Strengthening Rödl's theorem

Maria Chudnovsky[1]
Princeton University, Princeton, NJ 08544

Alex Scott[2]
Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK

Paul Seymour[3]
Princeton University, Princeton, NJ 08544

Sophie Spirkl[4]
University of Waterloo, Waterloo, Ontario N2L3G1, Canada

[1]This material is based upon work supported in part by the U. S. Army Research Office under grant number W911NF-16-1-0404, and supported by NSF grant DMS 1763817.

[2]Research supported by EPSRC grant EP/V007327/1.

[3]Supported by AFOSR grant A9550-19-1-0187, and by NSF grant DMS-1800053.

[4]We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number RGPIN-2020-03912]. Cette recherche a été financée par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), [numéro de référence RGPIN-2020-03912].

Abstract

What can be said about the structure of graphs that do not contain an induced copy of some graph H? Rödl showed in the 1980s that every H-free graph has large parts that are very dense or very sparse. More precisely, let us say that a graph F on n vertices is *ε-restricted* if either F or its complement has maximum degree at most εn. Rödl proved that for every graph H, and every $\varepsilon > 0$, every H-free graph G has a linear-sized set of vertices inducing an ε-restricted graph. We strengthen Rödl's result as follows: for every graph H, and all $\varepsilon > 0$, every H-free graph can be partitioned into a bounded number of subsets inducing ε-restricted graphs.

1 Introduction

What can be said about the structure of graphs that do not contain an induced copy of some graph H? In the 1980s, Rödl [5] showed that every H-free graph has large parts that are very dense or very sparse. More precisely: for a graph G, let us say $X \subseteq V(G)$ is *weakly ε-restricted* if one of $G[X], \overline{G}[X]$ has at most $\varepsilon|X|^2$ edges; and that X is *ε-restricted* if one of the graphs $G[X]$, $\overline{G}[X]$ has maximum degree at most $\varepsilon|X|$. Rödl [5] proved the following:

1.1 *For every graph H, and all $\varepsilon > 0$, there exists $\delta > 0$ such that for every H-free graph G, there is a weakly ε-restricted set $X \subseteq V(G)$ with $|X| \geq \delta|G|$.*

Every ε-restricted graph is weakly ε-restricted, and every weakly ϵ-restricted graph has an induced subgraph of linear size that is 2ε-restricted. Thus an equivalent version of Rödl's theorem is the following:

1.2 *For every graph H, and all $\varepsilon > 0$, there exists $\delta > 0$ such that for every H-free graph G, there is an ε-restricted set $X \subseteq V(G)$ with $|X| \geq \delta|G|$.*

Rödl's theorem is an easy consequence of Szemerédi's regularity lemma, and has proved extremely useful. For example, it is now a standard tool in approaching the Erdős-Hajnal conjecture (see for instance the breakthrough paper [1], where it was crucial, and much subsequent work). A proof of 1.2 not using the regularity lemma (and consequently with much better constants) was given by Fox and Sudakov [2].

In this paper, we are concerned with *partitions* of H-free graphs such that *every* vertex class is either sparse or dense. It is easy to prove that H-free graphs can be partitioned into a bounded number of weakly ε-restricted subsets:

1.3 *For every graph H, and all $\varepsilon > 0$, there is an integer N such that for every H-free graph G, there is a partition of $V(G)$ into at most N weakly ε-restricted subsets.*

Indeed, it is enough to apply 1.3 repeatedly to partition most of the vertices into weakly $\varepsilon/2$-restricted subsets, and then add the remaining vertices into the largest set. (See Nikiforov [3] for a strengthening of 1.3.)

But what about partitions into sets that satisfy the stronger property of being ε-restricted? This is much harder, and the main result of this paper is the following:

1.4 *For every graph H, and all $\varepsilon > 0$, there is an integer N such that for every H-free graph G, there is a partition of $V(G)$ into at most N ε-restricted subsets.*

This is significantly stronger than 1.3.

Some remarks: first, it is sometimes necessary to use some ε-restricted subsets of cardinality at most two, even in graphs G with $|G|$ large. For example, let G be a star $K_{1,n}$ with n large, and let $\varepsilon < 1/2$: then every ε-restricted subset containing the centre of the star has cardinality at most two. (Note that this is not the case for 1.3; for example, a large star is already weakly ϵ-restricted.)

Second, as far as we know, 1.4 does not follow from 1.3. There is a different weakening of 1.4: under the same hypotheses, $V(G)$ is the union of at most N ε-restricted subsets (not necessarily pairwise disjoint). Again, this does not imply 1.4 as far as we know, but in this case we do not know an easy proof for it.

Third, our proof of 1.4 (and the proof in [4] of 2.2, which we will need to apply) does not use the regularity lemma. Thus we anticipate that the number N in 1.4 is significantly smaller (as a function of $1/\varepsilon$) than numbers that are produced via the regularity lemma, but we have not made an estimate for it.

If $A, B \subseteq V(G)$ are disjoint, we say that B is *ε-sparse to A* (in G) if every vertex in B has at most $\varepsilon|A|$ neighbours in A; and B is *ε-dense to A* if B is ε-sparse to A in $\overline{G}$. The method of proof of 1.4 is via the following statement:

1.5 *For every graph H, and all $\varepsilon, \eta, \theta > 0$, there exists an integer N such that, for every H-free graph G, there is a partition of $V(G)$ into nonempty sets $A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_n$, where $k \leq |H|^2$ and $n \leq N$, such that:*

- *$A_1, \ldots, A_k$ and $C_1, \ldots, C_n$ are ε-restricted sets;*
- *for $1 \leq i \leq k$, $|B_i| \leq \eta|A_i|$;*
- *for $1 \leq i \leq k$, B_i is either θ-sparse or θ-dense to A_i.*

We will prove this in section 2, and apply it to deduce 1.4 in section 3.

Graphs in this paper are finite and without loops or parallel edges. If G is a graph and $X \subseteq V(G)$, we denote the subgraph of G induced on X by $G[X]$, and $\overline{G}$ denotes the complement graph of G. If G, H are graphs, we say that G is *H-free* if no induced subgraph of G is isomorphic to H.

