The Global Structure of a Typical Graph without H as an (induced) subgraph.

B. Reed

July 16, 2019

Abstract

1 An Overview

A graph is H-Free if it does not contain H as a subgraph and *H-free* if it does not contain H as an induced subgraph. In these two lectures we will study the structure of typical H-Free and *H-free* subgraphs.

An early result in this area was obtained by Erdos, Kleitman, and Rothschild over 35 years ago[ErKR76]. They characterized the structure of typical graphs without a cycle of length three precisely. Obviously every bipartite graph is triangle free. EKR showed the vertex set of almost all such C_3-Free(equivalently C_3-*free*) graphs is bipartite.

In the same vein, every $(t-1)$-partite graph is K_t-free, and in 1987 Kolatis, Prummel, and Rothschild (ref. missing sorry) showed that almost every K_t-Free (equivalently K_t-*free*) graph is $(t-1)$-partite.

Now, if H has chromatic number c then every $(c-1)$-partite graph is H-Free. In 1987 using the newly minted Regularity Lemma, Erdos, Frankl, and Rodl showed that for such H, almost every H-Free graph can be made $c-1$ partite by the deletion of $o(n^2)$ edges.

Promel and Steger[PrS91, PrS92a, PrS92b, PrS93] adapted the approach of EFR so that it could be applied to the study of *H-free* graphs. Using this approach, they showed that almost every C_4-*free* graph is a split graph, i.e. its vertex set can be partitioned into a clique and a stable set.They also showed that the vertex set of almost every C_5-*free* graph can be partitioned into either (i) a clique and the disjoint union of cliques or (ii) a stable set and a complete multipartite graph. In addition, again following EFR, they obtained, for all H, upper and lower bounds on the number of H-*free* graphs on n vertices whose exponents differed by a multiplicative factor of $1 + o(1)$.

Recently, Balogh and Butterfield[BB11] proved that for $k \geq 4$ almost every C_{2k+1}-*free* graph can be partitioned into k cliques while almost every C_7-*free* graph can be partitioned either into 3 cliques or a stable set and 2 cliques.

Reed and Scott (manuscript) have shown that almost every C_6-free graph is the disjoint union of a stable set and a graph containing no stable set of size three and no induced matching with two edges. They obtained similar results for longer even cycles, for cycles of length at least 12, the same results were obtained independently by Kim et al.[KKOT15].

Theorem 1 *For $l > 5$, almost every C_{2l}-free graph can be partitioned into $l - 2$ cliques and the complement of a graph which is the disjoint union of stars and C_3s.*

We discuss these and further results and conjectures along these lines.

By $Forb_H$ we mean the family of H-Free graphs, and by $Forb_H^n$ the family of H-Free graphs on the vertex set $\{1, ..., n\}$, which we denote V_n. By $IForb_H$ we mean the family of H-*free* graphs, and by $IForb_H^n$ the family of H-*free* graphs on the vertex set $\{1, ..., n\}$, which we denote V_n.

The *witnessing partition number* of H, denoted $wpn(H)$, is the maximum k such that for some pair (c_H, s_H) which sum to k, there is no partition of $V(H)$ into c_H cliques and s_H stable sets.

We define an *H-freeness witnessing partition* of a graph G to be a partition of $V(G)$ into sets $S_1, .., S_{wpn(H)}$ such that (i) for any partition of $V(H)$ into $X_1, ..., X_{wpn(H)}$ there is an i such that $H[X_i]$ is not an induced subgraph of $G[S_i]$, and (ii) for all i, $|S_i| - \frac{n}{wpn(H)} = o(n)$. (to make this definition precise we should be more precise about the $o(n)$ term, but we decline to do so).

The following conjecture made by Reed and Scott was disproven by Norine in 2018.

Conjecture 2 *For every H, falmost every graph G in $IForb_H^n$ has an H-freeness witnessing partition.*

The results discussed above suggest a conjecture along these lines. As we show now, they prove it whenever H is a cycle or the complement of a cycle. Reed and Yuditsky (manuscript,phd thesis) have verified the conjecture whenever H is a tree or the complement of the tree. It would be interesting to characterize those H for which it holds and to propose and prove weakenings of it.

A cycle of length three can be partitioned into a clique, and into three stable sets but not into two stable sets. Thus $wpn(C_3)$ is 2 and the set consisting of two copies of the family of all stable sets is a C_3-freeness witnessing set. Thus, the result of Erdos, Kleitman, and Rothschild mentioned above tells us that our conjecture holds when H is C_3.

For l at least four, the cycle of length $2l+1$ can be partitioned into three stable sets, into two stable sets and a clique, into a stable stable of size 3 and $l-1$ cliques, and into $l+1$ cliques. However, it cannot be partitioned into l cliques. So, $wpn(C_{2l+1})$ is l and the multiset consisting of l copies of the family of all cliques is a C_{2l+1}-freeness witnessing set. Thus, a result of Balogh and Butterfiled mentioned above tells us that our conjecture holds when H is C_{2l+1}.

