size n, whose edge set is $\bigcup_{\{x,y\} \in E(H)} E_{x,y}$, where $E_{x,y} \subseteq V_x \times V_y$ and $|E_{x,y}| = m$. Let $\mathcal{G}(H,n,m,\varepsilon) \subseteq \mathcal{G}(H,n,m)$ denote the set of graphs in $\mathcal{G}(H,n,m)$ satisfying that each $(V_x \cup V_y, E_{x,y})$ is an ε-regular graph.

A graph $G \in \mathcal{G}(H,n,m,\varepsilon)$ looks like H, in which every vertex has been replaced by an independent set of size n, and every edge – by a set of m edges which form an ε-regular bipartite graph. Kohayakawa, Łuczak, and Rödl [55] conjectured that, whenever these bipartite graphs are dense enough, only a small fraction of graphs in $\mathcal{G}(H,n,m,\varepsilon)$ does not contain a copy of H.

Conjecture 4.12. *Let H be a fixed graph. For any positive β, there exist positive constants ε, C, and n_0 such that for all m and n satisfying $m \ge Cn^{2-1/d_2(H)}$ and $n \ge n_0$, we have*

$$|\{G \in \mathcal{G}(H,n,m,\varepsilon) : H \nsubseteq G\}| \le \beta^m \binom{n^2}{m}^{|E(H)|}.$$

Sor far, Conjecture 4.12 has been resolved in the affirmative when H is a tree, a cycle [44] or a complete graph on three [44, 60], four [46], or five vertices [47]. Some partial results are also known for larger complete graphs [45]. A straightforward application of Theorem 4.4 gives the following relaxed version of Conjecture 4.12 when H is a complete bipartite graph.

Corollary 4.13 ([15]). *Let s and t be integers satisfying $2 \le s \le t$, and let $\mu_{s,t}$ be as in the statement of Theorem 4.4. For any positive β and ε, there exist positive constants C and $n_0 0$ such that for all n and m satisfying $m \ge Cn^{2-\mu_{s,t}}(\log n)^{3t/s+2}$ and $n \ge n_0$, we have*

$$|\{G \in \mathcal{G}(K_{s,t}, n, m, \varepsilon) : K_{s,t} \not\subseteq G\}| \le \beta^m \binom{n^2}{m}^{|E(K_{s,t})|}. \tag{4.4}$$

Note that in order to prove Conjecture 4.12, one would have to show that (4.4) is still true if we only assume that $m \ge Cn^{2-\frac{s+t-2}{st-1}}$.

4.2.4 Random Ramsey graphs

A graph G is *Ramsey with respect to H*, $G \to H$, if every two-coloring of the edges of G results in a monochromatic subgraph isomorphic to H. Unsurprisingly, the smallest graphs which are Ramsey with respect to the four-cycle, are 'saturated' by C_4's. Erdős and Faudree asked (see [41]) whether this is always the case, i.e., if there exists a graph G such that $G \to C_4$, but G does not contain a $K_{2,3}$. Answering this question, Füredi [41] proved a much stronger result – whenever m is large enough, there are $K_{2,3}$-free graphs with m edges, whose largest C_4-free subgraph has only m^{1-c} edges, where $c \ge 1/51 + o(1)$. Clearly, all such graphs are Ramsey with respect to C_4. He also asked if similar results can be proved for other pairs of graphs. Using the random graph argument from [41], combined with Theorem 4.4, we can provide an answer to this question.

Corollary 4.14 ([15]). *For all integers s and t with $2 \le s \le t$, there exist an integer u with $u > t$ and positive constant c such that for all large enough m, there exists a $K_{s,u}$-free graphs G with m edges, whose largest $K_{s,t}$-free subgraph has only m^{1-c} edges. In particular, if $s = t = 3$, then one can take $u = 4$.*

4.3 Technical lemmas

One of the key ingredients in the proof of Theorem 4.2 is the following lemma, whose proof is a double counting argument in the spirit of Kövári, Sós, and Turán [58].

Lemma 4.15 ([14]). *Fix two integers s and t with $1 \le s \le t$ and a positive real ε. Let G be an n-vertex graph with minimum degree at least d, and A be any set of a vertices of G, where $a \ge (1+\varepsilon)(t-1)\binom{n}{s}/\binom{d}{s}$. Then the number of copies of $K_{s,t}$ in G with the larger partite set completely contained in A, denoted $N_{s,t}(A)$, satisfies*

$$N_{s,t}(A) \ge \beta \cdot a^t,$$

where

$$\beta = \beta(s,t,d,\varepsilon) = \frac{\varepsilon^t}{t!}\binom{d}{s}^t \Big/ \binom{n}{s}^{t-1}.$$

Proof. Let U be an s-set of vertices of G and assume that $U = \{u_1, \ldots, u_s\}$. Let $c(U)$ be the number of common neighbors of $u_1, \ldots, u_s$ in the set A, i.e.

$$c(U) = \left|N_G^*(U) \cap A\right|.$$

Clearly,

$$\sum_U c(U) = \sum_{w\in A} \binom{d_G(w)}{s} \geq a\binom{\delta(G)}{s} \geq a\binom{d}{s}.$$

The number of copies of $K_{s,t}$ in G with the larger partite set contained in A satisfies

$$N_{s,t}(A) = \sum_U \binom{c(U)}{t} \geq \binom{n}{s}\binom{a\binom{d}{s}/\binom{n}{s}}{t},$$

where the above inequality follows from convexity of the function B_t defined by

$$B_t(x) = \begin{cases} 0 & \text{if } x \leq t-1, \\ \binom{x}{t} & \text{if } x > t-1, \end{cases}$$

and the assumption that $a\binom{d}{s}/\binom{n}{s} > t-1$. It follows that

$$\begin{aligned} N_{s,t}(A) &\geq \binom{n}{s} \cdot \frac{1}{t!} \prod_{i=0}^{t-1} \left(\frac{a\binom{d}{s}}{\binom{n}{s}} - i \right) = \binom{n}{s} \cdot \left(\frac{a\binom{d}{s}}{\binom{n}{s}} \right)^t \cdot \frac{1}{t!} \prod_{i=0}^{t-1} \left(1 - i \frac{\binom{n}{s}}{a\binom{d}{s}} \right) \\ &\geq \frac{a^t}{t!} \binom{d}{s}^t / \binom{n}{s}^{t-1} \cdot \prod_{i=0}^{t-1} \left(1 - \frac{i}{(1+\varepsilon)(t-1)} \right) \\ &\geq \frac{a^t}{t!} \binom{d}{s}^t / \binom{n}{s}^{t-1} \cdot \left(1 - \frac{1}{1+\varepsilon} \right)^t \geq \frac{\varepsilon^t}{t!} \binom{d}{s}^t / \binom{n}{s}^{t-1} \cdot a^t. \end{aligned}$$

□

The next Lemma formalizes the following intuition. A random partition of the vertex set of a graph into two sets of sizes a and b splits the neighborhood of each vertex roughly in proportion $a : b$.

Lemma 4.16 ([13])**.** *For all $\beta, \rho \in (0,1)$, there exists an n_0 such that the following holds. If $n \geq n_0$ and G is an n-vertex graph with $\delta(G) \geq \log^2 n$, then there exists an $A \subseteq V(G)$ with $|A| \in \big((1-\rho)\beta n, (1+\rho)\beta n\big)$ such that for all $v \in V(G)$,*

$$(1-\rho)\beta \cdot \deg(v) \leq \deg(v, A) \leq (1+\rho)\beta \cdot \deg(v).$$

Proof. Let us pick, randomly and independently, each vertex of G with probability β. Let A be the set of selected vertices. By Theorem 1.2,

$$P\big(\big||A| - \beta n\big| \geq \rho\beta n\big) \leq 2e^{-2\rho^2\beta^2 n},$$

Similarly, for every vertex v,

$$P(|\deg(v,A) - \beta \deg(v)| \geq \rho\beta\deg(v)) \leq 2e^{-2\rho^2\beta^2\deg(v)} \leq 2e^{-\rho^2\beta^2\delta(G)} \leq 2e^{-\rho^2\beta^2\log^2 n}.$$

By the union bound, the set A has all the required properties with probability tending to 1 as n tends to infinity. Hence, provided that n is large enough, there exists a set A satisfying all the required conditions. □

Finally, we need the following well-known estimate relating binomial coefficients with the binary entropy function (see, e.g., [61, Lemma 9]).

Lemma 4.17. *If $0 \le k \le n$, then*

$$\frac{1}{n+1} \cdot 2^{n \cdot H(k/n)} \le \binom{n}{k} \le 2^{n \cdot H(k/n)}.$$

4.4 Proof of Theorem 4.2

Let G be a $K_{s,t}$-free graph of order n, let v be a vertex of minimum degree in G, and let $G' = G - \{v\}$. Clearly, the graph G' is $K_{s,t}$-free and $\delta(G') \ge \delta(G) - 1 = d_G(v) - 1$. It easily follows that one can find an ordering $v_1, \ldots, v_n$ of $V(G)$ such that if we let $G_i = G[\{v_1, \ldots, v_i\}]$, then

$$\delta(G_i) \ge d_{G_{i+1}}(v_{i+1}) - 1 \quad \text{for all} \quad i \in \{1, \ldots, n-1\}.$$

In other words, every n-vertex $K_{s,t}$-free graph can be obtained from a single vertex by successively adjoining a vertex of degree $d+1$ to a graph with minimum degree at least d, for some d. The general idea of the proof is showing that the number of ways in which one can obtain a $K_{s,t}$-free graph of order $i+1$ from some i-vertex $K_{s,t}$-free graph in the above process of adjoining vertices of minimum degree is $2^{O(i^{1-1/s})}$, and therefore the number of labeled $K_{s,t}$-free graphs of order n is at most

$$f_n(K_{s,t}) \le n! \cdot \prod_{i=1}^{n-1} 2^{O(i^{1-1/s})} = 2^{O(n^{2-1/s})}.$$

We start by introducing some notation. For a fixed n-vertex $K_{s,t}$-free graph G, let $f(G; K_{s,t})$ denote the number of ways we can extend G to a $K_{s,t}$-free graph of order $n+1$ by adjoining to G a new vertex of degree at most $\delta(G)+1$. Then, we let

$$f(n; K_{s,t}) = \sup_G f(G; K_{s,t}),$$

where the supremum is taken over all $K_{s,t}$-free graphs with n vertices.

The core of the proof is description and analysis of an algorithm that encodes the aforementioned one-vertex extensions in an economical way, i.e., using only few bits. Precisely, we will achieve the following goal.

Goal. Construct an algorithm $\mathcal{A}$ meeting the following specification:

- INPUT: An n-vertex $K_{s,t}$-free graph G and a set $N \subseteq V(G)$ of size at most $\delta(G)+1$ such that the addition of a new vertex v with $N(v) = N$ yields a $K_{s,t}$-free graph of order $n+1$.
- OUTPUT: A bitstring of length at most $(1+o(1))(t-1)^{1/s} C_s \cdot n^{1-1/s}$ that uniquely encodes N.