2 Proving the main lemma

In this section we prove 1.5. Let $A, B \subseteq V(G)$ be disjoint, and let $c, \varepsilon > 0$. We say that (A, B) is *(c, ε)-full* if for all $A' \subseteq A$ with $|A'| \geq c|A|$ and $B' \subseteq B$ with $|B'| \geq c|B|$, the number of edges between A', B' is at least $\varepsilon|A'| \cdot |B'|$. Similarly, (A, B) is *(c, ε)-empty* if it is (c, ε)-full in the complement graph. Thus if (A, B) is (c, ε)-full, and $A' \subseteq A$ and $B' \subseteq B$ with $|A'|/|A|, |B'|/|B| \geq c' > c$ then (A', B') is $(c/c', \varepsilon)$-full.

We need a version of a standard result called the "embedding lemma":

2.1 *Let G, H be graphs, let $0 < \varepsilon \leq 1/2$, and let A_v ($v \in V(H)$) be pairwise disjoint nonempty subsets of $V(G)$, such that for all distinct $u, v \in V(H)$, if u, v are adjacent in H then (A_u, A_v) is $(\varepsilon^{|H|}, \varepsilon)$-full, and if u, v are nonadjacent then (A_u, A_v) is $(\varepsilon^{|H|}, \varepsilon)$-empty. Then for each $v \in V(H)$ there exists $a_v \in A_v$ such that the map sending v to a_v for each $v \in V(H)$ is an isomorphism from H to an induced subgraph of G.*

Proof. We proceed by induction on $|H|$. If $|H| \leq 1$ the result is true, so we assume $|H| > 1$. Let $v \in V(H)$, and let N, M be the sets of neighbours of v in H and in $\overline{H}$ respectively. Let $c = \varepsilon^{|H|}$. For each $u \in N$ there are fewer than $c|A_v|$ vertices in A_v with fewer than $\varepsilon|A_u|$ neighbours in A_u, since (A_v, A_u) is (c, ε)-full; and similarly for each $u \in M$ there are fewer than $c|A_v|$ vertices in A_v with fewer than $\varepsilon|A_u|$ non-neighbours in A_u. Since $(|H| - 1)c < 1$ (because $\varepsilon \leq 1/2$), there exists $a_v \in A_v$ with at least $\varepsilon|A_u|$ neighbours in A_u for each $u \in N$, and at least $\varepsilon|A_u|$ non-neighbours in A_u for each $u \in M$. For each $u \in N$ let B_u be the set of neighbours of $v \in A_u$, and for each $u \in M$ let B_u

be the set of non-neighbours of v in A_u. Thus each $B_u \neq \emptyset$, since $|B_u| \geq \varepsilon|A_u|$. Let H' be obtained from H by deleting v.

Thus for all distinct $u, w \in V(H')$, if u, w are adjacent then (B_u, B_w) is $(c\varepsilon^{-1}, \varepsilon)$-full, and if u, w are nonadjacent then (B_u, B_w) is $(c\varepsilon^{-1}, \varepsilon)$-empty. From the inductive hypothesis, for each $u \in V(H')$ there exists $a_u \in A_u$ such that the map sending u to a_u for each $u \in V(H')$ is an isomorphism from H' to an induced subgraph of G. But then the theorem holds. This proves 2.1. ∎

The following is proved in [4]:

2.2 *For all $c, \varepsilon, \tau > 0$ with $\varepsilon < \tau \leq 8/9$, there exists $\gamma > 0$ with the following property. Let G be a bipartite graph with a bipartition (A, B), with at least $\tau|A| \cdot |B|$ edges and with $A, B \neq \emptyset$. Then there exist $A' \subseteq A$ and $B' \subseteq B$ with $|A'|/|A|, |B'|/|B| \geq \gamma$, such that (A', B') is (c, ε)-full.*

Now we can prove 1.5, which we restate:

2.3 *For every graph H, and all $0 < \varepsilon, \eta, \theta < 1$, there exists an integer N such that, for every H-free graph G, there is a partition of $V(G)$ into nonempty sets $A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_n$, where $k \leq |H|^2$ and $n \leq N$, such that:*

- *$A_1, \ldots, A_k$ and $C_1, \ldots, C_n$ are ε-restricted sets;*
- *for $1 \leq i \leq k$, $|B_i| \leq \eta|A_i|$;*
- *for $1 \leq i \leq k$, B_i is either θ-dense or θ-sparse to A_i.*

Proof. We may assume that $\varepsilon, \eta, \theta < 1/3$, by reducing them if necessary. For each $\varepsilon' > 0$, let $\delta_{\varepsilon'}$ satisfy 1.2 with ε, δ replaced by $\varepsilon', \delta_{\varepsilon'}$.

Let $c_{|H|} = (\theta/4)^{|H|}$. For $m = |H| - 1, |H| - 2, \ldots, 0$ in turn:

- for $i = m, m-1, \ldots, 1$ in turn, let $\Gamma(m, i) = \gamma_{m,i+1}\Gamma(m, i+1)$ (or 1 if $i = m$); and choose $\gamma_{m,i}$ such that 2.2 holds, with $c, \varepsilon, \tau, \gamma > 0$ replaced by $\Gamma(m, i)c_{m+1}/3, \theta/4, \theta/2, \gamma_{m,i}$ respectively (by decreasing $\gamma_{m,i}$ if necessary we may assume that $\gamma_{m,i} \leq 1/3$ and $\gamma_{m,i} \leq \gamma_{m,i+1}$);
- let $c_m = \gamma_{m,0}c_{m+1}$;
- let $\varepsilon_m = 3^{|H|-m}\varepsilon$.

For each $\gamma > 0$, let n_γ be the smallest nonnegative integer that satisfies $(1 - \delta_\varepsilon)^{n_\gamma} \leq \gamma$. For $1 \leq m \leq |H|$, define $k_m = m(m-1)/2$ and

$$\ell_m = \sum_{1 \leq i \leq m} n_{q_i},$$

where $p_i = \varepsilon_{i+1}\Gamma(i, 0)$ and $q_i = \eta\delta_{p_i}\Gamma(i, 0)/2$. Let $N = \ell_{|H|} + k_{|H|} + |H|$; we claim that N satisfies the theorem.

Let G be H-free.