The cycle of length 4 can be partitioned into two stable sets and into two cliques, However, it cannot be partitioned into a clique and a stable set. So, $wpn(C_4)$ is 2 and the family of all cliques together with the family of all stable sets form a C_4-freeness witnessing set. Thus a result of Promel and Steger mentioned above tells us that our conjecture holds when H is C_4.

The cycle of length 5 can be partitioned into three stable sets,into three cliques, into two stable sets and a clique, and into two cliques and a stable set, However, it cannot be partitioned into two stable sets. So, $wpn(C_5)$ is 2. Furthermore, C_5 cannot be partitioned into a clique and a graph containing no P_3. Thus the multiset consisting of the family of all cliques together with the family of all disjoint unions of cliques form a C_5-freeness witnessing set. Since C_5 is self-complementary so does the set consisting of the family of all stable sets and the family of all stable sets and complete multipartite graphs. Thus a result of Promel and Steger mentioned above tells us that our conjecture holds when H is C_5.

The cycle of length 7 can be partitioned into three stable sets,into four cliques, into a stable set and three cliques, and into a clique and two stables sets. However, it cannot be partitioned into three cliques or into two cliques and a stable set. So, $wpn(C_7)$ is 3 and a result of Balogh and Butterfield mentined above shows that our conjecture holds when H is C_7.

The cycle of length 6 can be partitioned into two stable sets,into three cliques, and into a stable set and two cliques. However, it cannot be partitioned into two cliques. So, $wpn(C_6)$ is 2. Furthermore, C_6 cannot be partitioned into a stable set and the complement of a graph of girth 5.Thus the results of Reed and Scott mentioned above shows that our conjecture holds when H is C_6.

For $l > 5$, the cycle of length $2l$ can be partitioned into two stable sets, into l cliques, and into a stable set of size 4 and $l-2$ cliques, However, it cannot be partitioned into $l-1$ cliques. So, $wpn(C_{2l})$ is $l-1$. Furthermore, C_{2l} cannot be partitioned into $l-2$ cliques and a graph which has at most 3 vertices or which has at least four vertices and either has a disconnected complement or contains C_3. So Theorem 2 above, shows that our conjecture holds when H is C_{2l}.

Thus, our conjecture holds whenever H is a cycle. Reed and Yuditsky have shown that it also holds whenever H is a tree, Since the complement of a hereditary family is a hereditary family and the $\overline{H}$-free graphs are precisely the complements of the H-free graphs, it also holds when H is the complement of a cycle or the complement of a tree.

2 A Correct Weakening of the Conjecture

We prove the conjecture is true, if we permit a small exceptional set Z and ask only for an H–freeness witnessing partition of G-Z. Our starting point is a recent beautiful paper of Alon, Balogh, Bollobas, and Morris[AlBBM11]. Their discussion is in terms of partitions such that no part contain a specific bipartite graph $U(k)$ as a (not neccesarily induced) subgraph. Their Corollary 8 (whose proof is only a few lines and will be discussed in the lecture) is

Corollary 3 *For every k there is a positive ϵ such that for every sufficiently large l, the number of graphs with l vertices which are $U(k)$-Free is $\leq 2^{l^{2-\epsilon}}$.*

The following is essentially an immediate corollary of their Theorem 1.

Corollary 4 *For every H, sufficiently large k and positive δ which is sufficiently small in terms of H and k, there are positive ϵ and b such that the following holds:*

For almost every H-free graph G on V_n there is a partition of V_n into $S_1, ...S_{wpn(H)}, A_1, ..., A_{wpn(H)}$ such that for some set B of at most b vertices the following holds:

(a) $G[S_i]$ is $U(k)$-free for every i between 1 and $wpn(H)$,

(b) $|A_1 \cup A_2 ... \cup A_{wpn(H)}| \leq n^{1-\epsilon}$, and

(c) for every vertex v of $S_i \cup A_i$ there is a vertex b of B such that $|(N(v) - N(b)) \cap (S_i \cup A_I)| + |(N(b) - N(v)) \cap (S_i \cup A_i)|$ is at most δn.

Proof. We only need to prove this result for δ sufficiently small as it then follows for all δ. We essentially follow the ABBM proof of their Theorem 1 where the hereditary family $\mathcal{P}$ is $Forb_H$ and α is $\frac{\delta}{3}$. We omit the details, simply sketching the very very minor modifications. We do not expect school participants to read or understand this.

