Excluding induced subgraphs: critical graphs

József Balogh[*], Jane Butterfield[†]

September 30, 2010

Abstract

Determining the cardinality and describing the structure of H-free graphs is well-investigated for many graphs H. In the nineties, Prömel and Steger proved that for a graph H with chromatic number $k+1$ almost all graphs not containing H as a subgraph are k-colorable if and only if H contains a color-critical edge. We strengthen the concept of H-free to induced subgraph containment, proving that if H has coloring number $k+1$ then almost all H-free graphs can be covered by k graphs that are cliques or independent sets if and only if H is in some well-defined sense critical. The family of critical graphs includes C_4 and C_{2k+1} for all $k \geq 3$.

Keywords: extremal graphs, structure of H-free graphs, graph counting, critical graphs.

Primary MSC: 05C35 Secondary: 05C30, 05D40, 05C75

1 Introduction

1.1 Definitions and motivation

The *complete graph* on n vertices, denoted K_n, is the graph containing every possible edge. If each vertex of K_n is blown up to an independent set of size k, we call that the *complete n-partite graph*, denoted $K_n(k)$. Let $T_{m,k}$ denote the **Turán graph**, formed by partitioning m vertices into k subsets whose sizes are within one of each other and then connecting two vertices by an edge exactly when they are in different subsets; notice that $T_{kn,n} = K_n(k)$. We will denote by K_n^- the graph on n vertices missing exactly one edge. Call a graph on n vertices **non-trivial** if it is neither K_n or its complement. Let $R(G,H)$ denote the **Ramsey**

[*]Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA; and Department of Mathematics, University of Illinois, 1409 W Green Street, Urbana, IL 61801, USA. E-mail address: jobal@math.uiuc.edu. This material is based upon work supported by NSF CAREER Grant DMS-0745185 and DMS-0600303, UIUC Campus Research Board Grants 09072 and 08086, and OTKA Grant K76099.

[†]Department of Mathematics, University of Illinois, 1409 W Green Street, Urbana, IL 61801, USA. E-mail address: jbutter2@uiuc.edu. The author acknowledges support from NSF grant DMS 08-38434 "EMSW21-MCTP: Research Experience for Graduate Students" and the Dr. Lois M. Lackner Mathematics Fellowship.

number, which is the minimum n such that any red-blue coloring of the edges of K_n results in either a red copy of G or a blue copy of H. If $G = H$ we write $R(G)$ instead of $R(G, G)$, and if $G = H = K_m$ we write $R(m)$ instead of $R(K_m)$. If $V_1, V_2 \subseteq V(G)$ let $e(V_1, V_2)$ denote the number of edges in G between V_1 and V_2.

Definition 1.1 ([4])**.** The **binary entropy function** is $H(x) = -x \log(x) - (1-x)\log(1-x)$, where $\log(x)$ denotes $\log_2(x)$. It satisfies the following property: for x sufficiently small,

$$\sum_{i \le xn} \binom{n}{i} \le 2^{H(x)n}. \tag{1}$$

Lemma 1.2 (Chernoff bound [4])**.** *For $X_1, X_2, \ldots, X_\ell \in \{0,1\}$ with $Pr(X_i = 1) = p$,*

$$Pr\left(\sum_{i=1}^{\ell} X_i < p\ell/2\right) < \exp(-p\ell/8).$$

To avoid unnecessary technicalities, we will omit ceiling and floor functions whenever they are not crucial.

The structure of graphs not containing a fixed graph as a subgraph has been extensively investigated. Erdős, Kleitman, and Rothschild [10] and Kolaitis, Prömel and Rothschild [12] studied K_n-free graphs. More generally, much work has been done to characterize for a fixed graph H the structure of graphs that do not contain H as a subgraph [15] or as an *induced* subgraph [6], [8], and [16]. Prömel and Steger [17], Alekseev [1], and Bollobás and Thomason [9] defined the following parameter, known as the 'coloring number' of a graph, which will be important in what follows.

