Journal of Combinatorial Theory, Series B 101 (2011) 67–84

Contents lists available at ScienceDirect

Journal of Combinatorial Theory, Series B

www.elsevier.com/locate/jctb

The fine structure of octahedron-free graphs

József Balogh[a,b,1], Béla Bollobás[c,d,2], Miklós Simonovits[e,3]

[a] Department of Mathematics, University of Illinois, Urbana, IL 61801, USA
[b] Department of Mathematics, U.C. California at San Diego, 9500 Gilmann Drive, La Jolla, USA
[c] Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge CB3 0WB, UK
[d] The University of Memphis, Memphis, TN 38152, USA
[e] Rényi Institute, Budapest, Hungary

A R T I C L E   I N F O

Article history:
Received 7 January 2009
Available online 20 November 2010

Keywords:
Extremal graphs
Structure of H-free graphs
Graph counting

A B S T R A C T

This paper is one of a series of papers in which, for a family $\mathcal{L}$ of graphs, we describe the typical structure of graphs not containing any $L \in \mathcal{L}$. In this paper, we prove sharp results about the case $\mathcal{L} = \{O_6\}$, where O_6 is the graph with 6 vertices and 12 edges, given by the edges of an octahedron. Among others, we prove the following results.

(a) The vertex set of almost every O_6-free graph can be partitioned into two classes of almost equal sizes, U_1 and U_2, where the graph spanned by U_1 is a C_4-free and that by U_2 is P_3-free.

(b) Similar assertions hold when $\mathcal{L}$ is the family of *all* graphs with 6 vertices and 12 edges.

(c) If H is a graph with a color-critical edge and $\chi(H) = p + 1$, then almost every sH-free graph becomes p-chromatic after the deletion of some $s - 1$ vertices, where sH is the graph formed by s vertex disjoint copies of H.

These results are natural extensions of theorems of classical extremal graph theory. To show that results like those above do not hold in great generality, we provide examples for which the analogs of our results do not hold.

E-mail addresses: jobal@math.uiuc.edu (J. Balogh), bb12@cam.ac.uk (B. Bollobás), miki@renyi.hu (M. Simonovits).

[1] This material is based upon work supported by NSF CAREER Grants DMS-0745185 and DMS-0600303, UIUC Campus Research Board Grants 06139, 07048 and 08086, and OTKA grants 049398 and K76099.

[2] Partially supported by NSF grants DMS-0505550, DMS-0906634, ITR grants CCR-0225610, CNS-0721983 and CCF-0728928 and DARPA grant F33615-01-C-1900, ARO grant W911NF-06-1-0076.

[3] Partially supported by OTKA grants T026069, T038210, T069062.

0095-8956/$ – see front matter
doi:10.1016/j.jctb.2010.11.001

1. Introduction

1.1. Notation

Our notation is standard; in particular, the suffix is often the order of the graph in question: K_n denotes the complete graph, C_n the cycle and P_n the path with n vertices. Also, G_n stands for any graph on n vertices and M_{2s} is the 1-regular graph, i.e., the complete matching, on $2s$ vertices. For $X \subseteq V(G)$ we denote by $G[X]$ the subgraph of G spanned by X.

We define $U \otimes W$ to be the graph obtained by taking vertex disjoint copies of U and W and joining each vertex of U to each vertex of W.

Given a family $\mathcal{L}$, $\mathbf{ex}(n, \mathcal{L})$ denotes the maximum number of edges a graph G_n can have without containing a graph in $\mathcal{L}$ as a not necessarily induced subgraph. We call $\mathcal{L}$ the family of *excluded* or *forbidden* graphs. For $\mathcal{L} = \{L\}$ we simply write $\mathbf{ex}(n, L)$, and use analogous abbreviations elsewhere. $\mathcal{P}(n, \mathcal{L})$ denotes the set of $\mathcal{L}$-free graphs on $[n] := \{1, \ldots, n\}$. We write $T_{n,p}$ for the p-partite Turán graph on n vertices: this is the complete p-partite graph of order n in which the p classes are as equal as possible. More generally, $K(n_1, \ldots, n_p)$ is the complete p-partite graph with class sizes $n_1, \ldots, n_p$. The crucial property of $T_{n,p}$ is that it is the unique p-chromatic n-vertex graph with the maximum number of edges. Also, $H(n, p, s) := K_{s-1} \otimes T_{n-s+1,p}$, and let $\mathcal{A}(n, p, s)$ be the class of graphs G_n from which one can delete fewer than s vertices to obtain a graph G with $\chi(G) \leqslant p$. Note that $H(n, p, s)$ has the most edges among the n-vertex graphs having $s - 1$ vertices whose deletion yields a p-colorable graph.

Denote by sL the graph obtained by taking s vertex-disjoint copies of a graph L, by $\Gamma(u)$ the neighborhood of a vertex x, and write

$$\Gamma^*(X) := \bigcap_{u \in X} \Gamma(u) \tag{1}$$

for the set of common neighbors of a set X. We shall write $H(x) = x \log_2 \frac{1}{x} + (1 - x) \log_2 \frac{1}{1-x}$ for the binary entropy function; note that

$$\binom{n}{xn} \leqslant 2^{H(x)n}.$$

This bound is useful because $H(x) \approx -x \log_2 x \to 0$ as $x \to 0$.

1.2. Erdős–Frankl–Rödl type results

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

$$\left|\mathcal{P}(n, \mathcal{L})\right| \geqslant 2^{\mathbf{ex}(n, \mathcal{L})}. \tag{2}$$

If $\mathcal{L}$ consists of a single star: $\mathcal{L} = \{L\}$, with $L = K(1, s)$, then $\mathbf{ex}(n, \mathcal{L}) = \lfloor (s - 1)n/2 \rfloor$ and

$$\left|\mathcal{P}(n, L)\right| = 2^{(1+o(1))(\log_2 n)(s-1)n/2} = 2^{(1+o(1))(\log_2 n)\mathbf{ex}(n, L)},$$

so $|\mathcal{P}(n, L)|$ is considerably larger than $2^{\mathbf{ex}}(n, L)$. However, Erdős conjectured that if L contains a cycle then

$$\left|\mathcal{P}(n, L)\right| = 2^{(1+o(1))\mathbf{ex}(n, L)}. \tag{3}$$

Erdős, Frankl and Rödl [10] proved this conjecture when $\chi(L) > 2$. The case when $\chi(L) = 2$ is still wide open.

The main purpose of this and our earlier papers ([2] and [3]), is to establish sharp forms of Erdős–Frankl–Rödl [10] type results: to prove that for a given family $\mathcal{L}$ of forbidden graphs almost all $\mathcal{L}$-free graphs G_n look very similar to the subgraphs of $\mathcal{L}$-extremal graphs. Of course, in each case we have to specify, what we mean by being similar. It is worth emphasizing that structural results are much deeper than quantitative ones. Instead of going into technicalities, we explain this through an example. Every bipartite graph is triangle-free, but the number of bipartite graphs is larger than the number of the subgraphs of the K_3-extremal $T_{n,2}$, roughly by a factor $c\sqrt{n}$. The reason is simple: non-balanced bipartite graphs are not necessarily subgraphs of the extremal graph $K(n/2, n/2)$. So the theorem that "almost all triangle-free graphs are bipartite" is more natural and sharper than the assertion that "the number of triangle-free graphs of order n is $2^{(1+o(1))n^2/4}$".

Throughout this paper, for a given family $\mathcal{L}$, we define

$$p = p(\mathcal{L}) := \min\bigl\{\chi(L) - 1:\ L \in \mathcal{L}\bigr\}. \tag{4}$$

We shall need the notion of the "decomposition family" of $\mathcal{L}$, since in extremal graph theory often this governs the finer error terms.

Definition 1 *(Decomposition family).* Denote by I_ν the ν-vertex graph with no edges. Given a family $\mathcal{L}$, let $\mathcal{M} := \mathcal{M}(\mathcal{L})$ be the family of minimal graphs M for which there exist an $L \in \mathcal{L}L$ and a $t = t_L$ such that $L \subseteq M' \otimes K_{p-1}(t, \dots, t)$, where $M' = M'(t)$ is the graph obtained by adding t isolated vertices to M. We call $\mathcal{M}$ the *decomposition family* of $\mathcal{L}$.

If $L \in \mathcal{L}$ with minimum chromatic number, then $L \subset K_{p+1}(t, \dots, t)$ for some $t \geqslant 1$, therefore the decomposition family $\mathcal{M}$ always contains some bipartite graphs.

Example. Denote by O_6 the edge-graph of the octahedron; equivalently, $O_6 = K(2,2,2)$. As $O_6 = K(2,2,2) = C_4 \otimes K_1(2)$ we have that $\mathcal{M}(O_6) = \{C_4\}$. Note also that $\mathcal{M}(C_{2\ell+1}) = \{K_2\}$ showing that for $\ell > 1$ the role of M' is important in Definition 1.

In [3] our main result was the following.