We note that their $\chi_c(\mathcal{P})$ is exactly $wpn(H)$. We want to use the strengthening of their Lemma 23 obtained by replacing $\alpha = \alpha(k, \mathcal{P}) > 0$ in its statement with $\alpha > 0$ sufficiently small in terms of k. Their proof of the lemma actually proves this strengthening (without any modification whatsoever).

Now while following their (two paragraph) proof of their Theorem 1, we again replace $\alpha = \alpha(k, \mathcal{P})$ by $\alpha > 0$ sufficiently small in terms of k. Then we consider the adjustment $S'_1, ..., S'_r$ and exceptional set A they obtain and set $A_i = S'_i \cap A, S_i = S'_i - A$. Now, in their proof, they consider a maximal 2α bad set B. They implicitly use the fact, which is a consequence of their Lemmas 17 and 18, that for almost every graph in $Forb^n_H$, the size of B is at most some constant c. We set b to be this c. Now, (a) is their Theorem 1(b), (b) is their Theorem 1 (a), and (c) follows immediately from the fact that $S'_1, S'_{wpn(H)}$ is an α-adustment and the definition of γ-adjustment. $\square$

We show momentarily that we can strengthen this theorem in two ways. In our strengthening we use X_i in place of S_i and Z_i in place of A_i to avoid confusion. We will see that we can insist that every partition element has approximately the same size. Secondly we can insist that $X_1, ..., X_n$ is an H-freeness witnessing partition of $G - Z$.

Theorem 5 *For every H, and $\delta > 0$ sufficiently small in terms of H, there are $\gamma, b > 0$ such that the following holds:*

For almost every H-free graph G on V_n there is a partition of V_n into $X_1, ...X_{wpn(H)}, Z_1, ..., Z_{wpn(H)}$ such that for some set B of at most b vertices the following hold:

(I) $X_1, ..., X_{wpn(H)}$ is an H-freeness witnessing partition of $G - Z_1 - Z_2 - ..., -Z_{wpn(H)}$,

(II) $|Z_1 \cup Z_2 ... \cup Z_{wpn(H)}| \leq n^{1-\gamma}$, and

(III) for every vertex v of $X_i \cup Z_i$ there is a vertex b of B such that $|(N(v) - N(b)) \cap (X_i \cup Z_I)| + |(N(b) - N(v)) \cap (X_i \cup Z_i)|$ is at most δn.

(IV) For every i, we have that $|Z_i \cup X_i| - \frac{n}{wpn(H)} \leq n^{1-\frac{\gamma}{2}}$.

Proof. Most particpants can skip the proof.

We choose k sufficiently large and then $\delta < \frac{1}{10wpn(H)}$ sufficiently small in terms of H and k. We choose $\epsilon, b > 0$ such that Corollary 4 holds for this choice of k and δ and set $\gamma = \frac{\epsilon}{10}$. We consider n large enough to saisfy certain implicit inequalities below. We know that for almost every graph in $Forb_H^n$ there is a set B of at most b vertices and a partition into S_i and A_i satisfying (a),(b), and (c) set out in that corollary. Since S_i is $U(k)$-free and n is large, Corollary 8 of [AlBBM11] tells us that there are only $2^{n^{2-\epsilon}}$ choices for $G[S_i]$, . The number of choices for the edges out of each vertex of A_i is 2^{n-1}. So, since $|A_i|$ has size at most $n^{1-\epsilon}$, we know there are at most $2^{n^{2-\epsilon}}$ choices for the edges out of A_i. It follows that there are at most $2^{O(n^{2-\epsilon})}$ choices for the partition $S_1, ..., S_{wpn(H)}, A_1, .., A_{wpn(H)}$ and the graphs $G[S_1 \cup A_1], ..., G[S_{wpn(H)} \cup A_{wpn(H)}]$. This implies that we can actually show that almost every graph in $Forb_H^n$ has a partition satisfying (a),(b),(c) such that in addition $|S_i \cup A_i| - \frac{n}{wpn(H)}$ is at most $n^{1-\gamma}$. We will show that for almost every graph in $Forb_H^n$ with such a partition, we can obtain a partition satisfying (I),(II),(III), and (IV).

For each i, we let J_i consist of those L which are induced subgraphs of H such that there do not exist $n^{1-\frac{\epsilon}{2}}$ disjoint sets of vertices in S_i which induce L. For each L in J_i we choose a maximal family of disjoint induced copies of L in S_i. We let Z_i be the union of A_i and all the vertices in all these copies for all graphs in J_i. We let $X_i = S_i - Z_i$. We note that for some constant C_H which depends on H, Z_i has at most $C_H n^{1-\frac{\epsilon}{2}}$ vertices. So, for large n, (II) holds.

Now, $S_i \cup A_i = X_i \cup Z_i$. So we have that (IV) holds, and since (c) holds, so does (III).

