Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Journal of Combinatorial Theory, Series B 91 (2004) 1–24

Journal of Combinatorial Theory
Series B

http://www.elsevier.com/locate/jctb

The number of graphs without forbidden subgraphs

József Balogh, Béla Bollobás,[1] and Miklós Simonovits[2]

Mathematical Institute, Hungarian Academy of Sciences, H-1364 Budapest, Pf. 127, Hungary

Received 7 November 2001

Abstract

Given a family $\mathscr{L}$ of graphs, set $p = p(\mathscr{L}) = \min_{L \in \mathscr{L}} \chi(L) - 1$ and, for $n \geqslant 1$, denote by $\mathscr{P}(n, \mathscr{L})$ the set of graphs with vertex set $[n]$ containing no member of $\mathscr{L}$ as a subgraph, and write $\mathbf{ex}(n, \mathscr{L})$ for the maximal size of a member of $\mathscr{P}(n, \mathscr{L})$. Extending a result of Erdős, Frankl and Rödl (Graphs Combin. 2 (1986) 113), we prove that

$$|\mathscr{P}(n, \mathscr{L})| \leqslant 2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2 + O(n^{2-\gamma})}$$

for some constant $\gamma = \gamma(\mathscr{L}) > 0$, and characterize γ in terms of some related extremal graph problems. In fact, if $\mathbf{ex}(n, \mathscr{L}) = O(n^{2-\delta})$, then any $\gamma < \delta$ will do. Our proof is based on Szemerédi's Regularity Lemma and the stability theorem of Erdős and Simonovits. The bound above is essentially best possible.

Keywords: Extremal graphs; Speed function; Erdős–Kleitman–Rothschild theory

1. Introduction

1.1. Notation

Our notation in this paper is standard but for the sake of completeness, we review it briefly. Readers familiar with extremal graph theory may skip this section.

E-mail addresses: jobal@sol.cc.u-szeged.hu (J. Balogh), bollobas@msci.memphis.edu (B. Bollobás), miki@renyi.hu (M. Simonovits).

[1] Partially supported by NSF Grant DMS-9971788, ITR Grant 0225610, and DARPA Grant F33615-01-C-1900.

[2] Partially supported by OTKA Grants T026069, T038210.

0095-8956/$ - see front matter
doi:10.1016/j.jctb.2003.08.001

In this paper we restrict our attention to undirected graphs without loops and multiple edges. Given a graph G and a subset $X \subseteq V(G)$, we write $G[X]$ for the subgraph of G induced by X. For $X \subseteq V(G)$, we mostly shorten $e(G[X])$ to $e(X)$. We write G_n for a graph of order n; in fact much of the time, the first suffix in our notation is the order of the graph, like in K_p, $T_{n,p}$ and H_k. The chromatic number of a graph L will be denoted by $\chi(L)$. We write $\Gamma(x)$ for the set of neighbors of a vertex x, $d(x) = |\Gamma(x)|$ is the degree of x and $d(x, A) = |\Gamma(x) \cap A|$ is the degree of x into a set $A \subseteq V(G)$.

As usual, we write K_p for the complete graph on p vertices, and $T_{n,p}$ for the *p-class Turán graph*. Thus in $T_{n,p}$ the n vertices are partitioned into p classes so that their sizes are as equal as possible, and two vertices in the graph are joined iff they belong to different classes. It is easy to see that if $n \equiv r \pmod p$, $0 \leqslant r < p$, then

$$e(T_{n,p}) = \frac{1}{2}\left(1 - \frac{1}{p}\right)(n^2 - r^2) + \binom{r}{2}.$$

We shall often make use of the facts that

$$e(T_{n,p}) \approx \left(1 - \frac{1}{p}\right)\frac{n^2}{2}$$

and[3]

$$\left(1 - \frac{1}{p}\right)\binom{n}{2} \leqslant e(T_{n,p}) \leqslant \left(1 - \frac{1}{p}\right)\frac{n^2}{2}.$$

Furthermore, we shall use the abbreviation

$$\mathbb{A}(n) := 2^{\frac{1}{2}\left(1 - \frac{1}{p}\right)n^2}.$$

Many of our inequalities hold only for $n > n_0$ and occasionally we shall remind the reader of this. (Further, the value of n_0 will vary from place to place.)

We say that a pair of vertex sets (A, B) is *completely joined* in a graph G_n if $A, B \subset V(G_n)$, $A \cap B = \emptyset$, and for all $x \in A, y \in B$ we have $xy \in E(G_n)$. If we have two vertex-disjoint graphs M and Q, we denote by $M \otimes Q$ the graph obtained by joining each vertex of M to each vertex of Q.

1.2. Turán-type extremal problems

Given a family $\mathscr{L}$ of graphs, we say that G is *$\mathscr{L}$-free* if $L \nsubseteq G$ for every $L \in \mathscr{L}$, where $L \subseteq G$ denotes the *not necessarily induced* containment. We call $\mathscr{L}$ the family of *forbidden graphs*; to avoid trivialities, we shall always assume that $\mathscr{L}$ is non-trivial, i.e., $e(L) > 0$ for $L \in \mathscr{L}$. We write $\mathscr{P}(n, \mathscr{L})$ for the class of $\mathscr{L}$-free graphs with vertex

[3] Clearly, $e(T_{n,2}) = \lfloor \frac{n^2}{4} \rfloor$. As Füredi observed, this extends to $p \leqslant 7$:

$$e(T_{n,p}) = \left\lfloor \left(1 - \frac{1}{p}\right)\frac{n^2}{2} \right\rfloor.$$

set $[n] \coloneqq \{1, \dots, n\}$.[4] We shall use the customary notation

$$\mathbf{ex}(n, \mathscr{L}) = \max\{e(G) : G \in \mathscr{P}(n, \mathscr{L})\}.$$

When $\mathscr{L}$ consists of a single graph L, we use the shorthand $\mathbf{ex}(n, L)$ instead of the pedantic notation $\mathbf{ex}(n, \{L\})$. The basic Turán-type extremal problem is as follows:

> For a given family $\mathscr{L}$, determine or estimate $\mathbf{ex}(n, \mathscr{L})$, and describe the (asymptotic) structure of extremal graphs, as $n \to \infty$.

The theory started with Turán's classical theorem [30], see also [31,32]. For a more detailed description of this field, see the book of Bollobás [5] or the surveys of Simonovits [26,27], or Füredi [14].

1.3. Erdős–Kleitman–Rothschild-type results

Since all subgraphs of an $\mathscr{L}$-free graph are $\mathscr{L}$-free, we have

$$|\mathscr{P}(n, \mathscr{L})| \geqslant 2^{\mathbf{ex}(n, \mathscr{L})}. \tag{1}$$

Erdős [9] conjectured that (1) is essentially best possible, namely for 'most' graphs L we have

$$|\mathscr{P}(n, L)| = 2^{(1+o(1))\mathbf{ex}(n,L)}. \tag{2}$$

If L is a tree then (2) fails, and if L is bipartite containing a cycle, then proving (2) seems to be difficult, even for $L = C_4$ (see [16]). Erdős, Kleitman and Rothschild [13] were the first to study the function $|\mathscr{P}(n, L)|$ in detail, by proving Erdős' conjecture for $L = K_{p+1}$.

Theorem 1.1.

$$|\mathscr{P}(n, K_{p+1})| \leqslant 2^{\left(1-\frac{1}{p}\right)\binom{n}{2}+o(n^2)}.$$

In the case $\chi(L) \geqslant 3$, the conjecture was proved by Erdős, Frankl and Rödl [12].

Theorem 1.2. *Let L be a graph with $\chi(L) \geqslant 3$. Then*[5]

$$|\mathscr{P}(n, L)| = 2^{(1+o(1))\mathbf{ex}(n,L)} = 2^{\left(1-\frac{1}{\chi(L)-1}\right)\binom{n}{2}+o(n^2)}.$$

Kolaitis, Prömel and Rothschild [17] sharpened Theorem 1.1: they proved that, in fact, almost every K_{p+1}-free graph is p-colorable.

[4] The vertices of our graphs are fixed, labeled and, for the sake of simplicity, we shall assume that $V(G_n) = \{1, \dots, n\}$.

[5] Mostly it is irrelevant if we use $2^{\left(1-\frac{1}{p}\right)\binom{n}{2}}$ or $2^{\left(1-\frac{1}{p}\right)\frac{n^2}{2}}$, because of the additional error terms in the exponent.

Theorem 1.3. *Let* $\mathscr{C}_n(p)$ *be the set of labeled* p*-colorable graphs on* $[n]$. *Then*

$$\frac{|\mathscr{P}(n, K_{p+1})|}{|\mathscr{C}_n(p)|} \to 1 \quad \text{as } n \to \infty.$$

Later, Prömel and Steger [21] extended Theorem 1.3 to graphs with critical edges, where an edge e of L *color-critical* if $\chi(L - e) < \chi(L)$. Results in a similar vein have been proved by Hundack, Prömel and Steger [15] for a larger family of graphs.

2. New results

There are many beautiful theorems generalizing the results mentioned above. Unless $\mathscr{L} = \{K_{p+1}\}$, we have two distinct problems: estimating the number of n-vertex graphs not containing

- *induced* subgraphs isomorphic to any $L \in \mathscr{L}$;
- or *not necessarily induced* subgraphs.

Here we refer the reader to the papers of Alekseev [1], Prömel and Steger [20,22], Bollobás and Thomason [7,8], Scheinerman and Zito [24] and Balogh, Bollobás and Weinreich [2–4], and restrict ourselves to the not necessarily induced case. Our starting point is Theorem 1.2, due to Erdős, Frankl and Rödl [12]. It is trivial to rephrase Theorem 1.2 for a family $\mathscr{L}$ of forbidden graphs:

$$|\mathscr{P}(n, \mathscr{L})| = 2^{(1+o(1))\mathbf{ex}(n,\mathscr{L})} = 2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2+o(n^2)}, \tag{3}$$

where

$$p := \min_{L \in \mathscr{L}} \chi(L) - 1. \tag{4}$$

The problem we study in this paper is how much the 'error term' $o(n^2)$ in the exponent in (3) can be improved. Our main result is that $o(n^2)$ can be replaced by $O(n^{2-\gamma})$ for some $\gamma = \gamma(\mathscr{L}) > 0$.