Definition 1.3. If G is a graph on n vertices and $H_1, H_2, \ldots, H_k$ are induced subgraphs of G such that every vertex of G appears in exactly one H_i then we say that $\{H_1, H_2, \ldots, H_k\}$ **covers** G.

For a graph G, let the **coloring number** of G, denoted by $\chi_c(G)$, be the minimum k such that for any pair of non-negative integers (s,t) with $s + t = k$ there exists a coloring $f : V(G) \to [k]$ for which the following holds. Let $V_i := \{v | f(v) = i\}$; then $G[V_i]$ is a clique for $1 \le i \le s$ and $G[V_i]$ is an independent set for $s+1 \le i \le s+t$. Call any pair (s,t) whose sum is $\chi_c(G) - 1$ and for which no such coloring exists a **witnessing pair**. Call any pair (s,t) such that G can be covered by s cliques and t independent sets a **covering pair**. If $\mathcal{H}$ is a family of graphs, then we define $\chi_c(\mathcal{H})$ to be the maximum k such that for any pair of non-negative integers (s,t) with $s + t = k - 1$ there exists an $H \in \mathcal{H}$ such that (s,t) is not a covering pair for H.

For example, C_7 cannot be covered by either three cliques or by two cliques and one independent set, so $\chi_c(C_7) > 3$. On the other hand, each of $(0,4), (1,3), (2,2), (3,1)$, and $(4,0)$ is a covering pair, so $\chi_c(C_7) = 4$. Because $(1,2)$ and $(0,3)$ are covering pairs, the only witnessing pairs are $(3,0)$ and $(2,1)$. In general, $\chi_c(C_{2k+1}) = k + 1$ for $k \ge 1$.

Definition 1.4. For two families of graphs on n vertices, $\mathcal{A}(n)$ and $\mathcal{B}(n)$, such that $\mathcal{B}(n) \subseteq \mathcal{A}(n)$, we say that **almost all** $G \in \mathcal{A}(n)$ are in $\mathcal{B}(n)$ if

$$\lim_{n\to\infty} \frac{|\mathcal{A}(n)|}{|\mathcal{B}(n)|} = 1.$$

Definition 1.5. For a family $\mathcal{H}$ of graphs with $\chi_c(\mathcal{H}) = k+1$, let $\mathcal{P}(n, \mathcal{H})$ denote the family of graphs G on n vertices such that for every $H \in \mathcal{H}$, G is H-free. Let $\mathcal{Q}(n, \mathcal{H})$ denote the family of graphs G on n vertices that can be covered by s cliques and t independent sets for some $s+t = k$ such that (s,t) is a witnessing pair for every $H \in \mathcal{H}$. If $\mathcal{H}$ consists of a single graph H, we use the notation $\mathcal{P}(n, H)$ and $\mathcal{Q}(n, H)$.

It is obvious that if $\chi_c(H) = k+1$ and (s,t) is a witnessing pair for H, then any graph G that can be covered by s cliques and t independent sets is H-free. Therefore $\mathcal{Q}(n, H) \subset \mathcal{P}(n, H)$. One might ask the following natural question: for what graphs H is almost every H-free graph in $\mathcal{Q}(n, H)$? A similar question was answered in the non-induced case by Prömel and Steger in [15]:

Theorem 1.6. *Suppose $\chi(H) = k+1$. Then almost every graph G that does not contain H as a (not necessarily induced) subgraph is k-colorable $\Leftrightarrow$ there exists an edge e such that $\chi(H - e) = k$.*

In other words, H is in some sense "critical;" such a graph H is sometimes called "weakly-edge-color-critical." The aim of our paper is to extend Theorem 1.6 to induced containment. An analogous notion of critical for induced containment is the following.