To complete the theorem we need only show that the proportion of graphs with the property that $X_1, ..., X_{wpn(H)}$ is not an H-witnessing partition of

$G - Z_! - ... - Z_{wpn(H)}$ is $o(|Forb_H^n|)$.In order to do so, we sum over all possible choices for (i) the original S_i and A_i satisfying (a),(b),(c), (ii) the new partition X_i, Z_i of $S_i \cup A_i$ for each i, and (ii) the subgraphs $G[S_i \cup A_i]$ which have the property of interest.

Clearly the number of choices for the partitions of V_n is at most $(4wpn(H))^n$. For each partition of interest,since (a) states that $G[S_i]$ contains no $U(k)$, Corollary 8 of the ABBM paper tells us that the number of choices for $G[S_i]$ is at most $n^{2-\epsilon}$. Now, for each vertex v of A_i there are fewer than 2^n choices for $N(v) \cap (S_i \cup A_i)$. So, the number of choices for $G[S_i \cup A_i]$ is at most $2^{n^{2-\epsilon}} 2^{|A_i|n}$. Applying (b) we obtain that the number of choices for $(G[S_1 \cup A_1], ..., G[S_{wpn(H)} \cup A_{wpn(H)}])$ is at most $2^{(wpn(H)+1)n^{2-\epsilon}}$.

We now count, for a specific choice of partition and $(G[S_1 \cup A_1], ..., G[S_{wpn(H)} \cup A_{wpn(H)}])$, the number of G corresponding to this choice for which $X_1, ..., X_{wpn(H)}$ is not an H-freeness certifying partition of $G - Z_! - Z_2 - ... - Z_{wpn(H)}$. This implies there is a partition $Y_1, ..., Y_{wpn(H)}$ of $V(H)$ such that $H[Y_i]$ is a subgraph of $G[X_i]$ and hence not in J_i. Thus, for each i, we can find $n^{1-\frac{\epsilon}{2}}$ disjoint sets of vertices in S_i which induce $H[Y_i]$. We consider the complete $wpn(H)$ partitite graph whose vertices are these sets and where two sets are joined if they are in different S_i A standard argument (reference to be filled in) tells us that we can find $\frac{(n^{1-\frac{\epsilon}{2}})^2}{100}$ edge disjoint cliques in this graph. For each such clique, there is one choice of edges between the sets corresponding to its vertices such that if G makes this choice than G contains H as an induced subgraph. It follows that for some γ_H which depends on H, the number of G corresponding to our choice of partition and $(G[S_1 \cup A_1], ..., G[S_{wpn(H)} \cup A_{wpn(H)}])$ is at most $2^{(1-\frac{1}{r})\binom{n}{2}} 2^{-\gamma_H \frac{(n^{1-\frac{\epsilon}{2}})^2}{100}}$. Summing over all choices of partitions and subgraphs they induce the desired result follows

□

It is natural to ask for each H, for the slowest growing function f_H which can be used to bound the exceptional set Z. In particular we would like to know when $f_H = 0$, i.e. for which H the Reed-Scott conjecture holds.

3 The number of H-*free* and H-Free graphs.

If H is c-chromatic then by considering a fixed partition of $V_n = 1, ..., n$ into $c-1$ parts each of size $\lceil \frac{n}{c-1} \rceil$ or $\lfloor \frac{n}{c-1} \rfloor$ there are more than $2^{(1-\frac{1}{c-1})\binom{n}{2}}$ $c-1$-partite H-Free graphs. On the other hand, by the result of EFR, for almost every H-free graph G there is a set F of $o(n^2)$ edges of G such that $G-F$ is $c-1$-partite. There are $\binom{n^2}{|F|} = 2^{o(n^2)}$ choices for F, $(c-1)^n = 2^{o(n^2)}$ choices for the partition of G-F, so it follows that the number of H-Free graphs is $2^{(1+\frac{1}{c-1}+o(1))\binom{n}{2}}$.

By considering a fixed a partition of V_n into $wpn(H)$ parts each of size $\lceil \frac{n}{t} \rceil$ or $\lfloor \frac{n}{t} \rfloor$ we see there are more than $2^{(1-\frac{1}{wpn(H)})\binom{n}{2}}$ *H-free* graphs which can be partitioned into c_H cliques and s_H stable sets. Applying the results of ABBM, we see that almost every *H-free* graph permits a partition into a set Z of size $o(n)$, and $wpn(H)$ sets each of which is $U(k)$-Free. Since the number of $U(k)$-Free graphs on n vertices is $2^{o(n^2)}$, it follows that the number of *H-free* graphs is $2^{(1+\frac{1}{wpn(H)}+o(1))\binom{n}{2}}$.

4 Some Remarks on the structure of H-Free graphs

The natural analog to the Reed-Scott Conjecture would be the following. As discussed below it is known to be false.