By saying that $\mathcal{A}$ uniquely encodes N, we mean that there is another algorithm $\mathcal{B}$, which given G and $\mathcal{A}(G, N)$ – the code of N in G produced by $\mathcal{A}$ – outputs N. Although we will not explicitly construct such $\mathcal{B}$, it will become clear that one can obtain such an algorithm by slightly modifying $\mathcal{A}$. In particular, the existence of such coding and decoding procedures implies that for a fixed $K_{s,t}$-free graph G, the map $N \mapsto \mathcal{A}(G, N)$ is an injection of the set of all possible $K_{s,t}$-free extensions of G by a single vertex of degree at most $\delta(G)+1$ into a set of size $2^{(1+o(1))(t-1)^{1/s}n^{1-1/s}}$. It then follows that for every n-vertex $K_{s,t}$-free graph G,

$$f(G; K_{s,t}) \le 2^{(1+o(1))(t-1)^{1/s}C_s \cdot n^{1-1/s}},$$

and hence

$$\begin{aligned}\log_2 f_n(K_{s,t}) &\le \log_2 \left(n! \cdot \prod_{i=1}^{n-1} f(i; K_{s,t}) \right) \le (1+o(1))(t-1)^{1/s}C_s \cdot \sum_{i=1}^{n-1} i^{1-1/s} \\ &= (1+o(1))\frac{s(t-1)^{1/s}}{2s-1} C_s \cdot n^{2-1/s}.\end{aligned}$$

In the remainder of the proof we will describe and analyze an algorithm that meets our requirements. To begin with, let us fix some valid input for $\mathcal{A}$, i.e., an n-vertex $K_{s,t}$-free graph G and a set $N \subseteq V(G)$ with $|N| \le \delta(G)+1$ such that making a new vertex v adjacent to all of N yields a $K_{s,t}$-free graph G' of order $n+1$. Furthermore, let $d = |N|-1$ and note that by our assumption $\delta(G) \ge d$.

Since $|V(G)| = n$, we may clearly assume that there is an injective mapping of $V(G)$ into the set $\{0,1\}^{\lceil \log_2 n \rceil}$ or, in other words, a distinct $\lceil \log_2 n \rceil$-bit code for each vertex of G. To simplify notation, from now on we will generally not distinguish vertices of G from their codes.

For the most part, the output of our algorithm will be a sequence of vertex codes interwined with numbers and short 'control sequences' (strings like `LOW DEGREE VERTEX`, `PREPROCESSING`, etc.) coming from a constant sized set. Since all the numbers involved will come from the set $\{0, \dots, n-1\}$, to avoid confusion, let us agree that by outputting a number we will mean outputting its unique code of fixed length $\lceil \log_2 n \rceil$, e.g., the binary representation of the number. Same convention applies to 'control sequences' – each of them will be assigned a unique code of length $\lceil \log_2 n \rceil$.

Recall that $d+1 = |N|$ and $\delta(G) \ge d$. If $d \le n^{1-1/s}/\lceil \log_2 n \rceil$, then $\mathcal{A}$ will simply output `LOW DEGREE VERTEX`, followed by the list of all elements of N in an arbitrary order. Clearly, the length of the output string is precisely

$$(d+2)\lceil \log_2 n \rceil \le (1+o(1))n^{1-1/s}.$$

After having handled the easy case, for the remainder of this section we will restrict our attention to the more interesting case $d > n^{1-1/s}/\lceil \log_2 n \rceil$. Since G', which we recall is the graph obtained from G by adjoining v to the vertices in N, is $K_{s,t}$-free, whenever a t-set $D \subseteq V(G)$ is the larger partite set in a copy of $K_{s-1,t}$ in G, N does not contain D, i.e., $|N \cap D| \le t-1$. Since $d \ge n^{1-1/s}/(2\log n) \gg n^{1-1/(s-1)}$, Lemma 4.15 implies that G contains many copies of $K_{s-1,t}$. In vague terms, this means that N cannot be an arbitrary subset of $V(G)$, but is very restricted, and hence its entropy is much lower than $\log_2 \binom{n}{d+1}$. Below we try to make this intuition precise. For the sake of brevity, let us first introduce the following definition.

Definition 4.18. A t-set $D \subseteq V(G)$ is *dangerous* if $|N^*(D)| = s-1$, i.e., D is the larger partite set in a copy of $K_{s-1,t}$ in G. In other words, a t-set D is *dangerous* if and only if $D \subseteq N^*(U)$ for some $(s-1)$-set $U \subseteq V(G)$.

The starting point in designing of the algorithm are the following three simple observations and an estimate on the number of dangerous sets.

Observation 4.19. *No dangerous set is fully contained in N.*

Observation 4.20. *Let $U \subseteq V(G)$ be an arbitrary $(s-1)$-set of vertices. Then $|N \cap N^*(U)| \leq t-1$.*

Observation 4.21. *Let $W \subseteq V(G)$ be an arbitrary s-set of vertices. Then $|N^*(W)| \leq t-1$ and hence $N^*(W)$ contains at most $\binom{t-1}{s-1}$ different $(s-1)$-subsets.*

Lemma 4.22. *Fix some positive ε and let A be any set of a vertices in G with $a \geq (1+\varepsilon)(t-1)\binom{n}{s-1}/\binom{d}{s-1}$. There is a d_0 such that for all d with $d \geq d_0$, the number $D(A)$ of dangerous t-sets in A satisfies*

$$D(A) \geq \alpha \cdot a^t,$$

where

$$\alpha = \alpha(s,t,d,\varepsilon) = \frac{\varepsilon^t}{s!t!} \cdot \frac{d^{(s-1)t}}{n^{(s-1)(t-1)}}.$$

Proof. Since G is $K_{s,t}$-free, every dangerous t-set is the larger partite set of exactly one copy of $K_{s-1,t}$ in G, and therefore by Lemma 4.15,

$$D(A) = N_{s-1,t}(A) \geq \beta(s-1,t,d,\varepsilon) \cdot a^t,$$

where $\beta(s-1,t,d,\varepsilon)$ is defined in the statement of Lemma 4.15. It suffices to prove that $\beta \geq \alpha$. First let us observe that

$$\lim_{d\to\infty} (1 - s/d)^{(s-1)t} = 1,$$

and hence there is a d_0 such that if $d \geq d_0$, then

$$s \cdot (d-s)^{(s-1)t} \geq d^{(s-1)t}.$$

It follows that if $d \geq d_0$, then

$$\begin{aligned} \beta = \frac{\varepsilon^t}{t!}\binom{d}{s-1}^t \bigg/ \binom{n}{s-1}^{t-1} &\geq \frac{\varepsilon^t}{t!} \cdot \left(\frac{(d-s)^{s-1}}{(s-1)!}\right)^t \cdot \left(\frac{(s-1)!}{n^{s-1}}\right)^{t-1} \\ &\geq \frac{\varepsilon^t}{t!} \cdot \frac{d^{(s-1)t}}{s(s-1)!n^{(s-1)(t-1)}} = \alpha. \end{aligned}$$

□

Next, let us sketch the rough idea of how our algorithm works. Although this description is not very formal or precise, and misses out a lot of technical details, we hope that it will make the understanding of the pseudocode of $\mathcal{A}$ somewhat clearer.

At all times $\mathcal{A}$ will maintain a list of already encoded elements of N (neighbors of v), denoted by E, and a set A containing the remaining neighbors – the set $N - E$. We will refer to A as the set of *eligible* vertices and E – the set of *already encoded* vertices. Our goal will be to shrink the eligible set A as much as we can, without growing E too much at the same time. Since, as we will later see, encoding one element of E requires approximately $\log_2 n$ bits, at all times we can encode the entire set N using roughly $|E|\log_2 n + \log_2 \binom{|A|}{|N-E|}$ bits. Once we are done shrinking A, this number will be small enough for our purposes.

Before we proceed with the explanation, let us define a few parameters. Let

$$\varepsilon = \varepsilon(n) = 1/\log n, \quad \omega = \omega(n) = (\log n)^3 \quad \text{and} \quad b = d^{\frac{t-s}{t-s+1}}.$$

The target size of the eligible set A, i.e., the maximum number of elemets we would like A to have at the very end, is a_0, which we define by

$$a_0 = (1+\varepsilon)(t-1)\binom{n}{s-1}\Big/\binom{d}{s-1}. \tag{4.5}$$

Note that a_0 is the lower bound on the cardinality of a set A that surely contains a lof of dangerous t-subsets (see Lemma 4.22).

The algorithm works in steps. Each step starts with preprocessing of the eligible set A, i.e., a procedure which makes sure that A is 'well-behaved' in terms of the sizes of its intersections with common neighborhoods of $(s-1)$-element sets of vertices. In step 3a we simply remove from A all such common neighborhoods that are larger than $\omega|A|/d$, and encode all neighbors of v (elements of N) which those large neighborhoods contain (Observation 4.20 says that there are at most $t-1$ neighbors of v in each such neighborhood). This will be of extremely high importance later.

Next, in step 3c, we pick out a carefully chosen sequence of subsets $E_t, \ldots, E_{s+1} \subseteq N - E$ of size b each that we encode and move to E. At the same time we construct a sequence of hypergraphs $\mathcal{H}_t, \ldots, \mathcal{H}_s$, where each $\mathcal{H}_r$ is an r-uniform hypergraph on the vertex set A with

$$E(\mathcal{H}_r) \subseteq \big\{D \subseteq A : \{w_t, \ldots, w_{r+1}\} \cup D \text{ is dangerous for some } w_t \in E_t, \ldots, w_{r+1} \in E_{r+1}\big\}.$$

Our ultimate goal in the second part of each step, the for loop 3c, is to maximize the number of edges in $\mathcal{H}_s$. Since edges of the r-uniform hypergraph $\mathcal{H}_r$ are neighbors in the $(r+1)$-uniform hypergraph $\mathcal{H}_{r+1}$ of vertices from E_{r+1}, i.e., $D \in E(\mathcal{H}_r)$ if $D \cup \{w\} \in E(\mathcal{H}_{r+1})$ for some $w \in E_{r+1}$, we try to achieve this goal by maximizing $e(\mathcal{H}_r)$ in turn for all $r \in \{t-1, \ldots, s\}$. In order to do that, we try to add to E_{r+1} vertices with highest degree in $\mathcal{H}_{r+1}$. Since $E_{r+1} \subseteq N$, our choices are quite limited, as it might happen that very few of the high-degree vertices in $\mathcal{H}_{r+1}$ are members of N. If this is the case, we will not be able to make $e(\mathcal{H}_r)$ very large, but this gives us some information about N that we can use to shrink the eligible set – we simply delete from A all the high-degree non-neighbors of v which we keep listed in the set Y. Finally, notice that having a few vertices outside of N which cover most of the edges of $\mathcal{H}_{r+1}$ would get us into trouble – deleting them from A would not shrink the eligible set enough. We overcome this obstacle by keeping the maximum degree of $\mathcal{H}_r$ bounded – the auxilliary set X serves that purpose.

By definition, the edges of $\mathcal{H}_s$ will have the nice property that none of them is fully contained in N. At the end of each step, in the for loop 3d, we will exploit this fact to shrink the eligible set by working with $\mathcal{H}_s$. The rough idea is the following. Either many vertices of N have high degree in $\mathcal{H}_s$, and hence some $(s-1)\cdot b$ of them almost-cover many edges (meaning that $|N \cap D| = s-1$) – this information allows us to remove all the uncovered vertices in these almost-covered edges from A, or very few of the high-degree vertices in $\mathcal{H}_s$ are members of N – we delete from A all the high-degree non-neighbors of v.