Theorem 2. *Let $\mathcal{L}$ be an arbitrary finite family of graphs. Then there exists a constant $h_{\mathcal{L}}$ such that for almost all $\mathcal{L}$-free graphs G_n we can delete $h_{\mathcal{L}}$ vertices of G_n and partition the remaining vertices into p classes, $U_1, \dots, U_p$ such that each $G[U_i]$ is $\mathcal{M}$-free.*

Recently, in [1] and [4], similar results were obtained about characterizing the structure of almost all $\mathcal{L}$-free graphs G_n, when G_n has no induced subgraph $L \in \mathcal{L}$.

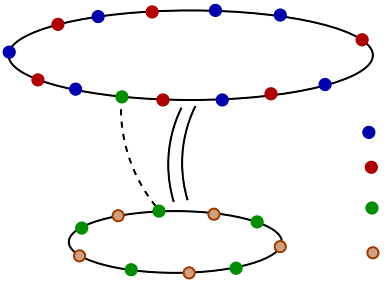
Note. Here and elsewhere, "almost always" means that all but $o(|\mathcal{P}(n, \mathcal{L})|)$ of the considered graphs (i.e. $G_n \in \mathcal{P}(n, \mathcal{L})$) have the claimed property.

Remark. The main motivation for this paper is to investigate to what extent Theorem 2 is sharp. As we shall see, in several instances it is indeed sharp. However, we have to be cautious: as we shall see, we have to avoid certain pitfalls; we shall return to this in Section 6.

2. New results

In this paper we prove several sharp results, and will discuss some of the limitations of our methods.

We say that L is a *weakly edge-color-critical* graph if there is an edge $e \in E(L)$ for which $\chi(L - e) < \chi(L)$. The edge e itself is called *critical*. We shall call these graphs shortly *weakly critical*. A graph L is *edge-color-critical* if each edge e of L is critical. Important examples of edge-color-critical graphs are the complete graphs and the odd cycles. In extremal graph theory, if an assertion can be proved for K_{p+1}, then usually it can be proved for weakly $(p+1)$-critical graphs L as well. To illustrate the differences between the notion of weakly critical and edge-color-critical, consider $L := C_\ell \otimes C_m$ for $\ell, m \geqslant 3$. If both ℓ and m are odd, then L is 6-chromatic and edge-color-critical. If ℓ is odd, m is even, then L is 5-chromatic and the edges of the C_ℓ are critical, but the edges of C_m are not, neither are the cross edges. If both ℓ and m are even, L is 4-chromatic and has no critical edges: deleting any edge leaves some $K_4 \subset L$ untouched.



For every weakly edge-color-critical graph L having a critical edge, Prömel and Steger [19] proved that almost all L-free graphs have chromatic number $\chi(L) - 1$. Their result is clearly sharp, since no graph with chromatic number $\chi(L) - 1$ contains L as a subgraph. To demonstrate the power of our methods we prove a generalization of their result: we consider the case when the excluded graph is $L = sH$, where H is weakly critical, and $\chi(H) = p + 1 \geqslant 3$. Note that Simonovits [21] proved that for n sufficiently large, the unique L-extremal graph is $H(n, p, s)$. Observe that if one can delete $s - 1$ vertices of a graph G_n to obtain a p-partite graph, then G_n is L-free. We shall prove that almost all L-free graphs have this property.

Theorem 3. *Let p and s be positive integers and H be a weakly edge-color-critical graph of chromatic number $p + 1$. Then almost every sH-free graph G_n on n vertices has a set S of $s - 1$ vertices for which $\chi(G_n - S) = p$.*

In our main result below we describe the structure of almost all octahedron-free graphs. We say that a graph G has property $\mathcal{Q} = \mathcal{Q}(C_4, P_3)$ if its vertices can be partitioned into two sets, U_1 and U_2, such that $C_4 \nsubseteq G[U_1]$ and $P_3 \nsubseteq G[U_2]$. As we remarked earlier, if $G \in \mathcal{Q}$ then G does not contain O_6. It was proved by Erdős and Simonovits [12] that for n sufficiently large every O_6-extremal G_n has property $\mathcal{Q}$. Here we prove the following.

Theorem 4. *The vertices of almost every O_6-free graph can be partitioned into two classes, U_1 and U_2, so that U_1 spans a C_4-free graph and U_2 spans a P_3-free graph.*

A similar, slightly simpler, result is the following. Denote by $\mathcal{P}(n; a, b)$ the family of graphs G_n for which no a vertices of G_n span at least b edges. In some sense, G.A. Dirac [8] started investigating such problems. Several results of Erdős and Simonovits are related to this topic, and they became very important for hypergraphs, see, e.g., Brown, Erdős and V.T. Sós [7], or Ruzsa and Szemerédi [20]. Much later, Griggs, Simonovits and Thomas [15] proved that, for n sufficiently large, the vertex set of any extremal $(6, 12)$-free graph G_n can be partitioned into U_1 and U_2 so that the induced subgraphs, $G[U_1]$ is $\{C_3, C_4\}$-free and $G[U_2]$ is an independent set. Note that if G_1 is $\{C_3, C_4\}$-free and $e(G_2) = 0$ then $G_1 \otimes G_2$ is $(6, 12)$-free.

Theorem 5. *The vertex set of almost every graph in $\mathcal{P}(n; 6, 12)$ can be partitioned into two classes, U_1 and U_2, so that U_1 spans a $\{C_3, C_4\}$-free graph and U_2 is an independent set.*

Note that what we actually prove is that almost every $\mathcal{L}$-free graph G can be partitioned into two classes, U_1 and U_2, so that U_1 spans a $\{C_3, C_4\}$-free graph and U_2 is an independent set, where $\mathcal{L} = \{O_6, C_3 \otimes I_3, P_3 \otimes (P_2 + I_1)\}$.

Along the lines of the proofs of Theorems 4 and 5, the following can also be proved. We leave the details to the reader.

Theorem 6. *Let p and $2 \leqslant a_2 \leqslant \cdots \leqslant a_p$ be integers. Then for $L = K(2, 2, a_2, \ldots, a_p)$, almost every L-free graph G has a partition $(U_1, \ldots, U_p)$ where $G[U_1]$ is C_4-free, and $P_3 \nsubseteq G[U_i]$ for $i > 1$.*

Theorem 7. *For $p \geqslant 2$, almost every $(3p, 9\binom{p}{2} + 3)$-free G has a vertex-partition $(U_1, \ldots, U_p)$ for which $G[U_1]$ is $\{C_3, C_4\}$-free, and $e(U_i) = 0$ for $i > 1$.*

One important difference between the octahedron problem and the $(6, 12)$-problem is that we know much about the C_4-extremal graphs, including asymptotically sharp bounds on the extremal number $\mathbf{ex}(n, C_4)$, while concerning the $\{C_3, C_4\}$-extremal problem there is the following tantalizing unsolved conjecture of Erdős (see [13]).

Conjecture 8.

$$\mathbf{ex}\big(n, \{C_3, C_4\}\big) = \frac{1}{2\sqrt{2}} n^{3/2} + o\big(n^{3/2}\big).$$

What this conjecture claims is that if we exclude C_3 in addition to C_4, then we are not far from having excluded all the odd cycles.

In our proofs below we shall make use of the following lower bounds on $\mathbf{ex}(n, \{C_4\})$ and $\mathbf{ex}(n, \{C_3, C_4\})$, see Kővári, V.T. Sós and Turán [18], Erdős [9] and Erdős, Rényi and V.T. Sós [11].

Theorem 9.

$$\mathbf{ex}\big(n, \{C_4\}\big) = \frac{1}{2} n^{3/2} + o\big(n^{3/2}\big)$$

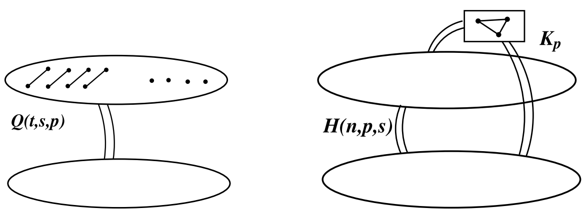
and

$$\mathbf{ex}\big(n, \{C_3, C_4\}\big) \geqslant \frac{1}{2\sqrt{2}} n^{3/2} + o\big(n^{3/2}\big). \tag{5}$$

Our results are, in some sense, pseudo-generalizations of the corresponding results in extremal graph theory: formally they do not imply the theorems: knowing the structure of almost all $\mathcal{L}$-free graph yields $\mathbf{ex}(n, \mathcal{L})$ only asymptotically. However, our results do show that the structural description of the extremal graphs, provided by us, is really the crucial one.

One would think that for every classical theorem of extremal graph theory it would not be too difficult to obtain natural generalizations to counting results. Our previous papers contain several such theorems. However, here we consider finer descriptions of almost all $\mathcal{L}$-free graphs. These finer extensions cannot be taken granted: there are cases, where the finer extremal results do not extend to counting versions. This is shown by our next example.

Definition 10. Let $s, p > 0$ be integers. For $t \geqslant 2s$, let $Q(t, p, s)$ be the complete p-partite graph with t vertices in each class and with additional s independent edges, say, in the first class.



Simonovits [22] proved that if $2s \leqslant t$, and n is sufficiently large, then $H(n, p, s)$ is the unique extremal graph for $Q(t, p, s)$.

Originally we thought that the following conjecture should hold.