Conjecture 6 *For every H, almost every graph G in $Forb_H^n$ can be partitioned into $X_1, ..., X_{\chi(G)-1}$ such that for any partition of $V(H)$ into $Y_1, ..., Y_{\chi(G)-1}$ there is an i such that $H[Y_i]$ is not a subgraph of $G[X_i]$.*

Again a natural weakening of this conjecture would be to ask for the slowest growing function f_H, bounding the size of an exceptional set whose deletion leaves a graph with the desired partition. However, work here seems to have centered around obtaining partitions which do not guarantee the H-Freeness of the partitioned graph. Specifically we let $M(H)$ consist of those graphs F such that H is a subgraph of the join of (i) the union of F and a stable set with (ii) a $\chi(G)-2$- partite graph.

Ballogh, Bollobas, and Simonovits[BaBS09]have shown that for every H there is a b_H such that almost every H-Free graph can be partitioned into $\chi(H)-1$ parts $X_1, ..., X_{wpn(H)}$ such that for every i $G[X_i]$ contains no subgraph in $M(H)$.

They also gave an example (Example 13 in [BaBS11] which shows that this constant cannot always be zero, and also shows that Conjecture 6 is false.

In this example, H is obtained from a $K_{2s,2s}$ by adding a matching of size s on the vertices on one side of the bipartition. Now $\chi(H) = 3$, and if we partition G into X_1 and X_2 such that neither $G[X_1]$ nor $G[X_2]$ contains an element of $M(H)$ then both $G[X_1]$ and $G[X_2]$ contain no matching of size s. Thus each $G[X_i]$ has a vertex cover of size at most $2s-2$ and at most $s-1$ vertices of degree greater than $2s$. We can choose such a graph by choosing the $2s-1$ vertices in a cover in order so their degrees are nonincreasing then choosing the neighbourhood of each vertex of the cover. We have at most $k^{2s-2}2^{(s-1)k}(k^{2s})^{s+1} = 2^{(s-1+o(1))k}$ choices for such a graph with k vertices. So, the total number of graphs on n vertices which have such a partition is at most $2^n 2^{(s-1+o(1))n+\frac{n^2}{4}}$.

On the other hand choosing s large and $\lambda = 2s + \lceil s^{1/4} \rceil$, we can partition G into a stable set S of size λ, a stable set U of size $\lceil \frac{n-\lambda}{2} \rceil$ and a stable set W of size $\lfloor \frac{n-\lambda}{2} \rfloor$ and then choose any set of edges between U and W, and for each vertex v in $V-S$ any choice of at most $s-1$ vertices of S as neighbours. If such a graph were to contain a copy of H then every triangle of H has to intersect each of S, U and W. This means that each of S, U, W contains at least s vertices of the induced copy of H all lying in the same side of the oriiginal bipartition of H. But then some vertex of V-S sees s vertices of S, a contradiction.

Now, for a fixed partition into S, U, W each vertex of $V-S$ has $\sum_{i=0}^{s-1} \binom{\lambda}{s} > 2^{\lambda-2}$ choices for its neighbourhood in S. so there are at least $2^{|U||W|+(\lambda-2)n-n} > 2^{\frac{n^2}{4}+\frac{(\lambda-6)n}{2}} > 2^{\frac{n^2}{4}+(s+6)n}$. This is far greater than those permitting a partition as discussed in the last paragraph.

it would be interesting to explore the situation here further. I warn the reader that I am not as on top of the developments with respect to H-Free graphs as I am those concerning *H-free* graphs and there may well be relevant further results.

5 An Overview of Lecture 2

In the second lecture we will sketch the proof of a weakening of Theorem 5, to illustrate the crucial role that Szemeredi's Regularity Lemma plays in this domain. We will then sketch how we can use Theorem 5 to almost prove the Reed-Scott Conjecture for C_4. Specifically we sketch how to replace the bound on the size of the exceptional set by a constant. In the lecture we discussed how to complete the proof of the conjecture for C_4 but we omit those details here.

Our weakening of Theorem 5, relaxes the upper bound on the size of the exceptional set Z to $o(n)$ and removes any mention of the set B. I.e. we show:

Theorem 7 *For every H and $\delta > 0$, the following holds:*

For almost every H-free graph G on V_n there is a partition of V_n into $X_1, ...X_{wpn(H)}, Z_1, ..., Z_{wpn(H)}$ such that the following hold:

(I) $X_1, ..., X_{wpn(H)}$ is an H-freeness witnessing partition of $G - Z_1 - Z_2 - ..., -Z_{wpn(H)}$,

(II) $|Z_1 \cup Z_2 ... \cup Z_{wpn(H)}| \leq \delta n$, and

(III) For every i, we have that $|Z_i \cup X_i| - \frac{n}{wpn(H)} \leq \delta n$.