(1) *For all γ with $0 < \gamma < 1$, and for every $X \subseteq V(G)$, there is a partition of X into at most $n_\gamma + 1$ sets, so that one of them has cardinality at most $\gamma|X|$ and the others are all ε-restricted.*

Let $X \subseteq V(G)$. Choose an ε-restricted set $A_1 \subseteq X$ with $|A_1| \geq \delta_\varepsilon |X|$; and inductively for each $i > 1$, choose an ε-restricted set $A_i \subseteq X \setminus (A_1 \cup \cdots \cup A_{i-1})$ with $|A_i| \geq \delta_\varepsilon |X \setminus (A_1 \cup \cdots \cup A_{i-1})|$. It follows that $|X \setminus (A_1 \cup \cdots \cup A_{i-1})| \leq (1-\delta_\varepsilon)^i |X|$ for each $i \geq 0$, and in particular when $i = n_\gamma$. This proves (1).

Let $V(H)$ have vertices $v_1, \ldots, v_{|H|}$. For $0 \leq m \leq |H|$, we are interested in partitions of $V(G)$ into sets $A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_\ell, D_1, \ldots, D_m$, and E, with the following properties:

- $k \leq k_m$ and $\ell \leq \ell_m$;
- $A_1, \ldots, A_k, C_1, \ldots, C_\ell$ and $D_1, \ldots, D_m$ are all nonempty;
- $A_1, \ldots, A_k, C_1, \ldots, C_\ell$ are ε-restricted;
- for $1 \leq i \leq k$, $|B_i| \leq \eta |A_i|$, and B_i is either $(1-\theta)$-dense or $(1-\theta)$-sparse to A_i;
- $D_1, \ldots, D_m$ are ε_m-restricted;
- for $1 \leq i < j \leq m$, if v_i, v_j are adjacent in H then (D_i, D_j) is $(c_m, \theta/4)$-full, and if v_i, v_j are nonadjacent then (D_i, D_j) is $(c_m, \theta/4)$-empty;
- $|E| \leq (\eta/2) \min(|D_1|, \ldots, |D_m|)$ if $m > 0$.

Let us call such a thing a *partition of type* (k, ℓ, m). To make clear which set plays which role in the partition, we will write them as:

$$(A_1, B_1), \ldots, (A_k, B_k)$$
$$C_1, \ldots, C_\ell$$
$$D_1, \ldots, D_m$$
$$E.$$

Choose such a partition, of type (k, ℓ, m) say, with $m \leq |H|$ maximum. (This is possible, since G admits a partition of type $(0,0,0)$, setting $E = V(G)$.) Suppose that $E \neq \emptyset$.

Since $c_{|H|} = (\theta/4)^{|H|}$, and $D_1, \ldots, D_m$ are nonempty, it follows from 2.1 that $m \leq |H| - 1$. For $1 \leq i \leq m$, choose $E_i \subseteq E$ such that E_i is $\theta/2$-sparse to D_i if v_{m+1}, v_i are adjacent in H, and E_i is $\theta/2$-dense to D_i if v_{m+1} are nonadjacent, with $E_1 \cup \cdots \cup E_m$ maximal. Let $E_0 = E \setminus (E_1 \cup \cdots \cup E_m)$. Thus for $1 \leq i \leq m$, E_0 is $(1-\theta/2)$-dense to D_i if v_{m+1}, v_i are adjacent in H, and E_i is $(1-\theta/2)$-sparse to D_i if v_{m+1}, v_i are nonadjacent.

We recall that $|E| \leq \eta \min(|D_1|, \ldots, |D_m|)$, and since $E \neq \emptyset$, it follows that $|D_i| \geq \eta^{-1} > 1$ for $1 \leq i \leq m$. Thus $\lfloor |D_i|/2 \rfloor \geq |D_i|/3$, for $1 \leq i \leq m$.

Let $\varepsilon' = \varepsilon_{m+1}\Gamma(m, 0)$, and let $\delta' = \delta_{\varepsilon'}$. From 1.2 there exists $F \subseteq E_0$ with $|F| \geq \delta' |E|$ such that F is ε'-restricted. Let $F_0 = F$, and for $1 \leq i \leq m$ define $F_i \subseteq F_{i-1}$ with $|F_i| \geq \gamma_{m,i} |F_{i-1}|$, and $H_i \subseteq E_i$ with $|D_i|/2 \geq |H_i| \geq \gamma_{m,i} |D_i|$, as follows. Let us assume that v_{m+1}, v_i are adjacent (if they are non-adjacent, the construction is the same in the complement). Thus F_{i-1} is $(1-\theta/2)$-dense to D_i. (We remark that this is a weak assertion: it means that each vertex in F_{i-1} has at most $(1-\theta/2)|D_i|$ non-neighbours in D_i.) From the definition of $\gamma_{m,i}$, there exist $F_i \subseteq F_{i-1}$ and $H'_i \subseteq D_i$, with $|F_i| \geq \gamma_{m,i} |F_{i-1}|$ and $|H'_i| \geq \gamma_{m,i} |D_i|$, such that (F_i, H'_i) is $(\Gamma(m,i) c_{m+1}/3, \theta/4)$-full.

Let $H_i \subseteq H'_i$ of cardinality $\min(|H'_i|, \lfloor |D_i|/2 \rfloor)$. In particular, $|H_i| \geq \gamma_{m,i}|F_{i-1}|$, since $\gamma_{m,i} \leq 1/3$. Also, $|H_i| \geq |H'_i|/3$, and so (F_i, H_i) is $(\Gamma(m,i)c_{m+1}, \theta/4)$-full. Thus for $1 \leq i \leq m$, (F_i, H_i) is $(\Gamma(m,i)c_{m+1}, \theta/4)$-full if v_i, v_{m+1} are adjacent, and $(\Gamma(m,i+1)c_{m+1}, \theta/4)$-empty if v_i, v_{m+1} are non-adjacent. Also

$$|F_m| \geq \Gamma(m, m-1)|F_{m-1}| \geq \cdots \geq \Gamma(m,i)|F_i|.$$

Consequently (F_m, H_i) is $(c_{m+1}, \theta/4)$-full if v_i, v_{m+1} are adjacent, and $(c_{m+1}, \theta/4)$-empty if v_i, v_{m+1} are non-adjacent.