Theorem 2.1. *For every non-trivial family* $\mathscr{L}$ *of graphs there exists a constant* $\gamma = \gamma_{\mathscr{L}} > 0$ *such that*

$$|\mathscr{P}(n, \mathscr{L})| \leqslant 2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2+O(n^{2-\gamma}\log n)}, \tag{5}$$

for $p = \min_{L \in \mathscr{L}} \chi(L) - 1$.[6]

In fact, we shall prove a considerably sharper result: we shall determine the exact order of the error term in the exponent. To this end, we define a new family $\mathscr{M} = \mathscr{M}(\mathscr{L})$ of graphs.

[6] Of course, $\log n$ can be deleted by slightly decreasing γ, but we shall see that (5) is a better form.

Definition 2.2 (Decomposition). Given a family $\mathscr{L}$, let $\mathscr{M} := \mathscr{M}(\mathscr{L})$ be the family of graphs M for which there exist an $L \in \mathscr{L}$ and a $t = t(L)$ such that $L \subseteq M \otimes K_{p-1}(t, \dots, t)$. We call $\mathscr{M}$ the *decomposition family* of $\mathscr{L}$.

Thus, a graph M belongs to $\mathscr{M}$ if whenever M is placed into a class of $T_{n,p}$ for n sufficiently large, then the new graph contains a forbidden $L \in \mathscr{L}$.

If $\mathscr{L}$ is finite, then $\mathscr{M}$ is also finite but the converse is not necessarily true. For example, if $\mathscr{L}$ is the family of all odd cycles, then $\mathscr{M} = \{K_2\}$.

We would be in a strong position to give a precise estimate for $|\mathscr{P}(n, \mathscr{L})|$ if we could prove the following conjecture about the structure of most $\mathscr{L}$-free graphs. In fact, a good description of a typical $\mathscr{L}$-free graph should be even more interesting than a good estimate of the function $|\mathscr{P}(n, \mathscr{L})|$.

Conjecture 2.3 (Sharp form). *Assume that $\mathscr{L}$ is finite. Then for almost all $\mathscr{L}$-free graphs G_n we can delete $h = O_{\mathscr{L}}(1)$ vertices of G_n and partition the remaining vertices into p classes $U_1, \dots, U_p$ such that each $G[U_i]$ is $\mathscr{M}$-free, $i = 1, \dots, p$.*

Remark 2.4. We shall see that if we take an *optimal* p-partition $(U_1, \dots, U_p)$ of a typical G_n, where "optimal" means that $\sum e(U_i)$ is as small as possible, then the number of vertices which are joined to each U_i by more than εn edges is bounded; we believe that these are the vertices that should be deleted to make the remaining parts of $G[U_i]$ to be $\mathscr{M}$-free. This would imply

$$2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2+\mathbf{ex}(\lfloor\frac{n}{p}\rfloor,\mathscr{M})} \leqslant |\mathscr{P}(n, \mathscr{L})| \leqslant 2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2+\mathbf{ex}(n,\mathscr{M})}.$$

The lower bound is trivial.

The problem in proving Conjecture 2.3 is that we do not know in general whether $\mathscr{P}(n, \mathscr{M})$ is a "smooth" function or it oscillates wildly. Although we cannot prove that $\mathscr{P}(n, \mathscr{M})$ is smooth, we have the following result.

Theorem 2.5. *For every $\mathscr{L}$, if $\mathscr{M}$ is the decomposition family, $\mathscr{M}$ is finite, then*

$$|\mathscr{P}(n, \mathscr{L})| \leqslant n^{\mathbf{ex}(n,\mathscr{M})+c\cdot n} 2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2}, \tag{6}$$

for some appropriate constant $c > 0$.

Applying the theorem we obtain that the γ of Theorem 2.1 is essentially the lim sup of those α for which $\mathbf{ex}(n, \mathscr{M}) = O(n^{2-\alpha})$.

2.1. The case $p = 1$

The crucial step in our proofs of Theorems 2.1 and 2.5 is the reduction of the general case to the case $p = 1$. For this we shall need a lemma asserting that the number of graphs from which we can delete $\Psi = o(n^2)$ edges to get p-chromatic graphs is not much larger than $2^{\mathbf{ex}(n,\mathscr{L})}$.

Lemma 2.6. *Let $\mathscr{C}_\Psi(n)$ denote the class of graphs for which there is a partition $(U_1, \dots, U_p)$ with $\sum_i e(U_i) \leqslant \Psi$, where $\Psi > e \cdot pn$. Then*

$$|\mathscr{C}_\Psi(n)| \leqslant n^\Psi \cdot 2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2} = n^\Psi \mathbb{A}(n). \tag{7}$$

Proof. We shall use the simple inequality $\binom{a}{b} \leqslant \left(\frac{e \cdot a}{b}\right)^b$. There are at most p^n partitions of $V(G_n)$ into p distinguishable sets $(U_1, \dots, U_p)$. Given a partition, there are at most $\mathbb{A}(n)$ choices for the cross-edges between the classes and at most

$$\left(\begin{pmatrix} n \\ 2 \end{pmatrix} \atop \Psi\right) \leqslant \left(\frac{en^2}{2\Psi}\right)^\Psi \leqslant \frac{n^\Psi}{(2p)^n} \tag{8}$$

choices for the edges within the sets U_i. These prove the lemma. □

In proving Theorem 2.5, we may and shall assume that $p \geqslant 2$. Indeed, for $p = 1$ we have $\mathscr{M} = \mathscr{L}$ and Theorem 2.5 becomes trivial. To get Theorem 2.1 for $p = 1$, we use that, by the Kővári–T. Sós–Turán theorem [18], for the complete bipartite graph $K(p, q)$ we have

$$\mathbf{ex}(n, K(p, q)) \leqslant \tfrac{1}{2} \sqrt[p]{q-1}\, n^{2-1/p} + \tfrac{1}{2} pn. \tag{9}$$

Hence $\mathbf{ex}(n, K(t, t)) \leqslant n^{2-1/t}$, for $n > n_0(t)$. Since $\mathscr{L}$ contains some bipartite L_0,

$$\mathbf{ex}(n, \mathscr{L}) \leqslant \mathbf{ex}(n, K(t, t)) \leqslant n^{2-(1/t)}$$

for some t. So Theorems 2.5 and 2.1 are trivial for $p = 1$, by the above lemma and (9).

Theorem 2.5 easily implies Theorem 2.1 since for every $\mathscr{L}$ we can find a $(p+1)$-chromatic $L_0 \in \mathscr{L}$ and a $K_{p+1}(t, \dots, t) \supseteq L_0$: with this t we have $K(t, t) \in \mathscr{M}$.

Remark 2.7. In our proofs, in most cases we do not have to consider all of $\mathscr{L}$ but one $L_0 \in \mathscr{L}$, as above. We fix now such an L_0 and a $K_{p+1}(t, \dots, t) \supseteq L_0$ (and refer to this t) we shall need all of $\mathscr{L}$ and $\mathscr{M}$ only in the last step of our proof.

3. Almost-Turán graphs

Let $\mathscr{L}$ and p be fixed. From now on, we shall frequently suppress the dependence of various functions on p and $\mathscr{L}$.

We plan to prove Theorem 2.5 in the following way. We shall try to prove in various ways that a typical $\mathscr{L}$-free G_n looks like a random subgraph of an $\mathscr{L}$-extremal graph. Furthermore, we think of an $\mathscr{L}$-extremal graph as one obtained from a $T_{n,p}$ by putting $\mathscr{M}$-extremal graphs into some of its classes. Let us describe this plan in more details.

1. $(U_1, \dots, U_p)$ is an *optimal partition* if $\sum e(U_i)$ is as small as possible. We assign to each G_n an optimal partition, denoted by $\Pi(G_n)$.[7]
2. First, with the aid of Szemerédi's Regularity Lemma (see Section 4.1), using some variants of previous techniques (primarily of Erdős, Frankl and Rödl [12] and of Bollobás and Thomason [7]) we show that an optimal partition $(U_1, \dots, U_p)$ of $V(G_n)$ satisfies
$$\sum_i e(U_i) \leqslant \vartheta n^2, \tag{10}$$
for almost all graphs in $\mathscr{P}(n, \mathscr{L})$, for any fixed $\vartheta > 0$ and $n > n_0(\vartheta)$. The graphs having such partitions will be called ϑ-*Turán graphs* and the class of $\mathscr{L}$-free ϑ-Turán graphs will be denoted by $\mathscr{P}_\vartheta(n, \mathscr{L})$.
3. We define $\mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L}) \subseteq \mathscr{P}_\vartheta(n, \mathscr{L})$, as follows.
   Let $h(x)$ denote the so called entropy function:
$$h(x) \coloneqq x \log_2 \frac{1}{x} + (1-x)\log_2 \frac{1}{(1-x)}.$$
   Given a graph G_n, call a pair (A, B) of disjoint sets of vertices with $|A| = |B|$ *sparse* if $e(A, B) < \frac{1}{4}|A| \cdot |B|$. Let $\mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L})$ denote the set of graphs in $\mathscr{P}_\vartheta(n, \mathscr{L})$ containing no sparse pairs (A, B) with $A \subset U_i, B \subset U_j, |A| = |B| = \lfloor \delta n \rfloor$ for $i \neq j$, where
$$\delta = 10\sqrt{h(\vartheta)} \tag{11}$$
   and $(U_1, \dots, U_p)$ is an optimal partition of G_n.
   We shall show that almost all graphs of $\mathscr{P}_\vartheta(n, \mathscr{L})$ belong to $\mathscr{P}^\delta_{\mathbf{UNIF}}$.
4. Next we fix two constants, ℓ and μ, and define $\mathscr{P}^\ell_{\mathbf{GOOD}}(n, \mathscr{L})$ as the family of those $G_n \in \mathscr{P}_\vartheta(n, \mathscr{L})$ for which we can delete μ vertices such that in the remaining graph $G_{n-\mu}$, if we choose ℓ vertices in one class, then each other class contains $\lceil n/(p2^{\ell+2}) \rceil$ vertices completely joined to these ℓ vertices. Then we show that, for a properly chosen $\vartheta > 0$, almost all graphs in $\mathscr{P}^\ell_{\mathbf{GOOD}}(n, \mathscr{L})$ satisfy
$$e(U_i) \leqslant \mathbf{ex}(|U_i|, \mathscr{M}) + O(n). \tag{12}$$
   More precisely, for each $i = 1, \dots, p$, there are only $O_\varepsilon(1)$ vertices in $G[U_i]$ of degree $\geqslant \varepsilon n$ (i.e., joined to their own class U_i by more than εn edges) and deleting them from $G[U_i]$ we get an $\mathscr{M}$-free graph.