Conjecture 11 *(Disproved below). Fix three integers $p > 1$ and $t \geqslant 2s > 0$. Let $L \subset Q(t,p,s)$ with $\chi(L) = p+1$. Then almost all L-free graphs G_n contain a vertex set S with at most $s-1$ vertices such that $G - S$ is p-chromatic.*

We shall see that this conjecture often holds, however, not always. To our surprise, we have found some counterexamples. Conjecture 11 holds for $s = 1$ by Prömel and Steger [19], since then L has a color-critical edge. Also, Theorem 3 yields it for $L = sH$, if H is a weakly critical graph. However, for $s \geqslant 2$ there are graphs for which Conjecture 11 is false. Below we describe such a counterexample.

Let $\mathcal{B}(U,V,x)$ be the family of graphs G_n with vertex set $[n]$, and $x \in [n]$ for which $G_n - x$ is bipartite with bipartition (U,V).

Example 12. Let $t \geqslant 4$ and $L = Q(t,2,2)$. Conjecture 11 asserts that almost all L-free graphs belong to one of the classes $\mathcal{B}(U,V,x)$. Since

$$\left|\mathcal{B}(U,V,x)\right| \leqslant 2^{|U||V|+n-1},$$

that would imply the following bounds:

$$\left|\mathcal{P}(n,\mathcal{L})\right| \leqslant \left(1+o(1)\right) \sum_{\{U,V,x\}} 2^{|U||V|+n-1} \approx 2^{n^2/4+3n/2}. \tag{6}$$

Note that in the above summation the main contribution is coming from triplets with $||U|-|V|| = o(n)$, which can be assumed when $|\mathcal{P}(n,\mathcal{L})|$ is estimated.

However, we can generate many more L-free graphs: let $\mathcal{B}(U',V',x,y)$ be the family of graphs G_n with vertex set $[n]$, and $x, y \in [n]$ for which $G_n - \{x,y\}$ is bipartite with bipartition (U',V'), and no vertex $v \in [n] - \{x,y\}$ joined to both x and y. One can check that these graphs are L-free, and

$$\left|\mathcal{B}\left(U',V',x,y\right)\right| = 2^{|U'||V'|}3^{n-2}.$$

Some of $G_n \in \mathcal{P}(n,\mathcal{L})$ are in more than four families $\mathcal{B}(U',V',x,y)$, however, almost all of them uniquely determine x, y, U', V', apart from that x and y can be switched and U' and V' can also be switched. But this results in a loss of a factor at most 4, i.e., we obtain the following lower bound on $|\mathcal{P}(n,\mathcal{L})|$:

$$\left|\mathcal{P}(n,\mathcal{L})\right| \geqslant 2^{n^2/4+3n/2-o(n)}\left(3/(2\sqrt{2})\right)^n.$$

Comparing it with (6) yields a contradiction with Conjecture 11. (Here we used that there are $2^{(1+o(1))n}$ pairs of U', V' with $||U'|-|V'|| \leqslant 1$.)

Example 13. Note that in Example 12 both p and s could have been chosen larger, providing counterexamples to Conjecture 11 for many $p, s \geqslant 2$. More importantly, we obtain an example (actually family of examples) which shows that the "condition of deletion of h vertices" cannot be omitted from Theorem 2. To spell out, fix $p = 2$ and some $s > 1$.

If Theorem 2 were true with $h = 0$, then the vertices of almost all $Q(2s,2,s)$-free graphs could be partitioned into two classes, each spanning an M_{2s}-free graph. Denote by $D(k,s)$ the number of M_{2s}-free graphs on k vertices. As an M_{2s}-free graph cannot contain more than $s-1$ vertices with degrees at least $2s$, and $2s-2$ vertices cover all the edges,

$$2^{(s-1)(k-s)} = 2^{(s-1)k+o(k)} \leqslant D(k,s) \leqslant 2^{2s-2}\binom{k}{2s-2}\binom{k}{2s}^s 2^{(s-1)k} = 2^{(s-1)k+o(k)},$$

for fixed s as k is tending to infinity. So for a fixed partition U, V of $[n]$, the number of such graphs is

$$D\big(|U|,s-1\big)D\big(|V|,s-1\big)2^{|U||V|}=2^{|U||V|+(s-1)n+o(n)}.$$

However, we can construct for a fixed partition U, V a larger family of $Q(2s,2,s)$-free graphs. Assume that $|U|\geqslant|V|$, and fix an ℓ-set $S\subset U$. We let $U-S$ and V to span independent sets, there is no restriction for the edges between $U-S$ and V, and each vertex of S is joined to at most $s-1$ vertices of $U\cup V-S$. It is easy to check that these graphs are $Q(2s,2,s)$-free, and for a fixed partition U, V we constructed

$$\Omega(1)\left(\binom{\ell}{0}+\cdots+\binom{\ell}{s-1}\right)^n 2^{(|U|-\ell)|V|} \tag{7}$$

graphs. Note that for most of the graphs constructed that way, the partition $(U-S,V,S)$ can be reconstructed, a factor of 2 might be lost if $U-S$ and V cannot be distinguished.

If we manage to find an ℓ and an s such that $\binom{\ell}{0}+\cdots+\binom{\ell}{s-1}>2^{s-1+\ell/2}$ then the second family is exponentially larger than the first, justifying our claim. In general, if ℓ is around $2s+o(s)$ then this should be true: an example (with the smallest possible s) is $s=8$ and $\ell=19$.

The organization of the paper is as follows: in Section 3 we lay the groundwork for our proofs, in Section 4 we prove Theorem 3, in Section 5 we prove Theorems 4 and 5, and in Section 6 we make some concluding remarks.

3. Notation, parameters, lemmas

In this section we state a sharper version of Theorem 2, needed in many applications. We shall define a system of parameters. Assume that a finite family $\mathcal{L}$ is given. We have already defined p in (4): $p=p(\mathcal{L}):=\min\{\chi(L)-1:\ L\in\mathcal{L}\}$.

Definition 14 *(Parameters).* Let $t:=\max\{v(L):\ L\in\mathcal{L}\}$. Fix an $\varepsilon>0$, satisfying

$$H(\varepsilon)<\big(100e^{10^{20}}\cdot 4^{t+1}p^5t^5\big)^{-1}. \tag{8}$$

This choice of ε ensures that

$$(511/512)^{10^{-9}/2}\big(27/\varepsilon^3\big)^{\varepsilon}<1-10^{-14}. \tag{9}$$

Now, depending on this ε we fix a (small) positive δ (actually depending on Lemma 14 in [3]) and to be sufficiently small to satisfy

$$\big(1-10^{-14}\big)2^{4H(4\delta p)}<1. \tag{10}$$

Now we define ϑ by $\delta=2\sqrt{H(\vartheta)}$. Note that δ will be chosen to be much smaller than ε, so that we can find appropriate embeddings as in the proof of Claim 20 needed.

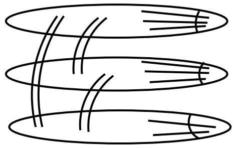
For this ε and δ we choose a (large) integer h (again by Lemma 14 in [3]). Note that $\varepsilon,\delta,\vartheta>0$ in general are assumed to be smaller than any fixed numerical positive constant (for example they satisfy (9) and (10)).

Note that we need two upper bounds on ε: we wanted (9) to be satisfied, and in [3], from which we use several results here, we had $H(\varepsilon)<(4^{t+1}p^5t^5)^{-1}$.

Using these parameters, we define several graph (sub)-classes of $\mathcal{P}(n,\mathcal{L})$, whose intersection, $\mathcal{P}^*(n,\mathcal{L})$, is the most important for us.

Definition 15 *(Optimal partition).* Given a graph G, a vertex partition $(U_1,\dots,U_p)$ of $V(G)$ is *an optimal p-partition of* G if $\sum_i e(G[U_i])$ is the minimum possible. Given $(U_1,\dots,U_p)$, we shall call the edges joining vertices in the same U_i *horizontal*, vertices from different classes *vertical edges*.

Note that the names “horizontal” and “vertical” were motivated by the way the classes are represented in the figure below.



Definition 16 *(GOOD graphs).* Given a graph G with a vertex partition $(U_1, \dots, U_p)$, call an r-tuple $X \subset V(G)$ GOOD if, for every $U_i \subseteq V(G) - X$,

$$\left|\Gamma^*(X) \cap U_i\right| = \left|\bigcap_{x \in X} \Gamma(x) \cap U_i\right| > \frac{1}{4^{r+1}}|U_i|. \tag{11}$$

An r-tuple is BAD if it is not GOOD. We say that X is a BAD *r-tuple* for a class U_i, if (11) is violated.[4]

Note that a set X may be BAD for several classes in a partition, and whether X is GOOD or BAD depends on $(U_1, \dots, U_p)$.

Definition 17. $\mathcal{P}^*(n, \mathcal{L})$ is the family of graphs $G_n \in \mathcal{P}(n, \mathcal{L})$ satisfying the following conditions: for every optimal p-partition $(U_1, \dots, U_p)$ of G_n

(i) $\sum_i e(U_i[G_n]) < \vartheta n^2$.