After this lengthy introduction, we present the algorithm in the 'high-degree' case, i.e., when $d > n^{1-1/s}/(2\log n)$:

1. Output "`HIGH DEGREE VERTEX`".
2. Set $A = V(G)$ and $E = \emptyset$.
3. While $|A| > a_0$, do the following:
   (a) If there exists an $(s-1)$-set $U \subseteq V(G)$ with $|N^*(U) \cap A| > \omega|A|/d$, do the following:
      i. Let $U = \{u_1, \dots, u_{s-1}\}$ and $N^*(U) \cap N = \{w_1, \dots, w_k\}$.
      ii. Set $A = A - N^*(U)$ and $E = E \cup \{w_1, \dots, w_k\}$.
      iii. Output "`PREPROCESSING :` u_1, ..., u_{s-1}, k, w_1, ..., w_k" and go to step 3.
   (b) Let $\mathcal{H}_t = \{D \subseteq A : |D| = t$ and D is dangerous$\}$.
   (c) For $r = t-1, \dots, s$, do the following:
      i. Set $E_{r+1} = \emptyset, X = \emptyset$ and $Y = \emptyset$.
      ii. Let $\mathcal{H}_r$ be an empty r-uniform hypergraph on A.
      iii. For $i = 1, \dots, b$, do the following:
         - List all vertices in $A - X - Y$ as $w_1^i, \dots, w_{|A-X-Y|}^i$, so that for each j, if we let $W_j^i = \{w_1^i, \dots, w_j^i\}$, then the vertex w_{j+1}^i is the vertex with the minimum label among all vertices in $A - X - Y - W_j^i$ maximizing $\deg_{\mathcal{H}_{r+1}[A-X-Y-W_j^i]}(w_{j+1}^i)$.
         - Let j_i be the smallest j such that $w_j^i \in N$.
         - $\mathcal{H}_r = \mathcal{H}_r \cup \big\{D : \{w_{j_i}^i\} \cup D \in \mathcal{H}_{r+1}[A - X - Y - W_{j_i-1}^i]\big\}$.
         - Set $E_{r+1} = E_{r+1} \cup \{w_{j_i}^i\}$ and $Y = Y \cup W_{j_i}^i$.
         - Set $X = X \cup \{w \in A : \deg_{\mathcal{H}_r}(w) > b^{t-r} d^{s-t} |A|^{r-1}\}$.
      iv. Set $E = E \cup E_{r+1}$.
      v. Suppose the vertices added to E_{r+1} were $w_1, \dots, w_b$. Output "w_1, ..., w_b".
      vi. If $|Y| \geq \sigma(\mathcal{H}_{r+1})/2$, then set $A = A - Y$, output "`SKIP`" and go to step 3.
   (d) For $i = 1, \dots, b$, do the following:
      i. For $r = s-1, \dots, 1$, do the following:
         - List all vertices in A as $w_1^r, \dots, w_{|A|}^r$, so that for each j, if we let $W_j^r = \{w_1^r, \dots, w_j^r\}$, then the vertex w_{j+1}^r is the vertex with the minimum label among all vertices in $A - W_j^r$ maximizing $\deg_{\mathcal{H}_{r+1}[A-W_j^r]}(w_{j+1}^r)$.
         - Let j_r be the smallest j such that $w_j \in N$.
         - Set $A = A - W_{j_r}^r$ and $E = E \cup \{w_{j_r}^r\}$.
         - Set $\mathcal{H}_r = \big\{D \subseteq A : \{w_{j_r}^r\} \cup D \in \mathcal{H}_{r+1}\big\}$.
      ii. Let $A = A - \big\{w : \{w\} \in E(\mathcal{H}_1)\big\}$.
      iii. Output "$w_{j_{s-1}}^{s-1}$, ..., $w_{j_1}^1$".
4. Let $N' = N - E$. Clearly $N' \subseteq A$. The set N' is one of the $\binom{|A|}{|N'|}$ different $|N'|$-subsets of A. Output "`REMAINDER :` $|A|$, $|N'|$", followed by a $\lceil \log_2 \binom{|A|}{|N'|} \rceil$-bit code of N' in A.

For the remainder of this discussion, let us fix G and N with $d = |N| - 1 \geq n^{1-1/s}/(2\log n)$ and assume that we run $\mathcal{A}$ on the pair (G, N). Note that given G and the output $\mathcal{A}(G, N)$, one can reconstruct N. The key observation that reassures us that it is possible, is noting that the final sets A and E can be recomputed step-by-step in the exact same way as they were computed by $\mathcal{A}$ – all the necessary information about N appears in $\mathcal{A}(G, N)$. Once we reconstruct A and E, we can easily decode $N = N' \cup E$ using the last fragment of $\mathcal{A}(G, N)$ starting with `REMAINDER`.

The non-trivial part of the analysis is proving an $O(n^{1-1/s})$ bound on the size of the output of $\mathcal{A}$. Recall that our aim is to prove that

$$|\mathcal{A}(G, N)| \leq (1 + o(1))(t-1)^{1/s} C_s \cdot n^{1-1/s}.$$

We start by looking at the preprocessing stage. Let p denote the total number of times $\mathcal{A}$ preprocesses the eligible set A, i.e., the number of times an appropriate $(s-1)$-set U is found in step 3a.

Claim 4.23. *The total number of preprocessing steps satisfies* $p \leq \frac{d \log n}{\omega} + 1$.

Proof. Each time $\mathcal{A}$ preprocesses the eligible set, A loses more than $\omega|A|/d$ elements. Hence, preprocessing the eligible set q times shrinks it by a factor of at most

$$\left(1 - \frac{\omega}{d}\right)^q \leq e^{-q\frac{\omega}{d}}.$$

Since $\mathcal{A}$ starts with $|A| = n$ and after $p - 1$ preprocessing steps A is still non-empty, $(p-1)\frac{\omega}{d} \leq \log n$. ∎

Let α be as in the definition in Lemma 4.22. Moreover, for each $r \in \{t, \ldots, s-1\}$, let

$$B_r = \left(4t\binom{t-1}{s-1}\right)^{r-t} \quad \text{and} \quad D_r = 3(t-s)(t-r). \tag{4.6}$$

The core of our analysis is the following lemma.

Lemma 4.24. *Suppose that during some iteration of the main* while *loop, step* 3*,* $\mathcal{A}$ *does not preprocess the eligible set. Then during that iteration, the eligible set* A *loses at least*

$$\frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \cdot |A|$$

elements.

Let z be the total number of times $\mathcal{A}$ does not preprocess the eligible set A in an iteration of the main while loop. The following corollary is an immediate consequence of Lemma 4.24.

Corollary 4.25. *The total number of times* $\mathcal{A}$ *executes the main* while *loop without preprocessing* A *satisfies*

$$z \leq \frac{(\log n)^{D_{s-1}+1}}{B_{s-1}} \cdot d^{s-t}\alpha^{-1} + 1.$$

Proof. By Lemma 4.24, during each iteration of the main while loop, in which $\mathcal{A}$ does not preprocess the eligible set, A loses at least

$$\frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \cdot |A|$$

elements. Hence, as a result of q such iterations, the eligible set shrinks by a factor of at most

$$\left(1 - \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha\right)^q \leq \exp\left(-q \cdot \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha\right).$$

Since $\mathcal{A}$ starts with $|A| = n$ and after $z-1$ such iterations A is still non-empty,

$$(z-1) \cdot \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \leq \log n.$$

□

Before we dive into the proof of Lemma 4.24, let us show how Corollary 4.25 implies that $\mathcal{A}$ outputs short codes.

Lemma 4.26. *For every input pair (G, N), the length of the output produced by $\mathcal{A}$ does not exceed*

$$(1 + o(1))(t-1)^{1/s} C_s \cdot n^{1-1/s}. \tag{4.7}$$

Proof. Note that by Observation 4.20, the k in the preprocessing step 3a never exceeds $t-1$. Hence the total length of the output produced by $\mathcal{A}$ in step 3a is at most

$$p \cdot (1 + (s-1) + 1 + (t-1)) \cdot \lceil \log_2 n \rceil \leq \left(\frac{d \log n}{\omega} + 1\right) \cdot (s+t)\lceil \log_2 n \rceil, \tag{4.8}$$

where the bound on the number p of preprocessing steps comes from Claim 4.23. Since $\omega = (\log n)^3 \gg \log n \cdot \lceil \log_2 n \rceil$, the quantity in the right-hand side of (4.8) is clearly $o(d)$.

Each of the z executions of the main while loop with no preprocessing outputs either codes of at most $(t-1)b$ vertices or codes of at most $(t-s)b$ vertices and the SKIP control sequence. Either way, this is never more than $tb\lceil \log_2 n \rceil$ bits. Therefore the total length of the output produced by $\mathcal{A}$ in steps 3c and 3d is at most

$$z \cdot tb\lceil \log_2 n \rceil \leq \left(\frac{(\log n)^{D_{s-1}+1}}{B_{s-1}} \cdot d^{s-t}\alpha^{-1} + 1\right) \cdot tb\lceil \log_2 n \rceil, \tag{4.9}$$

where the above inequality comes from Corollary 4.25. Recall that $\varepsilon = 1/\log n$, and we are in the 'high-degree' case, i.e., $d > n^{1-1/s}/(2\log n)$. Therefore,

$$d^{s-t}\alpha^{-1} = s!t! \cdot (\log n)^t \cdot \frac{n^{(s-1)(t-1)}}{d^{s(t-1)}} \leq 2^{s(t-1)} s!t! \cdot (\log n)^{(s+1)t-s}, \tag{4.10}$$

and hence the right-hand side of (4.9) is bounded above by $g(n) \cdot b$, where g is polylogarithmic in n, and this clearly is $o(d)$.