Definition 1.7. Let $\mathcal{F}(H, s, t)$ denote the set of minimal (by induced containment) graphs F such that H can be covered by s cliques, t independent sets, and F. Call a graph G **critical** if $\chi_c(G) = k+1$ and for all $s+t = k-1$ and large enough n,

$$|\mathcal{P}(n, \mathcal{F}(H, s, t))| \leq 2.$$

Example: C_{2k+1} is critical for every $k \geq 3$. Recall that $\chi_c(C_{2k+1}) = k+1$. It is easy to see that $\mathcal{F}(C_7, 2, 0)$ consists of K_3^- and $\overline{K_3^-}$, $\mathcal{F}(C_7, 1, 1) = \{\overline{K_2}\}$, and $\mathcal{F}(C_7, 0, 2) = \{K_1\}$. For each $s+t = 2$, therefore, $\mathcal{P}(n, \mathcal{F}(C_7, s, t)) \subseteq \{K_n, \overline{K_n}\}$, implying that C_7 is critical. For $k > 3$, we find that $\mathcal{F}(C_{2k+1}, k-1, 0) = \{K_3^-, \overline{K_3^-}, \overline{K_3}\}$, which means that $\mathcal{P}(n, \mathcal{F}(C_{2k+1}, k-1, 0)) = \{K_n\}$; because $\mathcal{F}(C_{2k+1}, k-2, 1) = \{K_1\}$ and for every other $s+t = k-1$ the pair (s,t) is a covering pair this means that C_{2k+1} is critical.

Let G be the graph formed by adding two pendant edges to K_3, called the "bull graph." It is easy to check that $\chi_c(G) = 3$, with witnessing pairs $(2,0)$ and $(0,2)$. We need only check $\mathcal{F}(G, s, t)$ for $s+t = 1$. Because $\mathcal{F}(G, 1, 0)$ contains $\overline{K_2}$ (take the triangle to be the clique) the only graph in $\mathcal{P}(n, \mathcal{F}(G, 1, 0))$ is K_n. Because $\mathcal{F}(G, 0, 1)$ contains K_2 (take the two leaves and the degree two vertex to be the independent set) the only graph in $\mathcal{P}(n, \mathcal{F}(G, 0, 1))$ is $\overline{K_n}$.

It may seem like the following is a natural definition for a "critical" graph. For each $u, v \in V(H)$ let $H(u,v)$ be the graph formed from H by removing the edge uv if $uv \in E(H)$ and by adding the edge uv if $uv \notin E(H)$. Then we might consider H critical if $\chi_c(H(u,v)) < \chi_c(H)$ for some $u, v \in V(H)$. This is not an equivalent definition; it is easy to check that K_4^- is critical by Definition 1.7 and is not critical by this definition.

The following lemma is crucial in the proof of our main result.

Lemma 1.8. *If H is a critical graph then for any $s + t = k$ such that (s,t) is a witnessing pair of H:*

1. *If $s \geq 1$, then $\mathcal{F}(H, s-1, t)$ contains a graph that is the disjoint union of a clique and a vertex.*
2. *If $t \geq 1$, then $\mathcal{F}(H, s, t-1)$ contains a graph that is the join of an independent set and a vertex.*

Proof. Because H is critical, $\mathcal{P}(n, \mathcal{F}(s-1, t, H)) \subseteq \{K_n, \overline{K_n}\}$ for all $n > n_0$. Consequently, any graph consisting of K_{n-1} and a vertex v such that v has no edges to the clique must contain some graph in $\mathcal{F}(s-1, t, H)$ as an induced subgraph. Since (s,t) is a witnessing pair, that graph cannot be a clique and so must contain the vertex v.

Similarly, any graph consisting of $\overline{K_{n-1}}$ and a vertex v such that v is completely connected to the independent set must contain some graph in $\mathcal{F}(s-1, t, H)$ as an induced subgraph. Since (s,t) is a witnessing pair, that graph cannot be an independent set and so must contain the vertex v.

□

1.2 New results

Our main result is Theorem 1.9.

