A Turán theorem for random graphs

by

Mathias Schacht
Master of Science

Department of Mathematics and Computer Science

Vojtěch Rödl
Advisor

Yoshiharu Kohayakawa
Advisor

Dwight Duffus
Committee Member

Andrzej Ruciński
Committee Member

Accepted:

Dean of the Graduate School

Date

A Turán theorem for random graphs

by

Mathias Schacht

Advisors : Vojtěch Rödl, Ph.D.
Yoshiharu Kohayakawa, Ph.D.

Abstract of
a thesis submitted to the Faculty of the Graduate School
of Emory University in partial fulfillment
of the requirements of the degree of
Master of Science

Department of Mathematics and Computer Science

2002

Abstract

Turán's theorem is one of the cornerstones of extremal graph theory today. The aim of this thesis is to prove a Turán type theorem for sparse random graphs.

For $0 < \gamma \leq 1$ and graphs G and H, write $G \rightarrow_\gamma H$ if any γ-proportion of the edges of G contains at least one copy of H in G. In this thesis, we prove that for every d-degenerate graph H and every fixed real $\gamma > 1-1/(\chi(H)-1)$ asymptotically almost surely a graph G in the binomial random graph model $\mathcal{G}(n,q)$ with $q = q(n) \gg ((\log n)^4/n)^{1/d}$ satisfies $G \rightarrow_\gamma H$, where as usual $\chi(H)$ denotes the chromatic number of H.

As a corollary we immediately derive that for every $l \geq 2$ and every fixed real $\gamma > 1 - 1/(l-1)$ asymptotically almost surely a graph G in $\mathcal{G}(n,q)$ with $q = q(n) \gg ((\log n)^4/n)^{1/(l-1)}$ satisfies $G \rightarrow_\gamma K_l$, where K_l is the complete graph on l vertices.

A Turán theorem for random graphs

by

Mathias Schacht

Advisors : Vojtěch Rödl, Ph.D.
Yoshiharu Kohayakawa, Ph.D.

A thesis submitted to the Faculty of the Graduate School
of Emory University in partial fulfillment
of the requirements of the degree of
Master of Science

Department of Mathematics and Computer Science

2002

Acknowledgments

I would like to thank both of my advisors Vojtěch Rödl and Yoshiharu Kohayakawa for their time and encouragement. I would also like to thank to Dwight Duffus and Andrzej Ruciński for serving on my thesis committee.

Moreover, I would like to express my appreciation to the people who established and continue the exchange program of the Technical University of Berlin, Germany and Emory University. In particular, I would like to thank Vladimir Oliker and Ronald J. Gould from Emory University and Udo Simon from the Technical University of Berlin.

Contents

List of Figures

Chapter 1

Introduction

A classical area of extremal graph theory investigates numerical and structural problems concerning H-free graphs, namely, graphs that do not contain a copy of a given fixed graph H as a subgraph. Let $\mathrm{ex}(n, H)$, the Turán number of H, be the maximal number of edges that an H-free graph on n vertices may have. A basic question is then to determine or estimate $\mathrm{ex}(n, H)$ for any given H. In the special case where $H = K_l$ is the complete graph on l vertices this question was answered precisely by Turán.

Theorem 1 (Turán [25]). *Given integers $n \geq l > 1$. Let $r = n \mod {}_{l-1}$, then*

$$\mathrm{ex}(n, K_l) = \frac{1}{2}\left(1 - \frac{1}{l-1}\right)\left(n^2 - r^2\right) + \binom{r}{2}.$$

An asymptotic solution to the general problem (for arbitrary graphs H) is given by the following celebrated theorem.

Theorem 2 (Erdős–Stone–Simonovits [5, 6]). *For every graph H with chromatic number $\chi(H)$*

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

where $o(1)$ is a function approaching zero as n goes to infinity.

Furthermore, as proved independently by Erdős [3, 4] and Simonovits [22], every H-free graph $G = G^n$ that has as many edges as in (1.1) is in fact 'very close' (in a certain precise sense) to the densest n-vertex $(\chi(H)-1)$-partite graph. For these and related results, see, for instance, Bollobás [1].

Here we are interested in a variant of the function $\mathrm{ex}(n, H)$. Let G and H be graphs, and write $\mathrm{ex}(G, H)$ for the maximal number of edges that an H-free subgraph of G may have. Formally, $\mathrm{ex}(G, H) = \max\{|E(F)|\colon H \not\subset F \subset G\}$. For instance, if $G = K_n$, the complete graph on n vertices, then $\mathrm{ex}(K_n, H) = \mathrm{ex}(n, H)$ is the usual Turán number of H.

Our aim here is to study $\mathrm{ex}(G, H)$ when G is a *random graph*. Let $0 < q = q(n) \le 1$ be given. The binomial random graph G in $\mathcal{G}(n, q)$ has as its vertex set a fixed set $V(G)$ of cardinality n, and two vertices are adjacent in G with probability q. All such adjacencies are independent. (For concepts and results concerning random graphs not given in detail below, see [2, 12].) As is usual in the theory of random graphs, we say that a property P holds *asymptotically almost surely* (abbreviated *a.a.s.*) if P holds with probability tending to 1 as $n \to \infty$.

Here we wish to investigate the random variable $\mathrm{ex}(\mathcal{G}(n, q), H)$. Since Theorem 2 can be viewed as a result for random graphs $\mathcal{G}(n, q)$ with $q = 1$, naturally, the question arises for which $q = q(n)$ the formula (1.1) remains true with K_n replaced by $\mathcal{G}(n, q)$ and $\binom{n}{2} = |E(K_n)|$ by $q\binom{n}{2}$ (the expected number of edges in the random graph $\mathcal{G}(n, q)$). We are interested in the probabilities $q = q(n)$ for which a.a.s.

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

holds. It follows from [12, Proposition 8.6] that if equality (1.2) holds a.a.s. for p_1, then it does so for each $p_2 \ge p_1$. Thus, we are interested in the smallest probability $q = q(n)$ such that (1.2) holds.

If $q = q(n)$ is such that the expected number of copies of H in $G \in \mathcal{G}(n, q)$

is much smaller than the expected number of edges of G, then it is not hard to show that (1.2) fails (see [12, Proposition 8.9]). Conjecture 3, stated below, demonstrates the belief that this is the only obstacle. The observation that containing H implies containing every subgraph H' of H, leads to the following definition.

Let H be a graph of order $|V(H)| \geq 3$. Let us write $m_2(H)$ for the 2-density of H, that is,

$$m_2(H) = \max\left\{\frac{|E(H')|-1}{|V(H')|-2} \colon H' \subset H,\, |V(H')| \geq 3\right\}.$$

A general conjecture concerning $\mathrm{ex}(\mathcal{G}(n,q),H)$, first stated in [15], is as follows.

Conjecture 3. *Let H be a non-empty graph of order at least 3, and let $0 < q = q(n) \leq 1$ be such that $qn^{1/m_2(H)} \to \infty$ as $n \to \infty$. Then a.a.s. G in $\mathcal{G}(n,q)$ satisfies*

$$\mathrm{ex}(G,H) = \left(1 - \frac{1}{\chi(H)-1} + o(1)\right)|E(G)|.$$

There are a few results in support of Conjecture 3. A simple application of Szemerédi's regularity lemma for sparse graphs (see Theorem 8 below), gives Conjecture 3 for H a forest. The cases in which $H = K_3$ and $H = C_4$ are essentially proved in Frankl and Rödl [7] and Füredi [8], respectively, in connection with problems concerning the existence of some graphs with certain extremal properties. The case for $H = K_4$ was proved by Kohayakawa, Łuczak and Rödl [15]. Recently Schickinger proved in his Ph.D. thesis [20] a somewhat stronger conjecture for $H = K_5$ (and $H = K_4$ as well, see also Gerke et al. [9]), which implies Conjecture 3 for this case. The case in which H is a general cycle was settled by Haxell, Kohayakawa, and Łuczak [10, 11] (see also Kohayakawa, Kreuter, and Steger [14]).

Our main result relates to Conjecture 3 in the following way: we deal with the case in which H is arbitrary and $q = q(n) \gg ((\log n)^4/n)^{1/d}$, where

$d = d(H)$ is the "degeneracy number" of H (defined below) and $q(n) \gg ((\log n)^4/n)^{1/d}$ means $\lim_{n\to\infty}((\log n)^4/n)^{1/d}/q(n) = 0$.

Definition 4 *(d-degenerate). A graph H of order h is called d-degenerate if there exists an ordering of the vertices $V(H) = \{w_1, \ldots, w_h\}$ such that each w_i ($1 \le i \le h$) has at most d neighbours in $\{w_1, \ldots, w_{i-1}\}$. Moreover, we denote the degeneracy number of H by the minimal integer $d = d(H)$ for which H is d-degenerate.*

For more details concerning d-degenerate graphs see [19, 21]).

The following theorem is the main result of this thesis.

Theorem 5. *Let d be a positive integer, H a d-degenerate graph of order h, and $q = q(n) \gg \left((\log n)^4/n\right)^{1/d}$. Then for every $1/(\chi(H) - 1) > \delta > 0$ a graph G in $\mathcal{G}(n, q)$ satisfies a.a.s. the following property: If F is an arbitrary, not necessarily induced subgraph of G with*

$$|E(F)| \ge \left(1 - \frac{1}{\chi(H) - 1} + \delta\right) q \binom{n}{2},$$

then F contains H as a subgraph. Moreover, there exists a constant $c = c(\delta, H)$ such that F contains at least $cq^{|E(H)|}n^h$ copies of H.

In this thesis we give a proof of Theorem 5. Since K_l, the complete graph on l vertices, is clearly $(l-1)$-degenerate and l-chromatic, the following result is an immediate consequence of Theorem 5.

Corollary 6. *Let $l \ge 2$, and $q = q(n) \gg \left((\log n)^4/n\right)^{1/(l-1)}$. Then for every $1/(l - 1) > \delta > 0$ a graph G in $\mathcal{G}(n, q)$ satisfies a.a.s. the following property: If F is an arbitrary, not necessarily induced subgraph of G with*

$$|E(F)| \ge \left(1 - \frac{1}{l - 1} + \delta\right) q \binom{n}{2},$$

then F contains K_l as a subgraph. Moreover, there exists a constant $c = c(\delta, l)$ such that F contains at least $cq^{\binom{l}{2}}n^l$ copies of K_l.

The main result discussed in this thesis (Theorem 5) was already announced by the author and his advisors in [18]. There a simpler proof for the case $H = K_l$ for $l \geq 2$ was given. In this thesis we give a proof for arbitrary graphs H, which is based on the ideas of [18].

Very recently Szabó and Vu proved in [23], independently from us, Corollary 6 under slightly weaker assumptions. Their proof is shorter than the proof of Theorem 5 presented here and does not require the regularity lemma. On the other hand, their approach does not seem to extend to arbitrary graphs H, whereas Theorem 5 gives nontrivial results for arbitrary H depending on the "degeneracy number" of the graph H.

This thesis is organized as follows. In Chapter 2 we describe a sparse version of Szemerédi's regularity lemma (Theorem 8) and we state the counting lemma (Lemma 10), both of which are crucial in our proof of Theorem 5. We prove Theorem 5 in Chapter 3. Chapter 4 is entirely devoted to the proof of Lemma 10. The proof of Lemma 10 relies on the 'Pick-Up Lemma' (Lemma 18) and on the 'k-tuple Lemma' (Lemma 22). We give these preliminary results in Section 4.1–4.2. In Section 4.3 we outline the proof of Lemma 10 in the case $H = K_4 - e$, the complete graph on four vertices minus an edge. Finally, the proof of Lemma 10 is given in Section 4.4.

For a general remark about the notation we use throughout this paper see Remark 9 in Section 2.3.

Chapter 2

Preliminary results

2.1 Preliminary definitions

Let a graph $G = G^n$ of order $|V(G)| = n$ be fixed. For U, $W \subset V = V(G)$, we write

$$E(U,W) = E_G(U,W) = \Big\{\{u,w\} \in E(G)\colon\ u \in U,\ w \in W\Big\}$$

for the set of edges of G that have one end-vertex in U and the other in W. Notice that each edge in $U \cap W$ occurs only once in $E(U,W)$. We set $e(U,W) = e_G(U,W) = |E(U,W)|$, i.e. for the complete graph K_l we have

$$e_{K_l} = |U||W| - \binom{|U \cap W| + 1}{2}.$$

Suppose $\xi > 0$, $C > 1$, and $0 < q \leq 1$.

Definition 7 ((ξ, C)-bounded). For $\xi > 0$ and $C > 1$ we say that $G = (V,E)$ is a (ξ, C)-bounded graph with respect to density q, if for all U, $W \subset V$, not necessarily disjoint, with $|U|$, $|W| \geq \xi|V|$, we have

$$e_G(U,W) \leq Cq\left(|U||W| - \binom{|U \cap W| + 1}{2}\right).$$

If G is a graph and $V_1, \ldots, V_t \subset V(G)$ are disjoint sets of vertices, we write $G[V_1, \ldots, V_t]$ for the t-partite graph naturally induced by $V_1, \ldots, V_t$.

2.2 The regularity lemma for sparse graphs

Our aim in this section is to state a variant of the regularity lemma of Szemerédi [24].

For any two *disjoint* non-empty sets U, $W \subset V$, let

$$d_{G,q}(U,W) = \frac{e_G(U,W)}{q|U||W|}. \tag{2.1}$$

We refer to $d_{G,q}(U,W)$ as the *q-density* of the pair (U,W) in G. When there is no danger of confusion, we drop G from the subscript and write $d_q(U,W)$.

Now suppose $\varepsilon > 0$, U, $W \subset V$, and $U \cap W = \emptyset$. We say that the pair (U,W) is *(ε, G, q)-regular*, or simply *(ε, q)-regular*, if for all $U' \subset U$, $W' \subset W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$ we have

$$|d_{G,q}(U',W') - d_{G,q}(U,W)| \leq \varepsilon. \tag{2.2}$$

Below, we shall sometimes use the expression *ε-regular with respect to density q* to mean that (U,W) is an (ε, q)-regular pair.