When $\mathcal{A}$ finally reaches step 4, $|A| \leq a_0$, and hence the total length of the output produced by $\mathcal{A}$ in step 4 is

$$\begin{aligned} 3\lceil \log_2 n \rceil + \left\lceil \log_2 \binom{a_0}{|N'|} \right\rceil &\leq 4\log_2 n + \log_2 \binom{a_0 + |E|}{|N'| + |E|} = 4\log_2 n + \log_2 \binom{a_0 + |E|}{|N|} \\ &\leq 5\log_2 n + \log_2 \binom{a_0 + |E|}{d}. \end{aligned} \tag{4.11}$$

Next, note that for n large enough,

$$a_0 = (1+\varepsilon)(t-1)\binom{n}{s-1}\Big/\binom{d}{s-1} \le (1+\varepsilon)(t-1)\frac{n^{s-1}}{(d-s)^{s-1}} \le (1+2\varepsilon)(t-1)\left(\frac{n}{d}\right)^{s-1}. \quad (4.12)$$

Since $\operatorname{ex}(n, K_{s,t}) = O(n^{2-1/s})$,

$$d \le \frac{2e(G')}{n+1} \le \frac{2\operatorname{ex}(n+1, K_{s,t})}{n+1} = O(n^{1-1/s}).$$

Recall that E gains at most $t-1$ elements in each of the p preprocessing steps, and at most $(t-1)b$ elements in each of the z non-preprocessing steps. From (4.8), (4.9) and (4.10) it follows that

$$|E| \le p(t-1) + z(t-1)b = O\left(\frac{d}{(\log n)^2}\right) = O\left(\frac{1}{(\log n)^2}\left(\frac{n}{d}\right)^{s-1}\right). \quad (4.13)$$

Recall again that $\varepsilon = 1/\log n$. Inequalities (4.12) and (4.13) imply that

$$a_0 + |E| \le (1+3\varepsilon)(t-1)\left(\frac{n}{d}\right)^{s-1}. \quad (4.14)$$

Using (4.11), (4.14), and Lemma 4.17, we further estimate

$$\begin{aligned} 3\lceil \log_2 n\rceil + \left\lceil \log_2 \binom{a_0}{|N'|}\right\rceil &\le 5\log_2 n + \log_2 \binom{(1+3\varepsilon)(t-1)\left(\frac{n}{d}\right)^{s-1}}{d} \quad (4.15)\\ &\le 5\log_2 n + (1+3\varepsilon)(t-1)(n/d)^{s-1}\cdot H\left(\frac{d^s}{(1+3\varepsilon)(t-1)n^{s-1}}\right). \end{aligned}$$

Substituting $x = d^s/\big((1+3\varepsilon)(t-1)n^{s-1}\big)$ in (4.15) yields

$$3\lceil \log_2 n\rceil + \left\lceil \log_2 \binom{a_0}{|N'|}\right\rceil \le 5\log_2 n + \big((1+3\varepsilon)(t-1)\big)^{1/s}\cdot \frac{H(x)}{x^{1-1/s}}\cdot n^{1-1/s}. \quad (4.16)$$

Recall that $C_s = \sup_x \left(H(x)/x^{1-1/s}\right)$. Since the total size of the output is bounded by the sum of the quantities in the right hand sides of (4.8), (4.9), and (4.16), we get (4.7). □

Before we are able to prove Lemma 4.24, we need to make some preparations. For the sake of brevity, by i^{th} *iteration* of any for loop, we will denote the iteration, where the loop variable takes value i. The following claim explains why $\mathcal{A}$ constantly preprocesses the eligible set and maintains the oddly defined set X.

Claim 4.27. *Assume that during some iteration of the main* while *loop, step* 3*, at the time we reach step* 3c*, the eligible set A has a elements. Then throughout this iteration, for all $r \in \{s, \dots, t\}$,*

$$\Delta(\mathcal{H}_r) \le \binom{t-1}{s-1} b^{t-r}\left(\frac{\omega}{d}\right)^{t-s}\cdot a^{r-1}.$$

Proof. First observe that at all times during any iteration of the main while loop, for all $r \in \{s, \dots, t\}$, the edges of $\mathcal{H}_r$ all come from the set

$$\big\{D \subseteq A : \{w_t, \dots, w_{r+1}\} \cup D \text{ is dangerous for some } w_t \in E_t, \dots, w_{r+1} \in E_{r+1}\big\}.$$

Consider first the case $r \geq t-s+1$. Fix some $w_r, \ldots, w_{t-s+1} \in A$, let $W = \{w_r, \ldots, w_{t-s+1}\}$ and note that by our assumption, $|W| = r+s-t \geq 1$. The set W is contained in some $D \in \mathcal{H}_r$ only if there are $w_t \in E_t, \ldots, w_{r+1} \in E_{r+1}$ and an $(s-1)$-set $U \subseteq V(G)$ such that $\{w_t, \ldots, w_{r+1}\} \cup D \subseteq N^*(U) \cap A$. Let $W' = \{w_t, \ldots, w_{t-s+1}\}$. Then clearly $|W'| = s$ and $N^*(U) \supseteq W'$, or equivalently, $N^*(W') \supseteq U$. Hence by Observation 4.21, with W' fixed, there are at most $\binom{t-1}{s-1}$ such sets U. Moreover, since $|E_{r'}| \leq b$ for all r', the number of such W' that contain our fixed set W is at most $b^{|W'|-|W|} = b^{t-r}$. Also, because A is preprocessed, for every $(s-1)$-set U, $|N^*(U) \cap A| \leq \omega a/d$. Putting all these inequalities together,

$$\deg_{\mathcal{H}_r}(W) \leq b^{t-r} \cdot \binom{t-1}{s-1} \cdot \left(\frac{\omega a}{d}\right)^{r-|W|} = \binom{t-1}{s-1} b^{t-r} \left(\frac{\omega a}{d}\right)^{t-s}. \tag{4.17}$$

Since the term in the right-hand side of (4.17) does not depend on a particular choice of W, it follows that

$$\Delta_{r+s-t}(\mathcal{H}_r) \leq \binom{t-1}{s-1} b^{t-r} \left(\frac{\omega}{d}\right)^{t-s} \cdot a^{t-s},$$

and hence by (1.1),

$$\Delta(\mathcal{H}_r) \leq \binom{t-1}{s-1} b^{t-r} \left(\frac{\omega}{d}\right)^{t-s} \cdot a^{r-1}.$$

The case $s \leq r \leq t-s$ is much more delicate. First, consider how much the $\mathcal{H}_r$-degree of a vertex $w \in A$ can increase in one particular, say i^{th}, iteration of the for loop 3(c)iii. Since all edges added in the i^{th} iteration contain $w^i_{j_i}$, $\deg_{\mathcal{H}_r}(w)$ increases by no more than the number of edges $D \in \mathcal{H}_{r+1}$ containing both w and $w^i_{j_i}$. In order for such an $(r+1)$-set D to be an edge of $\mathcal{H}_{r+1}$, there ought to be some $w_t \in E_t, \ldots, w_{r+2} \in E_{r+2}$ and an $(s-1)$-set $U \subseteq V(G)$ such that $\{w_t, \ldots, w_{r+2}\} \cup D \subseteq N^*(U) \cap A$. Note that $r+2 \leq t-s+3$ by our assumption on r, and let $W' = \{w_t, \ldots, w_{t-s+3}, w^i_{j_i}, w\}$. Then clearly, $|W'| = s$ and $N^*(U) \supseteq W'$, or equivalently, $N^*(W') \supseteq U$. Hence by Observation 4.21, with W' fixed, there are at most $\binom{t-1}{s-1}$ such sets U. Moreover, since $|E_{r'}| \leq b$ for all r', the number of such W' that contain both $w^i_{j_i}$ and w is at most $b^{|W'|-2} = b^{s-2}$. Also, because A is preprocessed, for every $(s-1)$-set U, $|N^*(U) \cap A| \leq \omega a/d$. Putting all these inequalities together, we see that $\deg_{\mathcal{H}_r}(w)$ cannot change by more than

$$\begin{aligned} b^{s-2} \cdot \binom{t-1}{s-1} \cdot \left(\frac{\omega a}{d}\right)^{|D|-|\{w,w^i_{j_i}\}|} &= b^{s-2} \cdot \binom{t-1}{s-1} \cdot \left(\frac{\omega a}{d}\right)^{r-1} \\ &= \omega^{r-1} \binom{t-1}{s-1} \frac{d^{t-s-r+1}}{b^{t-s-r+2}} \cdot b^{t-r} d^{s-t} \cdot a^{r-1}. \end{aligned} \tag{4.18}$$

Recall that $b = d^{\frac{t-s}{t-s+1}}$ and note that the right-hand side of (4.18) is $o(b^{t-r} d^{s-t} \cdot a^{r-1})$, as

$$\frac{d^{t-s-r+1}}{b^{t-s-r+2}} = \frac{1}{d}\left(\frac{d}{b}\right)^{t-s-r+2} = \frac{1}{d} \cdot \left(d^{\frac{1}{t-s+1}}\right)^{t-s-r+2} = \left(\frac{1}{d}\right)^{\frac{r-1}{t-s+1}},$$

$\frac{r-1}{t-s+1} \geq \frac{s-1}{t-s+1} > 0$, $d \gg n^{1/3}$, and $\omega^{r-1}\binom{t-1}{s-1}$ is only polylogarithmic in n.

Every time a vertex $w \in A$ lands in the set X of vertices with high degree in $\mathcal{H}_r$, no more edges containing w get added to $\mathcal{H}_r$, since $\mathcal{A}$ looks only at the edges of $\mathcal{H}_{r+1}[A - X - Y]$. Hence the degree of w cannot exceed $b^{t-r} d^{s-t} a^{r-1}$, i.e., the quantity from the definition of X, by more than the right-hand side of (4.18). It follows that

$$\Delta(\mathcal{H}_r) \le (1+o(1))b^{t-r}d^{s-t} \cdot a^{r-1} \le \binom{t-1}{s-1} b^{t-r} \left(\frac{\omega}{d}\right)^{t-s} \cdot a^{r-1}.$$

□

The second key step in proving Lemma 4.24 is the following simple estimate on the number of edges in $\mathcal{H}_r$.

Claim 4.28. *Assume that during some iteration of the main* while *loop, at the time we reach* 3b*, the eligible set A has a elements. Then for every r, with $t-1 \ge r \ge s$, either $\mathcal{A}$ outputs* `SKIP` *by the end of the r^{th} iteration of the* for *loop* 3c *or*

$$e(\mathcal{H}_r) \ge \frac{B_r}{(\log n)^{D_r}} b^{t-r} \alpha \cdot a^r. \tag{4.19}$$

Proof. By Lemma 4.22, (4.19) clearly holds when $r = t$. It suffices to prove that if (4.19) holds for $r+1$ and $\mathcal{A}$ does not output `SKIP` in the r^{th} iteration of the for loop 3c, then (4.19) holds for r.

Let x and y be the sizes of X and Y at the end of the r^{th} iteration. There are two cases to consider.

Case 1. $x + y \ge \sigma(\mathcal{H}_{r+1})$.
Since $\mathcal{A}$ did not output `SKIP`, $y < \sigma(\mathcal{H}_{r+1})/2$, and hence by (1.2),

$$x > \frac{\sigma(\mathcal{H}_{r+1})}{2} \ge \frac{e(\mathcal{H}_{r+1})}{4\Delta(\mathcal{H}_{r+1})} \ge \frac{B_{r+1}}{(\log n)^{D_{r+1}}} \cdot \frac{d^{t-s}}{4\binom{t-1}{s-1}\omega^{t-s}} \alpha \cdot a,$$

where the last inequality follows from Claim 4.27 and the assumption that (4.19) holds for $r+1$. Recall that for every $w \in X$, $\deg_{\mathcal{H}_r}(w) > b^{t-r}d^{s-t}a^{r-1}$. Clearly,

$$e(\mathcal{H}_r) \ge \frac{1}{r}\sum_{w\in X} \deg_{\mathcal{H}_r}(w) > \frac{x}{r} \cdot b^{t-r}d^{s-t}a^{r-1} \ge \frac{B_r}{(\log n)^{D_r}} \cdot b^{t-r}\alpha \cdot a^r,$$

since $B_r \le (4r)^{-1}\binom{t-1}{s-1}^{-1} \cdot B_{r+1}$, $D_r = D_{r+1} + 3(t-s)$ and $\omega = (\log n)^3$.