[7] We shall enumerate the graphs according to their optimal partitions. If we took an arbitrary optimal partition instead of a fixed one, then we would count the same graph several times. An upper bound obtained that way would be equally good: so this assignment is not too important but may make the proof more transparent. We can take, e.g., the lexicographically first partition.

Remark 3.1. The last class, $\mathscr{P}^{\ell}_{\mathbf{GOOD}}$, differs from the previous ones in that it is given by a local definition: the previous classes are not really influenced by changing $o(n^2)$ edges in G_n, but this is.

4. Tools

In the proof of Theorem 2.5 we shall have a complicated system of constants and several families of graphs on $[n]$. We assume only that $0<\vartheta<(ep)^{-12}$ is an arbitrary small, fixed constant. Later we shall take $\vartheta\to 0$.

Our "Main Lemma" below asserts that almost all graphs to be considered are ϑ-Turán graphs.

Main Lemma. *Given $\vartheta>0$, there is an integer $n_0(\vartheta)$ such that if $n>n_0(\vartheta)$ then*

$$|\mathscr{P}(n,\mathscr{L})-\mathscr{P}_\vartheta(n,\mathscr{L})|\leqslant 2^{\left(1-\frac{1}{p}\right)\frac{n^2}{2}-n}=\mathbb{A}(n)/2^n. \tag{13}$$

In fact, this lemma claims more than we need in its applications: it would suffice to have $o(2^{\left(1-\frac{1}{p}\right)\frac{n^2}{2}})$ on the right hand side. Also, as the lemma holds for all $\vartheta>0$, it can be reformulated as follows.

Main Lemma$'$. *One can delete from each $\mathscr{P}(n,\mathscr{L})$ at most $2^{\left(1-\frac{1}{p}\right)\frac{n^2}{2}-n}$ graphs so that for each remaining graph G_n, for its optimal partition $(U_1,\dots,U_p)$, $\sum e(U_i)=o(n^2)$ as $n\to\infty$.*

4.1. Regularity lemma

Given a graph G and two disjoint vertex sets $X,Y\subset V(G)$, the *edge-density* between X and Y is defined as

$$d(X,Y)=\frac{e(X,Y)}{|X||Y|},$$

where $e(X,Y)=e_G(X,Y)$ is the number of edges of G between X and Y. We call the pair (X,Y) ε-*regular* if, for all $X^*\subset X$ and $Y^*\subset Y$ with $|X^*|>\varepsilon|X|$ and $|Y^*|>\varepsilon|Y|$, we have

$$|d(X^*,Y^*)-d(X,Y)|<\varepsilon.$$

Furthermore, we say that a partition $V(G)=V_1\cup\cdots\cup V_k$ is ε-*regular* if $||V_i|-n/k|<1$, for every i, and all but at most εk^2 pairs (V_i,V_j) are ε-regular. The sets V_i are the *clusters* of this partition. In this terminology, Szemerédi's Lemma [28] can be stated as follows.

Regularity Lemma. *For every $\varepsilon > 0$ and integer κ there exist integers $n_0 = n_0(\varepsilon, \kappa)$ and $k_0 = k_0(\varepsilon, \kappa)$ such that every graph of order at least n_0 has an ε-regular partition with more than κ and fewer than k_0 clusters.*

4.2. The cluster graph, the estimate

In order to capture the global structure of a graph, we introduce the notion of cluster graphs.

Definition 4.1 (Cluster graph). Given $\eta > 0$ and an ε-regular partition of a graph G_n with k clusters $V_1, \dots, V_k$, let H_k be the graph with vertex set $V_1, \dots, V_k$ in which V_iV_j is an edge iff (V_i, V_j) in G_n is an ε-regular pair of density at least η. We call H_k an (ε, η)-*cluster graph* of G_n.

With a slight abuse of notation, we write $H_k = H_k(G_n)$ for an (ε, η)-cluster graph of G_n with k clusters, where $\kappa = 1/\varepsilon < k < k_0(\varepsilon, \kappa)$ and $n > n_0(\varepsilon, \kappa)$. (In principle there are several appropriate partitions and we should also indicate the partition.) In addition to Szemerédi's Regularity Lemma, we shall need the following Stability Theorem of Erdős and Simonovits [10,11,37].

Stability Theorem. *For any given $\lambda > 0$, there is an $\omega = \omega(p, \lambda) > 0$ such that if G_k does not contain K_{p+1} and*

$$e(G_k) > e(T_{k,p}) - \omega k^2,$$

then we can change G_k into $T_{k,p}$ by changing at most λk^2 edges.

Here "changing" means deleting or adding. Clearly, $\omega \leqslant \lambda$.[8]

Szemerédi's Regularity Lemma is frequently used to guarantee the existence of small subgraphs, see, e.g., [6,8,29] for applications nearest to ours, or the survey [19].

Lemma 4.2. *Let L be a fixed graph with $\chi(L) = p + 1$, and let $\eta > 0$. If $\varepsilon > 0$ is sufficiently small, n is sufficiently large and an (ε, η)-cluster graph H_k of G_n contains K_{p+1}, then $L \subset G_n$.*

In [8] it was shown that in Lemma 4.2 it suffices to take $\varepsilon < \eta/2^{|L|}$.

Lemma 4.3. *Using the notation of the Regularity lemma, let $\mathscr{C}_{\eta,\varepsilon,\omega}(n)$ be the class of graphs on $[n]$ such that, for some (η, ε)-cluster-graph H_k, we have*

$$e(H_k) \leqslant e(T_{k,p}) - \omega k^2.$$

If

$$h(\eta) < \frac{\omega}{4}, \quad \varepsilon < \frac{\omega}{8} \quad \text{and} \quad \kappa = \frac{1}{\varepsilon}, \tag{14}$$

[8] Deleting ωk^2 edges from $T_{k,p}$ we get a graph to which we have to add at least ωk^2 edges to get $T_{k,p}$.

where $h(x)$ is the entropy function, then for $n > n_0(\eta, \varepsilon, \omega)$,

$$|\mathscr{C}_{\eta,\varepsilon,\omega}(n)| < 2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2 - \frac{1}{2}\omega n^2} < \mathbb{A}(n)/2^n.$$

Proof. When using the Regularity lemma, we tend to ignore the fact that n is mostly not divisible by k: our estimates are too robust to be influenced by this.

The number of $\leqslant k$-partitions is at most k^n. We fix a partition $(V_1, \dots, V_k)$. Then we select the edges of the cluster graph H_k in at most $2^{\binom{k}{2}}$ ways. If the cluster-size is $m = n/k$ and the edges of H_k are already fixed, then we have at most $2^{e(H_k)m^2} \leqslant 2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2 - \omega n^2}$ choices for the edges along the cluster-graph edges, at most $2^{n^2/(2k)}$ choices within the partition classes V_i and $2^{\varepsilon k^2 m^2} = 2^{\varepsilon n^2}$ options along the (at most εk^2) non-ε-regular connections of clusters. Finally, the most significant "loss" in our estimates comes from the low-density pairs: (V_i, V_j) with $d(V_i, V_j) < \eta$, where we have at most $\binom{m^2}{\eta m^2} \leqslant 2^{h(\eta)m^2}$ choices for the edges, *for each of these pairs*. We shall use that

$$\binom{a}{xa} \leqslant 2^{h(x)a},$$

and therefore

$$\sum_{i \leqslant xa-1} \binom{a}{i} \leqslant \binom{a}{xa} < 2^{h(x)a}.$$

One can see that $x \log \frac{1}{x} \leqslant h(x) \leqslant x \log \frac{1}{x} + \frac{3x}{2}$ if $x \leqslant \frac{1}{2}$. Also, if $m < \frac{a}{4}$ then

$$\sum_{i<m} \binom{a}{i} \leqslant \binom{a}{m}.$$

This is why $\eta \leqslant \frac{1}{4}$ will always be assumed.

Thus, using the above estimates and (14), we have

$$\begin{aligned}|\mathscr{C}_{\eta,\varepsilon,\omega}(n)| &\leqslant \sum_{\kappa<k<k_0(\kappa,\varepsilon)} k^n \cdot 2^{\binom{k}{2}} \cdot 2^{e(H_k)m^2} \cdot 2^{\frac{n^2}{2k}} \cdot 2^{\varepsilon n^2} \cdot 2^{h(\eta)n^2}, \\ &\leqslant 2^{n \log k_0(\kappa,\varepsilon) + O(1) + \frac{1}{2}\left(1-\frac{1}{p}\right)n^2 - \omega n^2 + \frac{n^2}{2\kappa} + \varepsilon n^2 + h(\eta)n^2} \\ &< \mathbb{A}(n) \cdot 2^{-\omega n^2 + \varepsilon n^2 + \varepsilon n^2 + \frac{\omega}{4} n^2} < \mathbb{A}(n)/2^n.\end{aligned}$$

In the formula above, the $O(1)$ term in the exponent depends on $k_0(\eta, \varepsilon)$ and "represents" the number of possible H_k graphs. $\square$

5. Proof of the Main Lemma

Let $\varepsilon > 0$ be as small as required in Lemma 4.2. By Lemma 4.2

$$\text{if } G_n \in \mathscr{P}(n, L) \text{ then } K_{p+1} \nsubseteq H_k. \tag{15}$$

Hence

$$e(H_k) \leqslant \mathbf{ex}(k, K_{p+1}) \leqslant \left(1 - \frac{1}{p}\right)\frac{k^2}{2}. \tag{16}$$

We could use (16) and the above argument to deduce the Erdős–Frankl–Rödl theorem and then improve it to get our results but we rather go directly to the proof.

Let us embark on our estimate of $|\mathscr{P}(n, \mathscr{L})|$.