Theorem 1.9. *For $k \geq 2$ and $\chi_c(H) = k+1$, almost every graph in $\mathcal{P}(n, H)$ is in $\mathcal{Q}(n, H)$ if and only if H is critical.*

It is fairly easy to prove necessity. We prove sufficiency by finding constants $0 < \sigma < 1$ and C such that if H is critical then

$$|\mathcal{P}(n,H) - \mathcal{Q}(n,H)| < C\sigma^n |\mathcal{Q}(n,H)| \tag{2}$$

for all n. This we do by first showing in Section 2.2 that almost every H-free graph can be covered by k sets such that together they contain fewer than ϑn^2 "bad" edges, for some small $\vartheta > 0$. We also show that almost all H-free graphs have certain other random-like properties (Section 2.3).

In Section 2.4 we show that almost all H-free graphs have a partition into s almost-cliques and t almost-independent sets (for some $s + t = k$) such that no vertex has more than βn "bad" edges incident with it. Finally, in Section 2.5 we show that almost all H-free graphs with no "bad" vertex in fact contain no "bad" edge at all, and are consequently in $\mathcal{Q}(n, H)$.

1.3 Previous results

Prömel and Steger [18] estimated the number of H-free graphs in terms of $\chi_c(H)$.

Theorem 1.10. *If* $\chi_c(H) = k+1$ *then* $|\mathcal{P}(n,H)| = 2^{(1-\frac{1}{k}+o(1))\binom{n}{2}}$.

Bollobás and Thomason [9], and independently Alekseev [1], proved a more general result for a family $\mathcal{H}$ of graphs. Recent sharper results for families of graphs appear in [2].

Prömel and Steger proved [17] a Turán-type of statement in the induced case for the following definition of 'critical' graph.

Definition 1.11. A graph H is called **PS-critical**[1] if $\chi_c(H) = k+1$ and for all $0 \le \tilde{k} \le k$ the following is satisfied. Either $(\tilde{k}, k-\tilde{k})$ is a covering pair or there exist (possibly four different) partitions of $V(H)$, $\Pi = (P_1, P_2, \ldots, P_{k+1})$, such that $P_1, \ldots, P_{\tilde{k}}$ are cliques and $P_{\tilde{k}+1}, \ldots, P_{k+1}$ are independent sets and

1. $|P_{k+1}| = 1$ and $e(P_{k+1}, P_1) \le 1$,
2. $|P_{k+1}| = 1$ and $e(P_{k+1}, P_1) \ge |P_1| - 1$,
3. $|P_{k+1}| = 1$ and $e(P_{k+1}, P_k) \le 1$,
4. $|P_{k+1}| = 1$ and $e(P_{k+1}, P_k) \ge |P_k| - 1$,

and there exist (again, possibly two different) partitions of $V(H)$, $\Pi = (P_1, \ldots, P_{k+1})$, with either $P_1, \ldots, P_{\tilde{k}}$ or $P_1, \ldots, P_{\tilde{k}+1}$ cliques and the rest independent sets such that

5. (for $\tilde{k} \ge 1$) $|P_{\tilde{k}+1}| = 2$ and $P_{\tilde{k}+1}$ is not connected to P_1, and
6. (for $\tilde{k} \le k-1$) $|P_{\tilde{k}+1}| = 2$ and $P_{\tilde{k}+1}$ is completely connected to P_{k+1}.

Prömel and Steger also proved in [13] that almost all C_4-free graphs can be covered by a clique and an independent set, and in [14] characterized the structure of almost all C_5-free graphs as well. It would be natural to consider C_6-free graphs next, but C_6 is not critical. Although $\chi_c(C_6) = 3$, $\mathcal{F}(C_6, 1, 0)$ consists of the path on four vertices: besides K_n and $\overline{K_n}$, any star forest also avoids this path, as does the complement of a star forest. Consequently, $\mathcal{P}(n, C_6)$ contains any graph that can be covered by the complement of a star forest and a clique; this is a much larger family than $\mathcal{Q}(n, C_6)$.