5.1 Some Prelimanaires

We begin with the following which is a strengthening of Exercise 15 from the background notes and can be proved using the answers to Exercises 13 and 14:

Theorem 8 *For all H and $1 > \delta > 0$ there is an $\epsilon = \epsilon_{H,\delta} < \frac{\delta}{2}$ and $n_{H,\delta}$ such that if G contains sets $\{P_v, |v \in H\}$ each of size at least $n_{H,\delta}$ such that every pair (P_u, P_v) is ϵ-regular and has density (i) at least δ if uv if as edge, and (ii) at most $1 - \delta$ otherwise, then there is an induced copy of H in G where the copy v' of v is in P_v.*

Proof. The proof is by induction on $|V(H)|$. For some v in H, we choose a vertex v' of P_v to be the copy of v. We insist that v' has at least $\frac{\delta|P_u|}{2}$ neighbours in P_u if $uv \in N(v)$ and at least $\frac{\delta|P_u|}{2}$ nonneighbours in P_u otherwise. Since $\epsilon < \frac{\delta}{2}$,by our hypotheses on the density of the pairs and Exercise 14, we can do so provided $\epsilon < \frac{1}{2|V(H)|}$. We let P'_u be $P(u) \cap N(v)$ if $uv \in H$ and $P'_u = P_u - N(v)$ otherwise. We now apply induction to $H - v$ and $\{P'_u | u \in V(H) - v\}$. For this technique to work we need to set $\epsilon_{H,\Delta}$ to be $\frac{\delta}{2} min\{\epsilon_{H-v,\frac{\delta}{2}} | v \in V(H)\}$ and $n_{H,\delta} = \frac{2}{\delta} max(n_{H-v,\frac{\delta}{2}} | v \in V(H))$. □

We note also that we can strengthen Exercise 13 as follows, we omit the details

Theorem 9 *For all epsilon* $< \frac{1}{2}$, *if we are given a set of* $wpn(H)+1$ *disjoint subsets* $P_1, .., P_{wpn(H)}$ *of* G, *every two of which are* ϵ^2*-regular, then for any family of equipartitions of the* P_i *which splits each into at most* $\frac{1}{\epsilon}$ *subparts, the following holds. If the pair of parts* P *and* Q *had density* d *then for any subpart* P' *of* P *and subpart* Q' *of* Q, (P', Q') *is* ϵ*-regular with density between* $d - \epsilon$ *and* $d + \epsilon$.

5.2 The Under Card

The key to the proof of Theorem 7 is:

Theorem 10 *For all* H *and* $1 > \delta > 0$ *there is an* $\epsilon = \epsilon_{H,\delta} < \frac{\delta}{2}$ *and* $n_{H,\delta}$ *such that if* G *contains sets* $\{P_1, ..., P_{wpn(H)+1}\}$ *each of size at least* $n_{H,\delta}$ *such that every pair* (P_u, P_v) *is* ϵ*-regular and has density between* δ *and* $1 - \delta$ *then there is an induced copy of* H *in* G.

Proof. To prove Theorem 10, we apply Theorem 8 to a family of subpartitions of the partition given by the hypothesis of Theorem 10. The subpartition of a part P_i will be an ϵ'-regular partition of $G[P_i]$ for some small ϵ' which is still much bigger than ϵ.We will specify our requirements as to ϵ' and the minimum number m of parts we want the subpartitions to have (we will need this large in terms of $|V(H)|$). This determines the maximum number of parts M we may require. Having done so we can choose $\epsilon < \frac{1}{M}$ to ensure we obtain a partition into at most $\frac{1}{\epsilon}$ parts.

Now, in any subpartition of the part, we have a 3-colouring of the "edges" between the parts, where grey means "not regular", blue means regular with density $< \frac{1}{2}$ and red means regular with density at least $\frac{1}{2}$. Since only an ϵ' proportion of the edges are grey, by Ramsey theory if we choose m large enough in terms of H we can find in each subpartition a clique K_i of subparts of size $|V(H)|$ whose edges are either all red or all blue. If we find the latter, we say P_i is a clique part, otherwise it is a stable set part.