5.1. Weak partition

The expressions 'weak' and 'strong' partitions will be used only informally, to aid the reader. A partition $(U_1, \dots, U_p)$ is "stronger" if $\sum e(U_i)$ is smaller. We call a partition *weak* if we know only (10) or $\sum e(U_i) = o(n^2)$, and we call it *strong* if $\sum e(U_i) = O(\mathbf{ex}(n, \mathscr{M}))$.

Let

$$\vartheta \in \left(0, \frac{1}{4p}\right) \text{ be fixed, set } \lambda := \frac{\vartheta}{4}, \text{ and let } \omega = \omega(p, \lambda)$$

be the constant whose existence is guaranteed by Stability Theorem.

Then fix η and ε as described in (14) and Lemma 4.2.

Lemma 5.1.

$$\mathscr{P}(n, \mathscr{L}) - \mathscr{C}_{\eta,\varepsilon,\omega}(n) \subseteq \mathscr{P}_\vartheta(n, \mathscr{L}).$$

Proof. Let $G_n \in \mathscr{P}(n, \mathscr{L}) - \mathscr{C}_{\eta,\varepsilon,\omega}(n)$. Let H_k be a cluster graph of G_n, and $m = n/k$. Since $G_n \notin \mathscr{C}_{\eta,\varepsilon,\omega}(n)$, $e(H_k) \geqslant e(T_{k,p}) - \omega k^2$, see Lemma 4.3. By Lemma 4.2, $K_{p+1} \nsubseteq H_k$. So, by Stability Theorem, we can change H_k into $T_{k,p}$ by changing at most λk^2 "cluster-edges".

This yields a p-partition $(U_1, \dots, U_p)$ of $V(G_n)$: write C_ℓ for the ℓth class of this $T_{k,p}$ and put $U_\ell = \bigcup_{V_j \in C_\ell} V_j$. We show that $\sum e(U_i) \leqslant \vartheta n^2$ and the class-sizes are roughly the same:

$$\left| |U_\ell| - \frac{n}{p} \right| \leqslant 2\varepsilon n \quad \text{for } \ell = 1, \dots, p.$$

Indeed, to change G_n into a $T_{n,p}$ we "move" $\leqslant \lambda k^2$ "cluster-edges" of H_k corresponding to $\leqslant \lambda k^2 m^2 = \lambda n^2$ edges in G_n; next we delete the edges corresponding to "low-density" pairs (V_i, V_j) with V_i and V_j in the same partition class C_h of $T_{k,p}$ and $d(V_i, V_j) < \eta$. This yields $\frac{1}{2}\eta n^2$. Our no more than εk^2 "irregular pairs" means εn^2

edges in G_n. Finally, we have to delete the edges within the clusters V_j: this is $\frac{1}{2}km^2 = \frac{1}{2}n^2/k < \frac{1}{2}\varepsilon n^2$. Recalling that $\lambda = \frac{1}{4}\vartheta$ and $\varepsilon < \frac{1}{8}\omega < \frac{1}{8}\lambda$, $\eta < h(\eta) < \frac{1}{4}\omega$, we get

$$\sum e(U_i) \leqslant \lambda n^2 + \frac{1}{2}\eta n^2 + \varepsilon n^2 + \frac{1}{2}km^2 < 4\lambda n^2 = \vartheta n^2. \qquad \square$$

6. Typical optimal partitions

Take an optimal partition of a graph $G_n \in \mathscr{P}_\vartheta(n, \mathscr{L})$ into p classes, $(U_1, \dots, U_p)$. The optimality of the partition implies that

$$d(x, U_i) \leqslant d(x, U_\ell)$$

whenever $x \in U_i$, since otherwise moving x into U_ℓ would decrease $\sum_i e(U_i)$. We shall call the edge (x, y) of G_n *horizontal* if x, y are in the same U_i.[9] The *horizontal degree* of a vertex $x \in U_i$ is $d(x, U_i) = |\Gamma(x) \cap U_i|$. Clearly, these definitions depend on the partition. Mostly we may forget to indicate this dependence, but there will be a point where this dependence will become crucial.

Here we shall prove the following assertions:

- Lemma 6.1: The edges are uniformly distributed between the partition classes.
- Lemma 6.3: All **small**, i.e. of bounded size, p-chromatic subgraphs occur in G_n if the edges in the partition are uniformly distributed.
- Lemma 6.10: Stability of optimal partitions; and its consequence: by Lemma 6.11, there are only few optimal partitions.
- Lemma 6.6: The vertices are uniformly distributed in partition classes.

6.1. Super-regularity

Having a fixed partition, the edges (x, y) joining different classes will be called *vertical*. In a typical graph the vertical edges behave as random edges with probability $\frac{1}{2}$. This motivates the following easy lemma.

Lemma 6.1. *For all but $\mathbb{A}(n)/2^n$ $\mathscr{L}$-free graphs G_n, if $(U_1, \dots, U_p)$ is an optimal partition, $\delta > 0$ and $n > n_0(\mathscr{L}, \delta)$, then if $A \subset U_i, B \subset V(G_n) - U_i$ with $|A| = |B| = \lfloor \delta n \rfloor$, then $e(A, B) > \frac{1}{4}|A| \cdot |B|$.*

Remark 6.2. (a) Here $\geqslant \frac{1}{4}|A||B|$ means that the number of edges is at least half of what it is expected to be.

(b) In our proof it suffices to consider only graphs from $\mathscr{P}_\vartheta(n, \mathscr{L})$, since the two sets differ in no more than $\mathbb{A}(n)/2^n$ graphs and it does not matter if we get twice the error term.

[9] "Horizontal" comes from some figures showing $T_{n,p}$ where these edges are, indeed, "almost horizontal".

(c) To assure more symmetry, we could write that $A \subseteq U_i$, $B \subseteq U_j$ for some $i \neq j$. From our point of view the two forms are equivalent but this form is easier to use.

There is a symmetric, and stronger form of Lemma 6.1: with the conclusion that if $A \subseteq \bigcup_{i \in I} U_i$, $B \subseteq \bigcup_{j \in J} U_j$ for some partition of $[1, \dots, p]$ into I and J, $I \cap J = \emptyset$, then $e(A, B) > \frac{1}{4}|A| \cdot |B|$.

Proof of Lemma 6.1. We estimate the number of "bad graphs": graphs G_n for which there is an optimal partition $(U_1, \dots, U_p)$ and two sets, $A \subseteq U_i$ and $B \subseteq U_j$ with $e(A, B) \leqslant \frac{1}{4}|A||B|$. The number of partitions is at most p^n, the number of possibilities for (A, B) is at most

$$\binom{n}{\delta n}^2 < 2^{2h(\delta)n},$$

the number of choices for the edges between distinct classes is at most

$$\mathbb{A}(n) \cdot 2^{-(1-h(1/4))\delta^2 n^2}$$

(since the number of possible "connections" between (A, B) is only at most $2^{h(1/4)(\delta n)^2}$, instead of $2^{(\delta n)^2}$), and the number of choices for edges within the classes is at most $2^{h(\vartheta)n^2}$. Since, by (11), $h(\vartheta) < \frac{1}{3}\delta^2$ and, further, $n > n_0$, the number of "bad graphs" is at most

$$\mathbb{A}(n) \cdot p^n \cdot 2^{2h(\delta)n} \cdot 2^{h(\vartheta)n^2} \cdot 2^{-(1-h(1/4))\delta^2 n^2} < \mathbb{A}(n)/2^n,$$

proving the lemma. Note that $h(1/4)$ is about 0.3177. $\square$

Lemma 6.1 will be used in combination with Lemma 6.3.

Lemma 6.3 (Weak-regularity). *For $\delta > 0$, $p \geqslant 1$ and $t \geqslant 1$ there is an integer $n_0(\delta, p, t)$ such that the following assertion holds. Suppose that $n > n_0(\delta, p, t)$, $1 \leqslant q \leqslant p$, and $U_1, \dots, U_q$ are q disjoint vertex-sets in $V(G_n)$ with $|U_i| > \frac{n}{2p}$ and*

$$e(A, B) \geqslant \tfrac{1}{4}|A||B| \tag{17}$$

whenever

$$A \subset U_i, \quad B \subset U_j, \quad i \neq j \quad \text{and} \quad |A|, |B| \geqslant \frac{\delta n}{p}. \tag{18}$$

Then $K_q(t, \dots, t) \subseteq G_n$.

The key case above is $q = p$, when $(U_1, \dots, U_p)$ is a slight modification of the optimal partition of G_n.

We could refer to conditions (17) and (18) as the *δ-super-regularity conditions*. Similar conditions occurred in several works of Erdős and T. Sós, Rödl, Komlós, and others.

Remark 6.4. There are several results closely related to this lemma. One of these (which, in some sense, is much stronger) is that of Rödl [23].

Proof of Lemma 6.3. We apply induction on q, although a direct proof could be given as well. We shall make use of Lemma 7.2, whose proof does not depend on this lemma. For $q = 2$ this immediately follows from the Kővári–T. Sós–Turán theorem, or from Lemma 7.2.

Assume that the assertion of the lemma holds for $q - 1$. Applying Lemma 7.2 recursively, $t - 1$ times, we can choose t vertices in U_q completely joined to αn vertices of U_i $(i = 1, \ldots, q - 1)$. Then applying induction to the first $q - 1$ classes with a slightly different $\delta > 0$, the assertion follows. $\square$

6.2. Uniform class-sizes

Definition 6.5 (Well-partitioning). G_n is ϑ-*well partitioned* if all its optimal partitions $(U_1, \ldots, U_p)$ satisfy $\left||U_i| - \frac{n}{p}\right| < \left(\sqrt{\vartheta}\log\frac{1}{\vartheta}\right)n$ for all i. We denote by $\mathscr{P}^{\vartheta}_{\mathbf{WP}}(n, \mathscr{L})$ the family of graphs from $\mathscr{P}_{\vartheta}(n, \mathscr{L})$ which are ϑ-well-partitioned.