Case 2. $x + y < \sigma(\mathcal{H}_{r+1})$.
In particular, for all $i \in \{1, \ldots, b\}$, during the r^{th} iteration of the for loop 3c,

$$e(\mathcal{H}_{r+1}[A - X - Y - W^i_j]) \ge e(\mathcal{H}_{r+1})/2.$$

By the maximality of $\deg_{\mathcal{H}_{r+1}[A-X-Y-W^i_{j_i-1}]}(w^i_{j_i})$, in each iteration of the for loop 3(c)iii, $\mathcal{H}_r$ acquires at least

$$\frac{r \cdot e(\mathcal{H}_{r+1})/2}{|A - X - Y - W^i_{j_i-1}|} \ge \frac{e(\mathcal{H}_{r+1})}{a} \ge \frac{B_{r+1}}{(\log n)^{D_{r+1}}} \cdot b^{t-r-1}\alpha \cdot a^r$$

edges. Unfortunately, some edges may get added to $\mathcal{H}_r$ more than once. How many times can we add to $\mathcal{H}_r$ the same edge D? For each $i \in \{1, \ldots, b\}$, let $D_i = D \cup \{w^i_{j_i}\}$. The set D becomes an edge of $\mathcal{H}_r$ precisely when $D_i \in \mathcal{H}_{r+1}$ for some i. In particular, every such D_i is fully contained in some dangerous set, and hence there must be an $(s-1)$-set $U \subseteq V(G)$ such that $D \subseteq D_i \subseteq N^*(U)$.

Since $|D| = r \geq s$, by Observation 4.21, there are at most $\binom{t-1}{s-1}$ such U. Also, because for each i, $w^i_{j_i} \in N$, by Observation 4.20, for no $(s-1)$-set U, $N^*(U)$ contains more than $t-1$ different $w^i_{j_i}$s. It follows that the maximum number of times D can be added to $\mathcal{H}_r$ is $(t-1)\binom{t-1}{s-1}$. Therefore,

$$e(\mathcal{H}_r) \geq \frac{1}{(t-1)\binom{t-1}{s-1}} \cdot b \cdot \frac{B_{r+1}}{(\log n)^{D_{r+1}}} \cdot b^{t-r-1}\alpha \cdot a^r \geq \frac{B_r}{(\log n)^{D_r}} \cdot b^{t-r}\alpha \cdot a^r.$$

□

Lemma 4.29 and its immediate consequence, Corollary 4.30, are the last missing ingredients needed in the proof of Lemma 4.24.

Lemma 4.29. *For every fixed i and r satisfying $1 \leq i \leq b$ and $0 \leq r \leq s-1$, the following holds. Suppose that during the i^{th} iteration of the* for *loop* 3d*, at the beginning of the r^{th} iteration of the* for *loop* 3(d)i*, $e(\mathcal{H}_{r+1}[A]) \geq \gamma a^{r+1}$ for some γ and a with $0 < \gamma \leq 1$ and $a \geq |A|$. Then*

$$e(\mathcal{H}_1) + \sum_{q=1}^{r} j_q \geq \gamma a. \tag{4.20}$$

Proof. For a fixed i, we prove the claim by induction on r. The inequality (4.20) holds with equality when $r = 0$. Suppose that $r > 0$ and (4.20) holds for $r-1$. Each of $w^r_1, \ldots, w^r_{j_r-1}$ clearly belongs to no more than $|A|^r$ $(r+1)$-subsets of A, and hence

$$e(\mathcal{H}_{r+1}[A - W^r_{j_r-1}]) \geq e(\mathcal{H}_{r+1}[A]) - (j_r - 1)|A|^r \geq \gamma a^{r+1} - (j_r - 1)a^r. \tag{4.21}$$

If $j_r \geq \gamma a$, then (4.20) holds, so we may suppose that the reverse inequality is true, and therefore the rightmost term in (4.21) is positive. Since we have selected $w^r_{j_r}$ to maximize its degree in $\mathcal{H}_{r+1}[A - W^r_{j_r-1}]$, we have

$$\begin{aligned} e(\mathcal{H}_r) = \deg_{\mathcal{H}_{r+1}[A-W^r_{j_r-1}]}(w^r_{j_r}) &\geq \frac{r+1}{|A| - j_r + 1} \cdot e(\mathcal{H}_{r+1}[A - W^r_{j_r-1}]) \\ &\geq \frac{r+1}{a - j_r + 1} \cdot (\gamma a - j_r + 1) \cdot a^r \geq \frac{\gamma a - j_r + 1}{a - j_r + 1} \cdot a^r \geq \frac{\gamma a - j_r}{a - j_r} \cdot a^r, \end{aligned}$$

where the last inequality holds since $\gamma \leq 1$, and hence $\gamma a - j_r \leq a - j_r$. By the inductive assumption, with '$\gamma = \frac{\gamma a - j_r}{a - j_r}$',

$$e(\mathcal{H}_1) + \sum_{q=1}^{r-1} j_q \geq \frac{\gamma a - j_r}{a - j_r} \cdot a \geq \gamma a - j_r.$$

□

Corollary 4.30. *Assume that at the beginning of the 1^{st} iteration of the* for *loop* 3d*, A has a elements. If at the beginning of the i^{th} iteration, $e(\mathcal{H}_s[A]) \geq \beta a^s$ for some positive β, then in that iteration A loses at least βa elements.*

Proof. During the i^{th} iteration, we delete from A precisely $e(\mathcal{H}_1) + \sum_{q=1}^{s-1} j_q$ elements. Since certainly $a \geq |A|$, and we have assumed that $e(\mathcal{H}_s) \geq \beta a^s$, the statement of Corollary 4.30 is just a direct application of Lemma 4.29. □

Finally, we are ready to give the proof of Lemma 4.24.

Proof of Lemma 4.24. By Claim 4.28, either at the end of the s^{th} iteration of the for loop 3d,

$$e(\mathcal{H}_s) \geq \frac{B_s}{(\log n)^{D_s}} \cdot b^{t-s}\alpha \cdot |A|^s, \tag{4.22}$$

or for some r with $r \geq s$, $\mathcal{A}$ outputs `SKIP` at the end of the r^{th} iteration. In the latter case, at the end of the r^{th} iteration, $|Y| \geq \sigma(\mathcal{H}_{r+1})/2$. By Claims 4.27 and 4.28,

$$\begin{aligned}|Y| \geq \sigma(\mathcal{H}_{r+1})/2 &\geq \frac{e(\mathcal{H}_{r+1})}{2\Delta(\mathcal{H}_{r+1})} \geq \frac{B_{r+1}}{2\binom{t-1}{s-1}\omega^{t-s}(\log n)^{D_{r+1}}} \cdot d^{t-s}\alpha \cdot a \\ &\geq \frac{B_r}{(\log n)^{D_r}} \cdot d^{t-s}\alpha \cdot a \geq \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \cdot |A|,\end{aligned}$$

and since $\mathcal{A}$ outputs `SKIP`, the eligible set A loses exactly $|Y|$ elements.

Therefore we can assume that (4.22) is true and $\mathcal{A}$ executes the for loop 3d. Similarly as in the proof of Claim 4.28, there are two cases to consider.

Case 1. At the end of the b^{th} iteration of the for loop 3d, $e(\mathcal{H}_s[A]) \geq e(\mathcal{H}_s)/2$.
In particular, this is true in all the previous iterations. Hence, if a is the size of the eligible set A at the beginning of the step 3d, by Corollary 4.30, as a result of a single iteration, A loses at least $\frac{e(\mathcal{H}_s)}{2a^{s-1}}$ elements (apply Corollary 4.30 with $\beta = \frac{e(\mathcal{H}_s)}{2a^s}$). Since there are b iterations, altogether A loses at least (recall that $d^{t-s} = b^{t-s+1}$)

$$b \cdot \frac{e(\mathcal{H}_s)}{2a^{s-1}} \geq \frac{B_s}{2(\log n)^{D_s}} \cdot b^{t-s+1}\alpha \cdot a \geq \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \cdot |A|$$

elements, where the first inequality follows from (4.22).

Case 2. At the end of the b^{th} iteration of the for loop 3d, $e(\mathcal{H}_s[A]) < e(\mathcal{H}_s)/2$.
It means that in the step 3d, A must have lost at least

$$\sigma(\mathcal{H}_s) \geq \frac{e(\mathcal{H}_s)}{2\Delta(\mathcal{H}_s)} \geq \frac{B_s}{(\log n)^{D_s}} \cdot \frac{d^{t-s}}{\binom{t-1}{s-1}\omega^{t-s}}\alpha \cdot a \geq \frac{B_{s-1}}{(\log n)^{D_{s-1}}} d^{t-s}\alpha \cdot |A|$$

elements, where the second inequality follows from (4.22) and Claim 4.27. □

Since in Lemma 4.26 we have already shown how Lemma 4.24 implies that $\mathcal{A}$ outputs short codes, the proof of Theorem 4.2 is now complete.

4.5 Proof of Theorem 4.4

As it was remarked at the beginning of the proof of Theorem 4.2 (Section 4.4), every n-vertex graph G can be constructed from an isolated vertex v_1 by successively connecting a vertex v_{i+1} to some d_i vertices in $G[\{v_1, \ldots, v_i\}]$ in such a way that for all $i \in \{1, \ldots, n-1\}$,

$$d_i = \delta(G[\{v_1, \ldots, v_{i+1}\}]) \leq \delta(G[\{v_1, \ldots, v_i\}]) + 1.$$

Moreover, if G is $K_{s,t}$-free, so are all the intermediate graphs $G[\{v_1, \ldots, v_i\}]$. Call the sequence $(d_i)_{i=1}^{n-1}$ a *degeneracy sequence* of G and note that $e(G) = \sum_{i=1}^{n-1} d_i$.

Let $f(G; d, K_{s,t})$ be the number of ways one can adjoin to a $K_{s,t}$-free graph G, with $\delta(G) \geq d$, a new vertex of degree $d+1$, so that the graph remains $K_{s,t}$-free. If we let

$$f(n; d, K_{s,t}) = \sup_G f(G; d, K_{s,t}),$$

where the supremum is taken over all n-vertex $K_{s,t}$-free graphs whose minimum degree is at least d, then

$$f_{n,m}(K_{s,t}) \leq n! \cdot \sum_{(d_i)} \prod_{i=1}^{n-1} f(i; d_i - 1, K_{s,t}), \tag{4.23}$$

where the above sum is taken over all degeneracy sequences with sum m.