(ii) For every $1 \leqslant i < j \leqslant p$ and every pair of sets $A \subset U_i$, $B \subset U_j$ with $|A| = |B| \geqslant \lceil \delta n \rceil$ we have

$$e(A, B) > \frac{1}{4}|A| \cdot |B|.$$

(iii) For $i = 1, \dots, p$,

$$\left| |U_i| - \frac{n}{p} \right| < \left(\sqrt{\vartheta} \log \frac{1}{\vartheta} \right) n.$$

(iv) If W is the set of vertices with horizontal degree $\geqslant \varepsilon n$, then $G_n - W$ has no BAD $(\leqslant t)$-tuple with respect to this $(U_1, \dots, U_p)$. (By Lemma 14 in [3] we know that $|W| \leqslant h$.)[5]

Of course, $\mathcal{P}^*(n, \mathcal{L})$ strongly depends on the constants fixed in the first part of this paragraph. In [3] we introduced the GOOD graphs in several steps.[6] In a more technical form, our main result in [3] was the following.

Theorem 18 *(Good graphs). Let $\mathcal{L}$ be a finite family of forbidden graphs. Then almost all $\mathcal{L}$-free graphs are in $\mathcal{P}^*(n, \mathcal{L})$: there exist two positive constants, C and $\omega > 1$ such that*

$$\left|\mathcal{P}(n, \mathcal{L}) - \mathcal{P}^*(n, \mathcal{L})\right| \leqslant \frac{C}{\omega^n}\left|\mathcal{P}(n, \mathcal{L})\right|. \tag{12}$$

We shall use the following important lemma that a typical $\mathcal{L}$-free graph has not many optimal partitions. (The lemma has an easy proof, it is essentially the same as that of Lemmas 6.10–6.11 in [2] and Lemma 22 in [3].)

Lemma 19. *The number of optimal partitions of a $G_n \in \mathcal{P}^*(n, \mathcal{L})$ is at most $2^{4H(4\delta p)n}$.*

[4] We shall use CAPITALS for ordinary words, like GOOD/BAD to emphasize that we use them here in a predefined way.

[5] The set of graphs from $\mathcal{P}(n, \mathcal{L})$ satisfying (i) were called as $\mathcal{P}_\vartheta(n, \mathcal{L})$, (i) and (ii) $\mathcal{P}^{\delta}_{\mathbf{UNIF}}(n, \mathcal{L})$, (i) and (iii) $\mathcal{P}^{\vartheta}_{\mathbf{WP}}(n, \mathcal{L})$, and (i) and (iv) $\mathcal{P}^{t,h}_{\mathbf{GOOD}}(n, \mathcal{L})$.

[6] There

$$\mathcal{P}^*(n, \mathcal{L}) = \mathcal{P}^{t,h}_{\mathbf{GOOD}}(n, \mathcal{L}) \cap \mathcal{P}_\vartheta(n, \mathcal{L}) \cap \mathcal{P}^{\delta,\lambda}_{\mathbf{UNIF}}(n, \mathcal{L}) \cap \mathcal{P}^{\vartheta}_{\mathbf{WP}}(n, \mathcal{L}).$$

4. Proof of Theorem 3: $L = sH$, where H has a critical edge

Proof of Theorem 3. We shall deduce Theorem 3 as a simple corollary of the main result of [3], here Theorem 18. The decomposition family for sH consists of one graph: $M_{2s} := sK_2$, a graph of order $2s$ with s independent edges. Fix an $\varepsilon > 0$ and then δ and ϑ as in Definition 14. Let $m = |V(H)|$. Fix a p-coloring of H which is almost proper, the only "violation" is that the first color class contains the critical edge, badly colored. Denote the sizes of the color classes $m_1, \dots, m_p$. We can assume that H is a complete p-partite graph with $m_1, \dots, m_p$ vertices in the classes, with an extra edge in the first class, as this assumption just increases the number of sH-free graphs, and we still claim that almost all of them have property $\mathcal{A}(n, p, s)$.

Consider an arbitrary graph $G_n \in \mathcal{P}^*(n, sH)$. Fix an optimal partition $(U_1, \dots, U_p)$. Let W be the set of vertices with horizontal degrees higher than εn. Let F be a maximal set of independent horizontal edges in $G_n - W$.

Claim 20. *Every graph $G_n \in \mathcal{P}^*(n, sH)$ contains $s'H$ as a subgraph, where $s' := |W| + |F|$. In particular, $s' < s$.*

Proof. Write $\{v_1, \dots, w_{|W|}\} \in W$. Then for every i and j, $|N(v_j) \cap U_i| > \varepsilon n$, and by property (ii) of $\mathcal{P}^*(n, \mathcal{L})$ (and using $\varepsilon \gg \delta$), v can be easily extended to a copy of H, $H_j \subseteq G_n$. Moreover, for each $v_j \in W$ a new copy of H, $H_j \subseteq G_n$ can be chosen to be vertex disjoint from $H_{j'}$ for every $j' < j$, since the sets $N(v) \cap U_i - W$ satisfy property (ii) of Definition 17. This way $|W|$ vertex disjoint copies of H can be found in G_n.[7]

Let $uw \in F$ be a horizontal edge of $U_1 - W$. The set $\{u, w\}$ with any $m_1 - 2$ additional vertices forming H^1 in $U_1 - W$ is a GOOD m_1-tuple. So they have many common neighbors in $U_2 - W$. Denote an m_2-set of them by H^2. The set $H^1 \cup H^2$ is a GOOD $(m_1 + m_2)$-tuple, so they have many (linear in n) common neighbors in U_3, etc. Observe that during building H we always had many options to choose the next vertex, therefore the copies of H's can be chosen to be pairwise vertex disjoint from each other. This way we find $|F|$ additional copies of H, implying $s'H \subseteq G_n$. $\square$

Claim 21. *Let $G_n \in \mathcal{P}^*(n, sH)$. Given an optimal partition $(U_1, \dots, U_p)$ of G_n, set $\tilde{U}_i := U_i - W$, where W is the set of vertices of high horizontal degree. Then $\{\tilde{U}_1, \dots, \tilde{U}_p\}$ is independent of the optimal partition $(U_1, \dots, U_p)$.*

Proof. Assume that $(U'_1, \dots, U'_p)$ is an other optimal partition of G_n. By relabelling, if necessary, we may assume that $|U_i \Delta U'_i| \leqslant 2p\delta n$ for every i, $1 \leqslant i \leqslant p$, otherwise by Definition 17(ii) there were too many horizontal edges in G_n. It is sufficient to prove that if $u \in U_1 - W$ then $u \in U'_1 - W$ as well. If $u \in U_1 - W$ then $|N(u) \cap U_1 - W| \leqslant \varepsilon n$, so $|N(u) \cap U'_1 - W| \leqslant (\varepsilon + 2p\delta)n$. However u is a good 1-tuple with respect to the partition $(U_1, \dots, U_p)$, so $|N(u) \cap U_i - W| \geqslant 0.02n/p$, so $|N(u) \cap U'_2 - W| \geqslant 0.01n/p$ for every $2 \leqslant i \leqslant p$. Using $\varepsilon + 2p\delta \ll 0.01/p$ we obtain that $u \in U'_1$. $\square$

Now we turn to the proof of Theorem 3. For $s = 1$ we are instantly done by Claim 20, so we shall assume $s \geqslant 2$. For each p-partite graph F on $[n]$ with p-partition $(U_1, \dots, U_p)$, assign a subfamily $\Phi(F) \subset \mathcal{P}^*(n, \ell)$ of graphs $G_n \in \mathcal{P}^*(n, \ell)$ with optimal partition $(U_1, \dots, U_p)$ and vertical (cross)-edges of G_n spanning F. By Claim 20, for each $G_n \in \mathcal{P}^*(n, \ell)$ the set of vertices of high horizontal degree vertices has fever than s elements. So, by Claim 21, for each G_n there are at most p^{s-1} copies of F with $G_n \in \Phi(F)$. This implies that it is sufficient to prove that all but $2^{-n/20p}|\Phi(F)|$ graphs of $\Phi(F)$ are in $\mathcal{A}(n, p, s)$.

For a given F, if $G_n \in \Phi(F)$, then by Claim 20, in G_n there are at most $s' := 2(s - 1 - |W|)$ vertices covering every horizontal edge. This gives that the number of possible ways the horizontal edges not covered by W can be placed is at most

[7] This, or a similar, embedding algorithm can be found in several papers, see for more precise reference [2].

$$\binom{n}{s'}\binom{n}{\varepsilon n}^{s'} \leqslant 2^{s'H(\varepsilon)n+o(n)}.$$

However if $G_n \notin \mathcal{A}(n,p,s)$ then $s' > 0$. Consider $\Psi(G_n)$, defined as the subfamily of $\Phi(F)$, that consists of graphs obtained from G_n by removing all horizontal edges not covered by W, but to one arbitrary vertex $x \in [n] - W$ add arbitrarily $n/18p$ horizontal edges. (We have the upper bound $n/18p$ on the number of edges added, as in the constructed graphs we would like to keep the same optimal partition.) Then $\Psi(G_n) \subset \mathcal{A}(n,p,s)$,

$$\left|\Psi(G_n)\right| > \binom{n/2p}{n/18p} > 2^{n/18p} \quad \text{and} \quad \left|\Psi^{-1}(G_n)\right| < n2^{s'H(\varepsilon)n+o(n)}.$$

This implies that all but $n2^{-n/18p+s'H(\varepsilon)n+o(n)}|\Phi(F)| < 2^{-n/20p}|\Phi(F)|$ graphs from $\Phi(F)$ are in $\mathcal{A}(n,p,s)$, completing the proof. $\square$

5. Proof of Theorems 4 and 5

The proofs of the two theorems are the same till the very last step; up to that point we use only that $(6,12)$-free graphs are O_6-free and that our graphs are O_6-free.