We count up the number c of clique parts and the number $s = wpn(H) + 1 - c$ of stable set parts. We find a partition of $V(H)$ into c cliques and s stable sets (which must exist because of the definition of wpn(H)) and find a bijection between these sets and the parts, where if P_i is a clique part it corresponds to some clique C_i and otherwise it corresponds to some stable set S_i. We call C_i or S_i, T_i so we can treat the cases together. We index $|T_i|$ parts of K_i with the elements of T_i and now apply Theorem 8 to prove Theorem 10. $\square$

Now, for a $\delta > 0$, $\epsilon < \delta$ and some constant number p of parts, we want to count the number of graphs permitting an ϵ-regular equipartitions into $p > \frac{1}{\epsilon}$ parts, where a proportion g of the pairs of parts are well-behaved, in that they are ϵ-regular with density between 0 and 1. Having picked one of the at most p^N equipartitions, and specified which of the pairs of parts are well-behaved,there are at most $2^{g\binom{n}{2}}$ choices for pairs of edges joining these well behaved parts. For some α going to zero with δ there are at most $2^{\alpha\binom{n}{2}}$ choices for (i) the at most $2\epsilon\binom{n}{2}$ edges within parts or between irregular parts, (ii) the at most $\delta\binom{n}{2}$ edges between low-density part pairs, and (iii) the at most $\delta\binom{n}{2}$ non-edges between high-density pairs.

It follows that for every $\beta > 0$ there is a δ_β such that for all $\delta < \delta_\beta$ and ϵ-regular partition for an ϵ sufficiently small in terms of δ, the proportion of pairs of parts which are well-behaved must exceed $(1 - \frac{1}{wpn(H)}) - \beta$.

Now, we combine Theorem 10 with this fact and an application of the Erdos-Stone Stability Theorem to an auxiliary graph whose vertices are the parts and where two vertices are joined by an edge if the corresponding pair of parts is well-behaved. . We obtain that almost every H-*free* graph permits a Szemeredi partition where this auxiliary graph can be made $wpn(H)$-partitite by deleting a $o(1)$ proportion of its edges. We let $X_1, .., X_{wpn(H)}$ be the partition of G corresponding to this partition of the parts graph. For each such partition we consider the possible sets $\{F_1, ..., F_{wpn(H)}\}$ of graphs where $F_i = G[X_i]$ for some H-*free* G for which our partition of the parts graph yield this partition of $V(G)$. We note that since almost all the edges of the auxiliary graph go between the parts there are $2^{o(n^2)}$ choices of such a partition and corresponding set of subgraphs induced by the parts.

This implies that almost every H-*free* graph corresponds to a partition $X_1, ..., X_{wpn(H)}$ and corresponding family of graph $F_1, ..., F_{wpn(H)}$ where there are $2^{(1-\frac{1}{wpn(H)}+o(1))\binom{n}{2}}$ H-free graphs for which for all i, $G[X_i] = F_i$. Now, for any $Y_i, ..Y_{wpn(H)}$ such that Y_i is a subgraph of F_i and there is a partition of H into $D_1, ..., D_{wpn(H)}$ such that $Y_i = H[D_i]$ there is a choice of the edges between the Y_i which yield an induced copy of H in D. Thus, if there is a family of k such sets $\{Y_1, ..., Y_{wpn(H)}\}$ for which no edge joins two parts Y_i and Y_j of distinct sets of the family, then the total number of choices of edges between the parts which yields an H-free graph is at most $(1 - \frac{1}{2^{\binom{|V(H)|}{2}}})^k$.

This allows us to show that we can delete a set Z of $o(n)$ vertices, such that $X_1 - Z, ..., X_{wpn(H)} - Z$ is an H-witnessing partition of $G - Z$. We omit the details.

5.3 The Main Event

We now apply Theorem 5 to prove that there is a constant c such that almost every C_4-free graph contains a set Z of at most c vertices such that $G-Z$ is split.

To begin we bound the number of split graphs on V_n from below. For every partition (A,B) of V_n, there are $2^{|A||B|}$ split graphs which yield this partition. This tells us that there are $\Omega(\frac{2^{n+\lceil\frac{n}{2}\rceil\lfloor\frac{n}{2}\rfloor}}{\sqrt{n}})$ partitions of a split graph on V_n into a clique and a stable set whose size differs by at most 1. Now, given one partition of a graph G into a clique C and a stable set S, any other clique of G contains at most one vertex of S and any other stable set of G contains at most one vertex of C. So, a split graph on V_n has fewer than n^2 partitions and there are $\Omega(\frac{2^{n+\lceil\frac{n}{2}\rceil\lfloor\frac{n}{2}\rfloor}}{n^{3/2}})$ split graphs on V_n.

Theorem 5 tells us that for every $\delta > 0$ there is an α and a b such that almost every C_4-*free* graph on V_n has a partition into a clique X_1, a stable set X_2 both of size at least $\frac{n}{2}-n^{1-\frac{\alpha}{2}}$, Z_1 and Z_2 with $|Z_1|+|Z_2|<n^{1-\alpha}$ such that for some set B of at most b vertices, we have $\forall v \in Y_i = X_i \cup Z_i, \exists w \in B$ such that $|((N(v)-N(w)) \cup (N(w)-N(v))) \cap Y_i| \leq \delta n$.