Now $|E| \leq \eta \min(|D_1|, \ldots, |D_m|)$. Let $\eta' = \eta\delta'\Gamma(m,0)/2$. By (1) there exist at most $n_{\eta'}$ pairwise disjoint nonempty ε-restricted subsets of $E_0 \setminus F_m$ such that their union (J say) satisfies $|E_0 \setminus (F_m \cup J)| \leq \eta'|E_0 \setminus F_m|$. Let these sets be $J_1, \ldots, J_n$ say. Let $H_{m+1} = F_m$. We claim that the sets

$$(A_1, B_1), \ldots, (A_k, B_k), (D_1 \setminus H_1, E_1), \ldots, (D_m \setminus H_m, E_m)$$

$$C_1, \ldots, C_\ell, J_1, \ldots, J_n$$

$$H_1, \ldots, H_m, H_{m+1}$$

$$E_0 \setminus (F_m \cup J)$$

form a partition of $V(G)$ of type $(k+m, \ell+n, m+1)$. To show this, we must check the following conditions:

- Is it true that $k+m \leq k_{m+1}$ and $\ell+n \leq \ell_{m+1}$? The first holds since $k \leq k_m$ and $k_{m+1} = k_m + m$; and the second holds since $\ell + n \leq \ell_m + n_{\eta'} = \ell_{m+1}$.

- Is it true that $A_1, \ldots, A_k, D_1 \setminus H_1, \ldots, D_m \setminus H_m, C_1, \ldots, C_\ell, J_1, \ldots, J_n, H_1, \ldots, H_m$ and F_m are all nonempty? Certainly $A_1, \ldots, A_k, C_1, \ldots, C_\ell$ are nonempty from their definition, and so are $J_1, \ldots, J_n$. For $1 \leq i \leq m$, since $|E| \leq \eta|D_i|$ and $E \neq \emptyset$, it follows that $|D_i| \geq 2$; and so $D_i \setminus H_i \neq \emptyset$, since $|H_i| \leq |D_i|/2$. Also $|H_i| = \min(|H'_i|, \lfloor |D_i|/2 \rfloor)$, and $|H'_i| \geq \gamma_{m,i}|D_i| > 0$, and $\lfloor |D_i|/2 \rfloor > 0$, so H_i is nonempty. Finally, $|F_m| \geq \Gamma(m,0)|F_0|$, and $|F_0| \geq \delta'|E| > 0$; so $F_m \neq \emptyset$.

- Is it true that $A_1, \ldots, A_k, D_1 \setminus H_1, \ldots, D_m \setminus H_m, C_1, \ldots, C_\ell, J_1, \ldots, J_n$ are ε-restricted? $A_1, \ldots, A_k$, $C_1, \ldots, C_\ell$ and $J_1, \ldots, J_n$ are ε-restricted from their definition. For $1 \leq i \leq m$, D_i is ε_m-restricted, and since $|H_i| \leq |D_i|/2$, it follows that $D_i \setminus H_i$ is $2\varepsilon_m$-restricted and hence ε-restricted.

- Is it true that for $1 \leq i \leq k$, $|B_i| \leq \eta|A_i|$, and B_i is either $(1-\theta)$-dense or $(1-\theta)$-sparse to A_i; and for $1 \leq i \leq m$, $|E_i| \leq \eta|D_i \setminus H_i|$, and E_i is either $(1-\theta)$-dense or $(1-\theta)$-sparse to $D_i \setminus H_i$? The first is true from their definition. For the second, let $1 \leq i \leq m$. Then $|E_i| \leq |E| \leq \eta|D_i|$. Assume that v_i, v_{m+1} are adjacent (the other case is similar). Then E_i is $\theta/2$-sparse to D_i, and since $|H_i| \leq |D_i|/2$, it follows that E_i is is θ-sparse to $D_i \setminus H_i$.

- Is it true that $H_1, \ldots, H_m, H_{m+1}$ are ε_{m+1}-restricted? For $1 \leq i \leq m$, D_i is ε_m-restricted, and since $|H_i| \geq |D_i|/3$, H_i is $3\varepsilon_m$-restricted and hence ε_{m+1}-restricted. Also F is ε'-restricted, and $|F_m| \geq \Gamma(m,0)|F|$; and so F_m is $\varepsilon'/\Gamma(m,0)$-restricted and hence ε_{m+1}-restricted.

- Is it true that for $1 \le i < j \le m+1$, if v_i, v_j are adjacent in H then (H_i, H_j) is $(c_{m+1}, \theta/4)$-full, and if v_i, v_j are nonadjacent then (H_i, H_j) is $(c_{m+1}, \theta/4)$-empty? If $j = m+1$, we already saw that (F_m, H_i) is $(c_{m+1}, \theta/4)$-full if v_i, v_{m+1} are adjacent, and $(c_{m+1}, \theta/4)$-empty if v_i, v_{m+1} are non-adjacent. So we may assume that $j \le m$. Assume that v_i, v_j are adjacent (the other case is similar). Then (D_i, D_j) is $(c_m, \theta/4)$-full, and since $|H_i| \ge \gamma_{m,i}|F_{i-1}|$ and $|H_j| \ge \gamma_{m,j}|F_{j-1}|$, and $\gamma_{m,i}, \gamma_{m,j} \ge \gamma_{m,0}$, and $c_m = \gamma_{m,0}c_{m+1}$, it follows that (H_i, H_j) is $(c_{m+1}, \theta/4)$-full.
- Is it true that $|E_0 \setminus (F_m \cup J)| \le (\eta/2)\min(|H_1|, \dots, |H_m|, |H_{m+1}|)$? From the choice of J, $|J| \ge (1-\eta)|E_0 \setminus F_m|$, and so
$$|E_0 \setminus (F_m \cup J)| \le \eta'|E_0 \setminus F_m| \le |E_0|/3.$$
But $|E_0| \le |E| \le (\eta/2)|D_i| \le (3\eta/2)|H_i|$, and so $|E_0 \setminus (F_m \cup J)| \le (\eta/2)|H_i|$, for $1 \le i \le m$. Finally, to show that $|E_0 \setminus (F_m \cup J)| \le (\eta/2)|H_{m+1}|$, observe that
$$|E_0 \setminus (F_m \cup J)| \le \eta'|E_0| \le \eta'|E| \le \eta'|F|/\delta' = (\eta/2)\Gamma(m,0)|F| \le (\eta/2)|F_m|.$$