If $d \leq n^{1-\mu_{s,t}} (\log n)^{3t/s}$ and $n \geq n_0$, then we give a rather crude bound

$$f(i; d, K_{s,t}) \leq \binom{i}{d+1} \leq n \binom{n}{d} \leq n \left(\frac{en}{d}\right)^d \leq \exp\left(n^{1-\mu_{s,t}} (\log n)^{3t/s+1}\right). \tag{4.24}$$

Suppose now that $d > n^{1-\mu_{s,t}} (\log n)^{3t/s}$ and let $\alpha(s, t, d, 1/(3t-3))$ be as in Lemma 4.22. Suppose we run the 'high-degree' case in the algorithm $\mathcal{A}$ from the proof of Theorem 4.2 on some i-vertex $K_{s,t}$-free graph G and a set N of size $d+1$, where G and N satisfy our usual assumptions. Note that Claim 4.23 and Corollary 4.25 are still true, since in their proofs we have not used any assumptions on d. Reasoning along the lines of Lemma 4.26, we can see that the total length of the output produced by $\mathcal{A}$ in the preprocessing step 3a is still $o(d)$. Moreover, recall that $b = d^{\frac{s-t}{s-t+1}}$ and hence

$$\begin{aligned} d^{s-t}\alpha^{-1} b &= s!t!(3t-3)^t \cdot \frac{i^{(s-1)(t-1)}}{d^{s(t-1)}} \cdot d^{\frac{t-s}{t-s+1}} \leq s!t!(3t-3)^t \cdot \frac{n^{(s-1)(t-1)}}{d^{s(t-1)+\frac{1}{t-s+1}}} \cdot d \\ &\leq s!t!(3t-3)^t \cdot (\log n)^{-\frac{3t}{s}\cdot\left[s(t-1)+\frac{1}{t-s+1}\right]} \cdot d, \end{aligned} \tag{4.25}$$

where the last inequality follows because $\mu_{s,t}$ satisfies

$$(s-1)(t-1) = (1-\mu_{s,t})\left(s(t-1) + \frac{1}{t-s+1}\right).$$

By (4.9), the total length of the output produced by $\mathcal{A}$ in steps 3c and 3d is at most

$$tb\lceil \log_2 i \rceil + t\left(4t\binom{t-1}{s-1}\right)^{t-s+1} (\log i)^{3(t-s)(t-s+1)+1} \lceil \log_2 i \rceil \cdot d^{s-t}\alpha^{-1} b.$$

By (4.25), this is clearly $o(d)$, since $b\lceil \log_2 i \rceil = o(d)$ and

$$(\log i)^{3(t-s)(t-s+1)+1} \lceil \log_2 i \rceil \ll (\log n)^{\frac{3t}{s}\cdot\left[s(t-1)+\frac{1}{t-s+1}\right]}.$$

By inequality (4.15), the total length of the output produced by $\mathcal{A}$ in step 4 is at most

$$5\log_2 i + \log_2 \binom{t\left(\frac{i}{d}\right)^{s-1}}{d} \leq 5\log_2 n + \log_2 \binom{t\left(\frac{n}{d}\right)^{s-1}}{d} \leq 5\log_2 n + d\log_2\left(\frac{etn^{s-1}}{d^s}\right).$$

Hence the total length of the output of $\mathcal{A}$ in the case $d > n^{1-\mu_{s,t}}(\log n)^{3t/s}$ is

$$d\log_2\left(\frac{etn^{s-1}}{d^s}\right) + o(d). \tag{4.26}$$

Since with G fixed, $\mathcal{A}$ outputs a unique code for every N, by (4.24) and (4.26), this means that regardless of the order of d, the total number of valid $(d+1)$-sets N,

$$f(G; d, K_{s,t}) \leq \exp\left(n^{1-\mu_{s,t}}(\log n)^{3t/s+1} + d\log\left(\frac{etn^{s-1}}{d^s}\right) + o(d)\right). \tag{4.27}$$

Since the term in the right-hand side of (4.27) does not depend on G, it is also an upper bound on $f(i; d, K_{s,t})$ and hence for every degeneracy sequence $(d_i)_{i=1}^{n-1}$ with sum m,

$$\prod_{i=1}^{n-1} f(i; d_i - 1, K_{s,t}) \leq \exp\left(n^{2-\mu_{s,t}}(\log n)^{3t/s+1} + \sum_{i=1}^{n-1}(d_i-1)\log\left(\frac{etn^{s-1}}{(d_i-1)^s}\right) + o(m)\right). \tag{4.28}$$

The function $[0,\infty) \ni x \mapsto x\log x \in \mathbb{R}$ is convex, and so for every degeneracy sequence with sum m (putting $d_0 = 0$ and $m' = m - n + 1$),

$$\sum_{i=1}^{n-1}(d_i-1)\log(d_i-1) = \sum_{i=0}^{n-1}(d_i-1)\log(d_i-1) \geq n \cdot (m'/n)\log(m'/n).$$

This yields

$$\sum_{i=1}^{n-1}(d_i-1)\log\left(\frac{etn^{s-1}}{(d_i-1)^s}\right) \leq m'\log(etn^{s-1}) - m's\log(m'/n) = m'\log\left(\frac{etn^{2s-1}}{(m')^s}\right). \tag{4.29}$$

Since $\frac{d}{dx}(x\log(y/x)) = \log(y/x) - 1$, we can estimate

$$\left| m'\log\left(\frac{etn^{2s-1}}{(m')^s}\right) - m\log\left(\frac{etn^{2s-1}}{m^s}\right)\right| = O\big((m-m')\log n\big) = o(m),$$

which combined with (4.28) and (4.29) gives

$$\prod_{i=1}^{n-1} f(i; d_i, K_{s,t}) \leq \exp\left(n^{2-\mu_{s,t}}(\log n)^{3t/s+1} + m\log\left(\frac{etn^{2s-1}}{m^s}\right) + o(m)\right). \tag{4.30}$$

Since

$$m \gg n^{2-\mu_{s,t}}(\log n)^{3t/s+1}, \quad e < 3, \quad m \leq \mathrm{ex}(n, K_{s,t}) \leq \frac{1}{2}(t-1)^{1/s}n^{2-1/s} + O(n),$$

and there are at most $n!$ degeneracy sequences, combining (4.23) with (4.30) yields

$$f_{n,m}(K_{s,t}) \le \left(\frac{3tn^{2s-1}}{m^s}\right)^m,$$

whenever n is large enough.

4.6 Proof of Theorem 4.7

First, let us recall that for all $t \ge 2$, $\mathrm{ex}(n, K_{2,t}) = \left(\frac{\sqrt{t-1}}{2} + o(1)\right) \cdot n^{3/2}$ (see, e.g., [33, 42]). Suppose that G is a $K_{2,t}$-free graph on n vertices which has at least $(1-\varepsilon)\,\mathrm{ex}(n, K_{2,t})$ edges. If the minimum degree of G is smaller than $\frac{3}{4}\sqrt{t-1} \cdot n^{1/2}$, then by removing a vertex of smallest degree from G, we would increase the *relative edge density*, i.e., the ratio $e(G)/v(G)^{3/2}$, of the resulting graph. Since removing vertices cannot create a copy of $K_{2,t}$ in our graph, the relative edge density after any number of such removals will not exceed $\frac{1}{2}\sqrt{t-1}+o(1)$, and hence we cannot continue removing low degree vertices indefinitely. It follows that after deleting a relatively small set of low degree vertices from our graph G, we will obtain a graph whose minimum degree is at least about $\frac{3}{4}\sqrt{t-1} \cdot n^{1/2}$. We formalize the above discussion in Lemma 4.31, cf. [13, Lemma 3].

Recall that whenever H is a fixed graph, we denote the family of all labeled H-free graphs on the vertex set $\{1, \ldots, n\}$ by $\mathcal{F}_n(H)$. For a fixed real number $\varepsilon \in (0,1)$, let $\mathcal{F}_n^\varepsilon(H)$ denote the subfamily of $\mathcal{F}_n(H)$ consisting only of graphs that have at least $(1-\varepsilon) \cdot \mathrm{ex}(n, H)$ edges.

Lemma 4.31. *For every positive constant α, there is a positive ε such that the following holds. Let $t \ge 2$ and let $G \in \mathcal{F}_n^\varepsilon(K_{2,t})$, where n is large enough. Then there is a set $X \subseteq V(G)$ with $|X| \le \alpha n$ such that $\delta(G - X) \ge (\frac{3}{4} - \alpha)\sqrt{t-1} \cdot n^{1/2}$.*

Proof. Fix $\varepsilon = \alpha^2/3$. Define an ordering of the vertices of G as follows. Let v_1 be a vertex of minimum degree in G. Provided that $v_1, \ldots, v_i$ have already been chosen, we let v_{i+1} be a vertex of minimum degree in $G - \{v_1, \ldots, v_i\}$. Since every subgraph of G is $K_{2,t}$-free, and $\mathrm{ex}(n, K_{2,t}) \le \frac{2\sqrt{t-1}n^{3/2}+n}{4}$ (see, e.g., [58]), the function f defined by

$$f(k) = \frac{2\sqrt{t-1} \cdot (n-k)^{3/2} + (n-k)}{4} - e(G - \{v_1, \ldots, v_k\})$$

is non-negative for all k.

Let k_0 be the smallest number such that $\delta(G - \{v_1, \ldots, v_k\}) \ge (\frac{3}{4} - \alpha)\sqrt{t-1} \cdot n^{1/2}$, or $k_0 = n$ in case such a number does not exist. If $k_0 \le \alpha n$, then the set X defined by $X = \{v_1, \ldots, v_{k_0}\}$ is good for our purposes. Otherwise, all k with $k \le \alpha n$ satisfy

$$\begin{aligned} f(k) - f(k-1) &= \sqrt{t-1} \cdot \frac{(n-k)^{3/2} - (n-k+1)^{3/2}}{2} - \frac{1}{4} + \deg(v_k, V(G) - \{v_1, \ldots, v_{k-1}\}) \\ &\le -\frac{3}{4}\sqrt{t-1} \cdot (n-k)^{1/2} + \left(\frac{3}{4} - \alpha\right)\sqrt{t-1} \cdot n^{1/2} \\ &\le -\frac{3}{4}\sqrt{t-1} \cdot (1-\alpha)\, n^{1/2} + \left(\frac{3}{4} - \alpha\right)\sqrt{t-1} \cdot n^{1/2} \le -\frac{\alpha}{4}\sqrt{t-1} \cdot n^{1/2}, \end{aligned}$$

where the first inequality follows from the fact that $(a+1)^{3/2} - a^{3/2} \ge 3a^{1/2}/2$ for all non-negative

a, which can be proved using elementary calculus. Since f is non-negative, it follows that

$$-f(0) \leq f(\alpha n) - f(0) = \sum_{k=1}^{\alpha n} \big(f(k) - f(k-1)\big) \leq \alpha n \cdot \Big(-\frac{\alpha}{4}\sqrt{t-1} \cdot n^{1/2}\Big) = -\frac{\alpha^2}{4}\sqrt{t-1} \cdot n^{3/2},$$

and hence

$$e(G) = \frac{2\sqrt{t-1} \cdot n^{3/2} + n}{4} - f(0) \leq \frac{2-\alpha^2}{4}\sqrt{t-1} \cdot n^{3/2} + \frac{n}{4} < \left(1 - \frac{3}{2}\varepsilon + o(1)\right) \cdot \mathrm{ex}(n, K_{2,t}),$$

which is a contradiction. □

The key ingredient in the proof of Theorem 4.7 is Lemma 4.32, which builds on the argument from Theorem 4.2.