We say that a partition $P = (V_i)_0^t$ of $V = V(G)$ is *(ε, t)-equitable* if $|V_0| \leq \varepsilon n$, and $|V_1| = \cdots = |V_t|$. Also, we say that V_0 is the *exceptional* class of P. When the value of ε is not relevant, we refer to an (ε, t)-equitable partition as a *t-equitable* partition. Similarly, P is an *equitable* partition of V if it is a t-equitable partition for some t.

We say that an (ε, t)-equitable partition $P = (V_i)_0^t$ of V is *(ε, G, q)-regular*, or simply *(ε, q)-regular*, if at most $\varepsilon\binom{t}{2}$ pairs (V_i, V_j) with $1 \leq i < j \leq t$ are not (ε, q)-regular. We may now state a version of Szemerédi's regularity lemma for (ξ, C)-bounded graphs.

Theorem 8. *For any given $\varepsilon > 0$, $C > 1$, and $t_0 \geq 1$, there exist constants $\xi = \xi(\varepsilon, C, t_0)$ and $T_0 = T_0(\varepsilon, C, t_0) \geq t_0$ such that any sufficiently large graph G that is (ξ, C)-bounded with respect to density $0 < q \leq 1$ admits an (ε, G, q)-regular (ε, t)-equitable partition of its vertex set with $t_0 \leq t \leq T_0$.*

A simple modification of Szemerédi's proof of his lemma gives Theorem 8. For applications of this variant of the regularity lemma and its proof, see [13, 17].

2.3 The counting lemma

Let H be a fixed d-degenerate graph on h vertices and let the vertices of H be ordered $V(H) = \{w_1, \dots, w_h\}$ such that each w_i has at most d neighbours in $\{w_1, \dots, w_{i-1}\}$. Let $t \geq h$ be a fixed integer and n a sufficiently large integer. Let α and ε be constants greater than 0. Let $G \in \mathcal{G}(n,q)$ be the binomial random graph with edge probability $q = q(n)$, and suppose J is an h-partite subgraph of G with vertex classes $V_1, \dots, V_h$. For all $1 \leq i < j \leq h$ we denote by J_{ij} the bipartite graph induced by V_i and V_j. Consider the following assertions for J and q.

(I) $|V_i| = m = n/t$

(II) $q^d n \gg (\log n)^4$

(III) for all $1 \leq i < j \leq h$,

$$e(J_{ij}) = \begin{cases} T = pm^2 & \{w_i, w_j\} \in E(H) \\ 0 & \{w_i, w_j\} \notin E(H) \end{cases}$$

where $1 > \alpha q = p \gg 1/n$, and

(IV) J_{ij} is (ε, q)-regular.

Remark 9. Strictly speaking, in (I) we should have, say, $\lfloor n/t \rfloor$, because m is an integer. However, throughout this paper we will omit the floor and ceiling signs $\lfloor \ \ \rfloor$ and $\lceil \ \ \rceil$, since they have no significant effect on the arguments. Moreover, let us make a few more comments about the notation that we shall use. For positive functions $f(n)$ and $g(n)$, we write $f(n) \gg g(n)$ to mean that $\lim_{n\to\infty} g(n)/f(n) = 0$. Unless otherwise stated, we understand by $o(1)$ a function approaching zero as the number of vertices of a given random graph goes to infinity.

Finally, we observe that our logarithms are natural logarithms.

We are interested in the number of copies of the fixed graph H in such a subgraph J satisfying conditions (I)–(IV).

Lemma 10 (Counting lemma). *For all reals α, $\sigma > 0$, positive integer d and every d-degenerate graph H on h vertices, there exists a real $\varepsilon > 0$ such that for every fixed integer $t \geq h$ a random graph G in $\mathcal{G}(n, q)$ satisfies the following property with probability $1 - o(1)$: Every subgraph $J \subseteq G \in \mathcal{G}(n, q)$ satisfying conditions* (I)–(IV) *contains at least*

$$(1 - \sigma)p^{e(H)}m^h$$

copies of H.

We will prove Lemma 10 in Chapter 4.

Chapter 3

The main result

In this section we will prove the main result of this thesis, Theorem 5. This section is organised as follows. First, we state two properties that hold for almost every $G \in \mathcal{G}(n,q)$. Then, in Section 3.2, we prove a deterministic statement about the regularity of certain subgraphs of an (ε, q)-regular α-dense t-partite graph. Finally, we prove Theorem 5.

3.1 Properties of almost all graphs

We start with a well known fact of random graph theory which follows easily from the properties of the binomial distribution.

Fact 11. *For every real $\varrho > 0$, if G is a random graph in $\mathcal{G}(n,q)$, then*

$$(1-\varrho)\, q \binom{n}{2} \le e(G) \le (1+\varrho)\, q \binom{n}{2}$$

holds with probability $1 - o(1)$.

The next property refers to Definition 7 and will enable us to apply Theorem 8.

Lemma 12. *For every $C > 1$, $\xi > 0$ and $q = q(n) \gg 1/n$ a random graph G in $\mathcal{G}(n,q)$ is (ξ, C)-bounded with probability $1 - o(1)$.*

We will apply the following one-sided estimate of a binomially distributed random variable. For the next lemma, recall that all logarithms are to base e, see the Remark 9 in Section 2.3.

Lemma 13. *Let X be a binomially distributed random variable with expectation $\mathbb{E}X = Nq$ and let $C > 1$ be a constant. Then*

$$\mathbb{P}(X \geq C\mathbb{E}X) \leq \exp(-\tau C\mathbb{E}X),$$

where $\tau = \log C - 1 + 1/C > 0$ for $C > 1$.

Proof. The proof is given in [12] (see Corollary 2.4). □

Proof of Lemma 12. Let $G \in \mathcal{G}(n,q)$ and let U, $W \subseteq V(G)$ be two not necessarily disjoint sets such that $|U|$, $|W| \geq \xi n$. Clearly, $e(U,W)$ is a binomial random variable with

$$\mathbb{E}[e(U,W)] = q\left(|U||W| - \binom{|U \cap W| + 1}{2}\right).$$

Observe that $\mathbb{E}[e(U,W)] \gg n$ since $q \gg 1/n$. Set $\tau = \log C - 1 + 1/C$. Then Lemma 13 implies

$$\mathbb{P}\left(e(U,W) > C\mathbb{E}[e(U,W)]\right) \leq \exp\left(-\tau C\mathbb{E}[e(U,W)]\right).$$

We now sum over all choices for U and W to deduce that

$$\begin{aligned}\mathbb{P}(G \text{ is not } (\xi, C)\text{-bounded}) &\leq \\ &\sum_{|U|\geq\xi n}\sum_{|W|\geq\xi n}\binom{n}{|U|}\binom{n}{|W|}\exp\left(-\tau C\mathbb{E}[e(U,W)]\right) \\ &\leq 4^n \exp\left(-\tau C\mathbb{E}[e(U,W)]\right) = o(1),\end{aligned}$$

since $\tau C > 0$ and $\mathbb{E}[e(U,W)] \gg n$. □

3.2 A deterministic subgraph lemma

The next lemma states that every (ε, q)-regular, bipartite graph with at least $\alpha q m^2$ edges contains a $(3\varepsilon, q)$-regular subgraph with exactly $\alpha q m^2$ edges.

Lemma 14. *For every $\varepsilon > 0$, $\alpha > 0$, and $C > 1$ there exists m_0 such that if $H = (U, W; F)$ is a bipartite graph satisfying*

(i) $|U| = m_1, |W| = m_2$ *and* $m_1, m_2 > m_0$,

(ii) $Cqm_1m_2 \geq e_H(U, W) \geq \alpha q m_1 m_2$ *for some function* $q = q(m_0) \gg 1/m_0$, *and*

(iii) H *is* (ε, q)*-regular,*

then there exists a subgraph $H' = (U, W; F') \subseteq H$ *such that*

(ii′) $e_{H'}(U, W) = \alpha q m_1 m_2$ *and*

(iii′) H' *is* $(3\varepsilon, q)$*-regular.*

Proof. We select a set D of

$$|D| = e_H(U, W) - \alpha q m_1 m_2$$

edges in $E_H(U, W)$ uniformly at random and fix $H' = (U, W; F \setminus D)$. We naturally define the density in D with respect to q for sets $U' \subseteq U$ and $W' \subseteq W$ by

$$d_{D,q}(U', W') = \frac{|E_H(U', W') \cap D|}{q|U'||W'|}. \tag{3.1}$$

In order to check the $(3\varepsilon, H', q)$-regularity of (U, W), it is enough to verify the inequality corresponding to (2.2) for sets $U' \subseteq U$, $W' \subseteq W$ such that $|U'| = 3\varepsilon m_1$ and $|W'| = 3\varepsilon m_2$. Let (U', W') be such a pair. We distinguish three cases depending on $|D|$ and $e_H(U', W')$.

Case 1 *($|D| \le \varepsilon^3 q m_1 m_2$).* The graph H is (ε, H, q)-regular and thus

$$d_{H,q}(U', W') \ge d_{H,q}(U, W) - \varepsilon.$$

Since $d_{H',q}(U', W') \ge d_{H,q}(U', W') - d_{D,q}(U', W')$, we have

$$d_{H',q}(U', W') \ge d_{H,q}(U', W') - \frac{|D|}{9\varepsilon^2 q m_1 m_2} \ge d_{H,q}(U, W) - \frac{10}{9}\varepsilon,$$

which implies that H' is $(3\varepsilon, q)$-regular.

Case 2 *($e_H(U', W') \le \varepsilon^3 q m_1 m_2$).* Observe that $e_H(U', W') \le \varepsilon^3 q m_1 m_2$ implies

$$d_{H,q}(U', W') \le \frac{\varepsilon}{9}. \tag{3.2}$$

Since H is (ε, H, q)-regular

$$d_{H,q}(U, W) \le \varepsilon + d_{H,q}(U', W') \le \frac{10}{9}\varepsilon. \tag{3.3}$$

On the other hand, $d_{H',q}(X, Y) \le d_{H,q}(X, Y)$ for arbitrary $X \subseteq U$ and $Y \subseteq W$, which combined with (3.2) and (3.3) yields

$$|d_{H',q}(U, W) - d_{H',q}(U', W')| \le \frac{10}{9}\varepsilon + \frac{\varepsilon}{9} \le 3\varepsilon.$$

Up to now, we have not used the fact that D is chosen at random. To deal with the case that we are left with (that is, the case in which $|D| > \varepsilon^3 q m_1 m_2$ and $e_H(U', W') > \varepsilon^3 q m_1 m_2$), we will make use of this randomness. Before we start, we state the following two-sided estimate for the hypergeometric distribution.

Lemma 15. *Let sets $B \subseteq U$ be fixed. Let $|U| = u$ and $|B| = b$. Suppose we select a d-set D uniformly at random from U. Then, for $3/2 \ge \lambda > 0$, we have*

$$\mathbb{P}\left(\left||D \cap B| - \frac{bd}{u}\right| \ge \lambda \frac{bd}{u}\right) \le 2\exp\left(-\frac{\lambda^2}{3}\frac{bd}{u}\right).$$

Proof. For the proof we refer to [12] (Theorem 2.10). □

We continue with the proof of Lemma 14.

Case 3 *($|D| > \varepsilon^3 qm_1m_2$* **and** *$e_H(U', W') > \varepsilon^3 qm_1m_2$).* Recall that $U' \subseteq U$ and $V' \subseteq V$ are such that $|U'| = 3\varepsilon m_1$ and $|V'| = 3\varepsilon m_2$. First, we verify that

$$\left| d_{D,q}(U,W) \frac{d_{H,q}(U',W')}{d_{H,q}(U,W)} - d_{D,q}(U',W') \right| \leq \varepsilon \tag{3.4}$$

implies that

$$|d_{H',q}(U,W) - d_{H',q}(U',W')| \leq 3\varepsilon. \tag{3.5}$$

Indeed, straightforward calculation using the (ε, q)-regularity of H and (3.4) give

$$\begin{aligned}
&|d_{H',q}(U,W) - d_{H',q}(U',W')| \\
&\qquad = |(d_{H,q}(U,W) - d_{D,q}(U,W)) - (d_{H,q}(U',W') - d_{D,q}(U',W'))| \\
&\qquad \leq \varepsilon + |d_{D,q}(U,W) - d_{D,q}(U',W')| \\
&\qquad \leq \varepsilon + \left| d_{D,q}(U,W) - d_{D,q}(U,W)\frac{d_{H,q}(U',W')}{d_{H,q}(U,W)} \right| \\
&\qquad\qquad\qquad + \left| d_{D,q}(U,W)\frac{d_{H,q}(U',W')}{d_{H,q}(U,W)} - d_{D,q}(U',W') \right| \\
&\qquad \leq \varepsilon + \frac{d_{D,q}(U,W)}{d_{H,q}(U,W)} |d_{H,q}(U,W) - d_{H,q}(U',W')| + \varepsilon \\
&\qquad \leq \varepsilon + \frac{d_{D,q}(U,W)}{d_{H,q}(U,W)}\varepsilon + \varepsilon \\
&\qquad \leq 3\varepsilon.
\end{aligned}$$

Next, we will prove that (3.4) is unlikely to fail, because of the random choice of D. We set

$$\lambda = \min\left\{ \frac{9\varepsilon^3}{C}, \frac{3}{2} \right\}. \tag{3.6}$$

Then the two-sided estimate in Lemma 15 gives that

$$\left| |D \cap E_H(U',W')| - \frac{e_H(U',W')|D|}{e_H(U,W)} \right| < \lambda \frac{e_H(U',W')|D|}{e_H(U,W)}$$

fails with probability

$$\leq 2\exp\left(-\frac{\lambda^2}{3}\frac{e_H(U',W')|D|}{e_H(U,W)}\right). \tag{3.7}$$

Since

$$\left| d_{D,q}(U',W') - d_{D,q}(U,W)\frac{d_{H,q}(U',W')}{d_{H,q}(U,W)} \right| \\ = \frac{1}{9\varepsilon^2 q m_1 m_2}\left| |D \cap E_H(U',W')| - \frac{e_H(U',W')|D|}{e_H(U,W)} \right|,$$

and because of (ii) and (3.6), we have

$$\lambda \frac{e_H(U',W')}{9q\varepsilon^2 m_1 m_2}\frac{|D|}{e_H(U,W)} \leq \lambda \frac{e_H(U',W')}{9q\varepsilon^2 m_1 m_2} \leq \lambda \frac{e_H(U,W)}{9q\varepsilon^2 m_1 m_2} \leq \varepsilon,$$

we infer that (3.4) and consequently (3.5) fails with small probability given in (3.7).