This proves that G admits a partition of type $(k+m, \ell+n, m+1)$, contrary to the choice of m. Hence $E = \emptyset$. So $V(G)$ can be partitioned into the pairs $A_i \cup B_i$ with $B_i \ne \emptyset$, and the ε-restricted sets $C_1, \dots, C_\ell, D_1, \dots, D_m$, and the ε-restricted sets A_i with $1 \le i \le k$ for which $B_i = \emptyset$. This proves 1.5. ∎

3 Deducing the theorem

We need the following lemma (logarithms in this paper are to base e):

3.1 *Let $\varepsilon > 0$ with $\varepsilon \le 1/16$. Let G be a graph, and let A, B be nonempty disjoint subsets of $V(G)$, such that B is ε-sparse to A. Let p be an integer with $\log(2|B|)/\varepsilon \le p \le |A|/12$. Then there exists $P \subseteq A$ with $|P| = p$, such that P is 2ε-sparse to B, and B is 12ε-sparse to P.*

Proof. We may assume that some vertex in B has a neighbour in A, because otherwise the result holds, and since $\varepsilon \le 1/16$ it follows that $|A| \ge 16$. Let Q be the set of vertices in A with fewer than $2\varepsilon|B|$ neighbours in B, and let $q = |Q|$. There are at least $(|A| - q)(2\varepsilon|B|)$ and at most $\varepsilon|A| \cdot |B|$ edges between A and B, and so $q \ge |A|/2 \ge 8$. Let $k = \lceil 12\varepsilon p \rceil$.

Let $u_1, \dots, u_{2p} \in Q$, not necessarily all distinct. Let y be the number of subsets of Q of cardinality p that contain all of $u_1, \dots, u_{2p}$ (note that $p \le |A|/12 \le q$); and for each $v \in B$, let $z(v)$ be the number of subsets $I \subseteq \{1, \dots, 2p\}$ of cardinality k such that u_i is adjacent to v for all $i \in I$ (note that $k = \lceil 12\varepsilon p \rceil \le \lceil 2p \rceil = 2p$).

(1) *There is a choice of $u_1, \dots, u_{2p}$ with $y = 0$ and $z(v) = 0$ for all $v \in B$.*

Choose $u_1, \dots, u_{2p} \in Q$ uniformly and independently at random. Let $\overline{y}$ be the expectation of y, and $\overline{z}(v)$ the expectation of each $z(v)$. We will show that $\overline{y} < 1/2$, and $\overline{z}(v) \le 1/(2|B|)$ for each $v \in B$, from which the claim follows. First,
$$\overline{y} = \binom{q}{p}\left(\frac{p}{q}\right)^{2p} \le \frac{q^p}{p!}\left(\frac{p}{q}\right)^{2p},$$

and by Stirling's approximation, $p! \geq (p/e)^p$; so

$$\overline{y} \leq \left(\frac{eq}{p}\right)^p \left(\frac{p}{q}\right)^{2p} = \left(\frac{ep}{q}\right)^p.$$

Since $p \leq |A|/12$ and $q \geq |A|/2$, it follows that $ep/q < 1/2$, and so $\overline{y} < 1/2$.

For $v \in B$, since v has at most $\varepsilon|A| \leq 2\varepsilon|Q|$ neighbours in Q, it follows that

$$\overline{z}(v) \leq \binom{2p}{k}(2\varepsilon)^k \leq \frac{(2p)^k}{k!}(2\varepsilon)^k.$$

Since $k! \geq (k/e)^k$, we deduce that

$$\overline{z}(v) \leq \left(\frac{2ep}{k}\right)^k (2\varepsilon)^k = \left(\frac{4e\varepsilon p}{k}\right)^k \leq \left(\frac{e}{3}\right)^k \leq \left(\frac{e}{3}\right)^{12\varepsilon p}$$

since $k \geq 12\varepsilon p$ and $e < 3$. From the hypothesis, $\log(2|B|) \leq \varepsilon p \leq 12\varepsilon p \log(3/e)$, and so $(e/3)^{12\varepsilon p} \leq 1/(2|B|)$.

Hence $\overline{z}(v) \leq 1/(2|B|)$, and so the sum of $\overline{y}$ and all the $\overline{z}(v)$ $(v \in B)$ is less than one. This proves (1).

Choose $u_1, \ldots, u_{2p}$ as in (1). Since $y = 0$ it follows that $|\{u_1, \ldots, u_{2p}\}| \geq p$; choose $P \subseteq \{u_1, \ldots, u_{2p}\}$ with $|P| = p$. Each vertex in P has at most $2\varepsilon|B|$ neighbours in B, since $P \subseteq Q$; and each $v \in B$ has at most $12\varepsilon p$ neighbours in P, since $z(v) = 0$. This proves 3.1. ∎

Let G be a graph, let $k \geq 1$ be an integer, and let $\varepsilon > 0$. A (k, ε)*-path-partition* of G is a sequence $(W_0, W_1, \ldots, W_k)$ of subsets of $V(G)$, pairwise disjoint and with union $V(G)$, such that for $0 \leq i \leq k-1$:

- W_i is ε-restricted;
- $|W_{i+1} \cup \cdots \cup W_k| \leq |W_i|/12$;
- $W_{i+1} \cup \cdots \cup W_k$ is either $\varepsilon/12$-sparse or $\varepsilon/12$-dense to W_i.

If we are trying to partition $V(G)$ into ε-restricted sets, and G admits a (k, ε)-path-partition, then all but one of its sets are ε-restricted; the difficulty lies in handling the final set W_k.

3.2 *Let $0 < \varepsilon \leq 1/3$, and let G be a graph admitting a $(k, \varepsilon/2)$-path-partition $(W_0, \ldots, W_k)$, where $k = \lceil 2/\varepsilon \rceil$. Then $V(G)$ can be partitioned into at most $9217\varepsilon^{-2}$ ε-restricted subsets.*

Proof. Let $p = |W_k|$, and $\varepsilon' = \varepsilon/24$.

(1) *We may assume that* $\log(2kp) \leq \varepsilon' p$.

Suppose not; then $\log(2kp) > \varepsilon p/24$, and since $k \leq 4/\varepsilon$, it follows that

$$8p/\varepsilon \geq 2kp > e^{\varepsilon p/24} \geq (\varepsilon p/24)^3/6,$$

(because $e^x \geq x^3/3!$ for all $x > 0$). We deduce that $p^2 \leq 48(24^3/\varepsilon^4)$, and so $p < 815/\varepsilon^2$. Since $\varepsilon \leq 1/3$ and therefore $k \leq 2/\varepsilon + 1 \leq 1/\varepsilon^2$, the theorem holds, because $V(G)$ is the union of $W_0, \ldots, W_{k-1}$ and the p singletons $\{v\}$ ($v \in W_k$). This proves (1).