Lemma 4.32. *Let $t \geq 2$ and let G be an n-vertex $K_{2,t}$-free graph with $\delta(G) \geq \delta \geq \frac{1}{2}n^{1/2}$. Suppose that $d \geq \frac{1}{2}n^{1/2}$ and we want to add to G a new vertex v of degree d so that the resulting graph remains $K_{2,t}$-free, and moreover, we have already chosen pd neighbors of v, where $p \in [0,1]$. Then the number of ways we can select the remaining $(1-p)d$ neighbors of v is at most*

$$2^{o(n^{1/2})} \cdot \binom{(t-1)n/\delta - pd}{(1-p)d}.$$

Proof. We slightly modify the argument from the proof of Theorem 4.2. The key idea there was giving a bound on the number of ways we can attach to a $K_{2,t}$-free n-vertex graph G with $\delta(G) \geq d$ a vertex v od degree $d+1$. This quantity was bounded by $2^{o(d)}$ times the number of ways one can choose d neighbors of v from a set X of size $|X| = \big(1 + (\log n)^{-1}\big) \cdot a_0$, where a_0 is the lower bound on a in Lemma 4.15 for appropriate values of parameters – $s = 2$, $d = \delta(G)$, $\varepsilon = (\log n)^{-1}$; see the proof of Lemma 4.26.

Here we apply the same argument. The only difference is that in the end, since pd out of the d choices we are to make have already been made for us, we only get to choose $(1-p)d$ elements from a set of size $|X| - pd$. Consequently, the quantity we are interested in, i.e., the number of choices for the remaining $(1-p)d$ neighbors of v, can be bounded by

$$2^{o(d)} \cdot \binom{|X| - pd}{(1-p)d} \leq 2^{o(n^{1/2})} \cdot \binom{\big(1 + 3(\log n)^{-1}\big)\,(t-1)n/\delta - pd}{(1-p)d} \leq 2^{o(n^{1/2})} \cdot \binom{(t-1)n/\delta - pd}{(1-p)d},$$

where the first inequality follows from (4.12) and the fact that in every $K_{2,t}$-free graph G, $\Delta(G)\delta(G) \leq (t-1)n$, and hence $d = O(n^{1/2})$; the second inequality follows from the inequality $\binom{a+c}{b} \leq \binom{a}{b} \cdot \binom{a+c}{c}$ which holds whenever $a \geq b \geq 0$ and $c \geq 0$. □

Finally, we are ready to prove Theorem 4.7.

Proof of Theorem 4.7. Very vaguely, the idea of the proof can be summarized as follows. If G is a $K_{2,t}$-free graph with large minimum degree, then the number of ways in which we can remove a certain proportion of its edges is much larger than the number of ways we can add the same number of edges back so that the resulting graph remains $K_{2,t}$-free. In other words, every $G \in \mathcal{F}_n^{\varepsilon}(K_{2,t})$ has many more different subgraphs $F \in \mathcal{F}_n(K_{2,t})$ than the number of supergraphs in $\mathcal{F}_n^{\varepsilon}(K_{2,t})$ that any such F can possibly have. This implies that $|\mathcal{F}_n^\varepsilon(K_{2,t})| = o(|\mathcal{F}_n(K_{2,t})|)$. In the sequel we will formalize the above discussion. We would like to remark that a similar technique was also used in [10].

Consider an arbitrary mapping

$$\varphi : \mathcal{F}_n^\varepsilon(K_{2,t}) \to \mathcal{P}\big(\mathcal{F}_n(K_{2,t}) \times 2^{[n]} \times n^{[n]}\big).$$

For a triple $T \in \mathcal{F}_n(C_4) \times 2^{[n]} \times n^{[n]}$, let

$$\psi(T) = \big\{G \in \mathcal{F}_n^\varepsilon(C_4) : T \in \varphi(G)\big\}.$$

Counting appearances of all triples T in the images $\varphi(G)$, where G ranges over $\mathcal{F}_n^\varepsilon(K_{2,t})$ and T ranges over $\mathcal{F}_n(K_{2,t}) \times 2^{[n]} \times n^{[n]}$, yields

$$\sum_G |\varphi(G)| = \sum_T |\psi(T)|. \tag{4.31}$$

Equality (4.31) implies an obvious bound on the size of $\mathcal{F}_n^\varepsilon(K_{2,t})$, namely

$$\big|\mathcal{F}_n^\varepsilon(K_{2,t})\big| \le (2n)^n \cdot \frac{\sup_T |\psi(T)|}{\inf_G |\varphi(G)|} \cdot \big|\mathcal{F}_n(K_{2,t})\big|. \tag{4.32}$$

Now, inequality (4.32) combined with any $o((2n)^{-n})$ bound on the $\sup_T |\psi(T)|/\inf_G |\varphi(G)|$ ratio (for a carefully chosen φ) will imply that $|\mathcal{F}_n^\varepsilon(K_{2,t})| = o(|\mathcal{F}_n(K_{2,t})|)$.

Since the remainder of the proof gets somewhat technical, we will start by giving its short and informal outline. In Lemma 4.31, we have already noted that every $K_{2,t}$-free graph G with many edges, i.e., one with $e(G)$ close to the extremal number $\mathrm{ex}(n, K_{2,t})$, contains an almost spanning subgraph G_0 with minimum degree at least about $\frac{3}{4}\sqrt{t-1} \cdot n^{1/2}$. Next, using Lemma 4.16, we find a subset $A \subseteq V(G_0)$ of size about βn such that the minimum degree of the graph $G_0 - A$, denoted G', is still at least almost $\frac{3}{4}\sqrt{t-1} \cdot n^{1/2}$ and all the vertices in A have (approximately) at least $\frac{3}{4}\sqrt{t-1} \cdot n^{1/2}$ neighbors in $V(G')$. Let $p = 0.9$. Given such a set A, we define $\varphi(G)$ to be the set of all graphs obtained from G by deleting $1 - p$ cross-edges incident to each vertex in A, together with all the necessary information to identify the set A and reconstruct all relevant degrees after such a deletion. Finally, given a triple T consisting of a graph F, a set $A \subseteq V(F)$ and the list of degrees that the vertices in A had in the original $K_{2,t}$-free graph $G \supseteq F$, we prove, using Lemma 4.32, an upper bound on the number of supergraphs $G \supseteq F$ with $T \in \varphi(G)$. Combining this upper bound with a lower bound on $|\varphi(G)|$ and (4.32), we conclude that $\big|\mathcal{F}_n^\varepsilon(C_4)\big| \le 2^{-\Omega(n^{3/2})} \cdot \big|\mathcal{F}_n(C_4)\big|$.

Let us start by rigorously defining the mapping φ. Fix some very small constants α, β and ρ (we will specify them later), and let ε be as in the statement of Lemma 4.31. Recall that $p = 0.9$. Suppose that $n \ge n_0/(1-\alpha)$, where n_0 is as in the statement of Lemma 4.16. Finally, fix some $G \in \mathcal{F}_n^\varepsilon(K_{2,t})$. By Lemma 4.31, there is a subset $X \subseteq V(G)$ of size at most αn such that $\delta(G - X) \ge (3/4 - \alpha)\sqrt{t-1} \cdot n^{1/2}$. Now, by Lemma 4.16, we can find an $A \subseteq V(G) - X$ with

$$(1-\rho)(1-\alpha)\beta n \le |A| \le (1+\rho)\beta n \tag{4.33}$$

such that if we let $G' = G - X - A$ and $\gamma = (1-(1+\rho)\beta)(3/4-\alpha)\sqrt{t-1}$, then

$$\delta(G') \geq (1-(1+\rho)\beta)\cdot\delta(G-X) \geq \gamma n^{1/2}, \tag{4.34}$$

and for every vertex $v \in A$,

$$\deg(v, V(G')) \geq (1-(1+\rho)\beta)\cdot\deg_{G-X}(v) \geq \gamma n^{1/2}. \tag{4.35}$$

We define $\varphi(G)$ to be the set of all triples $(F, X\cup A, f)$, where

$$f(v) = \begin{cases} \deg(v, V(G')) & \text{if } v \in X \cup A, \\ 0 & \text{otherwise}, \end{cases}$$

and F is any subgraph of G obtained by deleting, for each vertex $v \in X \cup A$, a set of $(1-p)\cdot\deg(v, V(G'))$ edges connecting v to $V(G')$. Also, note that since G is $K_{2,t}$-free, at most $t-1$ paths of length 2 starting at some $v \in X\cup A$ can reach the same vertex, and hence $f(v)\cdot\delta(G') \leq (t-1)\cdot|V(G')| \leq (t-1)n$, which together with (4.34) implies that $f(v) \leq (t-1)n^{1/2}/\gamma$.

Claim 4.33. *For every $G \in \mathcal{F}_n^{\varepsilon}(K_{2,t})$,*

$$\left|\varphi(G)\right| \geq 2^{H(p)(1-\rho)(1-\alpha)\gamma\beta n^{3/2} - O(n\log n)}.$$

Proof. It suffices to count the number of subgraphs F appearing in the definition of $\varphi(G)$. By our bounds on the size of A and the degrees of vertices in A, this is at least

$$\begin{aligned}\prod_{v\in A}\binom{\deg(v,V(G'))}{p\deg(v,V(G'))} &\geq (n+1)^{-|A|}\cdot 2^{H(p)\sum_{v\in A}\deg(v,V(G'))} \\ &\geq (n+1)^{-(1+\rho)\beta n}\cdot 2^{H(p)(1-\rho)(1-\alpha)\gamma\beta n^{3/2}},\end{aligned}$$

where the first inequality follows from Lemma 4.17, and the second inequality follows from (4.33) and (4.35). □

Let $T = (F, S, f)$ be a triple from the image of φ. The way we defined φ guarantees that the set S has size at most $(1+\rho)\beta n + \alpha n$, and $F - S$ has minimum degree at least $\gamma n^{1/2}$. By Lemmas 4.32 and 4.17, we get the following bound on the size of $\psi(T)$:

$$\begin{aligned}\left|\psi(T)\right| &\leq 2^{o(n^{1/2})}\cdot\prod_{v\in S}\binom{(t-1)(n-|S|)/(\gamma n^{1/2}) - pf(v)}{(1-p)f(v)} \leq 2^{o(n^{3/2})}\cdot\prod_{v\in S}\binom{(t-1)n^{1/2}/\gamma - pf(v)}{(1-p)f(v)} \\ &\leq 2^{o(n^{3/2})}\cdot\prod_{v\in S} 2^{((t-1)n^{1/2}/\gamma - pf(v))H\left(\frac{(1-p)f(v)}{(t-1)n^{1/2}/\gamma - pf(v)}\right)} \leq 2^{o(n^{3/2})}\cdot 2^{((1+\rho)\beta+\alpha)sn^{3/2}},\end{aligned}$$

where (we let $d = f(v)n^{-1/2}$ and note that $\gamma \leq d \leq (t-1)\gamma^{-1}$)

$$s = \sup_d\left[((t-1)/\gamma - pd)\cdot H\left(\frac{(1-p)d}{(t-1)/\gamma - pd}\right)\right].$$

If we set $\alpha = \rho = 10^{-10}$, $\beta = 10^{-5}$ and $p = 0.9$, then $s \leq 0.3467$, and hence

$$\big|\psi(T)\big| \leq 2^{3468 \cdot 10^{-9} n^{3/2} + o(n^{3/2})}.$$

On the other hand, Claim 4.33 gives

$$\big|\varphi(G)\big| \geq 2^{3517 \cdot 10^{-9} n^{3/2} - o(n^{3/2})}.$$

It follows that

$$\frac{\sup_T |\psi(T)|}{\inf_G |\varphi(G)|} \leq 2^{-49 \cdot 10^{-9} \cdot n^{3/2} + o(n^{3/2})},$$

and therefore, by (4.32),

$$\big|\mathcal{F}_n^{\varepsilon}(K_{2,t})\big| \leq 2^{-4 \cdot 10^{-8} \cdot n^{3/2}} \cdot \big|\mathcal{F}_n(K_{2,t})\big|,$$

provided that n is large enough. □

Bibliography

[1] M. Ajtai, J. Komlós, and E. Szemerédi, *The longest path in a random graph*, Combinatorica **1** (1981), 1–12.