We now sum over all possible choices for U' and W' and use $|D| > \varepsilon^3 q m_1 m_2$, $e_H(U',W') > \varepsilon^3 q m_1 m_2$ and (ii). We have that

$$\mathbb{P}\left(H' \text{ is not } (3\varepsilon, q)\text{-regular}\right) \leq 2^{m_1+m_2} \cdot 2\exp\left(-\frac{\lambda^2\varepsilon^6}{3C} q m_1 m_2\right) < 1$$

for m_1, m_2 sufficiently large, since $q = q(m_0) \gg 1/m_0$. This implies that, for m_0 large enough, there is a set D such that H' is $(3\varepsilon, q)$-regular, as required.

□

3.3 Proof of the main result

The proof of Theorem 5 is based on Lemma 10, which we prove later in Chapter 4. The main idea is to "find" a subgraph J satisfying (I)–(IV) of the Counting Lemma, in the arbitrary subgraph F with

$$e(F) \geq \left(1 - \frac{1}{\chi(H)-1} + \delta\right) q \binom{n}{2}.$$

Proof of Theorem 5. Let H be a fixed d-degenerate graph on h vertices and let the vertices of H be ordered $V(H) = \{w_1, \dots, w_h\}$ such that each w_i has at most d neighbours in $\{w_1, \dots, w_{i-1}\}$. Let $1/(\chi(H) - 1) > \delta > 0$ be fixed and suppose $q = q(n) \gg ((\log n)^4/n)^{1/d}$. First we define some constants that will be used in the proof.

We start by setting

$$\begin{aligned} \alpha &= \frac{\delta}{8}, && (3.8) \\ \sigma &= 10^{-6}. && (3.9) \end{aligned}$$

(As a matter of fact, our proof is not sensitive to the value of the constant σ; in fact, as long as $0 < \sigma < 1$, every choice works.) We want to use the Counting Lemma, Lemma 10, in order to determine the value of ε. Set $\alpha^{\mathrm{CL}} = \alpha$ and $\sigma^{\mathrm{CL}} = \sigma$, then Lemma 10 yields $\varepsilon^{\mathrm{CL}}$. We set

$$\varepsilon = \min\left\{\frac{\varepsilon^{\mathrm{CL}}}{3}, \frac{\delta}{80}\right\} \tag{3.10}$$

and

$$C = 1 + \frac{\delta}{4}. \tag{3.11}$$

We then apply the sparse regularity lemma (Theorem 8) with $\varepsilon^{\mathrm{SRL}} = \varepsilon$, $C^{\mathrm{SRL}} = C$ and $t_0^{\mathrm{SRL}} = \max\{h, 40/\delta\}$. Theorem 8 then gives ξ^{SRL} and we define

$$\xi = \xi^{\mathrm{SRL}}.$$

Moreover, Theorem 8 yields

$$T_0^{\mathrm{SRL}} \geq t = t^{\mathrm{SRL}} \geq t_0^{\mathrm{SRL}} = \max\left\{h, \frac{40}{\delta}\right\}. \tag{3.12}$$

For the rest of the proof all the constants defined above (α, σ, ε, C, ξ, and t) are fixed.

Fact 11, Lemma 12, and Lemma 10 imply that a graph G in $\mathcal{G}(n, q)$ satisfies the following properties (P1)–(P3) with probability $1 - o(1)$:

(P1) $e(G) \geq (1 + o(1))\, q\binom{n}{2}$,

(P2) G is (ξ, C)-bounded, and

(P3) G satisfies the property considered in Lemma 10.

We will show that if a graph G satisfies (P1)–(P3), then any $F \subseteq G$ with $e(F) \geq (1 - 1/(\chi(H) - 1) + \delta) q\binom{n}{2}$ contains at least $cq^{e(H)}n^h$ (for some constant $c = c(\delta, H)$) copies of H, and Theorem 5 will follow.

To achieve this, we first regularise F by applying Theorem 8 with $\varepsilon^{\mathrm{SRL}} = \varepsilon$, $C^{\mathrm{SRL}} = C$ and $t_0^{\mathrm{SRL}} = \max\{h, 40/\delta\}$. Consequently F admits an (ε, q)-regular (ε, t)-equitable partition $(V_i)_0^t$. We set $m = n/t = |V_i|$ for $i \neq 0$.

Let F_{cluster} be the cluster graph of F with respect to $(V_i)_0^t$ defined as follows

$$\begin{aligned} V\left(F_{\mathrm{cluster}}\right) &= \{1, \dots, t\}, \\ E\left(F_{\mathrm{cluster}}\right) &= \Big\{ \{i, j\} \colon\ (V_i, V_j) \text{ is } (\varepsilon, q)\text{-regular } \wedge\ e_F(V_i, V_j) \geq \alpha q m^2 \Big\}. \end{aligned}$$

Our next aim is to apply Theorem 2 to guarantee the existence of a copy of H in F_{cluster}. For this we define a subgraph F' of F. Set

$$E(F') = \bigcup \{E_F(V_i, V_j) \colon \{i, j\} \in E(F_{\mathrm{cluster}})\}$$

We now want to find a lower bound for $e(F')$. There are four possible reasons for an edge $e \in E(F)$ not to be in $E(F')$:

(R1) e has at least one vertex in V_0,

(R2) e is contained in some vertex class V_i for $1 \leq i \leq t$,

(R3) e is in $E(V_i, V_j)$ for an (ε, q)-irregular pair (V_i, V_j), or

(R4) e is in $E(V_i, V_j)$ for sparse a pair (*i.e.*, $e(V_i, V_j) < \alpha q m^2$).

We bound the number of discarded edges of type (R1)–(R3) by applying that G is (ξ, C)-bounded (Property (P2)):

$$\begin{aligned}
\#\text{ of edges of type (R1)} &\leq Cq\varepsilon n^2, \\
\#\text{ of edges of type (R2)} &\leq Cq\left(\frac{n}{t}\right)^2 \cdot t, \\
\#\text{ of edges of type (R3)} &\leq Cq\left(\frac{n}{t}\right)^2 \cdot \varepsilon\binom{t}{2}.
\end{aligned}$$

Furthermore, we bound the number of discarded edges of type (R4), by

$$\#\text{ of edges of type (R4)} \leq \alpha q\left(\frac{n}{t}\right)^2 \cdot \binom{t}{2}.$$

This, combined with $n \geq 2$, (3.8), (3.10), (3.11), (3.12), and $\delta < 1$ implies that

$$\begin{aligned}
|E(F) \setminus E(F')| &\leq \left(C\left(\varepsilon + \frac{1}{t} + \frac{\varepsilon}{2}\right) + \frac{\alpha}{2}\right) qn^2 \\
&\leq \left(C\left(2\varepsilon + \frac{1}{t}\right) + \frac{\alpha}{2}\right) \cdot 4q\binom{n}{2} \\
&\leq \left((4+\delta)\left(\frac{\delta}{40} + \frac{\delta}{40}\right) + \frac{\delta}{4}\right) q\binom{n}{2} \\
&\leq \frac{\delta}{2}q\binom{n}{2},
\end{aligned}$$

and thus

$$e(F') \geq \left(1 - \frac{1}{\chi(H)-1} + \frac{\delta}{2}\right) q\binom{n}{2}.$$

We use the last inequality and once again (P2) to achieve the desired lower bound for $e(F_{\text{cluster}})$. Indeed,

$$e(F_{\text{cluster}}) \geq \frac{e(F')}{Cq(n/t)^2} \geq \left(1 - \frac{1}{\chi(H)-1} + \frac{\delta}{2}\right)\left(1 - \frac{1}{n}\right)\left(1 + \frac{\delta}{4}\right)^{-1} \frac{t^2}{2},$$

and then, for n large enough ($n > 16/\delta^2$), using $t \geq h$, we deduce that

$$\begin{aligned}
e(F_{\text{cluster}}) &> \left(1 - \frac{1}{\chi(H)-1} + \frac{\delta}{2}\right)\left(1 - \frac{\delta}{4}\right)\frac{t^2}{2} \\
&\geq \left(1 - \frac{1}{\chi(H)-1} + \frac{\delta}{8}\right)\binom{t}{2}
\end{aligned}$$

The last inequality implies, by Theorem 2, that F_{cluster} contains H as a subgraph. Let $\{i_1, \ldots, i_h\}$ be the vertex set of this H in F_{cluster}. Then we set $J_0 = F[V_{i_1}, \ldots, V_{i_h}] \subseteq F$. Now, for every edge $\{w_j, w_{j'}\} \in E(H)$ the pair $(V_{i_j}, V_{i_{j'}})$ satisfies the conditions of Lemma 14 with $\varepsilon^{\text{Lem14}} = \varepsilon$, $\alpha^{\text{Lem14}} = \alpha$, and $C^{\text{Lem14}} = C$. Thus there is a subgraph $J \subseteq J_0 \subseteq F$ that is $(3\varepsilon, q)$-regular and $e_J(V_{i_j}, V_{i'_j}) = \alpha q m^2$ for every $\{w_j, w_{j'}\} \in E(H)$. Observe $\varepsilon \leq \varepsilon^{\text{CL}}/3$ and J satisfies conditions (I)–(IV) of the Counting Lemma, Lemma 10, with the constants chosen above ($\alpha^{\text{CL}} = \alpha$, $\sigma^{\text{CL}} = \sigma$, and $\varepsilon^{\text{CL}} \geq 3\varepsilon$), and thus there are at least

$$(1-\sigma)p^{e(H)}m^h = \frac{(1-\sigma)\alpha^{e(H)}}{t^h} q^{e(H)} n^h \geq \frac{(1-\sigma)\alpha^{e(H)}}{\left(T_0^{\text{SRL}}\right)^h} q^{e(H)} n^h$$

different copies of H in $J \subseteq F$. Observe that α, σ, and T_0 depend on δ and H but not on n. Consequently, there are $c(\delta, H) q^{e(H)} n^h \gg 1$ (where $c(\delta, H) = (1-\sigma)\alpha^{e(H)} / \left(T_0^{\text{SRL}}\right)^h$) copies of H in F, as required by Theorem 5. $\square$

Chapter 4

The counting lemma

Our aim in this section is to prove Lemma 10. In order to do this, we will need two lemmas (Lemma 18 and 22). We introduce these in the first two sections. Then, in Section 4.3, we will illustrate the proof of the Counting lemma on the particular case $H = K_4 - e$. Finally, we give the proof of Lemma 10 in Section 4.4.

4.1 The pick-up lemma

Before we state the 'Pick-Up Lemma', Lemma 18, let us state a simple one-sided estimate for the hypergeometric distribution, which will be useful in the proof of Lemma 18.

Lemma 16 (A hypergeometric tail lemma). *Let b, d, and u be positive integers and suppose we select a d-set D uniformly at random from a set U of cardinality u. Suppose also that we are given a fixed b-set $B \subseteq U$. Then we have for $\lambda > 0$*

$$\mathbb{P}\left(|D \cap B| \geq \lambda \frac{bd}{u}\right) \leq \left(\frac{\mathrm{e}}{\lambda}\right)^{\lambda bd/u}. \tag{4.1}$$

Proof. For the proof we refer the reader to [16]. $\square$

We now state and prove the Pick-Up Lemma. Let $k \geq 2$ be a fixed integer and let m be sufficiently large. Let $V_1, \ldots, V_k$ be pairwise disjoint sets all of size m and let $\mathcal{B}$ be a subset of $V_1 \times \cdots \times V_k$. For $1 > p = p(m) \gg 1/m$ set $T = pm^2$ and consider the probability space

$$\Omega = \binom{V_1 \times V_k}{T} \times \cdots \times \binom{V_{k-1} \times V_k}{T},$$

where $\binom{V_i \times V_k}{T}$ denotes the family of all subsets of $V_i \times V_k$ of size T, and all the $R = (R_1, \ldots, R_{k-1}) \in \Omega$ are equiprobable, *i.e.*, have probability

$$\binom{m^2}{T}^{-(k-1)}.$$

For $1 \leq i < k$ and $R_i \in \binom{V_i \times V_k}{T}$ the *degree with respect to* R_i of a vertex v_k in V_k is

$$d_{R_i}(v_k) = |\{v_i \in V_i \colon (v_i, v_k) \in R_i\}|. \tag{4.2}$$

Definition 17 *($\Pi(\zeta, \mu, K, \mathcal{B})$). For ζ, μ, K with $1 > \zeta$, $\mu > 0$ and $K > 0$ and $\mathcal{B} \subseteq V_1 \times \cdots \times V_k$, we say that property $\Pi(\zeta, \mu, K, \mathcal{B})$ holds for $R = (R_1, \ldots, R_{k-1}) \in \Omega$ if*

$$\widetilde{V}_k = \widetilde{V}_k(K) = \{v_k \in V_k \colon \ d_{R_i}(v_k) \leq Kpm, \ \forall 1 \leq i \leq k-1\}$$

and

$$\mathcal{B}(R) = \{b = (v_1, \ldots, v_k) \in \mathcal{B} \colon \ v_k \in \widetilde{V}_k \text{ and } (v_j, v_k) \in R_j, \ \forall\, 1 \leq j \leq k-1\}$$

satisfy the inequalities

$$|\widetilde{V}_k| \ \geq \ (1-\mu)m, \tag{4.3}$$

$$|\mathcal{B}(R)| \ \leq \ \zeta p^{k-1} m^k. \tag{4.4}$$

We think of $\mathcal{B}(R)$ as the members of $\mathcal{B}$ that have been *picked-up* by the random element $R \in \Omega$. We will be interested in the probability that the property $\Pi(\zeta, \mu, K, \mathcal{B})$ fails for a fixed $\mathcal{B}$ in the uniform probability space Ω.