It can be seen similarly that larger even cycles are not critical. For all $k > 5$, $\mathcal{F}(C_{2k}, k-2, 0) = \{P_4, P_3 \cup K_1, K_2 \cup K_2, K_2 \cup \overline{K_2}, \overline{K_4}\}$, where $G \cup H$ indicates the disjoint union of the two graphs. The only graphs that avoid all of these induced subgraphs are those whose complements are disjoint copies of K_3 and a star forest. Every other $\mathcal{F}(C_{2k}, s, t)$ is either empty or consists of K_1, except that $\mathcal{F}(C_{12}, 3, 1) = \{\overline{K_2}\}$. Consequently, the following conjecture seems reasonable.

[1] We say "PS-critical" to distinguish it from the "critical" in Definition 1.7 and because it was defined by Prömel and Steger.

Conjecture 1.12. *For $k > 5$, almost every $G \in \mathcal{P}(n, C_{2k})$ can be covered by $k-2$ cliques and a graph whose complement is disjoint copies of K_3 and a star forest.*

One could formulate a similar conjecture for C_6, C_8, or C_{10}; the statements, however, seem rather technical. We should also note that our definition of critical and the definition of PS-critical are both technical enough that it is unclear to us whether one is a generalization of the other. In particular, we know of no graph that is critical but not PS-critical, or that is PS-critical but not critical.

1.4 Constants

This section may be useful to clarify size and dependence. In theory, we should define the constants from largest to smallest; one can check, however, that the following are all well-defined.

$$\frac{1}{n_0} \ll \gamma \ll \mu \ll \epsilon \ll \omega \ll \lambda \ll \alpha \ll \vartheta \ll \delta \ll \beta \ll 1,$$

$$\text{and } \frac{1}{n_0} \ll \rho \ll \vartheta.$$

1. We fix a graph H and let $h = |V(H)|$ and $k = \chi_c(H) - 1$.
2. We choose $\lambda > 0$ such that $\lambda < \frac{1}{(10kR(h))^2}$. This inequality is needed for Lemma 2.14.
3. We obtain $\omega = \omega(k, \lambda, R(h))$ from Theorem 2.7.
4. We choose $\epsilon > 0$ so that for all sufficiently large m, $\left(1 - \frac{1}{k}\right)\frac{m}{2} + \epsilon\binom{m}{2} < \omega m^2$. This will be needed in Lemma 2.14.
5. We choose $\mu > 0$ so that $H(2\mu) < \epsilon/6$. This is used in Lemma 2.11.
6. We obtain $\gamma_{2.8}(\mu, k)$ from Lemma 2.8, and then we choose $\gamma < \gamma_{2.8}$ such that $\gamma < \min\left\{\frac{\epsilon}{12}, \mu\right\}$. We need $\gamma < \frac{\epsilon}{12}$ for Lemma 2.11, and we need $\gamma < \mu$ for Lemma 2.6 and Lemma 2.11.
7. We obtain $m_0(k, R(h), \epsilon)$ from Theorem 2.10 and then if necessary make m_0 large enough to satisfy the condition in 1.4.4 and large enough that $m_0 \geq \frac{12}{\epsilon}$. We then obtain $m \geq m_0$ from Lemma 2.9. The inequality $m \geq \frac{12}{\epsilon}$ is needed in Lemma 2.11.
8. We choose $\alpha > 0$ such that $\alpha^2 > 100\lambda$ and $\alpha < \frac{1}{kR(h)}$ (this is possible because we chose λ small enough). Both inequalities are needed in Lemma 2.14.
9. Let $\vartheta = \frac{\gamma}{2} + \alpha h + \frac{\lambda}{m^2} + \frac{1}{2m}$. As a shorthand we let $\psi = \sqrt{\vartheta}\log(\frac{1}{\vartheta})$. We can choose γ, α, λ, and $1/m$ small enough that $2H(\vartheta) < \frac{1}{k^2}$ and $\vartheta < \frac{3}{8k^4}$ and $\psi < \min\{\frac{1}{4k(k-2)}, \frac{1}{12hk2^{h+3}}\}$. These inequalities are needed for Lemma 2.17, Lemma 2.20, Lemmas 2.25 and 2.27, and Lemma 2.28 respectively.