Lemma 6.6. *Let* $0 < \vartheta < (ep)^{-12}$. *Then*

$$\left|\mathscr{P}_{\vartheta}(n, \mathscr{L}) - \mathscr{P}^{\vartheta}_{\mathbf{WP}}(n, \mathscr{L})\right| < 2^{\frac{1}{2}\left(1-\frac{1}{p}\right)n^2 - n}.$$

Remark 6.7. We do not really need the fine notion of "uniform partition": to prove our theorems it would be enough to assume, e.g., that $|U_i| > \frac{1}{100p}n$ for all i. From the optimality we need only that an $x \in U_i$ cannot send many edges to its own class and few edges to the others. Yet, it is interesting to have these estimates which give profound information on the structure of typical $\mathscr{L}$-free graphs.

Proof of Lemma 6.6. The idea of the proof is almost the same as that of the Main Lemma. Let $G_n \notin \mathscr{P}^{\vartheta}_{\mathbf{WP}}(n, \mathscr{L})$. Then G_n has an optimal partition $(U_1, \ldots, U_p)$ with $\sum_i e(U_i) < \vartheta n^2$, and there is an index i_0 such that

$$\left||U_{i_0}| - \frac{n}{p}\right| > \left(\sqrt{\vartheta}\log\frac{1}{\vartheta}\right)n.$$

We need the following weak variant of a lemma on uneven vertex-distributions from [25, p. 290]. Let G_n be a p-partite graph with p-partition $(U_1, \ldots, U_p)$. Then

$$e(G_n) \leqslant e(T_{n,p}) - \sum_{i=1}^{p}\binom{s_i}{2},$$

where $s_i = \lfloor |n/p - |U_i|| \rfloor$ for $i = 1, \dots, p$. Using this, with $s_{i_0} = \lfloor (\sqrt{\vartheta} \log \frac{1}{\vartheta}) n \rfloor$ we have

$$\sum_{i<j} e(U_i, U_j) \leqslant e(T_{n,p}) - \binom{s_{i_0}}{2} < e(T_{n,p}) - \frac{\vartheta}{3}\left(\log \frac{1}{\vartheta}\right)^2 n^2.$$

Since $h(x) \approx x \log \frac{1}{x}$, we find that

$$\binom{\binom{n}{2}}{\vartheta n^2} < 2^{2\vartheta\left(\log \frac{1}{\vartheta}\right)n^2},$$

and therefore, if $\log_2 \frac{1}{\vartheta} > 12$,

$$\begin{aligned} |\mathscr{P}_\vartheta(n, \mathscr{L}) - \mathscr{P}^\vartheta_{\mathbf{WP}}(n, \mathscr{L})| &\leqslant p^n \mathbb{A}(n) \cdot 2^{-\frac{\vartheta}{3}\left(\log \frac{1}{\vartheta}\right)^2 n^2} \binom{\binom{n}{2}}{\vartheta n^2} \\ &\leqslant p^n \mathbb{A}(n) \cdot 2^{-\frac{\vartheta}{3}\left(\log \frac{1}{\vartheta}\right)^2 n^2 + 2\vartheta\left(\log \frac{1}{\vartheta}\right)n^2} \\ &\leqslant p^n \mathbb{A}(n) \cdot 2^{-\frac{\vartheta}{6}\left(\log \frac{1}{\vartheta}\right)^2 n^2} < \mathbb{A}(n)/2^{2\vartheta n^2} < \mathbb{A}(n)/2^n. \end{aligned}$$

These inequalities hold since $\vartheta < (ep)^{-12}$ and n is large enough. $\square$

6.3. Stability of optimal partitions

In most graphs any two optimal partitions may differ in at most $O(1)$ vertices: when large parts $\tilde{U}_i \subseteq U_i$ ($i = 1, \dots, p$) are already "found", the remaining vertices are classified according to their connection to the vertices already classified. We say that a vertex x has *low-degree* if $d(x, U_i) < \frac{n}{10p}$ for at least two values of i.[10] If the vertices of low-degree were already eliminated, then only the high horizontal degree vertices would create trouble.

The problem is that using this would be a "vicious circle": we wish to use that "typically there are not too many optimal partitions" to prove that "typically there are no low degrees".

Remark 6.8 (ℓ_∞-Distances between partitions). If we have two p-partitions of a set S, say, $V_1, \dots, V_p$ and $W_1, \dots, W_p$, then we can define their distance as the minimum of the $\max_i |V_i \Delta W_{\pi(p)}|$ taken over all the $p!$ permutations π of the indices.

Now, the stability of the optimal partitions means that any two optimal partitions are near to each other.

[10] Here $n/(10p)$ is fairly ad hoc...

Definition 6.9 (Pseudo-optimal partition). Let an integer ℓ be fixed. Given a graph $G_n \in \mathscr{P}_\vartheta(n, \mathscr{L})$ with an optimal partition $(U_1, \dots, U_p)$ and a vertex set X, with $|X| \leqslant \ell$, we see that $(U_1 - X, \dots, U_p - X)$ is a partition of $G_n - X$; we shall call this a *pseudo-optimal* partition of $G_n - X$.

In Section 3, paragraph 3 we have defined $\mathscr{P}^\delta_{\mathbf{UNIF}}$ and in Lemma 6.1 we proved that the typical $\mathscr{L}$-free graphs are in $\mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L})$.

Given $\alpha > 0$, we say that a graph G_n is α-*stable* if for any two optimal partitions $(U_1, \dots, U_p)$ and $(V_1, \dots, V_p)$, there is a permutation π of $\{1, \dots, p\}$ such that $|V_i \Delta U_j| > \alpha n$ if and only if $j = \pi(i)$. [11] As we shall see, for $\alpha = 2p\delta$, the graphs in $\mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L})$ are α-stable.

Recall that $\vartheta < (ep)^{-12}$. By $\omega \leqslant \lambda = \vartheta/4$ and $\vartheta \in \left(0, \frac{1}{4p}\right)$, we have $p\omega < \frac{1}{16}$.

Lemma 6.10. *Let $\ell \geqslant 0$ and $\lambda, \vartheta > 0$, and set $\delta = 10\sqrt{h(\vartheta)}$, as before. If n is sufficiently large, the following assertion holds. Let $(U_1, \dots, U_p)$ be an optimal partition of $G_n \in \mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L}) \cap \mathscr{P}_\vartheta(n, \mathscr{L})$, and $(V_1, \dots, V_p)$ a pseudo-optimal partition of $G_n - W$ for a vertex set with $|W| \leqslant \ell$.*

Then for every i, $1 \leqslant i \leqslant p$, there is a unique j, $1 \leqslant j \leqslant p$, such that $|U_i \Delta V_j| \leqslant 2\delta n$. In particular, all graphs $G_n \in \mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L})$ are 2δ-stable.

Proof. The second assertion is just the case $W = \emptyset$, so it suffices to prove the first claim. Assume that $\vartheta < \frac{1}{16p^4}$. Since $G_n \in \mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L}) \subseteq \mathscr{P}_\vartheta(n, \mathscr{L})$, we have $\sum e(U_i) \leqslant \vartheta n^2$ and $\sum e(V_i) \leqslant \vartheta n^2$. Let $1 \leqslant k \leqslant p$.

Each V_i is partitioned into p parts: $V_i \cap U_j$, $j = 1, \dots, p$. Let V_i^* be the largest of these parts so that $|V_i^*| > \frac{n}{2p^2}$, and let $U_{f(i)}$ be the set containing V_i^*. We show that $f(i)$ is a permutation.

All we have to check is that $f(i) \neq f(j)$ if $i \neq j$. This holds since otherwise

$$e(U_{f(i)}) \geqslant e(V_i^*, V_j^*) \geqslant \frac{1}{4} \cdot \left(\frac{n}{2p^2}\right)^2 \geqslant \vartheta n^2 > \sum_t e(U_t),$$

a contradiction. So $f(i)$ is a permutation and we may assume that $f(i) = i$, i.e., $V_i^* \subseteq U_i$. But now the same argument yields that $|V_i - U_i| \leqslant \delta n$, otherwise, for sufficiently small $\vartheta > 0$,

$$e(V_i) \geqslant e(V_i^*, V_i - U_i) \geqslant \frac{1}{4} \cdot \left(\frac{n}{2p^2}\right) \cdot (\delta n) \geqslant \vartheta n^2 > \sum_t e(V_t)$$

would hold which is a contradiction, since $\vartheta = o(\delta)$.

Now, $(U_1, \dots, U_p)$ and $(V_1, \dots, V_p)$ are almost interchangeable, so one can immediately copy the above proof to get that $|U_i - V_i| \leqslant \delta n$, i.e. $|U_i \Delta V_i| \leqslant 2\delta n$. This

[11] Π stands for the partition, π for the permutation.

completes the proof. W was unimportant, negligible, in the proof because $|W|$ is bounded. $\square$

Lemma 6.11 (Near-partitions). (a) *The number of partitions of* $[n]$ *into* p *classes whose distance from a given partition* $U_1, \dots, U_p$ *of* $[n]$ *is smaller than* ωn *is at most*

$$\left(\sum_{k<\omega n}\binom{n}{k}\right)^p \leqslant 2^{ph(\omega)n}. \tag{19}$$

(b) *For almost all* $\mathscr{L}$*-free graphs* G_n *there are* $2^{o(n)}$ *optimal (or pseudo-optimal) partitions.*

Proof. We shall need only (a): (b) is interesting on its own and follows from (a). The number of partitions $(U_1', \dots, U_p')$ having distance $<\omega n$ from $(U_1, \dots, U_p)$ is at most $\binom{n}{\omega n}^p$. We assume here that ω is very small: $2p\omega < \frac{1}{2}$. So (19) immediately follows. $\square$

7. Strong partitions

7.1. Large horizontal degrees

Our next lemma extends some well known facts of extremal graph theory (see e.g. [11,25]) to our "typical graphs".

Lemma 7.1 (Large horizontal degrees). *Given* $\mathscr{L}$, *let* p *and* t *be defined in the usual way, and let* $\varepsilon > 0$. *Then there exists a* $\delta_0(\varepsilon)$ *such that if* $\delta < \delta_0(\varepsilon)$, *then there are integers* h_0 *and* n_0 *so that if* $G_n \in \mathscr{P}^{\delta}_{\mathbf{UNIF}}(n, \mathscr{L})$, *with* $n > n_0$, *and* $(U_1 \cup \dots \cup U_p)$ *is an optimal partition of* G_n, *then*

$$|\{x \in U_1 : d(x, U_1) \geqslant \varepsilon n\}| \leqslant h_0.$$

The proof of this lemma is basically the same as that of a corresponding assertion in the extremal graph theory: assuming that there are too many vertices of high horizontal degrees, we have to build a $K_{p+1}(t, \dots, t) \subseteq G_n$ step by step, by finding sets in $U_1, \dots, U_p$ completely joined to each other, see e.g. [25]. To prove Lemma 7.1, we start with a generalization of a lemma of Erdős and Simonovits [25].