In the first part of the proof we can assume that $\mathcal{L} = \{O_6\}$ or $\mathcal{L}$ consists of all graphs with 6 vertices and 12 edges, the arguments are valid in both cases. We shall consider the typical $\mathcal{L}$-free graphs $G_n \in \mathcal{P}^*(n,\mathcal{L})$ with one of their optimal partitions (U_1,U_2). Now, the number of high horizontal degree vertices is bounded: $|W| = O(1)$. The decomposition family of $\mathcal{L}$, $\mathcal{M}(\mathcal{L})$, contains a C_4, so by Theorem 18, every at most 4-tuple is GOOD inside $U_i - W$, so $U_i - W$ must span a C_4-free graph.

The difference between the O_6-free and the $(6,12)$-free cases is that in the O_6-free case one of the classes (say U_1) spans a C_4-free graph and the other a graph with maximum degree at most 1, while in the second case U_1 spans a $\{C_3,C_4\}$-free graph, and U_2 should span an independent set. We prove this in several steps. First, in Section 5.1, we show that

$$e(U_1) + e(U_2) \geqslant \frac{n^{3/2}}{20000}.$$

Then we prove that in most $\mathcal{L}$-free graphs the horizontal edges are distributed unevenly, most of them are in one of the classes (say in U_1). Then we prove in Section 5.2 that in any optimal partition of the remaining graphs $W = \emptyset$. In the last step we complete the final structural description of the graphs spanned by U_2. At that point we shall separate the proofs of the two theorems from each other.

One idea used several times is that given the assumed typical graph structure, we fix the graph spanned by U_1 and compare the size of the set of graphs with this U_1 having the “required” structure with the size of the family of graphs not having nice structure. We show that for each fixed partition the number of “bad” graphs is “negligible” among the graphs with that partition, which implies that the total number of “bad” graphs is negligible.

5.1. The number of horizontal edges in one class is large and in the other is small

We need a result of Füredi [14].

Lemma 22. *For $T \geqslant 2n^{4/3}(\log n)^2$ the number of C_4-free graphs G_n with $e(G_n) = T$ is at most*

$$\left(4n^3/T^2\right)^T. \tag{13}$$

Note that $(4n^3/T^2)^T$ is monotone increasing for $1 \leqslant T < 0.01n^{3/2}$. This is used to prove the crucial step in the main result of this subsection.

Corollary 23. *The number of $\mathcal{L}$-free graphs G_n having a partition (U_1, U_2) with $e(G[U_1]) + e(G[U_2]) \leqslant n^{3/2}/10000$ is $o(|\mathcal{P}(n, \mathcal{L})|)$.*

Proof. Indeed, the number of optimal partitions of G_n is at most 2^n, the number of ways to add the vertical (cross-edges) is at most $2^{n^2/4}$, the number of ways to choose the at most h vertices of high horizontal degree and the adjacent edges is smaller than $n^h 2^{hn}$. The number of ways to add $T \leqslant 10^{-4} n^{3/2}$ edges avoiding a C_4 to the rest of the graph is, by Lemma 22, at most

$$\sum_{T \leqslant 2n^{4/3}\log^2 n} \binom{n^2}{T} + \sum_{T \leqslant 10^{-4} n^{3/2}} \left(\frac{4n^3}{T^2}\right)^T < n^2\left[\left(4 \times 10^8\right)^{1/10000}\right]^{n^{3/2}}$$
$$< 1.005^{n^{3/2}} < 2^{n^{3/2}/100},$$

where the first sum compensates that for small values of T we do not have Lemma 22, the second one estimates the number of ways to choose $G[U_i]$. So the number of graphs that we are counting, upper estimated by $2^{n^2/4 + n^{3/2}/100 + O(n)}$, is much smaller than the number of $\mathcal{L}$-free graphs, which is at least $2^{n^2/4 + ((1/4)+o(1))(n/2)^{3/2}}$. Here we used the lower bound of (5) of Theorem 9, due to Eszter Klein and Erdős, see [9]. $\square$

From now on we may and shall assume that an optimal partition (U_1, U_2) of a $G_n \in \mathcal{P}^*(n, \mathcal{L})$ satisfies

$$e\big(G[U_1]\big) \geqslant e\big(G_n[U_2]\big), \quad \text{thus} \quad e\big(G_n[U_1]\big) \geqslant \frac{1}{20000} n^{3/2}. \tag{14}$$

Lemma 24. *Let G_m be a C_4-free graph with $e > 20m$ edges. Then G_m contains at least $e^2/4m^2$ vertex disjoint P_3's, for m sufficiently large.*

Proof. Assume that the maximum number of vertex disjoint P_3's in G_m is t. Fix t vertex disjoint P_3's in G_m. The rest of the vertices spans no P_3. So the $3t$ vertices of these P_3's cover all but at most $(m-3t)/2$ edges. Since G_m is C_4-free, hence the total number of P_3's in G_m is at most $\binom{m}{2}$. On the other hand, the degree sum of these $3t$ vertices is at least $e - m$, so the number of P_3's in G_m is at least $3t\binom{e-m/3t}{2}$, hence we have

$$3t\binom{\frac{e-m}{3t}}{2} \leqslant \binom{m}{2}.$$

Using $m < e/20$ and $t \leqslant m/3$ this implies

$$t \geqslant \frac{(e-m)(e-m-3t)}{3m(m-1)} > \frac{19e \cdot 18e}{20 \cdot 20 \cdot 3m^2} > \frac{e^2}{4m^2}. \qquad \square$$

Now we eliminate the case when in an optimal partition of G_n both classes span at least $2n(\log n)^2$ edges. The idea is that in this case there are many vertex-disjoint P_3's in both classes, hence there are many pairs of P_3's, yielding restrictions for the cross-edges between P_3's in U_1 and P_3's in U_2. Indeed a $P_3 \otimes P_3$ has 13 edges and contains an octahedron.

Lemma 25. *Let (U_1, U_2) be an optimal partition of a $G_n \in \mathcal{P}^*(n, \mathcal{L})$, with*

$$e\big(G_n[U_1]\big) > \frac{1}{20000} n^{3/2} \tag{15}$$

and n large enough.

(i) *The number of graphs* $G_n \in \mathcal{P}^*(n, \mathcal{L})$ *with*

$$e\big(G_n[U_2]\big) > 2n(\log n)^2 \tag{16}$$

is at most $2^{-n\log^3 n}|\mathcal{P}^*(n, \mathcal{L})|$.

(ii) *The number of graphs* $G_n \in \mathcal{P}^*(n, \mathcal{L})$ *violating* (16) *but for which* $G[U_2]$ *contains at least* $\log^4 n$ *vertex disjoint* P_3*'s is at most* $2^{-n\log^3 n}|\mathcal{P}^*(n, \mathcal{L})|$.

Proof. (i) Consider a graph $G_n \in \mathcal{P}^*(n, \mathcal{L})$. Fix an optimal partition (U_1, U_2) of G_n satisfying (15), and (16) and a graph $H := G[U_1 - W]$ that is C_4-free in case of the octahedron case and $\{C_3, C_4\}$-free for the $(6, 12)$-case.

In the proof we compare the sizes of the following two families of graphs:

The first family, $\mathcal{P}_I(n, \mathcal{L}; H, U_1, W) \subset \mathcal{P}^*(n, \mathcal{L})$, is the collection of $\mathcal{L}$-free graphs G_n, where G_n has an optimal partition (U_1, U_2) with one class spanning[8] H on $U_1 - W$ and $|U_1 \cap W|$ is a set of isolated vertices, the other class, U_2 spans an independent set, and, besides conditions (ii) and (iv) of Definition 17, there is no restriction on the cross-edges. The number of such graphs is at least $(1 - o(1))2^{|U_1||U_2|}$, as standard application of Chernoff's inequality implies that only $o(1)2^{|U_1||U_2|}$ graphs violate these two conditions. (Here $W \subset V(G_n)$ with $O(1)$ elements.)

The second set, $\mathcal{P}_{II}(n, \mathcal{L}; H, U_1, W)$, is the family of $\mathcal{L}$-free graphs G_n, where in the optimal partition (U_1, U_2) of G_n, $U_1 - W$ spans H, and there is no restriction on the edges incident to W, and U_2 spans any C_4-free graph on $U_2 - W$ with $e(G[U_2]) > 2n(\log n)^2$.