(2) *For $0 \leq i \leq k$, there exists $C_i \subseteq W_i$ with $|W_i| = p$, such that for $0 \leq i \leq k-1$, either*

- *C_i is $2\varepsilon'$-sparse to $C_{i+1} \cup \cdots \cup C_k$, and $C_{i+1} \cup \cdots \cup C_k$ is $12\varepsilon'$-sparse to C_i, or*
- *C_i is $2\varepsilon'$-dense to $C_{i+1} \cup \cdots \cup C_k$, and $C_{i+1} \cup \cdots \cup C_k$ is $12\varepsilon'$-dense to C_i.*

The choice of C_i is inductive, as follows: let $C_k = W_k$, and now suppose that $0 \leq i \leq k-1$, and $C_{i+1}, \ldots, C_k$ are defined. Let $B = C_{i+1} \cup \cdots \cup C_k$. Thus $|B| = (k-i)p$ and B is either ε'-sparse or ε'-dense to W_i (because $(W_0, \ldots, W_k)$ is a $(k, 12\varepsilon')$-path-partition). Moreover,

$$p \leq |\bigcup_{i<j\leq k} W_j| \leq |W_i|/12.$$

Suppose first that B is ε'-sparse to W_i. By (1), $\log(2kp) \leq \varepsilon' p$. By 3.1, taking $A = W_i$, and replacing ε by ε', we deduce that there exists $C_i \subseteq W_i$ with $|C_i| = p$, such that C_i is $2\varepsilon'$-sparse to B, and B is $12\varepsilon'$-sparse to C_i. Similarly, if B is ε'-dense to W_i, then 3.1 applied in $\overline{G}$ implies that there exists $C_i \subseteq W_i$ with $|C_i| = p$, such that C_i is $2\varepsilon'$-dense to B, and B is $12\varepsilon'$-dense to C_i. In either case, this completes the inductive definition of $C_0, \ldots, C_k$, and so proves (2).

Now for $0 \leq i \leq k-1$, either C_i is $2\varepsilon'$-sparse to $C_{i+1} \cup \cdots \cup C_k$, or $2\varepsilon'$-dense to $C_{i+1} \cup \cdots \cup C_k$; choose $I \subseteq \{0, \ldots, k-1\}$ with $|I| \geq k/2$ such that either C_i is $2\varepsilon'$-sparse to $C_{i+1} \cup \cdots \cup C_k$ for all $i \in I$, or C_i is $2\varepsilon'$-dense to $C_{i+1} \cup \cdots \cup C_k$ for all $i \in I$. Let $C = \bigcup_{i \in I \cup \{k\}} C_i$.

(3) *C is ε-restricted.*

To see this, suppose first that C_i is $2\varepsilon'$-sparse to $C_{i+1} \cup \cdots \cup C_k$ for all $i \in I$. Let $v \in C_j$ where $j \in I \cup \{k\}$, and let $I_1 = \{i \in I : i < j\}$, and $I_2 = \{i \in I \cup \{k\} : i > j\}$. Since C_j is $2\varepsilon'$-sparse to $C_{j+1} \cup \cdots \cup C_k$, it follows that v has at most $2\varepsilon' p(k-j) \leq \varepsilon p(k-j)/2$ neighbours in $C_{j+1} \cup \cdots \cup C_k$ (and hence at most the same number in $\bigcup_{i \in I_2} C_i$). Since $C_{i+1} \cup \cdots \cup C_k$ (and hence C_j) is $12\varepsilon'$-sparse to C_i, it follows that v has at most $12\varepsilon' p = \varepsilon p/2$ neighbours in C_i, for each $i < j$; and therefore v has at most $\varepsilon pj/2$ neighbours in $\bigcup_{i \in I_1} C_i$. Since v has at most $p-1$ neighbours in C_j, and $\varepsilon \leq 1$, it follows that v has at most

$$\varepsilon p(k-j)/2 + \varepsilon pj/2 + p - 1 \leq \varepsilon pk/2 + p \leq \varepsilon p(k+1) = \varepsilon|C|$$

neighbours in C, and so C is ε-restricted. If C_i is $2\varepsilon'$-dense to $C_{i+1} \cup \cdots \cup C_k$ for all $i \in I$, we use the same argument in the complement. This proves (3).

For each $i \in I$, since $|C_i| = |W_k| \leq |W_i|/2$ and W_i is $\varepsilon/2$-restricted, it follows that $W_i \setminus C_i$ is ε-restricted. But then $V(G)$ admits a partition into the sets W_i ($i \in \{1, \ldots, k-1\} \setminus I$), the sets $W_i \setminus C_i$ ($i \in I$), and C, and these sets are all ε-restricted. This proves 3.2. ∎

A *rooted tree* is a tree T with a designated vertex r called the *root.* If $s,t \in V(T)$ we say s is a *descendant* of t if t belongs to the path of T between r and s. For $t \in V(T)$, let $d(t)$ be the length of the path of T between r,t. For $t \in V(T)$, let $T(t)$ be the set of all descendants of t different from t,

Now let G be a graph, let $h,\ell \geq 1$ be integers, and let $\varepsilon,\eta > 0$. An $(h,\ell,\varepsilon,\eta)$*-tree-partition* of G is a pair $(T,(W_t : t \in V(T)))$ such that T is a rooted tree (let its root be r) and $W_t \subseteq V(G)$ for each $t \in V(T)$, with the following properties:

- every vertex t of T satisfies $d(t) \leq \ell$ and is adjacent to at most h of its descendants; and if $d(t) = \ell - 1$ then t is adjacent to at most one of its descendants;
- the sets W_t ($t \in V(T)$) are nonempty and pairwise disjoint, and have union $V(G)$;
- for each $t \in V(T)$, if $d(t) < \ell$ then W_t is ε-restricted;
- for each $t \in V(T)$, $|\bigcup_{s\in T(t)} W_s| \leq \eta|W_t|$;
- for each $t \in V(T)$, $\bigcup_{s\in T(t)} W_s$ is either $\varepsilon/12$-sparse or $\varepsilon/12$-dense to W_t.