Although similar generalizations are implicit in [15,17,21], for the sake of clarity and completeness we shall prove Lemma 7.2 here.

Here we shall "reuse" $\varepsilon > 0$ (which until now was used in the Regularity Lemma) but from now on its meaning is a constant $\varepsilon = \varepsilon_{\mathscr{L}}$ depending on $\mathscr{L}$ but not on ϑ.

Lemma 7.2. *For every* $\ell \in \mathbb{N}$ *and* $\varepsilon > 0$, *there exist two integers,* $k, m_0 \in \mathbb{N}$ *and a* $c > 0$ *such that if* $m > m_0$ *and* $G = G(C, D)$ *is a bipartite graph with bipartition* (C, D) *where* $|C| = k$ *and* $|D| = m$ *and* $e(G) \geqslant \varepsilon km$, *then there are two vertex sets* $A \subset C, B \subset D$ *with* $|A| = \ell, |B| = \lceil cm \rceil$ *such that* (A, B) *is completely joined in* $G(C, D)$.

Proof. The lemma we need is a version of the Kővári–T. Sós–Turán Theorem [18]. We extend $\binom{x}{\ell}$ to $(-\infty, \infty)$ by

$$\binom{x}{\ell} := \begin{cases} \frac{x(x-1)\cdots(x-\ell+1)}{\ell!} & \text{for } x \geqslant \ell - 1, \\ 0 & \text{for } x \leqslant \ell - 1. \end{cases}$$

We claim that

$$k = \left\lceil \frac{2\ell}{\varepsilon} \right\rceil, \quad m_0 = 2k, \quad \text{and} \quad c = \binom{\varepsilon k}{\ell} \Big/ \binom{k}{\ell}$$

will do. Suppose that $G(C, D)$ satisfies the conditions. Call a pair (X, y) consisting of an ℓ-subset $X \subseteq C$ and a vertex $y \in D$ a *cap* if $\Gamma(y) \supseteq X$. Let us count the caps in our graph. If $d(y)$ is the degree of $y \in D$ then, by the convexity of $\binom{x}{\ell}$, there are

$$\sum_{y \in D} \binom{d(y)}{\ell} \geqslant m \binom{\varepsilon k}{\ell}$$

caps (X, y). Hence, some ℓ-subset $X_0 \subseteq C$ is in at least

$$m \binom{\varepsilon k}{\ell} \Big/ \binom{k}{\ell} = cm$$

caps, yielding the desired complete subgraph $G(X_0, Y)$, $|Y| \geqslant cm$. $\square$

Remark 7.3. In Lemma 7.2 we took $c = c_{\varepsilon,k,\ell} := \binom{\varepsilon k}{\ell} / \binom{k}{\ell}$ which is larger than $\frac{1}{2}\varepsilon^\ell$ if k is sufficiently large.

Proof of Lemma 7.1. We shall prove the lemma in a stronger form, replacing the condition that $(U_1, \dots, U_p)$ is an optimal partition by the condition that $|U_i| \geqslant \frac{n}{2p}$.

We may assume that $0 < \varepsilon < p^{-2}$. We have already fixed t and p. (See Remark 2.7.) Let $G_n \in \mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L})$, with n sufficiently large. We know the δ-super-regularity (17) + (18):

$$e(A, B) \geqslant \frac{1}{4} |A|\,|B| \quad \text{if } A \subset U_i,\ B \subset U_j,\ i \neq j \quad \text{and} \quad |A|, |B| \geqslant \frac{\delta n}{p}.$$

Define a sequence of integers $h_p, h_{p-1}, \dots, h_0$, by a backward recursion, as follows. Set $h_p = t$; having defined h_{i+1}, let $h_i = k$ be the integer whose existence is guaranteed by Lemma 7.2, when we take $\varepsilon/3$ instead of ε and $\ell = h_{i+1}$. We may assume that n is large enough so that we can take $m \in \left[\frac{n}{p}, n\right]$.

Let $n > n_0 = 3h_0/\varepsilon$. Let C^* be the set of those vertices of U_1 which are joined to *each* U_i by at least εn edges. We show that $|C^*| < h_0$. Assume the contrary, and fix h_0 vertices in C^*, forming a set $C \subseteq U_1$. For $D = U_1 - C$ we have $d(x, D) \geqslant \varepsilon n - h_0 \geqslant \frac{1}{2}\varepsilon n$.

Then by Lemma 7.2 there are sets $H_1 \subset C$ and $F_1 \subset D$ with $|H_1| = h_1$ and $|F_1| = cn$ such that H_1 is completely joined to F_1 in G_n. By the optimality of partition, each vertex $x \in H_1$ sends at least εn edges to U_2. Therefore, Lemma 7.2 implies that there are sets $H_2 \subset H_1$ and $F_2 \subset U_2$ with $|H_2| = h_2$ and $|F_2| = \lfloor cn \rfloor$ such that H_2 is completely connected to F_2 in G_n. Proceeding in this way, we find a nested sequence of sets $H_p \subset H_{p-1} \subset \cdots \subset H_2 \subset H_1$, and $F_i \subset U_i$ with $|H_i| = h_i$ and $|F_i| = \lfloor cn \rfloor$ for all $1 \leqslant i \leqslant p$ such that H_i is completely joined to $F_i \subseteq U_i$ in G_n. *A fortiori*, H_p is completely joined to $F_1, \dots, F_p$ in G_n. As $|H_p| = t$, it is sufficient to find a $K_p(t, \dots, t)$ in the p-partite subgraph of G_n spanned by the sets $F_1, \dots, F_p$. The existence of this immediately follows from Lemma 6.3, which is now applied to class-sizes $\approx cn$. To enable us to apply Lemma 6.3, δ must be much smaller here than there. Thus $K_{p+1}(t, \dots, t) \subseteq G_n$, a contradiction. □

7.2. Elimination of "low" degrees and "bad" subsets

We have assigned to each graph $G_n \in \mathscr{P}_\vartheta(n, \mathscr{L})$ an optimal partition $\Pi(G_n) = (U_1, \dots, U_p)$. We have defined (Section 6.3) a vertex x of a graph $G_n \in \mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L})$ to have *low degree* if G_n has an optimal partition $(U_1, \dots, U_p)$ such that $|\Gamma(x) \cap U_j| < \frac{n}{10p}$, where $1 \leqslant j \leqslant p$ and $x \notin U_j$. Note that if $x \in U_i$ is of low degree then by the optimality of the partition, $|\Gamma(x) \cap U_i| < \frac{n}{10p}$ holds as well.

Lemma 7.4. *In the notation of Theorem* 2.5, *at most* $n^{\mathbf{ex}(n,\mathscr{M}) + \tilde{c} \cdot n} \mathbb{A}(n)$ *graphs in* $\mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L})$ *contain vertices of low degree.*

Needless to say, we could write $o(n^{\mathbf{ex}(n,\mathscr{M}) + \tilde{c} \cdot n} \mathbb{A}(N))$ in our estimate as well, but that would not make any difference here, since increasing c immediately provides the same (seemingly stronger) result. Note that we chose $\tilde{c}$ here and in Lemma 7.6 large enough to be able to start the induction, and at least h_0.

We skip the proof since this lemma is in some sense a special case of the next one. To formulate the next lemma, we need a definition. The motivation of this definition is that in a random subgraph of $T_{n,p}$ with classes $(U_1, \dots, U_p)$, if an ℓ-tuple $X \subseteq U_i$ is fixed, a vertex $y \in U_j$ $(j \neq i)$ can be joined to X in 2^ℓ patterns, where a pattern means that $\Gamma(y) \cap X$ is fixed. The "expected number" of vertices of any fixed pattern is $|U_j|/2^\ell$. We denote a connection pattern by $[X_1/X_2]$ where $X_1 := X \cap \Gamma(y)$ and $X_2 := X - \Gamma(y)$.

Definition 7.5. Assume that $\varepsilon > 0$ is fixed. Given a set $X \subseteq U_i$ of size ℓ, having no vertices of horizontal degree $\geqslant \varepsilon n$, if for some $j \neq i$ the number of vertices $y \in U_j$ with a fixed $X_1 = \Gamma(y) \cap X$ is smaller than $|U_j|/2^{\ell+1}$, then we call X a *bad ℓ-tuple*. Denote by $\mathscr{P}^\ell_{\mathbf{GOOD}}(n, \mathscr{L})$ the family of those graphs which have no bad ℓ-tuples.

Lemma 7.6 (Bad ℓ-tuples). *Let $\mathscr{L}$ be a forbidden family and ℓ an integer. Then there are $\varepsilon = \varepsilon_{\mathscr{L}} > 0$ and $\delta = \delta(\varepsilon, \mathscr{L}) > 0$ such that*

$$|\mathscr{P}^\delta_{\mathbf{UNIF}}(n, \mathscr{L}) - \mathscr{P}^\ell_{\mathbf{GOOD}}(n, \mathscr{L})| \leqslant n^{\mathbf{ex}(n,\mathscr{M}) + \tilde{c} \cdot n} \mathbb{A}(n),$$

for some constant $\tilde{c} > 0$.

Proof of Lemma 7.6. We shall prove that if the assertion of Lemma 7.6 holds for $n > n_1$ and some sufficiently large $\tilde{c}$, then the assertion also holds for $n + \ell$. Then we choose a $\tilde{c} > 0$ so large that Lemma 7.6 holds for $n \leqslant n_1 + \ell$ and that will imply the assertion.

We can fix X in $\binom{n}{\ell}$ ways in $[n]$. Deleting X we get a $G_{n-\ell}$. The graph G_n generates a partition $(U_1 - X, U_2, \dots, U_p)$ of $V(G_{n-\ell}) = [n] - X$. Call this partition the *pseudo-optimal partition* of $G_{n-\ell}$.