Lemma 18 (Pick-Up Lemma). *For every β, ζ and μ with $1 > \beta, \zeta, \mu > 0$ there exist $1 > \eta = \eta(\beta, \zeta, \mu) > 0$, $K = K(\beta, \mu) > 0$ and m_0 such that if $m \geq m_0$, $\mathcal{B} \subseteq V_1 \times \cdots \times V_k$ and*

$$|\mathcal{B}| \leq \eta m^k, \tag{4.5}$$

then

$$\mathbb{P}(\Pi(\zeta, \mu, K, \mathcal{B}) \text{ fails for } R \in \Omega) \leq \beta^{(k-1)T}. \tag{4.6}$$

For the proof we need a few definitions. Suppose $\mathcal{B}$, β and μ are given. We define

$$\theta = \frac{1}{2}\beta^{k-1}, \tag{4.7}$$

$$K = \max\left\{\frac{3(k-1)\log 1/\theta}{\mu}, \mathrm{e}^2\right\}. \tag{4.8}$$

Since $p \gg 1/m$ the definition of $K \geq 3(k-1)\log(1/\theta)/\mu$ implies that

$$(k-1)\binom{m}{\mu m/(k-1)} \exp\left(-\frac{\mu T K \log K}{2(k-1)}\right) \leq \theta^T \tag{4.9}$$

holds for m sufficiently large.

Using the definition of d_{R_i} in (4.2) we construct for each $i = 1, \ldots, k-1$ a subset of V_k by putting

$$V_k^{(i)} = \{v_k \in V_k^{(i-1)} \colon\ d_{R_i}(v_k) \leq Kpm\},$$

where $V_k^{(0)} = V_k$. Observe that $V_k = V_k^{(0)} \supseteq V_k^{(1)} \supseteq \cdots \supseteq V_k^{(k-1)} = \widetilde{V}_k$. In the view of Lemma 18 we define the following "bad" events in Ω.

Definition 19 *(A_i, B). For each $i = 0, \ldots, k-1$ and K, $\mu > 0$, $\zeta > 0$, and $\mathcal{B} \subseteq V_1 \times \cdots \times V_k$ let $A_i = A_i(\mu, K)$, $B = B(\zeta, K) \subseteq \Omega$ be the events*

$$\begin{aligned} A_i\colon \quad |V_k^{(i)}| &< \big(1 - i\mu/(k-1)\big)m, \\ B\colon \quad |\mathcal{B}(R)| &> \zeta p^{k-1} m^k. \end{aligned}$$

Observe that the definition of $V_k^{(0)} = V_k$ implies

$$\mathbb{P}(A_0) = 0. \tag{4.10}$$

We restate Lemma 18 by using the notation introduced in Definition 19.

Lemma 18′ (Pick-up Lemma, event version)**.** *For every β, ζ and μ with $1 > \beta, \zeta, \mu > 0$ there exist $1 > \eta = \eta(\beta, \zeta, \mu) > 0$, $K = K(\beta, \mu) > 0$ and m_0 such that if $m \geq m_0$, $\mathcal{B} \subseteq V_1 \times \cdots \times V_k$ and*

$$|\mathcal{B}| \leq \eta m^k, \tag{4.11}$$

then

$$\mathbb{P}(A_{k-1}(\mu, K) \vee B(\zeta, K)) \leq \beta^{(k-1)T}. \tag{4.12}$$

We need some more preparation before we prove Lemma 18′. Suppose β, ζ, μ are given by Lemma 18′ and θ, K are fixed by (4.7) and (4.8). For each $i = 1, \ldots, k-1$ we consider the set $\mathcal{B}_i \subseteq \mathcal{B}$ consisting of those k-tuples $b \in \mathcal{B}$ which were partially "picked up" by edges of $R_1, \ldots, R_i$. For technical reasons we consider only those k-tuples containing vertices $v_k \in V_k^{(i-1)}$, *i.e.*, with $d_{R_j}(v_k) \leq Kpm$ for $j = 1, \ldots, i-1$. More formally, we let

$$\mathcal{B}_i = \{b = (v_1, \ldots, v_k) \in \mathcal{B}\colon\ v_k \in V_k^{(i-1)} \text{ and } (v_j, v_k) \in R_j,\ \forall\ 1 \leq j \leq i\}.$$

We also set $\mathcal{B}_0 = \mathcal{B}$.

The definitions of $\widetilde{V}_k = V_k^{(k-1)} \subseteq V_k^{(k-2)}$ and $\mathcal{B}_{k-1}$ imply

$$\mathcal{B}(R) \subseteq \mathcal{B}_{k-1}. \tag{4.13}$$

(Equality may fail in (4.13) because we may have $V_k^{(k-2)} \setminus V_k^{(k-1)} \neq \emptyset$.) For each $i = k, \ldots, 1$ define ζ_{i-1} by

$$\begin{aligned} \zeta_{k-1} &= \zeta, \\ \zeta_{i-1} &= \frac{k-1-(i-1)\mu}{4(k-1)K^{i-1}} \zeta_i^2 \theta^{4K^{i-1}/\zeta_i}. \end{aligned} \tag{4.14}$$

Furthermore, consider for each $i = 0, \dots, k-1$ the event $B_i = B_i(\zeta_i, K) \subseteq \Omega$ defined by

$$B_i\colon \quad |\mathcal{B}_i| > \zeta_i p^i m^k. \tag{4.15}$$

In order to prove Lemma 18′ we need two more claims, which we will prove later.

Claim 20. *For all $1 \le i \le k-1$, we have*

$$\mathbb{P}(A_i) = \mathbb{P}\left(|V_k^{(i)}| < \left(1 - \frac{i\mu}{k-1}\right) m\right) \le \theta^T.$$

Claim 21. *For all $1 \le i \le k-1$, we have*

$$\mathbb{P}(B_i \mid \neg A_{i-1} \wedge \neg B_{i-1}) \le \theta^T.$$

Assuming Claims 20 and 21, we may easily prove Lemma 18′.

Proof of Lemma 18′. Set $\eta = \zeta_0$ where ζ_0 is given by (4.14). The definition of $\mathcal{B}_0 = \mathcal{B}$ and (4.11) implies $|\mathcal{B}_0| \le \zeta_0 m^k$ and consequently by the definition of the event B_0 in (4.15)

$$\mathbb{P}(B_0) = 0. \tag{4.16}$$

Because of (4.13) and $\zeta_{k-1} = \zeta$ in (4.14) we have

$$\mathbb{P}(B) \le \mathbb{P}(B_{k-1}). \tag{4.17}$$

Using the formal identity

$$\mathbb{P}(B_i) = \mathbb{P}(B_i \wedge (\neg A_{i-1} \wedge \neg B_{i-1})) + \mathbb{P}(B_i \wedge (A_{i-1} \vee B_{i-1})),$$

we observe that

$$\mathbb{P}(B_i) \le \mathbb{P}(B_i \mid \neg A_{i-1} \wedge \neg B_{i-1}) + \mathbb{P}(A_{i-1}) + \mathbb{P}(B_{i-1}) \tag{4.18}$$

for each $i = 1, \dots, k-1$. It follows by applying (4.17) and (4.18) that

$$\begin{aligned}\mathbb{P}(A_{k-1} \vee B) &\leq \mathbb{P}(A_{k-1}) + \mathbb{P}(B_{k-1}) \\ &\leq \mathbb{P}(A_{k-1}) + \sum_{i=1}^{k-1} \Big(\mathbb{P}(B_i \mid \neg A_{i-1} \wedge \neg B_{i-1}) + \mathbb{P}(A_{i-1}) \Big) + \mathbb{P}(B_0).\end{aligned}$$

Claims 20 and 21, and (4.10), (4.16) and (4.7) finally imply

$$\mathbb{P}(A_{k-1} \vee B) \leq 2(k-1)\theta^T \leq 2(k-1) \left(\frac{\beta^{k-1}}{2} \right)^T \leq \beta^{(k-1)T}$$

for m sufficiently large, as required. $\square$

We now prove Claim 20 and then Claim 21.

Proof of Claim 20. Fix a set $V^* \subseteq V_k$ of size $\mu m/(k-1)$. For a fixed j ($1 \leq j \leq i$) assume that $d_{R_j}(v_k) > Kpm$ for every v_k in V^*. This clearly implies the event

$$E_j(V^*)\colon \quad |R_j \cap (V_j \times V^*)| > Kpm \frac{\mu m}{k-1} = K \frac{\mu T}{k-1}. \tag{4.19}$$

The T pairs of R_j are chosen uniformly in $V_j \times V_k$, so the hypergeometric tail lemma, Lemma 16, applies, and using the fact that $\mathrm{e} \leq K^{1/2}$ by (4.8) we get

$$\mathbb{P}\left(E_j(V^*)\right) \leq \left(\frac{\mathrm{e}}{K} \right)^{K\mu T/(k-1)} \leq \exp\left(-\frac{\mu T K \log K}{2(k-1)} \right). \tag{4.20}$$

Set $E_j = \bigvee E_j(V^*)$, where the union is taken over all $V^* \subseteq V_k$ of size $\mu m/(k-1)$. Then

$$\mathbb{P}(E_j) \leq \binom{m}{\mu m/(k-1)} \exp\left(-\frac{\mu T K \log K}{2(k-1)} \right) \tag{4.21}$$

holds for each $j = 1, \dots, i$, and this implies

$$\mathbb{P}\left(\bigvee_{j=1}^{i} E_j \right) \leq i \binom{m}{\mu m/(k-1)} \exp\left(-\frac{\mu T K \log K}{2(k-1)} \right).$$

Finally, the fact that $A_i \subseteq \bigvee_{j=1}^{i} E_j$ and the choice of K with (4.9) gives that

$$\mathbb{P}(A_i) \leq i \binom{m}{\mu m/(k-1)} \exp\left(-\frac{\mu T K \log K}{2(k-1)}\right) \leq \theta^T,$$

as required. □

Proof of Claim 21. Recall β, ζ and μ are given by Lemma 18′ and θ, K and ζ_i are fixed by (4.7), (4.8) and (4.14). In order to prove Claim 21 we fix i ($1 \leq i \leq k-1$) and we assume $\neg A_{i-1}$ and $\neg B_{i-1}$ occur. This means by Definition 19 and (4.15) that

$$|V_k^{(i-1)}| \geq \left(1 - \frac{(i-1)\mu}{k-1}\right) m = \left(\frac{k-1-(i-1)\mu}{k-1}\right) m, \tag{4.22}$$

$$|\mathcal{B}_{i-1}| \leq \zeta_{i-1} p^{i-1} m^k. \tag{4.23}$$

We have to show that

$$|\mathcal{B}_i| \leq \zeta_i p^i m^k \tag{4.24}$$

holds for R in the uniform probability space Ω with probability $\geq 1 - \theta^T$.

First we define the auxiliary constant

$$L_i = \left(\frac{1}{\theta}\right)^{4K^{i-1}/\zeta_i}. \tag{4.25}$$

The definition of θ in (4.7) and the facts that $0 < \zeta_i < 1$ for each $i = 1, \ldots, k-1$ and $K > 1$ imply that

$$L_i \geq \left(\frac{2}{\beta^{k-1}}\right)^4 > \mathrm{e}^2 \tag{4.26}$$

holds.