[2] N. Alon, J. Balogh, A. Kostochka, and W. Samotij, *Sizes of induced subgraphs of Ramsey graphs*, Combinatorics, Probability and Computing **18** (2009), 459–476.

[3] N. Alon, M. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński, and E. Szemerédi, *Universality and tolerance (extended abstract)*, 41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000), IEEE Comput. Soc. Press, 2000, pp. 14–21.

[4] N. Alon and G. Gutin, *Properly colored Hamilton cycles in edge-colored complete graphs*, Random Structures & Algorithms **11** (1997), 179–186.

[5] N. Alon and A. V. Kostochka, *Induced subgraphs with distinct sizes*, Random Structures & Algorithms **34** (2009), 45–53.

[6] N. Alon, M. Krivelevich, and B. Sudakov, *Embedding nearly-spanning bounded degree trees*, Combinatorica **27** (2007), 629–644.

[7] N. Alon, L. Rónyai, and T. Szabó, *Norm-graphs: variations and applications*, Journal of Combinatorial Theory B **76** (1999), 280–290.

[8] N. Alon and J. Spencer, *The probabilistic method*, third ed., John Wiley & Sons Inc., 2008.

[9] J. Balogh, B. Bollobás, and M. Simonovits, *The fine structure of octahedron-free graphs*, submitted.

[10] ______, *The typical structure of graphs without given excluded subgraphs*, Random Structures and Algorithms **34** (2009), 305–318.

[11] J. Balogh, B. Csaba, M. Pei, and W. Samotij, *Large bounded degree trees in expanding graphs*, Electronic Journal of Combinatorics **17** (2010), Research Paper 6.

[12] J. Balogh, B. Csaba, and W. Samotij, *Local resilience of almost spanning trees in random graphs*, to appear in Random Structures & Algorithms.

[13] J. Balogh and W. Samotij, *Almost all C_4-free graphs have less than $(1-\varepsilon)\operatorname{ex}(n, C_4)$ edges*, to appear in SIAM Journal on Discrete Mathematics.

[14] ______, *The number of $K_{m,m}$-free graphs*, submitted to Combinatorica.

[15] ______, *The number of $K_{s,t}$-free graphs*, submitted to Journal of London Mathematical Society.

[16] B. Bollobás and A. D. Scott, *Discrepancy in graphs and hypergraphs*, More sets, graphs and numbers, Bolyai Soc. Math. Stud., vol. 15, Springer, 2006, pp. 33–56.

[17] J. Böttcher, Y. Kohayakawa, and A. Taraz, *Almost spanning subgraphs of random graphs after adversarial edge removal*, Electronic Notes in Discrete Mathematics **35** (2009), 335–340.

[18] W. G. Brown, *On graphs that do not contain a Thomsen graph*, Canadian Mathematical Bulletin **9** (1966), 281–285.

[19] H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Annals of Mathematical Statistics **23** (1952), 493–507.

[20] F. Chung, *Open problems of Paul Erdős in graph theory*, Journal of Graph Theory **25** (1997), 3–36.

[21] F. R. K. Chung, R. L. Graham, and R. M. Wilson, *Quasi-random graphs*, Combinatorica **9** (1989), 345–362.

[22] D. Conlon and T. Gowers, *Combinatorial theorems relative to a random set*, manuscript.

[23] D. Dellamonica, Y. Kohayakawa, M. Marciniszyn, and A. Steger, *On the resilience of long cycles in random graphs*, Electronic Journal of Combinatorics **15** (2008), Research Paper 32.

[24] G. A. Dirac, *Some theorems on abstract graphs*, Proceedings of the London Mathematical Society (3) **2** (1952), 69–81.

[25] P. Erdős, *Graph theory and probability*, Canadian Journal of Mathematics **11** (1959), 34–38.

[26] ———, *Some of my favourite problems in various branches of combinatorics*, Matematiche **47** (1992), 231–240, Combinatorics 92 (Catania, 1992).

[27] ———, *Some recent problems and results in graph theory*, Discrete Mathematics **164** (1997), 81–85, The Second Krakow Conference on Graph Theory (Zgorzelisko, 1994).

[28] P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs and Combinatorics **2** (1986), 113–121.

[29] P. Erdős, M. Goldberg, J. Pach, and J. Spencer, *Cutting a graph into two dissimilar halves*, Journal of Graph Theory **12** (1988), 121–131.

[30] P. Erdős and A. Hajnal, *On spanned subgraphs of graphs*, Contributions to graph theory and its applications (Internat. Colloq., Oberhof, 1977) (German), Tech. Hochschule Ilmenau, 1977, pp. 80–96.

[31] P. Erdős and A. Rényi, *On random graphs. I*, Publicationes Mathematicae Debrecen **6** (1959), 290–297.

[32] ———, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **5** (1960), 17–61.

[33] P. Erdős, A. Rényi, and V. Sós, *On a problem of graph theory*, Studia Scientiarum Mathematicarum Hungarica **1** (1966), 215–235.

[34] P. Erdős and A. Stone, *On the structure of linear graphs*, Bulletin of the American Mathematical Society **52** (1946), 1087–1091.

[35] P. Erdös and G. Szekeres, *A combinatorial problem in geometry*, Compositio Mathematica **2** (1935), 463–470.

[36] P. Erdős and A. Szemerédi, *On a Ramsey type theorem*, Periodica Mathematica Hungarica **2** (1972), 295–299, Collection of articles dedicated to the memory of Alfréd Rényi, I.

[37] P. Erdős and P. Turán, *On some sequences of integers*, Journal of the London Mathematical Society **11** (1936), 261–164.

[38] J. Friedman and N. Pippenger, *Expanding graphs contain all small trees*, Combinatorica **7** (1987), 71–76.

[39] A. Frieze, *On large matchings and cycles in sparse random graphs*, Discrete Mathematics **59** (1986), 243–256.

[40] A. Frieze and M. Krivelevich, *On two Hamilton cycle problems in random graphs*, Israel Journal of Mathematics **166** (2008), 221–234.

[41] Z. Füredi, *Random Ramsey graphs for the four-cycle*, Discrete Mathematics **126** (1994), 407–410.

[42] ———, *New asymptotics for bipartite Turán numbers*, Journal of Combinatorial Theory A **75** (1996), 141–144.

[43] ———, *An upper bound on Zarankiewicz' problem*, Combinatorics, Probability and Computing **5** (1996), 29–33.

[44] S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger, *Small subsets inherit sparse ϵ-regularity*, Journal of Combinatorial Theory B **97** (2007), 34–56.

[45] S. Gerke, M. Marciniszyn, and A. Steger, *A probabilistic counting lemma for complete graphs*, Random Structures & Algorithms **31** (2007), 517–534.

[46] S. Gerke, H. J. Prömel, T. Schickinger, A. Steger, and A. Taraz, *K_4-free subgraphs of random graphs revisited*, Combinatorica **27** (2007), 329–365.

[47] S. Gerke, T. Schickinger, and A. Steger, *K_5-free subgraphs of random graphs*, Random Structures & Algorithms **24** (2004), 194–232.

[48] E. N. Gilbert, *Random graphs*, Annals of Mathematical Statistics **30** (1959), 1141–1144.

[49] P. Haxell, *Tree embeddings*, Journal of Graph Theory **36** (2001), 121–130.

[50] P. Haxell, Y. Kohayakawa, and T. Łuczak, *Turán's extremal problem in random graphs: forbidding even cycles*, Journal of Combinatorial Theory B **64** (1995), 273–287.

[51] ———, *Turán's extremal problem in random graphs: forbidding odd cycles*, Combinatorica **16** (1996), 107–122.

[52] D. Kleitman and D. Wilson, *On the number of graphs which lack small cycles*, manuscript, 1996.

[53] D. Kleitman and K. Winston, *On the number of graphs without* 4*-cycles*, Discrete Mathematics **41** (1982), 167–172.

[54] Y. Kohayakawa, *Szemerédi's Regularity Lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230.

[55] Y. Kohayakawa, Y. Łuczak, and V. Rödl, *On K_4-free subgraphs of random graphs*, Combinatorica **17** (1997), 173–213.

[56] Y. Kohayakawa, V.ěch Rödl, and M. Schacht, *The Turán theorem for random graphs*, Combinatorics, Probability and Computing **13** (2004), 61–91.

[57] J. Kollár, L. Rónyai, and T. Szabó, *Norm-graphs and bipartite Turán numbers*, Combinatorica **16** (1996), 399–406.

[58] T. Kövári, V. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloquium Mathematicum **3** (1954), 50–57.

[59] M. Krivelevich, C. Lee, and B. Sudakov, *Resilient pancyclicity of random and pseudo-random graphs*, SIAM Journal on Discrete Mathematics **24** (2010), 1–16.

[60] T. Łuczak, *On triangle-free random graphs*, Random Structures & Algorithms **16** (2000), 260–276.

[61] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*, Cambridge University Press, 2005.

[62] M. Pei, *List colouring hypergraphs and extremal results for acyclic graphs*, Ph.D. thesis, University of Waterloo, 2008.

[63] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Mathematics **14** (1976), 359–364.

[64] H. J. Prömel and V. Rödl, *Non-Ramsey graphs are $c \log n$-universal*, Journal of Combinatorial Theory. Series A **88** (1999), 379–384.

[65] F. P. Ramsey, *On a problem of formal logic*, Proceedings of the London Mathematical Society **30** (1930), 264–286.

[66] V. Rödl, *On universality of graphs with uniformly distributed edges*, Discrete Mathematics **59** (1986), 125–134.

[67] M. Schacht, *Extremal results for random discrete structures*, submitted.

[68] S. Shelah, *Erdős and Rényi conjecture*, Journal of Combinatorial Theory. Series A **82** (1998), 179–185.

[69] B. Sudakov and V. Vu, *Local resilience of graphs*, Random Structures & Algorithms **33** (2008), 409–433.

[70] T. Szabó and V. Vu, *Turán's theorem in sparse random graphs*, Random Structures & Algorithms **23** (2003), 225–234.

[71] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica **27** (1975), 199–245.

[72] A. Thomason, *Random graphs, strongly regular graphs and pseudorandom graphs*, Surveys in combinatorics 1987 (New Cross, 1987), London Math. Soc. Lecture Note Ser., vol. 123, Cambridge Univ. Press, 1987, pp. 173–195.

[73] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452.