An $(h,\ell,\varepsilon,\eta)$-tree-partition is *tight* if every maximal path of T with one end r has length ℓ. We need:

3.3 *Let $h > 0$ be an integer, let $0 < \varepsilon \leq 1$, let $K = \lceil 2/\varepsilon \rceil$, and let $0 < \eta \leq 1/(24h^K)$. Let G be a graph admitting a tight $(h,K,\varepsilon/(4h^K),\eta)$-tree-partition $(T,(W_t : t \in V(T)))$, where $K = \lceil 2/\varepsilon \rceil$. Then $V(G)$ can be partitioned into at most $9217h^K\varepsilon^{-2}$ ε-restricted subsets.*

Proof. For each $t \in V(T)$, let $S(t)$ be the set of all descendants s of t such that $d(t) = K$. (Thus $S(t) \neq \emptyset$, since the tree-partition is tight.) Let $t \in V(T)$ with $d(t) = i$. Since there exists $s \in S(t)$, and

$$|W_t| \geq \eta^{i-K}|W_s| \geq \eta^{i-K} \geq h^{K-i},$$

it follows that $|W_t|h^{i-K} \geq 1$, and so $\lfloor |W_t|h^{i-K} \rfloor \geq |W_t|h^{i-K}/2$. Consequently, and since $|S(t)| \leq h^{K-i}$, for each $s \in S(t)$ we may choose $A_t^s \subseteq W_t$, with $|A_t^s| \geq |W_t|h^{i-K}/2$, such that the sets A_t^s ($s \in S(t)$) are pairwise disjoint and have union W_t. (In particular, $A_s^s = W_s$.)

Let $s \in S(r)$, and let $t_0,\dots,t_K$ be the vertices in order of the path of T between r,s, where $t_0 = r$ and $t_K = s$. We claim that $(A_{t_0}^s, A_{t_1}^s,\dots,A_{t_K}^s)$ is a $(K,\varepsilon/2)$-path-partition of $G[A^s]$, where $A^s = A_{t_0}^s \cup A_{t_1}^s \cup \dots \cup A_{t_K}^s$. To see this, we must show that for $0 \leq i \leq K-1$:

- $A_{t_i}^s$ is $\varepsilon/2$-restricted;
- $|\bigcup_{i<j\leq K} A_{t_j}^s| \leq |A_{t_i}^s|/12$;
- $\bigcup_{i<j\leq K} A_{t_j}^s$ is either $\varepsilon/24$-sparse or $\varepsilon/24$-dense to $A_{t_i}^s$.

Let $0 \leq i \leq K-1$. Since $d(t_i) = i < K$, it follows that W_{t_i} is $\varepsilon/(4h^K)$-restricted, and since $|A_{t_i}^s| \geq |W_{t_i}|h^{i-K}/2 \geq |W_{t_i}|/(2h^K)$, we deduce that $A_{t_i}^s$ is $(\varepsilon/2)$-restricted.

To show that $|\bigcup_{i<j\leq K} A_{t_j}^s| \leq |A_{t_i}^s|/12$, observe that

$$|\bigcup_{i<j\leq K} A_{t_j}^s| \leq |\bigcup_{i<j\leq K} W_{t_j}| \leq \eta W_{t_i} \leq |A_{t_i}^s|(2\eta h^{K-i}) \leq |A_{t_i}^s|/12.$$

Finally, to show that $\bigcup_{i<j\le K} A^s_{t_j}$ is either $\varepsilon/24$-sparse or $\varepsilon/24$-dense to $A^s_{t_i}$, observe that, since $(T,(W_t : t \in V(T)))$ is an $(h,K,\varepsilon/(4h^K),\eta)$-tree-partition, it follows that $\bigcup_{i<j\le K} W_{t_j}$ is either $\varepsilon/(48h^K)$-sparse or $\varepsilon/(48h^K)$-dense to W_{t_i}, and hence so is $\bigcup_{i<j\le K} A^s_{t_j}$; and therefore $\bigcup_{i<j\le K} A^s_{t_j}$ is either $\varepsilon/24$-sparse or $\varepsilon/24$-dense to $A^s_{t_i}$, since $A^s_{t_i}| \ge |W_{t_i}|/(2h^K)$.

This proves that $(A^s_{t_0}, A^s_{t_1}, \dots, A^2_{t_K})$ is a $(K,\varepsilon/2)$-path-partition of $G[A^s]$. By 3.2, A^s is the disjoint union of at most $9217\varepsilon^{-2}$ ε-restricted subsets. But the sets A^s $(s \in S(r))$ form a partition of $V(G)$, and since $|S(r)| \le h^K$, it follows that $V(G)$ is the disjoint union of at most $9217h^K\varepsilon^{-2}$ ε-restricted subsets. This proves 3.3. ∎

Next we eliminate the "tight" hypothesis:

3.4 *Let $h > 0$ be an integer, let $0 < \varepsilon \le 1$, let $K = \lceil 2/\varepsilon \rceil$, and let $0 < \eta \le 1/(24h^K)$. Let G be a graph admitting an $(h,K,\varepsilon/(4h^K),\eta)$-tree-partition $(T,(W_t : t \in V(T)))$, Then $V(G)$ can be partitioned into at most $9218h^K\varepsilon^{-2}$ ε-restricted subsets.*

Proof. As before, for each $t \in V(T)$, let $S(t)$ be the set of all descendants s of t such that $d(s) = K$. If $t \in V(T)$ and $S(t) = \emptyset$, then W_t is ε-restricted; so if $S(t) = \emptyset$ for all $t \in V(T)$ then $V(G)$ can be partitioned into $|V(T)|$ ε-restricted sets, and the theorem is true. So we may assume that $S(t) \ne \emptyset$ for some t. Let T' be the subtree of T induced on the set of all vertices $t \in V(T)$ with $S(t) \ne \emptyset$. Let X be the union of the sets W_t $(t \in V(T'))$; then $(T',(W_t : t \in V(T)))$ is a tight $(h,K,\varepsilon/(4h^K),\eta)$-tree-partition of $G[X]$. Moreover, $V(G) \setminus X$ can be partitioned into at most $|V(T) \setminus X|$ ε-restricted sets, and by 3.3, X can be partitioned into at most $9217h^K\varepsilon^{-2}$ ε-restricted sets; and since $|V(T)| \le h^K$, this proves 3.4. ∎