We define the degree of a pair in $V_i \times V_k^{(i-1)}$ with respect to $\mathcal{B}_{i-1}$ by

$$d_{\mathcal{B}_{i-1}}(w_i, w_k) = \Big|\{b = (v_1, \ldots, v_k) \in \mathcal{B}_{i-1} \colon\ v_i = w_i \text{ and } v_k = w_k\}\Big|.$$

We can bound the value of the average degree by (4.22) and (4.23):

$$\operatorname{avg}\left\{d_{\mathcal{B}_{i-1}}(v_i,v_k)\colon\ (v_i,v_k)\in V_i\times V_k^{(i-1)}\right\} = \frac{|\mathcal{B}_{i-1}|}{m|V_k^{(i-1)}|} \le \frac{k-1}{k-1-(i-1)\mu}\zeta_{i-1}p^{i-1}m^{k-2}. \tag{4.27}$$

We also can bound $\Delta_{\mathcal{B}_{i-1}}(V_i,V_k^{(i-1)}) = \max\{d_{\mathcal{B}_{i-1}}(v_i,v_k)\colon\ (v_i,v_k)\in V_i\times V_k^{(i-1)}\}$ by the following observation. Let (v_i,v_k) be an arbitrary element in $V_i\times V_k^{(i-1)}$. Then, by the definition of $V_k^{(i-1)}$, we have

$$d_{\mathcal{B}_{i-1}}(v_i,v_k)\le d_{R_1}(v_k)\cdot\ldots\cdot d_{R_{i-1}}(v_k)\cdot m^{k-2-(i-1)}\le (Kpm)^{i-1}m^{k-i-1}. \tag{4.28}$$

Inequality (4.28) implies

$$\Delta_{\mathcal{B}_{i-1}}\left(V_i,V_k^{(i-1)}\right)\le K^{i-1}p^{i-1}m^{k-2}. \tag{4.29}$$

Let F be the set of pairs of "high degree". More precisely, set

$$F=\left\{(v_i,v_k)\in V_i\times V_k^{(i-1)}\colon\ d_{\mathcal{B}_{i-1}}>\frac{\zeta_i}{2}p^{i-1}m^{k-2}\right\}.$$

A simple averaging argument applying (4.27) yields

$$|F|\le\frac{2(k-1)\zeta_{i-1}}{(k-1-(i-1)\mu)\zeta_i}|V_i||V_k^{(i-1)}|\le\frac{2(k-1)\zeta_{i-1}}{(k-1-(i-1)\mu)\zeta_i}m^2. \tag{4.30}$$

On the other hand, if we set $\bar{F}=V_i\times V_k^{(i-1)}\setminus F$ then the definition of F and (4.29) imply

$$\begin{aligned}
|\mathcal{B}_i| &= \sum_{(v_i,v_k)\in R_i\cap\bar{F}} d_{\mathcal{B}_{i-1}}(v_i,v_k)+\sum_{(v_i,v_k)\in R_i\cap F} d_{\mathcal{B}_{i-1}}(v_i,v_k)\\
&\le \frac{\zeta_i}{2}p^{i-1}m^{k-2}|R_i\cap\bar{F}| \ +\ K^{i-1}p^{i-1}m^{k-2}|R_i\cap F|\\
&\le \frac{\zeta_i}{2}p^{i-1}m^{k-2}T \ +\ K^{i-1}p^{i-1}m^{k-2}|R_i\cap F|\\
&= \left(\frac{\zeta_i}{2}+\frac{K^{i-1}}{T}|R_i\cap F|\right)p^im^k.
\end{aligned} \tag{4.31}$$

Next we prove that

$$\mathbb{P}\left(|R_i \cap F| > \frac{\zeta_i T}{2K^{i-1}}\right) \leq \theta^T, \tag{4.32}$$

which, together with (4.31), yields our claim, namely, that

$$\mathbb{P}\left(|\mathcal{B}_i| > \zeta_i p^i m^k\right) \leq \theta^T. \tag{4.33}$$

We now prove inequality (4.32). Without loss of generality we assume equality holds in (4.30). Then the hypergeometric tail lemma, Lemma 16, implies that

$$\begin{aligned}
\mathbb{P}\left(|R_i \cap F| > L_i \frac{|F|T}{m^2}\right) &= \mathbb{P}\left(|R_i \cap F| > L_i \frac{2(k-1)\zeta_{i-1}}{(k-1-(i-1)\mu)\zeta_i} T\right) \\
&\leq \left(\frac{\mathrm{e}}{L_i}\right)^{L_i \frac{2(k-1)\zeta_{i-1}}{(k-1-(i-1)\mu)\zeta_i} T} \\
&\leq \exp\left(-\frac{L_i(\log L_i)(k-1)\zeta_{i-1} T}{(k-1-(i-1)\mu)\zeta_i}\right),
\end{aligned} \tag{4.34}$$

where in the last inequality we used that $L_i \geq \mathrm{e}^2$ (see (4.26)). The definitions of ζ_{i-1} and L_i in (4.14) and (4.25) yield

$$\frac{L_i(k-1)\zeta_{i-1}}{(k-1-(i-1)\mu)\zeta_i} = \frac{L_i \zeta_i}{4K^{i-1}} \theta^{4K^{i-1}/\zeta_i} = \frac{\zeta_i}{4K^{i-1}}.$$

We use the last inequality to derive

$$\begin{aligned}
\frac{L_i(\log L_i)(k-1)\zeta_{i-1}}{(k-1-(i-1)\mu)\zeta_i} &= \log \frac{1}{\theta}, \\
L_i \frac{2(k-1)\zeta_{i-1}}{(k-1-(i-1)\mu)\zeta_i} &= \frac{\zeta_i}{2K^{i-1}},
\end{aligned}$$

which, combined with inequality (4.34), gives (4.32). $\square$

4.2 The k-tuple lemma for subgraphs of random graphs

Let $G \in \mathcal{G}(n,q)$ be the binomial random graph with edge probability $q = q(n)$, and suppose $H = (U, W; F)$ is a bipartite, not necessarily induced subgraph of G with $|U| = m_1$ and $|W| = m_2$. Furthermore, denote the density of H by $p = e(H)/m_1 m_2$.

We now consider subsets of W of fixed cardinality $k \geq 1$, and classify them according to the size of their joint neighbourhood in H. For this purpose we define

$$\mathcal{B}^{(k)}(U, W; \gamma) = \left\{ b = \{v_1, \ldots, v_k\} \in W : \ \left| d_U^H(b) - p^k m_1 \right| \geq \gamma p^k m_1 \right\},$$

where $d_U^H(b)$ denotes the size of the joint neighbourhood of b in H, that is,

$$d_U^H(b) = \left| \bigcap_{i=1}^{k} \Gamma_H(v_i) \right|.$$

The following lemma states that in a typical $G \in \mathcal{G}(n,q)$ the set $\mathcal{B}^{(k)}(U, W; \gamma)$ is "small" for any sufficiently large (ε, q)-regular subgraph $H = (U, W; F)$ of a dense enough random graph G. Recall that if G is a graph and U, $W \subset V(G)$ are two disjoint sets of vertices, then $G[U, W]$ denotes the bipartite graph naturally induced by (U, W).

Lemma 22 (The k-tuple lemma). *For any constants $\alpha > 0$, $\gamma > 0$, $\eta > 0$, and $k \geq 1$ and function $m_0 = m_0(n)$ such that $q^k m_0 \gg (\log n)^4$, there exists a constant $\varepsilon > 0$ for which the random graph $G \in \mathcal{G}(n,q)$ satisfies the following property with probability $1 - o(1)$: If for a bipartite subgraph $H = (U, W; F)$ of G the conditions*

(i) $e(H) \geq \alpha e(G[U, W])$,

(ii) *H is (ε, q)-regular,*

(iii) $|U| = m_1 \geq m_0$ *and* $|W| = m_2 \geq m_0$

apply, then

$$|\mathcal{B}^{(k)}(U, W; \gamma)| \leq \eta \binom{m_2}{k} \tag{4.35}$$

also applies.

Proof. The proof of Lemma 22 is given in [16]. □

4.3 Illustration of the proof of the counting lemma for $H = K_4 - e$

The proof of the Lemma 10 contains some technical definitions. In order to make the reading more comprehensible, we first informally illustrate the basic ideas of the proof for the case where H is the 2-degenerate graph isomorphic to $K_4 - e$, before we give the proof for a general H in Section 4.4.

We fix an order of the vertices $\{w_1, w_2, w_3, w_4\}$ of $K_4 - e$ as pictured in Figure 4.1(a). Consider the following situation: Let V_1, V_2, V_3, and V_4 be pairwise disjoint sets of vertices of size m. Let J be a 4-partite graph with vertex set $V(J) = V_1 \cup V_2 \cup V_3 \cup V_4$. We think of J as a not necessarily induced subgraph of a random graph in $\mathcal{G}(n, q)$ with $T = pm^2$ edges between each V_i and V_j ($\{w_i, w_j\} \in E(H)$), where $p = \alpha q$. We will describe a situation in which we will be able to assert that J contains the "right" number of H's. Here and everywhere below by the "right" number we mean "as expected in a random graph of density p"; notice that, for the number of $H = K_4 - e$'s, this means $\sim p^5 m^4$. Observe that, however, J is a not necessarily induced subgraph of a graph in $\mathcal{G}(n, q)$, and this makes our task hard. As it turns out, it will be more convenient to imagine that J is generated in $h - 1 = 3$ stages. First we choose the edges from V_4 to $V_1 \cup V_3$ (since $\{w_4, w_2\}$ is not an edge in H). Then we choose the edges from V_3 to $V_1 \cup V_2$, and in the third stage we disclose the edges between V_2 and V_1.

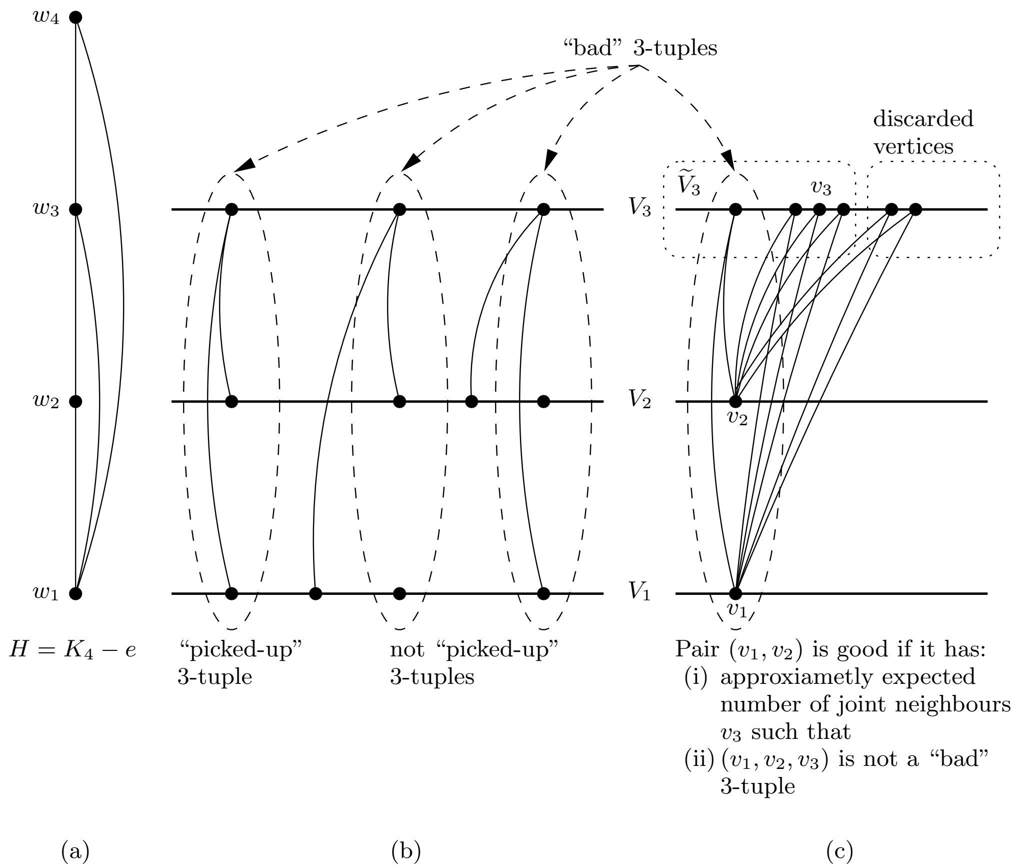

Figure 4.1: "bad" tuples

The key idea of the proof is to consider "bad" tuples, which we create in every stage. After we chose the edges from V_4 to $V_1 \cup V_3$, we define "bad" 3-tuples in $V_1 \times V_2 \times V_3$: a 3-tuple (v_1, v_2, v_3) is "bad" if the joint neighbourhood of v_1 and v_3 in V_4 is much smaller than expected. Then, with the right choice of constants, Proposition 26 for $k = 2$ and $J = J[V_4, V_1 \cup V_3]$ will ensure that there are not too many "bad" 3-tuples. (Proposition 26 is a corollary of the the k-tuple lemma, Lemma 22.)

We next generate the edges between V_3 and $V_1 \cup V_2$. We want to define "bad" pairs in $V_1 \times V_2$. Here it becomes slightly more complicated to distinguish “bad” from “good”. This is because there are two things that might go wrong for a pair in $V_1 \times V_2$. First of all, again the joint neighbourhood (now in V_3) of a pair in $V_1 \times V_2$ might be too small. On the other hand, it could have the right number of joint neighbours in V_3, but many of these neighbours “complete” the pair to a “bad” 3-tuple. Here the Pick-Up Lemma comes into play for $k = 3$ (see Proposition 25): this lemma will ensure that, given the set of “bad” 3-tuples (which was already defined in the first stage) is small, we will not “pick-up” too many of these (see Figure 4.1(b)), while choosing the edges between V_3 and $V_1 \cup V_2$. (We say that a triple (v_1, v_2, v_3) has been *picked-up* if (v_1, v_3) and (v_2, v_3) are in the edge set generated between V_3 and $V_1 \cup V_2$.)

Here the situation complicates somewhat. The Pick-Up Lemma forces us to discard a small portion (less or equal μ^{PU} fraction) of vertices in V_3. Thus, in order to avoid the first type of “badness” (too small joint neighbourhood) as a 2-tuple in $V_1 \times V_2$ it is not enough to have the right number of joint neighbours in V_3; we need the right number of joint neighbours in $\widetilde{V}_3$, which is V_3 without the $\mu^{\mathrm{PU}}m$ vertices (at most) we lose by applying the Pick-Up Lemma (see Figure 4.1(c)). This will be ensured by the the k-tuple lemma (to be more precise, Proposition 26), now for $k = 2$ and $J = J[\widetilde{V}_3, V_1 \cup V_2]$.

Later, in the general case, we will refer to the set of “bad” i-tuples in $V_1 \times \cdots \times V_i$ as $\mathcal{B}_i$ (see Definition 23 below). We define $\mathcal{B}_i$ as the union of the sets $\mathcal{B}_i^{(a)}$ and $\mathcal{B}_i^{(b)}$, defined as follows. Let $I_{i+1} = \{j \in [i] \colon \{w_{i+1}, w_j\} \in E(H)\}$ We put in $\mathcal{B}_i^{(a)}$ the i-tuples $(v_1, \dots, v_i)$ that are “bad” because the joint neighbourhood of $\{v_j \colon j \in I_{i+1}\}$ in $\widetilde{V}_{i+1}$ is too small; the set $\mathcal{B}_i^{(b)}$ is defined as the set of i-tuples in $V_1 \times \cdots \times V_i$ that “bad” because they extend to too many “bad” $(i+1)$-tuples (*i.e.*, $(i+1)$-tuples in $\mathcal{B}_{i+1}$).