Let $\mathscr{P}^X(n-\ell, \mathscr{L})$ be the set of $\mathscr{L}$-free graphs on $[n] - X$, and let

$$\phi_X : \mathscr{P}^{\delta}_{\mathbf{UNIF}}(n, \mathscr{L}) \to \mathscr{P}^X(n-\ell, \mathscr{L})$$

be the map given by $\phi_X(G_n) = G_n - X$. Clearly,

$$\begin{aligned} &|\mathscr{P}^{\delta}_{\mathbf{UNIF}}(n, \mathscr{L}) - \mathscr{P}^{\ell}_{\mathbf{GOOD}}(n, \mathscr{L})| \\ &\quad \leqslant \sum_{X} \sum_{G_{n-\ell}} |\phi^{-1}(G_{n-\ell})| \leqslant \binom{n}{\ell} p^2 |\mathscr{P}(n-\ell, \mathscr{L})| \cdot \max_{G_{n-\ell}} |\phi^{-1}(G_{n-\ell})|. \end{aligned} \tag{20}$$

We assumed that $|\mathscr{P}(n-\ell, \mathscr{L})| \leqslant (n-\ell)^{\mathbf{ex}(n-\ell, \mathscr{M}) + \tilde{c}\cdot(n-\ell)} \mathbb{A}(n-\ell)$. The problem is to estimate $|\phi^{-1}(G_{n-\ell})|$. We shall estimate this by

$$|\phi^{-1}(G_{n-\ell})| \leqslant \mathbb{N}_1(n) \cdot \mathbb{N}_2(n), \tag{21}$$

where

(a) $\mathbb{N}_1(n)$ is the number of *pseudo-optimal partitions* generated on a fixed $G_{n-\ell}$ by those G_n for which $G_n - X = G_{n-\ell}$;

(b) $\mathbb{N}_2(n)$ bounds the number of extensions between X and $G_n - X$, assumed that the induced partition is already fixed: this is the product of the number of connections of X to $G_n - X$, multiplied by the number of $G[X]$'s which can be estimated by $2^{\binom{\ell}{2}}$.

In order to make use of (20), we shall bound $N_1(n)$ and $N_2(n)$ separately. Clearly,

$$\mathbb{N}_1(n) \leqslant 2^{ph(6p\delta)n} \approx 2^{6p^2 h(\delta)n}, \tag{22}$$

since the optimal partitions of the graphs G'_n which lead to $G_{n-\ell}$ are at most at distance $5p\delta n$ from $(U_1, \dots, U_p)$.[12]

To bound $\mathbb{N}_2(n)$, we have to work a little harder. Let $\mathscr{P}^{\delta,[X]}_{\mathbf{UNIF}}(n, \mathscr{L})$ denote the family of the graphs $G_n \in \mathscr{P}^{\delta}_{\mathbf{UNIF}}(n, \mathscr{L})$ for which $X \subseteq U_i$, and another class, say U_j, $(i \neq j)$ has fewer than $|U_j|/2^{\ell+1}$ vertices completely joined to X.

Suppose that $X \subseteq U_1$ is a bad ℓ-tuple for an optimal partition $\Pi(G_n) = (U_1, \dots, U_p)$ for a $G_n \in \mathscr{P}^{\delta,[X]}_{\mathbf{UNIF}}(n, \mathscr{L})$.

[12] $h(px) \approx ph(x)$ if x is small.

Let us fix a $G_n \in \mathscr{P}^{\delta,[X]}_{\mathbf{UNIF}}(n, \mathscr{L})$. An optimal partition $\Pi(G_n) = (U_1, \dots, U_p)$ was already fixed. We may assume that $X \subseteq U_1$ and $j = 2$. Let $v := |U_2|$. We defined a connection pattern $[X_1/X_2]$ by fixing $X_1 := X \cap \Gamma(y)$. Then the number of connection-patterns between X and a $y \in U_2$ is 2^{ℓ}, the number of possible connections is $2^{v\ell}$. The number of those connections (i.e., graphs $G(X, U_2)$) where fewer than $v/2^{\ell+1}$ vertices $y \in U_2$ are joined to X according to the fixed $[X_1/X_2]$ is at most

$$2^{\ell v - \alpha_\ell v}$$

for some constant $\alpha_\ell > 0$. Indeed, this estimate is equivalent to showing that joining the two classes in a random way, with edge-probability $\frac{1}{2}$, the probability that one fixed connection pattern occurs at most $v/2^{\ell+1}$ times is at most $2^{-\alpha_\ell v}$. The expected number of vertices of a given connection is $v/2^{\ell}$. The probability that we get less than half of these connections, by Chernoff's Inequality, is smaller than $2^{-\alpha_\ell v}$. Therefore

$$\mathbb{N}_2(n) \leqslant 2^{\ell(n-|U_1|)-\alpha_\ell v+\ell h(\varepsilon)n}. \tag{23}$$

Here $2^{h(\varepsilon)\ell n}$ reflects the fact that each $x \in X$ may be joined in at most $2^{h(\varepsilon)n}$ ways to its own class. (This includes $2^{\binom{\ell}{2}}$ as well.)

We have arrived to the last stage of the proof of Lemma 7.6. Let $t = \max_{L\in\mathscr{L}} v(L)$. This maximum exists since $\mathscr{L}$ is finite. Fix $\varepsilon = \varepsilon_{\mathscr{L}}$ so that

$$h(\varepsilon) = \frac{1}{4t} \min_{\ell \leqslant t} \alpha_\ell. \tag{24}$$

Then fix ϑ and δ appropriately. This way, to estimate $|\mathscr{P}^{\delta,[X]}_{\mathbf{UNIF}}(n, \mathscr{L})|$ we gain $2^{\alpha_\ell v}$ because of the missing "connection pattern" and lose $2^{h(\varepsilon)\ell n}$ because of the possible horizontal edges.

In the "induction step" below we shall use that

$$\frac{(n-\ell)^{\mathbf{ex}(n-\ell,\mathscr{M})+\tilde{c}\cdot(n-\ell)}\,\mathbb{A}(n-\ell)}{n^{\mathbf{ex}(n,\mathscr{M})+\tilde{c}\cdot n}\,\mathbb{A}(n)} \leqslant 2^{-\left(1-\frac{1}{p}\right)n\ell}. \tag{25}$$

Here we used that $n^{\mathbf{ex}(n,\mathscr{M})+\tilde{c}\cdot n}$ is monotone increasing. Hence, using that by the "induction hypothesis" $|\mathscr{P}(n-\ell,\mathscr{L})| \leqslant \mathbb{A}(n-\ell)(n-\ell)^{\mathbf{ex}(n-\ell,\mathscr{M})+\tilde{c}\cdot(n-\ell)}$, and successively applying (20)–(23) and (25) we find that

$$\begin{aligned}
&|\mathscr{P}^{\delta}_{\mathbf{UNIF}}(n,\mathscr{L}) - \mathscr{P}^{\ell}_{\mathbf{GOOD}}(n,\mathscr{L})| \\
&\quad \leqslant \binom{n}{\ell} p^2 |\mathscr{P}(n-\ell,\mathscr{L})| \cdot \max_{G_{n-\ell}} |\phi^{-1}(G_{n-\ell})| \\
&\quad \leqslant \binom{n}{\ell} p^2 |\mathscr{P}(n-\ell,\mathscr{L})| \cdot \mathbb{N}_1(n) \cdot \mathbb{N}_2(n) \\
&\quad \leqslant n^{\ell} p^2 \cdot (n-\ell)^{\mathbf{ex}(n-\ell,\mathscr{M})+\tilde{c}\cdot(n-\ell)}\,\mathbb{A}(n-\ell) \cdot 2^{6p^2 h(\delta)n} \cdot 2^{\ell(n-|U_1|)-\alpha_\ell v+\ell h(\varepsilon)n} \\
&\quad \leqslant n^{\ell} p^2 \cdot n^{\mathbf{ex}(n,\mathscr{M})+\tilde{c}\cdot n}\,\mathbb{A}(n) 2^{-\left(1-\frac{1}{p}\right)n\ell} \cdot 2^{6p^2 h(\delta)n} \cdot 2^{\ell(n-|U_1)-\alpha_\ell v+\ell h(\varepsilon)n}.
\end{aligned}$$

All that remains is to estimate the factor of $\mathbb{A}(n) \cdot n^{\mathbf{ex}(n,\mathscr{M})+\tilde{c}\cdot n}$ above. This is

$$n^\ell p^2 \cdot 2^{-\left(1-\frac{1}{p}\right)n\ell} \cdot 2^{6p^2 h(\delta)n} \cdot 2^{\ell(n-|U_1|)-\alpha_\ell v\ell+\ell h(\varepsilon)n}.$$

Here $n^\ell p^2$ is negligible. Using that $|\ell(n-|U_1|)-\left(1-\frac{1}{p}\right)n\ell| \leqslant \sqrt{\vartheta}\log\frac{1}{\vartheta}n\ell = o(n)$, and δ and ε are such that $6p^2h(\delta)n \leqslant \alpha_\ell\frac{v}{4}$ and $\ell h(\varepsilon)n \leqslant \alpha_\ell\frac{v}{4}$, see (24),

$$n^\ell p^2 \times 2^{\sqrt{\vartheta}\log(1/\vartheta)n\ell} \times 2^{6p^2h(\delta)n} \cdot 2^{-\alpha_\ell v+\ell h(\varepsilon)n} \leqslant o(2^{-\alpha_\ell v/4}).$$

This yields the desired result. $\square$

8. Proof of Theorem 2.5

After all this preparation, we can easily prove Theorem 2.5: all we have to do is to apply the lemmas in appropriate order.

Until now we have fixed an $L = K_{p+1}(t,\dots,t)$, where $p,t \geqslant 2$, as described in Remark 2.7. Now we may slightly increase this t. It is important to emphasize here that until now we used only one graph, $K_{p+1}(t,\dots,t)$ as a forbidden graph containing some $L \in \mathscr{L}$. Now we use that $\mathscr{M}$ is finite: we fix for each $M \in \mathscr{M}$ a $K_{p-1}(\tau,\dots,\tau)$, so that $M \otimes K_{p-1}(\tau,\dots,\tau)$ contains some $L_M \in \mathscr{L}$. We may fix t as the maximum of these τ's and of $v(M)$'s. We may replace the original $\mathscr{L}$ by the finite $\mathscr{L}^*$ which is the set of the corresponding L_M's. Let ϑ, δ, η be sufficiently small positive constants satisfying the "corresponding" restrictions: we shall use below lemmas where we needed to build up some forbidden graphs, for which we needed that the number of vertices in our forbidden graphs were bounded. Therefore we could fix the constants appropriately.