We deduce:

3.5 *Let H be a graph, and let $h = |H|^2$. Let $\varepsilon > 0$, let $K = \lceil 2/\varepsilon \rceil$, and let $\eta = 1/(24h^K)$. Let N be as in 1.5, with $(\varepsilon,\eta,\theta)$ replaced by $\big(\varepsilon/(4h^K),\eta,\varepsilon/(48h^K)\big)$. Let G be an H-free graph admitting an $(h,k,\varepsilon/(4h^K),\eta)$-tree-partition $(T,(W_t : t \in V(T)))$, where $k \le K$. Then $V(G)$ can be partitioned into at most $(K-k)h^K N + 9218h^K\varepsilon^{-2}$ ε-restricted subsets.*

Proof. We proceed by induction on $K-k$. If $K-k = 0$ then the result follows from 3.4, so we assume that $k < K$, and the result holds for $k+1$. Let $(T,(W_t : t \in V(T)))$ be an $(h,k,\varepsilon/(4h^K),\eta)$-tree-partition of G, and let r be the root of T, and $S(r)$ the set of all $s \in V(T)$ with $d(s) = k$, where $d(s)$ denotes the distance in T between r,s.

Let $s \in S(r)$. By 1.5, there is a partition of W_s into nonempty sets

$$A^s_1,\dots,A^s_{m_s},B^s_1,\dots,B^s_{m_s},C^s_1,\dots,C^s_{n_s},$$

where $m_s \le h$ and $n_s \le N$, such that:

- $A^s_1,\dots,A^s_{m_s}$ and $C^s_1,\dots,C^s_{n_s}$ are $\varepsilon/(4h^K)$-restricted sets;
- for $1 \le i \le m_s$, $|B^s_i| \le \eta|A^s_i|$;
- for $1 \le i \le m_s$, B^s_i is either $\varepsilon/(48h^K)$-sparse or $\varepsilon/(48h^K)$-dense to A^s_i.

Let X be the union, over all $s \in S(r)$, of all the sets $C_1^s, \dots, C_{n_s}^s$. Since each of these sets is $\varepsilon/(4h^K)$-restricted and hence ε-restricted, it follows that X can be partitioned into at most $h^k N$ ε-restricted sets. For each $s \in S(r)$, let us delete s from T and add $2m_s$ new vertices $t_1^s, \dots, t_{m_s}^s, u_1^s, \dots, u_{m_s}^s$ to T, where $t_1^s, \dots, t_{m_s}^s$ are adjacent to the neighbour of s in T, and u_j^s is adjacent to t_j^s for $1 \le j \le m_s$. (Note that if t is the neighbour of s in T, then s is the only descendant of t.) Let T' be the tree formed by adding these new vertices (for all $s \in S(r)$). Define $W_{t_j^s} = A_j^s$ and $W_{u_j^s} = B_j^s$ for $1 \le j \le m_s$ and each $s \in S(r)$; then $(T', (W_t : t \in V(T')))$ is a $(h, k+1, \varepsilon/(4h^K), \eta)$-tree-partition of the subgraph $G[V(G) \setminus X]$. From the inductive hypothesis, $V(G) \setminus X$ can be partitioned into at most $(K-k-1)h^k N + 9218 h^K \varepsilon^{-2}$ ε-restricted sets; but X can be partitioned into at most $h^k N$ ε-restricted sets, and so $V(G)$ can be partitioned into at most $(K-k)h^k N + 9218 h^K \varepsilon^{-2}$ ε-restricted sets. This proves 3.5. ∎

Now we can deduce 1.4, which we restate:

3.6 *For every graph H, and all $\varepsilon > 0$, there is an integer M such that for every H-free graph G, there is a partition of $V(G)$ into at most M ε-restricted subsets.*

Proof. Let $K = \lceil 2/\varepsilon \rceil$, let $h = |H|^2$, let $\eta = 1/(24h^K)$, and let N be as in 1.5, with $(\varepsilon, \eta, \theta)$ replaced by

$$\left(\varepsilon/(4h^K), \eta, \varepsilon/(48h^K)\right).$$

Let $M = h^{K+1}((K-1)N + 9218\varepsilon^{-2}) + N$. We claim that M satisfies the theorem.

Let G be H-free. By 1.5, $V(G)$ can be partitioned into nonempty sets

$$A_1, \dots, A_k, B_1, \dots, B_k, C_1, \dots, C_n,$$

where $k \le h$ and $n \le N$, such that:

- $A_1, \dots, A_k$ and $C_1, \dots, C_n$ are $\varepsilon/(4h^K)$-restricted sets;
- for $1 \le i \le k$, $|B_i| \le \eta |A_i|$;
- for $1 \le i \le k$, B_i is either $(\varepsilon/(48h^K))$-sparse or $(\varepsilon/(48h^K))$-dense to A_i.

Thus for $1 \le i \le k$, $G[A_i \cup B_i]$ admits a $(1, 1, \varepsilon/(4h^K), \eta)$-tree-partition, and so, by 3.5, $A_i \cup B_i$ admits a partition into at most $(K-1)h^K N + 9218 h^K \varepsilon^{-2}$ ε-restricted sets. Since the sets $C_1, \dots, C_n$ are all ε-restricted, it follows that $V(G)$ admits a partition into at most

$$h((K-1)h^K N + 9218 h^K \varepsilon^{-2}) + N$$

ε-restricted sets. This proves 3.6. ∎

References

[1] N. Bousquet, A. Lagoutte and S. Thomassé, "The Erdős-Hajnal conjecture for paths and antipaths", *J. Combinatorial Theory, Ser. B* **113** (2015), 261–264.

[2] J. Fox and B. Sudakov, “Induced Ramsey-type theorems”, *Advances in Math.* **219** (2008), 1771–1800.

[3] V. Nikiforov, “Edge distribution of graphs with few copies of a given graph”, *Combin. Probab. Comput.* **15** (2006), 895–902.

[4] A. Scott, P. Seymour and S. Spirkl, “Pure pairs. IV. Trees in bipartite graphs”, submitted for publication, `arXiv:2009.09426`.

[5] V. Rödl, “On universality of graphs with uniformly distributed edges”, *Discrete Math.* **59** (1986), 125–134.