We give an upper bound on the number of graphs in the second family: the number of ways to choose the horizontal edges adjacent to W is at most $2^{|W|n}$, the number of ways to add e edges to $U_2 - W$ is at most $\binom{n^2}{e}$. We estimate the number of possible ways to put edges between $U_1 - W$ and $U_2 - W$ more precisely. By Lemma 24, there are at least $(n^{3/2}/20000)^2/(4n^2)$ vertex disjoint P_3's in U_1, and $e^2/(4n^2)$ in U_2. This gives

$$\frac{(n^{3/2}/20000)^2}{4n^2} \times \frac{e^2}{4n^2} = \frac{e^2}{64 \cdot 10^8 n}$$

pairs of edge-disjoint P_3's. Between each pair, to avoid an O_6, instead of the $2^9 = 512$ ways of adding the edges, there are only 511 ways (actually fewer). So the number of ways to add the cross edges is at most

$$2^{|U_1||U_2|}\left(\frac{511}{512}\right)^{e^2/(64\cdot 10^8 n)}.$$

This gives

$$\begin{aligned}\frac{|\mathcal{P}_{II}(n, \mathcal{L}, H, U_1, W) \cap \mathcal{P}^*(n, \mathcal{L})|}{|\mathcal{P}_I(n, \mathcal{L}, H, U_1, W)|} &\leqslant \frac{|\mathcal{P}_{II}(n, \mathcal{L}, H, U_1, W)|}{|\mathcal{P}_I(n, \mathcal{L}, H, U_1, W)|} \\ &\leqslant \big(1 + o(1)\big)2^{|W|n}\binom{n^2}{e}\left(\frac{511}{512}\right)^{e^2/(64\cdot 10^8 n)} \\ &\leqslant 2^{O(n)+e(O(\log n) - c\log^2 n)} \leqslant 2^{-e\log n},\end{aligned} \tag{17}$$

for some positive constant c, where we used that $e > n(\log n)^2$ and $|W| = O(1)$.

Note that $\mathcal{P}_I(n, \mathcal{L}, H, U_1, W)$ is not a partition of $\mathcal{P}^*(n, \mathcal{L})$, otherwise we would be done, but sufficiently "close". As for any $G_n \in \mathcal{P}^*(n, \mathcal{L})$ a partition (U_1, U_2) and a set W determines H spanned by $U_1 - W$, G_n is in at most $2^n n^h$ sets of $\mathcal{P}_I(n, \mathcal{L}, H, U_1, W)$. The number of ways to choose e is less than n^2, therefore we proved that for given graph spanned by U_1 there are many more $\mathcal{L}$-free graphs with U_2 spanning an independent set than graphs with $e(U_2) > 2n(\log n)^2$.

[8] Here we really mean that H is spanned, not only the graph spanned is isomorphic to H.

The proof of (ii) is exactly the same, the only difference is that in the computation in (17) we have to use that $e(G[U_2]) < n\log^2 n$ and that the graph spanned by U_2 contains at least $\log^4 n$ independent P_3's. $\square$

5.2. Elimination of vertices with large horizontal degrees

We shall use the following well-known lemma on the maximum number of edges of a C_4-free bipartite graph.

Lemma 26. *Let G be a bipartite C_4-free graph with bipartition X, Y. Then*

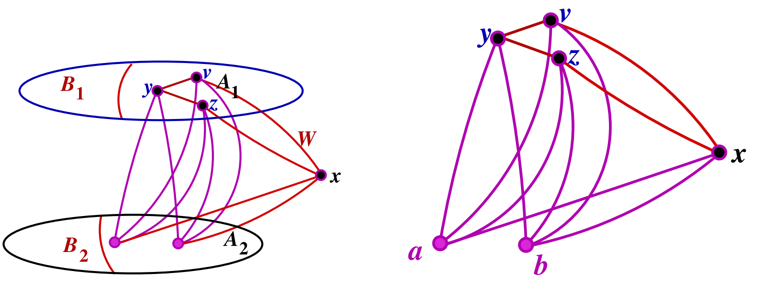
$$e(G) \leqslant |X|\sqrt{|Y|} + |X| + |Y|. \tag{18}$$

Denote by $W^*(n,\mathcal{L})$ the family of graphs $G_n \in P^*(n,\mathcal{L})$ with an optimal partition (U_1, U_2), satisfying (15) but violating (16), also assuming that W, the set of vertices with high horizontal vertices is non-empty. We shall estimate $|W^*(n,\mathcal{L})|$.

Lemma 27. *For n sufficiently large we have*

$$\left|W^*(n,\mathcal{L})\right| \leqslant \left|\mathcal{P}^*(n,\mathcal{L})\right| \cdot 2^{-(\varepsilon n)^{3/2}/10}.$$

Proof. For a $G_n \in P^*(n,\mathcal{L})$ fix a vertex $x \in W$ and denote $A_i := \Gamma(x) \cap U_i - W$ and $B_1 := U_1 - A_1 - W$. (Note that we do not care in which class is x located.) Let vyz be a P_3 in $U_1 - W$ with $v, z \in A_1$ and vy, yz edges. The key observation is that if $|\Gamma^*(\{v, y, z\}) \cap A_2| > 1$ then O_6 is a subgraph in G: x, y, z, v form a C_4 and the considered two points, $a, b \in A_2$ are completely joined to this C_4. Therefore the number of ways to place edges between $\{v, y, z\}$ and A_2 is at most $n7^{|A_2|}$, instead of $8^{|A_2|}$: there is a 'gain' of $n(7/8)^{|A_2|-1} < (7/8)^{(\varepsilon - o(1))n}$. Similarly to the proof of Lemma 25, this argument gives that at most $|\mathcal{P}^*(n,\mathcal{L})| \cdot 2^{-(n)^{3/2}}$ graphs contain more than $\sqrt{n}\log^2 n$ vertex independent P_3's like vyz.



As A_1 spans a C_4-free graph with at most $\sqrt{n}\log^2 n$ independent P_3's, by Lemma 24, we have that $e(A_1) \leqslant 2n^{5/4}\log n$. Let t denote the maximum number of independent P_3's with endpoints in A_1 and middle vertices in B_1. Fix a maximal collection of independent P_3's of this type, denote by A_3 the set of the endpoints of the paths, and by B_3 the set of middle vertices. Note that $A_3 \subset A_1$ and $B_3 \subset B_1$. Using Lemma 26, and $e(A_1 - A_3, B_1 - B_3) \leqslant 2|B_1| \leqslant 2n$, $|A_3| = 2t$, $|B_3| = t \leqslant \sqrt{n}\log^2 n$, we have that

$$\begin{aligned} e(A_1, B_1) &\leqslant e(A_3, B_3) + e(A_1 - A_3, B_3) + e(A_3, B_1 - B_3) + e(A_1 - A_3, B_1 - B_3) \\ &\leqslant \sqrt{n}3t + 4n \leqslant 5n\log^2 n. \end{aligned}$$

The conclusion is that there are only relatively few horizontal edges with at least one of the endpoints in A_1. This helps us to complete the proof. Indeed, the number of ways to choose an optimal partition, and x, A_1, A_2, W, and placing the horizontal edges incident to W, or inside U_2 or A_1 or between A_1 and B_1 is upper bounded by $2^{n^{5/4}\log^3 n}$.

Let $f(m,\mathcal{M})$ denote the number of $\mathcal{M}$-free graphs on $[m]$. Recall that when $\mathcal{L}=\{O_6\}$ then $\mathcal{M}=\{C_4\}$ and when $\mathcal{L}=\mathcal{P}(6,12)$ then $\mathcal{M}=\{C_3,C_4\}$. Trivially,

$$f(\varepsilon n/2,\mathcal{M})\geqslant 2^{((\varepsilon/2+o(1))n)^{3/2}/(2\sqrt{2})}\geqslant 2^{(\varepsilon n)^{3/2}/9}.$$

Using $|U_1|<n/2+o(n)<(1+\varepsilon)n/2$, $|B_1|\leqslant|U_1|-\varepsilon n<(1-\varepsilon)n/2$, we have

$$f\bigl(|B_1|,\mathcal{M}\bigr)f(\varepsilon n/2,\mathcal{M})\leqslant f\bigl(|B_1|+\varepsilon n/2,\mathcal{M}\bigr)\leqslant f(n/2,\mathcal{M}).$$

Putting together, we have that

$$\begin{aligned}\bigl|W^*(n,\mathcal{L})\bigr|&\leqslant f\bigl(|B_1|,\mathcal{M}\bigr)2^{n^2/4+n^{5/4}\log^3 n}\leqslant 2^{n^2/4+n^{5/4}\log^3 n}\frac{f(n/2,\mathcal{M})}{f(\varepsilon n/2,\mathcal{M})}\\&\leqslant\bigl(1+o(1)\bigr)\bigl|\mathcal{P}^*(n,\mathcal{L})\bigr|2^{n^{5/4}\log^3 n-(\varepsilon n)^{3/2}/9}\leqslant\bigl|\mathcal{P}^*(n,\mathcal{L})\bigr|2^{-(\varepsilon n)^{3/2}/10}.\qquad\Box\end{aligned}$$

5.3. *Completing the proof of the* (6, 12)*-theorem*

Now it is time to separate the proofs of Theorems 4 and 5. First we prove Theorem 5. For a partition (U_1,U_2) of $[n]$ and a graph H on U_1 we let $\mathcal{P}^*(n;6,12;U_1,H)\subset\mathcal{P}^*(n;6,12)$ be the family of graphs G_n, admitting an optimal partition (U_1,U_2), with U_1 spanning H and U_2 being an independent set, and not having vertices with high horizontal degree. Using Lemmas 25 and 27, we know that in an optimal partition (U_1,U_2) of a typical (6, 12)-free graph G_n we have $e(U_1)>n^{3/2}/20000$ and $e(U_2)<n(\log n)^2$, and every vertex of G_n has horizontal degree at most εn, U_1 spans an $\mathcal{M}$-free graph, and in U_2 there are no $\log^4 n$ independent P_3's.