As described above, we define $\mathcal{B}_i$ $(i = h-1, \dots, 1)$ by reverse induction, starting with $\mathcal{B}_{h-1}$, and going down to $\mathcal{B}_1$. With the right choice of constants, there will not be too many “bad” vertices in V_1.

Having ensured that most of the m vertices in V_1 are not "bad" (*i.e.*, do not belong to $\mathcal{B}_1$) we are now able to count the number of $H = K_4 - e$'s. We will use the following deterministic argument, which will later be formalised in Lemma 29. Consider a vertex v_1 in V_1 that is not "bad". This vertex has approximately the expected number of neighbours in $\widetilde{V}_2$ (*i.e.*, $\sim pm$), and not too many of these neighbours constitute, together with v_1, a "bad" 2-tuple. In other words, this means that v_1 extends to $\sim pm$ copies of $H_2 = H[\{w_1, w_2\}]$ in $(V_1 \times V_2) \setminus \mathcal{B}_2$. This implies that each such H_2 has the right number of joint neighbours in $\widetilde{V}_3$ (*i.e.*, $\sim p^2m$), and consequently extends to the right number of $H_3 = H[\{w_1, w_2, w_3\}]$'s in $(V_1 \times V_2 \times V_3) \setminus \mathcal{B}_3$. Repeating the last argument, each of these H_3's extends into $\sim p^2m$ (since w_4 is only adjacent to w_1 and w_3) different copies of $H = K_4 - e$. Since we have ensured that most of the m vertices in V_1 are not "bad", we have $\sim m \cdot pm \cdot p^2m \cdot p^2m = p^{e(H)}m^4$ copies of H.

4.4 Proof of the counting lemma

In this section we will prove Lemma 10. In Section 4.4.1, we introduce the key definitions and describe the logic of all important constants which will appear later in the proof. Afterwards we prove two technical propositions in Section 4.4.2. These propositions correspond to the lemmas in Sections 4.1 and 4.2, and their use will give a short proof of the Counting Lemma, to be presented in Section 4.4.3.

4.4.1 Concepts and constants

Let H be a fixed d-degenerate graph on h vertices and let the vertices of H be ordered $V(H) = \{w_1, \dots, w_h\}$ such that each w_i hast at most d neighbours in $\{w_1, \dots, w_{i-1}\}$. For every $1 \leq i \leq h$ we set I_i to the set of the indices of the neighbours of w_i in $\{w_1, \dots, w_{i-1}\}$

$$I_i = \Gamma_H(w_i) \cap \{w_1, \dots, w_{i-1}\}.$$

Let $t \geq h$ be a fixed integer and n a sufficiently large integer. Let α and ε be constants greater than 0. Let G be in $\mathcal{G}(n,q)$ with $q = q(n)$, and suppose J is an h-partite subgraph of G with vertex classes $V_1, \dots, V_h$. For all $1 \leq i < j \leq h$ we denote by J_{ij} the bipartite graph induced by V_i and V_j. Consider the following assertions for J and q.

(I) $|V_i| = m = n/t$

(II) $q^d n \gg (\log n)^4$

(III) for all $1 \leq i < j \leq h$,

$$e(J_{ij}) = \begin{cases} T = pm^2 & \{w_i, w_j\} \in E(H) \\ 0 & \{w_i, w_j\} \notin E(H) \end{cases}$$

where $1 > \alpha q = p \gg 1/n$, and

(IV) J_{ij} is (ε, q)-regular.

Let $\sigma > 0$ be given. We define the constants

$$\gamma = \mu = \nu = \frac{1}{3}\left(1 - (1-\sigma)^{1/h}\right), \tag{4.36}$$

and, for $1 \leq i \leq l-2$, we put

$$\beta_{i+1} = \begin{cases} \left(\frac{1}{2}\left(\frac{\alpha}{\mathrm{e}}\right)^{\sum_{j=i+1}^{h}|I_j|}\right)^{1/|I_{i+1}|} & I_{i+1} \neq \emptyset \\ 0 & I_{i+1} = \emptyset \end{cases}. \tag{4.37}$$

In order to prove Lemma 10 we need some definitions. These definitions always depend on a fixed subgraph J of our random graph $G \in \mathcal{G}(n,q)$

satisfying (I)–(IV). However, we will drop references to J because we want to simplify the notation (*e.g.*, we write V_i instead of V_i^J). Also, for each $i = 1, \dots, h$ we denote $V_1 \times \dots \times V_i$ by $\mathcal{W}_i$.

In the proof we consider for a fixed J sets of "bad" i-tuples $\mathcal{B}_i \subseteq \mathcal{W}_i$ ($1 \le i \le h-1$). We define these sets recursively from $\mathcal{B}_{h-1}$ to $\mathcal{B}_1$. As mentioned above in the discussion of the $H = K_4 - e$ case, there are two reasons that make a given i-tuple $(v_1, \dots, v_i)$ in $\mathcal{W}_i$ "bad". First of all, its joint neighbourhood of $\{v_j\colon\ j \in I_{i+1}\}$ in V_{i+1} might be too small (see the definition of $\mathcal{B}_i^{(a)}$ in Definition 23) and, secondly, it could extend into too many "bad" $(i+1)$-tuples in $\mathcal{B}_{i+1}$ (see the definition of $\mathcal{B}_i^{(b)}$ in Definition 23). Note that the "bad" $(i+1)$-tuples have already been defined, as we are using reverse induction in these definitions.

Next we apply the Pick-Up Lemma for $k = |I_{i+1}| + 1$ if $|I_{i+1}| > 0$ ($1 \le i \le h-2$) with $\mu_{i+1}^{\mathrm{PU}} = \mu$ and $\beta_{i+1}^{\mathrm{PU}} = \beta_{i+1}$ (and yet unspecified $\zeta_{i+1}^{\mathrm{PU}}$). As a result we obtain $K_{i+1}^{\mathrm{PU}} = K_{i+1}^{\mathrm{PU}}(\beta_{i+1}^{\mathrm{PU}}, \mu_{i+1}^{\mathrm{PU}})$ and the set

$$\widetilde{V}_{i+1} = \widetilde{V}_{i+1}^{\mathrm{PU}}(K_{i+1}^{\mathrm{PU}}) \subseteq V_{i+1}$$

of undiscarded vertices with

$$|\widetilde{V}_{i+1}| \ge (1-\mu)m$$

with probability bigger than

$$1 - \left(\beta_{i+1}^{\mathrm{PU}}\right)^{|I_{i+1}|T}.$$

For $2 \le i+1 \le h-1$ such that $|I_{i+1}| = 0$ we simply set $\widetilde{V}_{i+1} = V_{i+1}$ and, therefore, trivially $|\widetilde{V}_{i+1}| \ge (1-\mu)m$ holds.

We need a few more definitions before we define $\mathcal{B}_i$, $\mathcal{B}_i^{(a)}$ and $\mathcal{B}_i^{(b)}$ (recursively for $i = h-1, \dots, 1$). Let $\widetilde{\Gamma}_{i+1}(b)$ be the joint neighbourhood of $b = (v_1, \dots, v_i) \in \mathcal{W}_i$ in $\widetilde{V}_{i+1}$ with respect to J, more precisely

$$\widetilde{\Gamma}_{i+1}(b) = \{w \in \widetilde{V}_{i+1}\colon\ \{v_j, w\} \in E(J_{j,i+1}),\ \forall\, j \in I_{i+1}\}.$$

For a fixed set $\mathcal{B} \subseteq \mathcal{W}_{i+1}$ and $b = (v_1, \ldots, v_i) \in \mathcal{W}_i$ we denote the *degree* $d_{\mathcal{B}}(b)$ of b in $\mathcal{B}$ with respect to J by

$$d_{\mathcal{B}}(b) = \left|\left\{v \in \widetilde{\Gamma}_{i+1}(b)\colon \ (v_1, \ldots, v_i, v) \in \mathcal{B}\right\}\right|.$$

Next we define (still for a fixed J) the sets of "bad" i-tuples $\mathcal{B}_i = \mathcal{B}_i(\gamma, \mu, \nu) \subseteq \mathcal{W}_i$ mentioned earlier. Although we do not apply the Pick-Up Lemma for $k = h$, for the sake of convenience we consider the neighbourhood of elements in $\mathcal{W}_{h-1}$ in $\widetilde{V}_h$, instead of in V_h.

Definition 23 *($\mathcal{B}_{l-1}$, $\mathcal{B}_i^{(a)}$, $\mathcal{B}_i^{(b)}$, $\mathcal{B}_i$). Let γ, μ, ν be given by* (4.36)*. We define recursively the following sets of "bad" tuples for $i = h-1, \ldots, 1$:*

$$\begin{array}{rcccl}
\mathcal{B}_{h-1} & = & \mathcal{B}_{h-1}(\gamma, \mu) & = & \left\{b \in \mathcal{W}_{h-1}\colon \ \left|\widetilde{\Gamma}_h(b)\right| < (1-\gamma-\mu)p^{|I_h|}m\right\}, \\
\mathcal{B}_i^{(a)} & = & \mathcal{B}_i^{(a)}(\gamma, \mu) & = & \left\{b \in \mathcal{W}_i\colon \ \left|\widetilde{\Gamma}_{i+1}(b)\right| < (1-\gamma-\mu)p^{|I_{i+1}|}m\right\}, \\
\mathcal{B}_i^{(b)} & = & \mathcal{B}_i^{(b)}(\nu) & = & \left\{b \in \mathcal{W}_i\colon \ d_{\mathcal{B}_{i+1}}(b) \geq \nu p^{|I_{i+1}|}m\right\}, \\
\mathcal{B}_i & = & \mathcal{B}_i(\gamma, \mu, \nu) & = & \mathcal{B}_i^{(a)}(\gamma, \mu) \cup \mathcal{B}_i^{(b)}(\nu).
\end{array}$$

We also consider "bad" events in $\mathcal{G}(n, q)$ defined on the basis of the size of the sets $\mathcal{B}_{h-1}(\gamma, \mu)$, $\mathcal{B}_i^{(a)}(\gamma, \mu)$, $\mathcal{B}_i^{(b)}(\nu)$, and $\mathcal{B}_i(\gamma, \mu, \nu)$ defined above. In the following definition we mean by J an arbitrary subgraph of $G \in \mathcal{G}(n, q)$ satisfying conditions (I)–(IV).

Definition 24. *Let γ, μ, ν be given by* (4.36) *and let $\eta_i > 0$ ($i = h-1, \ldots, 1$) be fixed. We define the events*

$$\begin{aligned}
X_{h-1}(\gamma, \mu, \eta_{h-1})&\colon \exists\ J \subseteq G \text{ s.t. } |\mathcal{B}_{h-1}| > (\eta_{h-1}/2)m^{h-1}, \\
X_i^{(a)}(\gamma, \mu, \eta_i)&\colon \exists\ J \subseteq G \text{ s.t. } \left|\mathcal{B}_i^{(a)}\right| > (\eta_i/2)m^i, \\
X_i^{(b)}(\gamma, \mu, \nu, \eta_i, \eta_{i+1})&\colon \exists\ J \subseteq G \text{ s.t. } |\mathcal{B}_{i+1}| \leq \eta_{i+1}m^{i+1} \ \wedge \ |\mathcal{B}_i^{(b)}| > (\eta_i/2)m^i, \\
X_i(\gamma, \mu, \nu, \eta_i, \eta_{i+1}) &= X_i^{(a)}(\gamma, \mu, \eta_i) \ \vee \ X_i^{(b)}(\gamma, \mu, \nu, \eta_i, \eta_{i+1}).
\end{aligned}$$

For simplicity, we let

$$X_{h-1}^{(a)} = X_{h-1} = X_{h-1}(\gamma, \mu, \eta_{h-1}),$$
$$X_i^{(a)} = X_i^{(a)}(\gamma, \mu, \eta_i) \quad \text{for } i = 1, \ldots, l-1,$$
$$X_i^{(b)} = X_i^{(b)}(\gamma, \mu, \nu, \eta_i, \eta_{i+1}) \quad \text{for } i = 1, \ldots, l-2,$$

and

$$X_i = X_i(\gamma, \mu, \nu, \eta_i, \eta_{i+1}) \quad \text{for } i = 1, \ldots, l-1.$$

Owing to the special role of X_1 later in the proof, we let

$$X_{\text{bad}} = X_{\text{bad}}(\gamma, \mu, \nu, \eta_1, \eta_2) = X_1(\gamma, \mu, \nu, \eta_1, \eta_2).$$

We will now describe the remaining constants used in the proof. Notice that α and σ were given and we have already fixed γ, μ and ν in (4.36) and β_i for $2 \leq i \leq h-1$ in (4.37). The (yet unspecified) parameters η_i and ε will be determined by Propositions 25 and 26. First we set $\eta_1 = \nu$. Then Proposition 25 (PU_{i+1}) inductively describes $\eta_{i+1} = \eta_{i+1}(\beta_{i+1}, \gamma, \mu, \nu, \eta_i)$ for $i = 1, \ldots, h-2$ such that $\mathbb{P}(X_i^{(b)}) = o(1)$. Finally, for $i = 1, \ldots, h-1$, Proposition 26 (TL_i) implies the choice for $\varepsilon_i = \varepsilon_i(\alpha, \gamma, \mu, \eta_i)$ such that $\mathbb{P}(X_i^{(a)}) = o(1)$. We set

$$\varepsilon = \min\{\varepsilon_i \colon i = 1, \ldots, h-1\}.$$

A diagram illustrating the definition scheme for the constants above is given in Figure 4.2.

Thus, ε is defined for any given σ and α, as claimed in Lemma 10. From now on, these constants are fixed for the rest of the proof of Lemma 10.