We are now going to strenghten this corollary of Theorem 5, by showing that we can assume $|Z| = O(1)$. In order to do do so we let M be a maximum matching in $G[Y_2]$, and N be a maximum matching in $\overline{G[Y_1]}$. We let K be the clique $Y_1 - V(N)$ and S be the stable set $Y_2 - V(M)$. Since X_1 is a clique $|N| < |Y_1 - X_1| = |Z_1|$. In the same vein $|M| < |Z_2|$. Hence for sufficiently large n, $|K|, |S| > \frac{n}{3}$.

Now we can specify the edges within the Y_i by specifying (i) the choices for the neighbourhoods of the vertices of B on each side, (ii) the choices for $V(M) \cup V(N)$, (iii) for each v in M a choice of a neighbourhood of a vertex of B on Y_2 whose symmetric difference with the neighbourhood of v on Y_2 has size at most δn and this symmetric difference, and (iv) for each v in N a choice of a neighbourhood of a vertex of B on Y_1 whose symmetric difference with the neighbourhood of v on Y_1 has size at most δn and this symmetric difference. So the number of choices for these edges is at most:

$$2^{bn} n^{2|M|+2|N|} (b2^{\delta n})^{2|M|+2|N|}.$$

On the other hand we see that for every edge e of M and edge f of K there is a choice of the edges from e to f which creates a C_4. In the same vein, for every non-edge x of N and pair y of vertices of S if we add all the edges from x to y then we create a C_4. Hence partitioning K into $\frac{|K|}{2}$ disjoint edges and S into $\frac{|S|}{2}$ pairs of nonadjacent vertices we see that, for sufficiently large k, the number of choices for the edges between Y_1 and Y_2 is at most $2^{|Y_1||Y_2|}(\frac{15}{16})^{\frac{|M||K|+|N||S|}{2}} < 2^{\frac{n^2}{4}}(\frac{15}{16})^{\frac{(|M|+|N|)n}{6}}$. We see that we can strengthen Theorem 5, by replacing, for sufficiently small δ, our bound on the size of Z by some constant $c > b$ depending on δ.

I

References

[Al93] V. E. Alekseev. On the entropy values of hereditary classes of graphs, *Discrete Math. Appl.* **3** (1993), 191–199.

[AlBBM11] N. Alon, J. Balogh, B. Bollobás, and R. Morris. The structure of almost all graphs in a hereditary property, *J. Combin. Theory Ser. B* **101(2)** (2011), 85–110.

[BaBS04] J. Balogh, B. Bollobás, and M. Simonovits. On the number of graphs without forbidden subgraph, *J. Combin. Theory Ser. B* **91** (2004), 1–24.

[BaBS09] J. Balogh, B. Bollobás, and M. Simonovits. The typical structure of graphs without given excluded subgraphs, *Random Structures Algorithms* **34** (2009), 305–318.

[BaBS11] J. Balogh, B. Bollobás, and M. Simonovits. The fine structure of octahedron-free graphs, *J. Combin. Theory Ser. B* **101(2)** (2011), 67–84.

[BB11] J. Balogh and J. Butterfield. Excluding induced subgraphs: critical graphs, *Random Structures and Algorithms* **38** (2011), 100–120.

[BoT97] B. Bollobás and A. Thomason. Hereditary and monotone properties of graphs, *The Mathematics of Paul Erdős, Vol. II, R. L. Graham and J. Nešetřil (Eds.)* **14** (1997), 70–78.

[ErFR86] P. Erdős, P. Frankl, and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs Combin.* **2** (1986), 113–121.

[ErKR76] P. Erdős, D. J. Kleitman, B. L. Rothschild. Asymptotic enumeration of K_n-free graphs, *International Colloquium on Combinatorial Theory, Atti dei Convegni Lincei* **17** (1976), 19–27.

[KKOT15] J. Kim, D. Kuhn, D. Osthus, T. Townsend. Forbidding induced even cycles in a graph: typical structure and counting, http://arxiv.org/abs/1507.04944 (2015).

[PrS91] H. J. Prömel and A. Steger. Excluding Induced Subgraphs I: Quadrilaterals, *Random Structures and Algorithms.* **2** (1991), 53–79.

[PrS92a] H. J. Prömel and A. Steger. Almost all Berge graphs are perfect, *Combin., Probab. & Comp.* **1** (1992), 53–79.

[PrS92b] H. J. Prömel and A. Steger. Excluding induced subgraphs. III. A general asymptotic, *Random Structures Algorithms 3(1)* (1992), 19–31.

[PrS93] H. J. Prömel and A. Steger. Excluding induced subgraphs II: Extremal graphs, *Discrete Appl. Math.* **44** (1993) 283–294.

[ReSc17] B. Reed and A. Scott. The typical structure of an H-free graph when H is a cycle, *manuscript.*

[RY17] B. Reed and Y. Yuditsky. The typical structure of H-free graphs for H a tree, *manuscript.*