Assume that the maximum number of independent edges in U_2 is t. By Lemma 24 there are at least $(n^{3/2}/20000)^2/(4n^2)>n/(2\cdot 10^9)$ independent P_3s in U_1. If between any pair of vertices spanning an edge in U_2 and a triplet of vertices spanning a P_3 in U_1 there are 6 edges, then, as the triplet is a good 3-tuple in U_1, taking an extra vertex from their common neighborhood, we obtain 6 vertices spanning at least 12 edges. So not all the 6 cross-edges can be present at the same time, i.e., out of the 64 possibilities there are at most 63 realizable. Therefore we gain on the number of ways of putting the cross-edges a multiplicative factor at least

$$\left(\frac{63}{64}\right)^{tn/(2\cdot 10^9)}=\left[\left(\frac{63}{64}\right)^{10^{-9}/2}\right]^{nt}. \tag{19}$$

However the number of ways to have a graph in U_2 with maximum t independent edges and at most $n(\log n)^2$ edges is much fewer:

Clearly, if $t>\sqrt{n}\log^2 n$ then (19) is smaller than $2^{-n^{-3/2}\log n}$ and we are done, as the gain beats the number of choices for the graph spanned by U_1 and U_2.

Let us count the number of possible graphs G_n spanned by U_2, given the graph spanned by U_1 and t: there are $2t$ vertices covering all edges in U_2, each having horizontal degrees at most εn. For each, the number of ways choosing the neighborhood is at most

$$\sum_{i\leqslant\varepsilon n}\binom{n}{i}\leqslant 2\binom{n}{\varepsilon n}\leqslant\left(\frac{3}{\varepsilon}\right)^{\varepsilon n}.$$

Hence the number of the graphs spanned by U_2 is at most

$$\binom{n}{2t}\left(\frac{3}{\varepsilon}\right)^{2t\varepsilon n}\leqslant\left(2^{O(\log n/n)}\left(\frac{9}{\varepsilon^2}\right)^{\varepsilon}\right)^{tn}.$$

Using that for ε sufficiently small $(63/64)^{10^{-9}/2}(9/\varepsilon^2)^{\varepsilon}<1-10^{-12}$, we see that for each fixed graph H spanned by U_1 and t, the number of (6, 12)-free graphs is at most $(1-10^{-12})^n|\mathcal{P}^*(n;6,12;U_1,H)|$. The number of choices for t is $o(n)$, and more importantly, by Lemma 19 no $G_n\in\mathcal{P}^*(n;6,12)$ has more than $2^{4H(4\delta p)n}$ optimal partitions. So for each G_n, U_1 and H can be chosen in at most $2^{4H(4\delta p)n}$ ways, yielding that

$$\sum_{U_1,H} \left|\mathcal{P}^*(n;6,12;U_1,H)\right| \leqslant 2^{4H(4\delta p)n}\big(1+o(1)\big)\left|\mathcal{P}^*(n;6,12)\right|.$$

Here the $(1+o(1))$ factor is needed, as the gain is with respect to all possible ways of placing the cross-edges, but only $1-o(1)$ fractions of them satisfy the conditions of Definition 17(ii) and (iv). As δ was chosen in Definition 14 to satisfy (10), it is small enough to satisfy

$$\big(1-10^{-12}\big)2^{4H(4\delta p)} < 1, \tag{20}$$

implying the theorem.

5.4. *Completing the proof of the octahedron theorem*

The proof of Theorem 4 goes along the lines of the previous subsection. For a (U_1,U_2) partition of $[n]$ and a graph H spanned by U_1 let $\mathcal{P}^*(n,O_6;U_1,H) \subset \mathcal{P}^*(n,O_6)$ be the family of graphs G_n with optimal partition (U_1,U_2), where U_1 spans H and U_2 spans a graph with maximum degree at most 1, and having no high horizontal degree. Using Lemmas 25 and 27, we may assume that in an optimal partition (U_1,U_2) of a typical O_6-free graph G_n we have $e(U_1) > n^{3/2}/20000$ and $e(U_2) < n(\log n)^2$, and every vertex of G_n has horizontal degree at most εn, U_1 spans a C_4-free graph, and in U_2 there are no more than $\log^4 n$ independent paths P_3's. Now we fix the maximum number of independent P_3's in U_2, and then the rest of the graph spanned by U_2 is P_3-free.

Let $D(k)$ denote the number of labelled P_3-free graphs on $[k]$. Trivially, $D(k)$ is a monotone increasing function. Let t be the maximum number of independent P_3s in U_2. Like in the $(6,12)$-case in the previous subsection we gain a factor

$$\left(\frac{511}{512}\right)^{tn/(2\cdot 10^9)}$$

on the number of ways of choosing the cross-edges. Now we count the number of ways of choosing the graph spanned by U_2, given the vertex set U_2, and t. There are at most $\binom{n}{3t}$ ways of choosing $3t$ vertices spanning t paths P_3; there are at most $D(|U_2|-3t)$ ways of choosing the edges in U_2 not incident to these paths; finally, the number of ways of adding the horizontal edges to a vertex of a path is at most

$$\sum_{i\leqslant \varepsilon n}\binom{n}{i} \leqslant 2\binom{n}{\varepsilon n} \leqslant \left(\frac{3}{\varepsilon}\right)^{\varepsilon n}.$$

To summarize, for given H, U_1 and t, the number of graphs spanned by U_2 is at most

$$\left(\frac{3}{\varepsilon}\right)^{3\varepsilon tn}\binom{n}{3t}D\big(|U_2|-3t\big) \leqslant \left[2^{O(\log n/n)}\left(\frac{3^3}{\varepsilon^3}\right)^{\varepsilon}\right]^{tn} D\big(|U_2|\big).$$

Using that the ε chosen in Definition 14 is sufficiently small to satisfy

$$(511/512)^{10^{-9}/2}\big(27/\varepsilon^3\big)^{\varepsilon} < 1-10^{-14},$$

so for each fixed graph spanned by U_1 and t, the number of O_6-free graphs is at most $(1-10^{-14})^n(1+o(1))|\mathcal{P}^*(n;O_6;U_1,H)|$. Here the $(1+o(1))$ factor is needed, as the gain is with respect to all possible ways of placing the cross-edges, but only $1-o(1)$ fraction of them satisfy the conditions of Definition 17(ii) and (iv). The number of choices for t is $o(n)$ and, more importantly, by Lemma 19, no $G_n \in \mathcal{P}^*(n,O_6)$ has more than $2^{4H(4\delta p)n}$ optimal partitions (so for each G_n there are at most $2^{4H(4\delta p)n}$ choices of U_1 and H), yielding

$$\sum_{U_1,H} \left|\mathcal{P}^*(n; O_6; U_1, H)\right| \leqslant 2^{4H(4\delta p)n}\left|\mathcal{P}^*(n, O_6)\right|.$$

As δ has been chosen small enough to satisfy (10) our theorem follows.

6. Comments

Our main motivation in this paper was to investigate to what extent Theorem 2 is sharp.

(i) For many families $\mathcal{L}$, for almost all $\mathcal{L}$-free graphs G_n the vertex set $[n]$ can be partitioned into p $\mathcal{M}$-free classes, i.e. to achieve this we do not have to delete vertices.

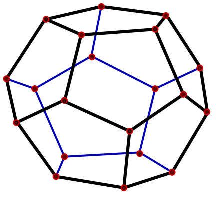
In principle it could happen that for all $\mathcal{L}$ the deletion is unnecessary. Example 12 disproves this, showing that there exist families $\mathcal{L}$ for which this vertex deletion is necessary.

(ii) There can also be another problem with the sharpness, namely, that the conclusion of the theorem is too weak. E.g., in Theorem 4 we do not have to delete vertices and only one $G[U_i]$ should be C_4-free, about the other we know a much stronger assertion that it is P_3-free. We are looking for structural descriptions that are almost necessary and sufficient, in the sense that we wish to have a property $\mathcal{Q}$ for which all the graphs of property $\mathcal{Q}$ are $\mathcal{L}$-free and almost all $\mathcal{L}$-free graphs have the property $\mathcal{Q}$. Our theorems are such statements.

Asserting that each $G[U_i]$ is $\mathcal{M}$-free is not sufficient: there are classes $\mathcal{L}$ where, if we take $\mathcal{M}$-free graphs $G^{(i)}$, the product $G^{(1)} \otimes \cdots \otimes G^{(p)}$ is not necessarily $\mathcal{L}$-free. For example, for $\mathcal{L} = \{O_6\}$ the decomposition family is $\mathcal{M} = \{C_4\}$, but if for $i = 1, 2$ $G^{(i)}$ is C_4-free and contains P_3, then $G^{(1)} \otimes G^{(2)}$ contains O_6. On the other hand if $G^{(1)}$ is C_4-free and $G^{(2)}$ is P_3-free, then $G^{(1)} \otimes G^{(2)}$ is O_6-free. We proved that for almost all O_6-free graphs G_n, $V(G_n)$ can be partitioned into U_1 and U_2 so that $G[U_1]$ is C_4-free and $G[U_2]$ has a more restricted structure, it is P_3-free.