1. We wish to estimate $|\mathscr{P}(n,\mathscr{L}^*)| \geqslant |\mathscr{P}(n,\mathscr{L})|$. (From now we shall not mention this difference between $\mathscr{L}$ and $\mathscr{L}^*$.) By the Main Lemma, we may restrict ourselves to estimating $|\mathscr{P}_\vartheta(n,\mathscr{L})|$, i.e., the family of $\mathscr{L}$-free graphs on $[n]$ for which, for the optimal partition $(U_1 \cup \dots \cup U_p)$, $\sum e(U_i) < \vartheta n^2$. Moreover, by Lemmas 6.1 and 6.6, we may restrict ourselves to $\mathscr{P}^\delta_{\mathbf{UNIF}}(n,\mathscr{L}) \cap \mathscr{P}^\vartheta_{\mathbf{WP}}(n,\mathscr{L})$.

2. By Lemma 7.1 there is a h_0 such that every $G_n \in \mathscr{P}^\delta_{\mathbf{UNIF}}(n,\mathscr{L})$ has at most h_0 vertices of horizontal degree $\geqslant \varepsilon n$.

3. By Lemmas 7.4 and 7.6, we may discard at most $2t \cdot n^{\mathbf{ex}(n,\mathscr{M})+\tilde{c}\cdot n} \cdot \mathbb{A}(n)$ graphs to get a part of $\mathscr{P}^\delta_{\mathbf{UNIF}}(n,\mathscr{L}) \cap \mathscr{P}^\vartheta_{\mathbf{WP}}(n,\mathscr{L})$ in which there are no low degrees, neither *bad* ℓ-tuples. $\mathscr{P}^\ell_{\mathbf{GOOD}}(n,\mathscr{L})$ is the set of graphs $G_n \in \mathscr{P}^\delta_{\mathbf{UNIF}}(n,\mathscr{L})$ that contain no vertices of low degree and no bad ℓ-tuples.

Take any of the remaining graphs, G_n with $\Pi(G_n) = (U_1,\dots,U_p)$.

We assert that $G[U_i]$ becomes $\mathscr{M}$-free, after the deletion of the vertices having horizontal degree $> \varepsilon n$. This will show that the deletion of $h_0 n$ edges provides a graph in $\mathscr{C}(n,\mathscr{M})$, where $\mathscr{C}(n,\mathscr{M})$ is the set of graphs from $\mathscr{P}^\ell_{\mathbf{GOOD}}(n,\mathscr{L})$, that after deletion of at most h_0 vertices, the rest of the vertices can be partitioned into p $\mathscr{M}$-free classes.

We know that

$$|\mathscr{C}(n,\mathscr{M})| < n^{\sum_i \mathbf{ex}(|U_i|,\mathscr{M})}\mathbb{A}(n) < n^{\mathbf{ex}(n,\mathscr{M})+c\cdot n}\mathbb{A}(n),$$

and that will complete the proof. Note that first $\tilde{c}$ is chosen large enough, at least h_0, to enable us to start the induction, then $c \geqslant \tilde{c}$ is chosen large enough to enable us to start the induction. So assume that Z is the set of vertices of U_1 of horizontal degrees $\geqslant \varepsilon n$. If $G[U_1 - Z]$ contains an $M \in \mathscr{M}$, then we can build up an $M \otimes K_{p-1}(t,\dots,t) \subseteq G_n$, first fixing an $M \subseteq G[U_1 - Z]$, using Lemma 7.6: we first fix $p-1$ sets of βn points, $V_i \subseteq U_i$ $(i > 1)$. Then, using Lemma 6.3 we find a $K_{p-1}(t,\dots,t)$ in $G_n - U_1$, and therefore an $M \otimes K_p(t,\dots,t) \supseteq L \in \mathscr{L}$ a contradiction. This completes the proof. $\square$

References

[1] V.E. Alekseev, Range of values of entropy of hereditary classes of graphs, Diskret. Mat. 4 (2) (1992) 148–157 (in Russian) (translation in Discrete Math. Appl. 3(2) (1993) 191–199).

[2] J. Balogh, B. Bollobás, D. Weinreich, The speed of hereditary properties of graphs, J. Combin. Theory Ser. B 79 (2000) 131–156.

[3] J. Balogh, B. Bollobás, D. Weinreich, The penultimate range of growth for graph properties, European J. Combin. 22 (3) (2001) 277–289.

[4] J. Balogh, B. Bollobás, D. Weinreich, Measures on monotone properties of graphs, Discrete Appl. Math. 116 (1–2) (2002) 17–36.

[5] B. Bollobás, Extremal Graph Theory, Academic Press, London, 1978.

[6] B. Bollobás. P. Erdős, M. Simonovits, E. Szemerédi, Extremal graphs without large forbidden subgraphs, in: B. Bollobás (Ed.), Advances in Graph Theory, Annals of Discrete Mathematics, Vol. 3, North-Holland, Amsterdam, 1978, pp. 29–41.

[7] B. Bollobás, A. Thomason, Projections of bodies and hereditary properties of hypergraphs, J. London Math. Soc. 27 (1995) 417–424.

[8] B. Bollobás, A. Thomason, Hereditary and monotone properties of graphs, in: R.L. Graham, J. Nešetřil (Eds.), The Mathematics of Paul Erdős, II, Algorithms Combinatorics, Vol. 14, Springer, Berlin, 1997, pp. 70–78.

[9] P. Erdős, On extremal problems of graphs and generalized graphs, Israel J. Math. 2 (1964) 183–190.

[10] P. Erdős, Some recent results on extremal problems in graph theory (Results), in: P. Rosenstiehl (Ed.), Theory of Graphs (International Symposium, Rome, 1966), Gordon and Breach, New York, Dunod, Paris, 1967, pp. 118–123.

[11] P. Erdős, On some new inequalities concerning extremal properties or graphs, in: P. Erdős, G. Katona (Eds.), Theory of Graphs, Academic Press, New York, 1968, pp. 77–81 MR38# 1026.

[12] P. Erdős, P. Frankl, V. Rödl, The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, Graphs and Combin. 2 (2) (1986) 113–121.

[13] P. Erdős, D.J. Kleitman, B.L. Rothschild, Asymptotic enumeration of K_n-free graphs, Theorie Combinatorie, Proceedings of the Conference, Vol. II, Rome, 1973, Roma, Acad. Nazionale dei Lincei, 1976, pp. 19–27.

[14] Z. Füredi, Turán type problems, in: A.D. Keedwell (Ed.), Surveys in Combinatorics (1991), Proceedings of the 13th British Combinatorial Conference, London Mathematical Society Lecture Note Series, Vol. 166, Cambridge University Press, Cambridge, 1991, pp. 253–300.

[15] C. Hundack, H.J. Prömel, A. Steger, Extremal graph problems for graphs with a color-critical vertex, Combin. Probab. Comput. 2 (4) (1993) 465–477.

[16] D.J. Kleitman, K.J. Winston, On the number of graphs without 4-cycles, Discrete Math. 41 (1982) 167–172.

[17] Ph.G. Kolaitis, H.J. Prömel, B.L. Rothschild, $K_{\ell+1}$-free graphs: asymptotic structure and a 0–1 law, Trans. Amer. Math. Soc. 303 (2) (1987) 637–671.
[18] T. Kővári, V.T. Sós, P. Turán, On a problem of Zarankiewicz, Colloq. Math. 3 (1954) 50–57.
[19] J. Komlós, M. Simonovits, Szemerédi Regularity lemma and its application in Graph Theory, Paul Erdős is 80, Proceedings of the Colloquium on Bolyai Mathematical Society, Vol. 2, Keszthely, 1993, pp. 295–352.
[20] H.J. Prömel, A. Steger, Excluding induced subgraphs. III. A general asymptotic, Random Struct. Algorithms 3 (1) (1992) 19–31.
[21] H.J. Prömel, A. Steger, The asymptotic number of graphs not containing a fixed color-critical subgraph, Combinatorica 12 (4) (1992) 463–473.
[22] H.J. Prömel, A. Steger, Excluding induced subgraphs. II. Extremal graphs, Discrete Appl. Math. 44 (1–3) (1993) 283–294.
[23] V. Rödl, On universality of graphs with uniformly distributed edges, Discrete Math. 59 (1986) 125–134.
[24] E.R. Scheinerman, J. Zito, On the size of hereditary classes of graphs, J. Combin. Theory Ser. B 61 (1) (1994) 16–39.
[25] M. Simonovits, A method for solving extremal problems in graph theory, stability problems, in: P. Erdős, G. Katona (Eds.), Theory of Graphs (Proceedings of the Colloquium, Tihany, 1966), Academic Press, New York, Akad. Kiadó, Budapest, 1968, pp. 279–319.
[26] M. Simonovits, Extremal graph theory, in: Beineke, Wilson (Eds.), Selected Topics in Graph Theory, Academic Press, London, New York, San Francisco, 1983, pp. 161–200.
[27] M. Simonovits, P. Erdős' influence on extremal graph theory, The Mathematics of Paul Erdős, II, Algorithms and Combinatorics, Vol. 14, Springer, Berlin, 1997, pp. 148–192.
[28] E. Szemerédi, Regular partitions of graphs, Colloques Internationaux C. N. R. S. Vol. 260, in Problèmes combinatoires et théorie des graphes. Proceedings of the Colloquium, Orsay, Éditions C. N. R. S, Paris, 1978, pp. 399–401.
[29] E. Szemerédi, On graphs containing no complete subgraphs with 4 vertices, Mat. Lapok 23 (1972) 111–116 (in Hungarian).
[30] P. Turán, Egy gráfelméleti szélsőérték feladatról (On an extremal problem in graph theory), Mat. Fiz. Lapok 48 (1941) 436–452.
[31] P. Turán, On the theory of graphs, Colloq. Math. 3 (1954) 19–30.
[32] Collected papers of Paul Turán, Vols. 1–3, Akadémiai Kiadó, Budapest, 1989.