4.4.2 Tools

We need some auxiliary results before we prove Lemma 10. For this purpose we state variants of the Pick-Up Lemma, Lemma 18, and of the k-tuple

$$\begin{tikzcd}
\alpha, \sigma, \gamma, \mu, \nu, \beta_2, \dots, \beta_{h-1} \arrow[d] & & & & & & \\
\eta_1 = \nu \arrow[r, "\mathrm{PU}_2"] \arrow[d, "\mathrm{TL}_1"] & \eta_2 \arrow[r, "\mathrm{PU}_3"] \arrow[d, "\mathrm{TL}_2"] & \cdots \arrow[r, "\mathrm{PU}_i"] & \eta_i \arrow[r, "\mathrm{PU}_{i+1}"] \arrow[d, "\mathrm{TL}_i"] & \cdots \arrow[r, "\mathrm{PU}_{h-1}"] & \eta_{h-1} \arrow[d, "\mathrm{TL}_{h-1}"] \\
\varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_i & \cdots & \varepsilon_{h-1}
\end{tikzcd}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\varepsilon = \min \varepsilon_i}$$

Figure 4.2: Flowchart of the constants

lemma, Lemma 22, in the form that we apply these later. These variants will be referred to as (PU_{i+1}) and (TL_i).

The next proposition follows from Lemma 18 for $k = |I_{i+1}| + 1$ ($1 \leq i \leq h-2$).

Proposition 25 **(PU_{i+1}).** *Fix $1 \leq i \leq h-2$. Let α, $\sigma > 0$ be arbitrary, let γ, μ, ν and β_{i+1} be given by (4.36) and (4.37), and let η_i be defined as stated in Section 4.4.1 (see Figure 4.2). Then there exists $\eta_{i+1} = \eta_{i+1}(\beta_{i+1}, \gamma, \mu, \nu, \eta_i) > 0$ such that for every $t \geq h$ a random graph G in $\mathcal{G}(n,q)$ satisfies the following property with probability $1 - o(1)$: If J is a subgraph of G satisfying (I)–(IV) and $\mathcal{B}_{i+1}(\gamma, \mu, \nu) \subseteq \mathcal{W}_{i+1}$ is such that*

$$|\mathcal{B}_{i+1}(\gamma, \mu, \nu)| \leq \eta_{i+1} m^{i+1}, \tag{4.38}$$

then the number of i-tuples b in $\mathcal{W}_i$ with

$$d_{\mathcal{B}_{i+1}}(b) \geq \nu p^{|I_{i+1}|} m$$

is less than

$$\frac{\eta_i}{2} m^i,$$

which means

$$\left|\mathcal{B}_i^{(b)}(\nu)\right| \leq \frac{\eta_i}{2} m^i. \tag{4.39}$$

Furthermore,

$$|\widetilde{V}_{i+1}| \geq (1-\mu)m$$

holds.

We restate Proposition 25, by using the events $X_i^{(b)}$ from Definition 24. Observe that inequalities (4.38) and (4.39) correspond to the event $X_i^{(b)}$, so that $\mathbb{P}(X_i^{(b)}) = o(1)$ is equivalent to the first part of Proposition 25′.

Proposition 25′ **(PU$_{i+1}$).** *Fix $1 \leq i \leq h-2$. Let α, $\sigma > 0$ be arbitrary, let γ, μ, ν and β_{i+1} be given by* (4.36) *and* (4.37)*, and let η_i be defined as stated in Section 4.4.1 (see Figure 4.2). Then there exists $\eta_{i+1} = \eta_{i+1}(\beta_{i+1}, \gamma, \mu, \nu, \eta_i) > 0$ such that for every $t \geq h$*

$$\mathbb{P}\left(X_i^{(b)}(\gamma, \mu, \nu, \eta_i, \eta_{i+1})\right) = o(1)$$

and

$$\mathbb{P}\left(|\widetilde{V}_{i+1}| < (1-\mu)m\right) = o(1).$$

Proof. If $|I_{i+1}| = 0$ we simply set $\eta_{i+1} = \eta_i\nu/2$ and $\widetilde{V}_{i+1} = V_{i+1}$. Suppose (4.38) holds and (4.39) fails. Then we derive

$$|\mathcal{B}_{i+1}| > \frac{\eta_i}{2}m^i \cdot \nu p^{|I_{i+1}|}m = \eta_{i+1}p^0 m^{i+1},$$

which contradicts (4.38).

Therefore, we assume $|I_{i+1}| > 0$. We apply Lemma 18 for $k = |I_{i+1}| + 1$ and with the following choice of β^{PU}, ζ^{PU}, μ^{PU}:

$$\begin{aligned}
\beta^{\mathrm{PU}} &= \beta_{i+1}, \\
\zeta^{\mathrm{PU}} &= \frac{\eta_i\nu}{2}, \\
\mu^{\mathrm{PU}} &= \mu.
\end{aligned}$$

Lemma 18 then gives η^{PU}, from which we define the constant η_{i+1} we are looking for by putting

$$\eta_{i+1} = \eta^{\mathrm{PU}}.$$

We assume inequality (4.38) holds. In other words, the number of the "bad" $(i+1)$-tuples in $\mathcal{W}_{i+1}$ is

$$|\mathcal{B}_{i+1}| \le \eta_{i+1} m^{i+1} = \eta^{\mathrm{PU}} m^{i+1}. \tag{4.40}$$

On the other hand, if we assume that (4.39) does not hold (*i.e.*, the event $X_i^{(b)}$ occurs), then the number of $(i+1)$-tuples in $\mathcal{B}_{i+1}$ that have been "picked-up" has to exceed

$$\frac{\eta_i}{2} m^i \cdot \nu p^{|I_{i+1}|} m = \zeta^{\mathrm{PU}} p^{|I_{i+1}|} m^{i+1}. \tag{4.41}$$

In particular at least $\zeta^{\mathrm{PU}} p^{|I_{i+1}|} m^{|I_{i+1}|+1}$ different $(|I_{i+1}|+1)$-tuples of $\mathcal{B}'_{i+1}$ were "picked-up", where

$$\mathcal{B}'_{i+1} = \mathcal{B}_{i+1}\Big|_{\left(\prod_{j\in I_{i+1}} V_j\right)\times V_{i+1}}$$

is the restriction of $\mathcal{B}_{i+1}$ on $\left(\prod_{j\in I_{i+1}} V_j\right)\times V_{i+1}$. The Pick-Up Lemma bounds the number of these configurations in

$$\prod_{j\in I_{i+1}} \binom{V_j \times V_{i+1}}{T}$$

by

$$\left(\beta^{\mathrm{PU}}\right)^{|I_{i+1}|T} \cdot \binom{m^2}{T}^{|I_i+1|} = (\beta_{i+1})^{|I_{i+1}|T} \binom{m^2}{T}^{|I_i+1|}. \tag{4.42}$$

We now estimate the number of all possible graphs J satisfying (I)–(IV) for which (4.40) holds but the number of members in $\mathcal{B}_{i+1}$ that have been "picked-up" exceeds (4.41). There are less than $\binom{n}{m}^h$ different ways to fix the h vertex classes of J. Furthermore, observe that $\mathcal{B}_{i+1}$ and, therefore, $\mathcal{B}'_{i+1}$ are determined by all the edges in $J_{jj'}$ ($i+1 < j' \le h$, $1 \le j < j' \le h$, which gives $L = \sum_{l=i+2}^{h} |I_l|$ different pairs (j,j') with $e(J_{jj'}) \ne 0$. Thus we have at most $\binom{m^2}{T}^L$ possibilities to determine $\mathcal{B}_{i+1}$ (*i.e.*, $\mathcal{B}'_{i+1}$). This, combined with (4.42), (III), and (4.37) yields that

$$\begin{aligned}\mathbb{P}\left(X_i^{(b)}\right) &\le \binom{n}{m}^h \binom{m^2}{T}^L \cdot (\beta_{i+1})^{|I_{i+1}|T} \binom{m^2}{T}^{|I_{i+1}|} \cdot q^{(L+|I_{i+1}|)T} \\ &\le 2^{nh} \left(\frac{\mathrm{e}m^2 q}{T}\right)^{(L+|I_{i+1}|)T} (\beta_{i+1})^{|I_{i+1}|T} \\ &\le 2^{nh} \left(\left(\frac{\mathrm{e}}{\alpha}\right)^{\sum_{j=i+1}^{h}|I_j|} (\beta_{i+1})^{|I_{i+1}|}\right)^T \\ &\le 2^{nh-T}.\end{aligned}$$

Since h is fixed and $T \gg m = n/t$, we have

$$\mathbb{P}\left(X_i^{(b)}\right) = o(1).$$

Note that the set $\widetilde{V}_{i+1}$ was determined by the application of the Pick-Up Lemma. Therefore, the second assertion in Proposition 25′ also follows from the proof above. $\square$

The following is an easy consequence of Lemma 22 for $k = |I_{i+1}|$ $(1 \le i \le h-1)$.

Proposition 26 (TL$_i$)**.** *Fix $1 \le i \le h-1$. Let α, $\sigma > 0$ be arbitrary, let γ, μ be given by (4.36), and let η_i be defined as stated in Section 4.4.1 (see Figure 4.2). Then there exists $\varepsilon_i = \varepsilon_i(\alpha, \gamma, \mu, \eta_i) > 0$ such that for every $t \ge h$ a random graph G in $\mathcal{G}(n, q)$ satisfies the following property with probability $1 - o(1)$: If $\varepsilon \le \varepsilon_i$ and J is a subgraph of G satisfying (I)–(IV), then the number of i-tuples b in $\mathcal{W}_i$ with*

$$\left|\widetilde{\Gamma}_{i+1}(b)\right| < (1-\gamma-\mu)p^{|I_{i+1}|}m$$

is less than

$$\frac{\eta_i}{2}m^i,$$

which means that

$$\left|\mathcal{B}_i^{(a)}(\gamma, \mu)\right| \le \frac{\eta_i}{2}m^i. \tag{4.43}$$

We can reformulate Proposition 26 in a shorter way by using the event $X_i^{(a)}$ (see Definition 24).

Proposition 26′ **(**TL_i**).** *Fix* $1 \le i \le h-1$*. Let* α*,* $\sigma > 0$ *be arbitrary, let* γ*,* μ *be given by* (4.36) *and let* η_i *be defined as stated in Section 4.4.1 (see Figure 4.2). Then there exists* $\varepsilon_i = \varepsilon_i(\alpha, \gamma, \mu, \eta_i) > 0$ *such that for every* $t \ge h$ *and* $\varepsilon \le \varepsilon_i$

$$\mathbb{P}\left(X_i^{(a)}(\gamma, \mu, \eta_i)\right) = o(1).$$

Proof. The proposition is trivial if $|I_{i+1}| = 0$. Therefore, without loss of generality assume $|I_{i+1}| > 0$.

We apply the k-tuple lemma, Lemma 22, with $k = |I_{i+1}|$, $\alpha^{\mathrm{TL}} = \alpha/3$, $\gamma^{\mathrm{TL}} = \gamma$ and

$$\eta^{\mathrm{TL}} = \frac{\eta_i}{(2i^i)}. \tag{4.44}$$

The k-tuple lemma gives an $\varepsilon^{\mathrm{TL}}$ and without loss of generality we may assume

$$\varepsilon^{\mathrm{TL}} \le \frac{2}{7}. \tag{4.45}$$

We set

$$\varepsilon_i = \min\left\{\left(\varepsilon^{\mathrm{TL}}\right)^3, 1-\mu, \frac{\alpha}{2}\right\}.$$

Let $\varepsilon \le \varepsilon_i$ and J be a subgraph of $G \in \mathcal{G}(n, q)$ satisfying (I)–(IV). Set $U = \widetilde{V}_{i+1}$ and $W = \bigcup_{j \in I_{i+1}}^{i} V_j$. By (IV), the graph $J_{jj'}$ $(1 \le j < j' \le i)$ is (ε, q)-regular. Due to Lemma 12 without loss of generality we may assume G is $(\xi, 3/2)$-bounded for some $\xi < \varepsilon$ Below, we verify that condition (i) and (ii) of Lemma 22 hold for $J[U, W]$ with respect to G.

Claim 27.

(i) $e(J[U, W]) \ge \alpha^{\mathrm{TL}} e(G[U, W])$*,*

(ii) $J[U, W]$ *is* $(\varepsilon^{\mathrm{TL}}, q)$*-regular,*

Proof of Claim 27. First we show (i). Since G is $(\xi, 3/2)$-bounded

$$e(G[U,W]) \leq \frac{3}{2}q(1-\mu)|I_{i+1}|m^2.$$

On the other hand, using the (ε, q)-regularity of $J_{j,i+1}$ for every $j \in I_{i+1}$ we derive

$$e(J[U,W]) \geq (\alpha - \varepsilon)q(1-\mu)|I_{i+1}|m^2$$

and, therefore, applying the choice of ε and α^{TL} gives

$$\frac{e(J[U,W])}{e(G[U,W])} \geq \frac{2(\alpha-\varepsilon)}{3} \geq \frac{\alpha}{3} \geq \alpha^{\mathrm{TL}},$$

which yields (i).