Remark 28. Generally we wish to prove that almost all graphs not containing the considered L are very similar to subgraphs of the extremal graphs. In many cases, where the extremal graph for $\mathcal{L}$ is $H(n, p, s)$, we wish to prove that from almost all $\mathcal{L}$-free graphs on n vertices belong to $\mathcal{A}(n, p, s)$.

To get some insight into this phenomenon, it would be useful to resolve some special cases of Conjecture 11, e.g., when L is the dodecahedron graph, D_{12}.



Observe that $D_{12} \subseteq Q(n, 2, 6)$, but one cannot delete 5 vertices from D_{12} to get a bipartite graph. Thus the graphs in $\mathcal{A}(n, 2, 6)$ do not contain D_{12}. In 1974 Simonovits [22] proved that there is an $n_0 = n_0(D_{12})$ such that if $n > n_0$, then $H(n, 2, 6)$ is the unique extremal graph for D_{12}.

Problem 29. Is it true that in almost all D_{12}-free graphs we can delete 5 appropriate vertices to get a bipartite graph?

The previous problem is a special case of the next one.

Problem 30. Let L be fixed, $K_3 \nsubseteq L$. Is it true that if $\mathcal{M}(L)$ contains s independent edges and $H(n, p, s)$ is an extremal graph for L, then almost all L-free graphs are in property $\mathcal{A}(n, p, s)$?

Clearly, $O_6 = K(2, 2, 2)$. One could ask that what can be said about the structure of almost all $K(a_0, a_1, \dots, a_p)$-free graphs. It was proved by Erdős and Simonovits [12] that if $a_0 \leqslant a_1 \leqslant \cdots \leqslant a_p$, and $a_0 = 2, 3$, and n is large enough, then every extremal graph G_n of the $(p + 1)$-chromatic

$K(a_0, a_1, a_2, \dots, a_p)$ has the following structure: $V(G_n)$ can be p-partitioned so that one of the classes spans a $K(a_0, a_1)$-free graph, and the others span graphs with maximum degree $a_1 - 1$. It can easily be checked that no graph with such a structure contains a $K(a_0, a_1, a_2, \dots, a_p)$. This motivates the following conjecture.

Conjecture 31. *Let p and $2 \leqslant a_0 \leqslant \cdots \leqslant a_p$ be fixed integers. Then for $L = K(a_0, a_1, \dots, a_p)$, almost every L-free graph G_n has a partition $(U_1, \dots, U_p)$ where $G[U_1]$ is $K(a_0, a_1)$-free, and $G[U_i]$ is a graph with maximum degree less than a_1 for $i > 1$.*

What are the main obstacles to proving Conjecture 31? The first one is that the order of magnitude of the number of edges of the extremal graph of $K(a_0, a_1)$ is known only if $a_0 = 2, 3$ or $a_0 \ll a_1$. Here, probably it would be sufficient to start our proof if we knew only that

$$\mathbf{ex}\big(n, K(a_0 - 1, a_1)\big) = o\big(\mathbf{ex}\big(n, K(a_0, a_1)\big)\big).$$

There is another, more important step missing from our proof. We would need the following statement.

Conjecture 32. *If $\chi(L) = 2$ and L contains a cycle, then there is a $c = c(L) > 0$ such that almost all L-free graphs of order n have at least $c \cdot \mathbf{ex}(n, L)$ edges.*

This was known only for $L = C_4$ by Kleitman and Winston [17] and C_6 by Kleitman and Wilson [16]. We expect that one could prove results for $L = C_6 \otimes K(a_2, \dots, a_p)$ that are similar to what we are aiming for in Conjecture 31.

After our work was submitted, Balogh and Samotij [6] obtained the following result, partially proving Conjecture 32.

Theorem 33. *Let s and t be integers satisfying $s \in \{2, 3\}$ and $t \geqslant s$, or $s > 3$ and $t > (s-1)!$. There exists a positive constant $c_{s,t}$ such that almost all $K(s,t)$-free graphs of order n have at least $c_{s,t}\mathbf{ex}(n, K(s,t))$ edges. Moreover, if $t \geqslant 2$, then we may choose $c_{2,t} = 1/12$.*

We checked carefully, an argument identical to the proof of Theorem 4 gives the following: if the pair (s, t) satisfies the conditions in Theorem 33, then Conjecture 31 is true for with $a_0 = s$, $a_1 = t$ and any $p, a_2, \dots, a_p$.

Finally, we conjecture the following variant of Conjecture 32.

Conjecture 34. *If $\chi(L) = 2$ and L contains a cycle, then there is a $c = c(L) > 0$ such that almost all L-free graphs of order n have at most $(1 - c) \cdot \mathbf{ex}(n, L)$ edges.*

This conjecture was recently resolved by Balogh and Samotij in [5] for $L = C_4$ and in [6] for $L = K(2, t)$ for $t \geqslant 2$.

Acknowledgments

We are indebted to the anonymous referee for carefully reading the manuscript and for the numerous valuable comments and suggestions that helped to improve the presentation of the paper. Also, we thank W. Samotij for his careful reading of the manuscript.

References

[1] N. Alon, J. Balogh, B. Bollobás, R. Morris, The structure of almost all graphs in a hereditary property, J. Combin. Theory Ser. B 101 (2) (2011) 85–110.
[2] J. Balogh, B. Bollobás, M. Simonovits, The number of graphs without forbidden subgraphs, J. Combin. Theory Ser. B 91 (2004) 1–24.

[3] J. Balogh, B. Bollobás, M. Simonovits, The typical structure of graphs without given excluded subgraphs, Random Structures Algorithms 34 (2009) 305–318.
[4] J. Balogh, J. Butterfield, Excluding induced subgraphs: critical graphs, Random Structures Algorithms, doi:10.1002/rsa.20353, in press.
[5] J. Balogh, W. Samotij, Almost all C_4-free graphs have fewer than $(1-\varepsilon)\mathbf{ex}(n, C_4)$ edges, SIAM J. Discrete Math. 24 (2010) 1011–1018.
[6] J. Balogh, W. Samotij, The number of $K_{s,t}$-free graphs, J. Lond. Math. Soc., in press.
[7] W.G. Brown, P. Erdős, V.T. Sós, On the existence of triangulated spheres in 3-graphs and related problems, Period. Math. Hungar. 3 (1973) 22–228.
[8] G.A. Dirac, Extensions of Turán's theorem on graphs, Acta Math. Acad. Sci. Hungar. 14 (1963) 417–422.
[9] P. Erdős, On sequences of integers no one of which divides the product of two others and some related problems, Izvestiya Naustno-Issl. Inst. Mat. i Meh. Tomsk 2 (1938) 74–82 (Mitteilungen des Forshungsinstitutes für Math. und Mechanik, Tomsk).
[10] 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 Combin. 2 (1986) 113–121.
[11] P. Erdős, A. Rényi, V.T. Sós, On a problem of graph theory, Studia Sci. Math. Hungar. 1 (1966) 215–235.
[12] P. Erdős, M. Simonovits, An extremal graph problem, Acta Math. Acad. Sci. Hungar. 22 (1971/1972) 275–282.
[13] P. Erdős, M. Simonovits, Compactness results in extremal graph theory, Combinatorica 2 (1982) 275–288.
[14] Z. Füredi, Random Ramsey graphs for the four-cycle, Discrete Math. 126 (1994) 407–410.
[15] J.R. Griggs, M. Simonovits, G.R. Thomas, Extremal graphs with bounded densities of small subgraphs, J. Graph Theory 29 (1998) 185–207.
[16] D.J. Kleitman, D.B. Wilson, On the number of graphs which lack small cycles, unpublished manuscript, 1996.
[17] D.J. Kleitman, K.J. Winston, On the number of graphs without 4-cycles, Discrete Math. 41 (1982) 167–172.
[18] T. Kővári, V.T. Sós, P. Turán, On a problem of Zarankiewicz, Colloq. Math. 3 (1954) 50–57.
[19] H.J. Prömel, A. Steger, The asymptotic number of graphs not containing a fixed color-critical subgraph, Combinatorica 12 (1992) 463–473.
[20] I. Ruzsa, E. Szemerédi, Triple systems with no six points carrying 3 triangles, in: Combinatorics, Keszthely, Hungary, 1976, in: Colloq. Math. Soc. J. Bolyai, vol. 18, 1978, pp. 939–945.
[21] M. Simonovits, A method for solving extremal problems in graph theory, stability problems, in: P. Erdős, G. Katona (Eds.), Theory of Graphs, Proc. Colloq., Tihany, 1966, Academic Press/Akad. Kiadó, New York/Budapest, 1968, pp. 279–319.
[22] M. Simonovits, Extremal graph problems with symmetrical extremal graphs, additional chromatic conditions, Discrete Math. 7 (1974) 349–376.