Exploiting the (ε, q)-regularity of $J_{j,i+1}$ for every $j \in I_{i+1}$ again, we observe

$$\alpha - \varepsilon \leq d_{J,q}(U,W) \leq \alpha + \varepsilon. \tag{4.46}$$

In order to verify the $(\varepsilon^{\mathrm{TL}}, J, q)$-regularity of (U,W) it suffices to show

$$|d_{J,q}(U',W') - d_{J,q}(U,W)| \leq \varepsilon^{\mathrm{TL}} \tag{4.47}$$

for set $U' \subseteq U$, $W' \subseteq W$ satisfying

$$\begin{aligned} |U'| &= \varepsilon^{\mathrm{TL}}|U| &= \varepsilon^{\mathrm{TL}}(1-\mu)m \\ |W'| &= \varepsilon^{\mathrm{TL}}|W| &= \varepsilon^{\mathrm{TL}}|I_{i+1}|m. \end{aligned} \tag{4.48}$$

For $j \in I_{i+1}$ we bound the number of edges in $J[U', W' \cap V_j]$ depending on the order of $W' \cap V_j$. If $|W' \cap V_j| \geq \varepsilon m$ we are enabled to use the (ε, q)-regularity of $J_{j,i+1}$ and derive

$$(\alpha-\varepsilon)q|U'||W' \cap V_j| \leq e(J[U', W' \cap V_j]) \leq (\alpha+\varepsilon)q|U'||W' \cap V_j|. \tag{4.49}$$

On the other hand, if $|W' \cap V_j| < \varepsilon m$ we use the $(\xi, 3/2)$-boundedness and infer

$$0 \leq e(J[U', W' \cap V_j]) \leq \frac{3}{2}q|U'||W' \cap V_j|. \tag{4.50}$$

Clearly, we get a lower and an upper bound for $e(J[U', W'])$ if we assume the 'worst case scenario': $|W' \cap V_j| < \varepsilon m$ for as many as possible $j \in I_{i+1}$. But, since $\varepsilon^{\mathrm{TL}} \geq \sqrt[3]{\varepsilon} > \varepsilon$, at least for one $j' \in I_{i+1}$ the order of $W' \cap V_j$ is bounded from below by $|I_{i+1}|\varepsilon^{\mathrm{TL}} m - (|I_{i+1}| - 1)\varepsilon m \geq \varepsilon m$. This observation accompanied by (4.48), (4.49), and (4.50) implies

$$
\begin{aligned}
(\alpha - \varepsilon)q \cdot \varepsilon^{\mathrm{TL}}(1-\mu)m \cdot \big(|I_{i+1}|\varepsilon^{\mathrm{TL}} m - (|I_{i+1}| - 1)\varepsilon m\big)& \\
\leq e(J[U', W']) \leq& \\
(|I_{i+1}| - 1) \cdot \frac{3}{2} q \cdot \varepsilon^{\mathrm{TL}}(1-\mu)m \cdot \varepsilon m \ +& \\
(\alpha + \varepsilon)q \cdot \varepsilon^{\mathrm{TL}}(1-\mu)m \cdot \big(|I_{i+1}|\varepsilon^{\mathrm{TL}} m - (|I_{i+1}| - 1)\varepsilon m\big)&,
\end{aligned}
$$

which yields

$$
\begin{aligned}
(\alpha - \varepsilon) - \frac{|I_{i+1}| - 1}{|I_{i+1}|} \frac{\varepsilon}{\varepsilon^{\mathrm{TL}}}(\alpha - \varepsilon)& \\
\leq d_{J,q}(U', W') \leq& \\
(\alpha + \varepsilon) + \frac{|I_{i+1}| - 1}{|I_{i+1}|} \frac{\varepsilon}{\varepsilon^{\mathrm{TL}}} \left(\frac{3}{2} - \alpha - \varepsilon\right)&.
\end{aligned}
$$

Finally, we compare the lower (upper) bound from above with the upper (lower) bound from (4.46) to verify (4.47). Therefore, with our choice of

$\varepsilon \leq (\varepsilon^{\mathrm{TL}})^3$ and (4.45) we observe

$$\begin{aligned}
&|d_{j,q}(U',W') - d_{j,q}(U,W)| \\
&\qquad \leq \max\left\{ \left| (\alpha-\varepsilon) - \frac{|I_{i+1}|-1}{|I_{i+1}|}\frac{\varepsilon}{\varepsilon^{\mathrm{TL}}}(\alpha-\varepsilon) - (\alpha+\varepsilon) \right|, \right. \\
&\qquad\qquad \left. \left| (\alpha+\varepsilon) + \frac{|I_{i+1}|-1}{|I_{i+1}|}\frac{\varepsilon}{\varepsilon^{\mathrm{TL}}}\left(\frac{3}{2}-\alpha-\varepsilon\right) - (\alpha-\varepsilon) \right| \right\} \\
&\qquad \leq \left\{ 2\varepsilon + \frac{\varepsilon}{\varepsilon^{\mathrm{TL}}}, \quad 2\varepsilon + \frac{3\varepsilon}{2\varepsilon^{\mathrm{TL}}} \right\} \\
&\qquad \leq \frac{7}{2}\left(\varepsilon^{\mathrm{TL}}\right)^2 \\
&\qquad \leq \varepsilon^{\mathrm{TL}}.
\end{aligned}$$

□

Since H is d-degenerate

$$|I_{i+1}| \leq d$$

and, thus assertion (II) for q and Claim 27 (i) and (ii) show that all assumptions of the k-tuple lemma are satisfied for $J[U,W]$.

Therefore, the k-tuple lemma implies that, with probability $1-o(1)$, we have

$$\left| \left\{ b \in \mathcal{W}_i \colon \; \left| \widetilde{\Gamma}_{i+1}(b) \right| \leq (1-\gamma)p^{|I_{i+1}|}(1-\mu)m \right\} \right| \leq \eta^{\mathrm{TL}} \binom{im}{i}.$$

The choice of η^{TL} in (4.44) gives

$$\left| \left\{ b \in \mathcal{W}_i \colon \; \left| \widetilde{\Gamma}_{i+1}(b) \right| \leq (1-\gamma-\mu+\gamma\mu)p^{|I_{i+1}|}m \right\} \right| \leq \frac{\eta_i}{2}m^i,$$

and hence (4.43) holds with probability $1-o(1)$, by the simple observation that

$$\left| \widetilde{\Gamma}_{i+1}(b) \right| \leq (1-\gamma-\mu)p^{|I_{i+1}|}m \quad \text{implies} \quad \left| \widetilde{\Gamma}_{i+1}(b) \right| \leq (1-\gamma-\mu+\gamma\mu)p^{|I_{i+1}|}m.$$

□

4.4.3 Main proof

Our proof of the Counting Lemma, Lemma 10, follows immediately from Lemmas 28 and 29 below. Lemma 28 is a probabilistic statement and asserts that the probability of the occurrence of the event $X_{\text{bad}} = X_1 \subseteq \mathcal{G}(n,q)$ is $o(1)$. On the other hand, Lemma 29 is deterministic and claims that if a graph G is not in X_{bad} and J is a not necessarily induced subgraph of G satisfying (I)–(IV), then J contains the "right" number of copies of H. We apply the technical propositions from the last section in the proof of the probabilistic Lemma 28 below.

Lemma 28. *For arbitrary α and $\sigma > 0$, let γ, μ, ν be given by (4.36), and let ε and η_i $(i = 2, \ldots, h-1)$ be defined as stated in Section 4.4.1. Let G be a random graph in $\mathcal{G}(n,q)$. Then*

$$\mathbb{P}(G \in X_{\text{bad}}(\gamma, \mu, \nu, \eta_1, \eta_2)) = o(1).$$

Proof. Formal logic implies

$$\begin{array}{cccccccc}
X_{\text{bad}} & \subseteq & X_1^{(a)} & \vee & (X_1^{(b)} \wedge \neg X_2) & \vee & \\
 & \vee & X_2^{(a)} & \vee & (X_2^{(b)} \wedge \neg X_3) & \vee & \\
 & \vee & \vdots & \vee & \vdots & \vee & \\
 & \vee & X_{h-2}^{(a)} & \vee & (X_{h-2}^{(b)} \wedge \neg X_{h-1}) & \vee & X_{h-1},
\end{array}$$

and thus, by Propositions 25 and 26 (notice $X_{h-1} = X_{h-1}^{(a)}$ by Definition 24), we have

$$\mathbb{P}\left(X_{\text{bad}}\right) \leq \sum_{i=1}^{h-2} \left(\mathbb{P}(X_i^{(a)}) + \mathbb{P}(X_i^{(b)})\right) + \mathbb{P}(X_{h-1}) = o(1).$$

$\square$

Lemma 29. *For arbitrary α and $\sigma > 0$, let γ, μ, ν be given by (4.36), and let ε and η_i for $(i = 2, \ldots, h-1)$ be defined as stated in Section 4.4.1. Then* *every subgraph J of a graph $G \notin X_{\mathrm{bad}}(\gamma, \mu, \nu)$ satisfying conditions (I)–(IV) contains at least*

$$(1-\sigma)p^{e(H)}m^h$$

copies of H.

Proof. We shall prove by induction on i that the following statement holds for all $1 \leq i \leq h$:

($\mathcal{S}_i$) Let J be a subgraph of $G \notin X_{\mathrm{bad}}$ such that (I)–(IV) apply. Then there are at least $(1-\gamma-\mu-\nu)^i p^{\sum_{j=1}^{i}|I_j|}m^i$ different i-tuples in $\mathcal{W}_i \setminus \mathcal{B}_i$ that induce $H_i = H[\{w_1, \ldots, w_i\}]$ in $J[V_1, \ldots, V_i]$.

Suppose $i = 1$. Note that $\neg X_{\mathrm{bad}}$ implies that $|V_1 \cap \mathcal{B}_1| \leq \eta_1 m = \nu m$. Therefore $V_1 \setminus \mathcal{B}_1$ contains at least $(1-\nu)m \geq (1-\gamma-\mu-\nu)p^0 m^1$ copies of H_1.

We now proceed to the induction step. Assume $i \geq 2$ and ($\mathcal{S}_{i-1}$) holds. Therefore, $\mathcal{W}_{i-1} \setminus \mathcal{B}_{i-1}$ contains at least $(1-\gamma-\mu-\nu)^{i-1}p^{\sum_{j=1}^{i-1}|I_j|}m^{i-1}$ different $(i-1)$-tuples $b = (v_1, \ldots, v_{i-1})$, each constituting the vertex set of a H_{i-1} in $J[V_1, \ldots, V_{i-1}]$. For every $b \in \mathcal{W}_{i-1} \setminus \mathcal{B}_{i-1}$, we have

(i) $|\widetilde{\Gamma}_i(b)| \geq (1-\gamma-\mu)p^{|I_i|}m$, and

(ii) $d_{\mathcal{B}_i}(b) < \nu p^{|I_i|}m$.

Therefore, every such b extends to at least $(1-\gamma-\mu-\nu)p^{|I_i|}m$ different $b' \in \mathcal{W}_i \setminus \mathcal{B}_i$ that correspond to a $H_i \subseteq J[V_1, \ldots, V_i]$. This implies ($\mathcal{S}_i$), and hence our induction is complete.

Assertion ($\mathcal{S}_h$) and the choice of γ, μ, and ν in (4.36) give at least

$$(1-\gamma-\mu-\nu)^h p^{\sum_{j=1}^{h}|I_j|}m^h = (1-\sigma)p^{e(H)}m^h$$

copies of $H_h = H$ in J. □

Clearly, Lemmas 28 and Lemma 29 together imply the Counting Lemma, Lemma 10.

Bibliography

[1] B. Bollobás, *Extremal graph theory*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1978. 1

[2] ———, *Random graphs*, second ed., Cambridge University Press, Cambridge, 2001. 1

[3] P. Erdős, *Some recent results on extremal problems in graph theory. Results*, Theory of Graphs (Internat. Sympos., Rome, 1966), Gordon and Breach, New York, 1967, pp. 117–123 (English); pp. 124–130 (French). 1

[4] ———, *On some new inequalities concerning extremal properties of graphs*, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, pp. 77–81. 1

[5] P. Erdős and M. Simonovits, *A limit theorem in graph theory*, Studia Sci. Math. Hungar **1** (1966), 51–57. 2

[6] P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091. 2

[7] P. Frankl and V. Rödl, *Large triangle-free subgraphs in graphs without K_4*, Graphs Combin. **2** (1986), no. 2, 135–144. 1

[8] Z. Füredi, *Random Ramsey graphs for the four-cycle*, Discrete Math. **126** (1994), no. 1-3, 407–410. 1

[9] S. Gerke, H.-J. Prömel, T. Schickinger, A. Steger, and A. Taraz, *K_4-free subgraphs of random graphs revisited*, submitted for publication. 1

[10] P. E. Haxell, Y. Kohayakawa, and T. Łuczak, *Turán's extremal problem in random graphs: forbidding even cycles*, J. Combin. Theory Ser. B **64** (1995), no. 2, 273–287. 1

[11] ______, *Turán's extremal problem in random graphs: forbidding odd cycles*, Combinatorica **16** (1996), no. 1, 107–122. 1

[12] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience, New York, 2000. 1, 1, 3.1, 3.2

[13] Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997) (F. Cucker and M. Shub, eds.), Springer, Berlin, January 1997, pp. 216–230. 2.2

[14] Y. Kohayakawa, B. Kreuter, and A. Steger, *An extremal problem for random graphs and the number of graphs with large even-girth*, Combinatorica **18** (1998), no. 1, 101–120. 1

[15] Y. Kohayakawa, T. Łuczak, and V. Rödl, *On K^4-free subgraphs of random graphs*, Combinatorica **17** (1997), no. 2, 173–213. 1, 1

[16] Y. Kohayakawa and V. Rödl, *Regular pairs in sparse random graphs I*, submitted for publication, 2001. 4.1, 4.2

[17] ______, *Szemerédi's regularity lemma and quasi-randomness*, Summer School on Algorithms and Combinatorics (tentative title), Springer, Berlin, 2001, to appear. 2.2

[18] Y. Kohayakawa, V. Rödl, and M. Schacht, *The Turán theorem for random graphs*, accepted for publication in Combin. Probab. Comput. 1

[19] D. R. Lick and A. T. White, *k-degenerate graphs*, Canad. J. Math. **22** (1970), 1082–1096. 1

[20] T. Schickinger, *Complete subgraphs of random graphs*, Ph.D. thesis, Department of Computer Science, Techical University Munich, Germany, 2002. 1

[21] J. M. S. Simões-Pereira, *A survey of k-degenerate graphs*, Graph Theory Newslett. **5** (1975/76), no. 6, 1–7. 1

[22] M. Simonovits, *A method for solving extremal problems in graph theory, stability problems*, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, pp. 279–319. 1

[23] T. Szabó and V. H. Vu, *Turán's theorem in sparse random graphs*, submitted for publication, 2002. 1

[24] E. Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), CNRS, Paris, 1978, pp. 399–401. 2.2

[25] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452. 1