10. We find $\delta_0(h-1)$ and $\delta_0(kh2^h)$ from Lemma 2.22. We choose $\delta > 0$ so that $\delta < \min\{\frac{1}{k^2}, \delta_0(h-1), \delta_0(kh2^h)\}$ and $H(2\delta) < \min\{\frac{1}{4k^2}, \frac{1}{3k^2 2^{h+3}}\}$ and $\delta > \max\{2H(\vartheta), \frac{8k^2}{3}\vartheta, \sqrt{\frac{4\vartheta}{3}}\}$. The inequality $\delta < \frac{1}{k^2}$ is needed for Lemma 2.20. We need $\delta < \min\{\delta_0(h-1), \delta_0(k2^h)\}$ for Lemmas 2.23 and 2.27. Lemma 2.25 requires that $H(2\delta) < \frac{1}{4k^2}$ and Lemma 2.28 requires that $H(2\delta) < \frac{1}{3k^2 2^{h+3}}$. The final set of inequalities are needed for Lemmas 2.17, 2.20, and 2.23 respectively.

11. We obtain ρ_0 from Lemma 2.17. In Corollary 2.19 we find $\rho' > 0$ such that $\rho' < \min\{\rho_0, \vartheta\}$ and then find $\rho > 0$ such that $\rho < \rho' - \frac{1}{n_0^2}$. Apart from $\frac{1}{n_0} \ll \rho$, we do not need a lower bound on ρ.

12. We choose $\beta > 0$ so that $\frac{\beta}{2^h} \leq 1 - \frac{1}{4k-8}$ and $\beta \gg \max\left\{\frac{8k^2h}{3}\vartheta, \frac{8k^3h2^h}{3}\vartheta\right\}$ and $H(\beta k) < \min\{\frac{1}{8k}, \frac{1}{3hk2^{h3}}\}$. We need $\frac{\beta}{2^h} \leq 1 - \frac{1}{4k-8}$ for Lemma 2.27. The second set of inequalities are needed for applications of Lemma 2.22 in Lemmas 2.23 and 2.27. Lemma 2.25 requires that $H(\beta k) < \frac{1}{8k}$ and Lemma 2.28 requires that $H(\beta k) < \frac{1}{3hk2^{h3}}$.

13. We obtain $\hat{\rho} = \frac{1}{k} - H(2\delta)k - 2H(k\beta) - (k-2)\psi - \frac{2\log(n)}{n}$ from Lemma 2.25. Previous bounds guarantee that $\hat{\rho}$ is positive. We obtain $-\rho'' = H(2\delta) + H(\beta)h + 4\psi h - \frac{1}{2^{h+3}} - \frac{\psi}{2^{h+3}}$ during the proof of Lemma 2.28, and then let $\tilde{\rho} = \frac{9}{10}\rho''$. Previous bounds guarantee that $\tilde{\rho}$ is positive.

14. We obtain σ by induction such that $\max\{2^{-\hat{\rho}}, 2^{-\tilde{\rho}}\} < \sigma < 1$. The proof of Theorem 1.9 requires $\sigma > \max\{2^{-\hat{\rho}}, 2^{-\tilde{\rho}}\}$, and the statement of Theorem 1.9 requires $\sigma < 1$.

15. We choose n_0 an integer such that $\frac{1}{n_0}$ is smaller than every preceding constant and larger than the n_0 in Lemma 2.22 for $m = n-1$ and for $m = kh2^h$. That is needed for Lemma 2.23 and for Lemma 2.27. We also require n_0 to be large enough that whenever $n > n_0$ the following inequalities are all true:

$$\begin{array}{rcll} n2^{(1-\frac{1}{k}-\frac{\epsilon}{6})\frac{n^2}{2}} & < & 2^{(1-\frac{1}{k}-\frac{\epsilon}{5})\frac{n^2}{2}} & \text{(for Lemma 2.6)},\\ 2^{2\log(m)n+\log(3)\frac{m^2}{2}+m^2} & < & 2^{\epsilon n^2/4} & \text{(for Lemma 2.11)},\\ \frac{1}{n}2\log(n) & < & \frac{1}{4k} & \text{(for Lemma 2.23)},\\ (1-\frac{1}{k})\binom{h}{2} - \rho' n & < & -\frac{9}{10}\rho' n & \text{(for Lemma 2.28)},\\ \log(n) & < & \psi n & \text{(for Lemma 2.28)}, \end{array}$$

and $\min\{|\mathcal{Q}(n,H)|2^{-\hat{\rho}n}, |\mathcal{Q}(n,H)|2^{-\tilde{\rho}n}\} > \max\{2^{(1-\frac{1}{k})\frac{n^2}{2}-\frac{\epsilon}{10}n^2}, 2^{(1-\frac{1}{k})\frac{n^2}{2}-\rho n^2}\}$ (for Theorem 1.9).

16. We choose $C > 8$ large enough that for $n \leq n_0$ Equation (2) trivially holds.

2 H-free graphs

2.1 Observations

One part of Theorem 1.9 is easy to see. If H is not a critical graph, then

$$|\mathcal{P}(n,H)| \gg |\mathcal{Q}(n,H)|. \tag{3}$$

Sketch of proof, Theorem 1.9 necessity. Suppose $\mathcal{P}(m,\mathcal{F}(H,\tilde{s},\tilde{t}))$ contains more than two graphs for some $\tilde{s}$ and $\tilde{t}$ with $\tilde{s}+\tilde{t}=k-1$; then one is neither K_n nor $\overline{K_n}$. Then it must contain at least m labelled graphs (see Balogh-Bollobás-Weinreich [7] and Scheinerman-Zito [19]). Fix some partition Π of $[n]$ into k classes, and suppose there are c possible edges between vertices in different classes of the partition. For each witnessing pair (s,t) there are $\binom{k}{s}2^c$ possible ways to form a graph in $\mathcal{Q}(n,H)$ that respects this partition, and each graph in $\mathcal{Q}(n,H)$ respecting this partition must be formed in such a way. Consequently, the number of graphs in $|\mathcal{Q}(n,H)|$ respecting Π is at most $\sum_{s=0}^{k}\binom{k}{s}2^c \le 2^k 2^c$.

On the other hand, if H is not critical then there exist $\tilde{s}$ and $\tilde{t}$ such that $\tilde{s}+\tilde{t}=k+1$ and for sufficiently large n there are at least n/k labelled graphs $F \in \mathcal{P}(n/k,\mathcal{F}(H,\tilde{s},\tilde{t}))$. We may therefore generate many graphs that are in $\mathcal{P}(n,H)$ and not in $\mathcal{Q}(n,H)$ by replacing the largest class in the partition with some such F and then choosing $\tilde{s}$ of the remaining classes to be cliques, letting the remaining $\tilde{t}$ be independent sets, and choosing the cross-edges arbitrarily. The number of graphs respecting Π that are in $\mathcal{P}(n,H)$ but not $\mathcal{Q}(n,H)$ is therefore at least $\frac{n}{k}\binom{k}{\tilde{s}}2^c$.

Since k is a constant, the ratio of graphs in $\mathcal{P}$ to graphs in $\mathcal{Q}$, that respect Π, is

$$\frac{\frac{n}{k}\binom{k}{\tilde{s}}2^c}{2^k 2^c} \ge \frac{n\binom{k}{\tilde{s}}}{k2^k},$$

which tends to infinity. If $|\mathcal{P}(n/k,\mathcal{F}(H,\tilde{s},\tilde{t}))| \gg k^n$ then this is more than there are choices of Π and we are done. Otherwise, "most" graphs in $\mathcal{P}$ respect not many partitions of $[n]$; we omit the technical details. □

For the remainder of the paper, assume that H is a critical graph on h vertices with coloring number $k+1$. It remains to show that there exist constants $\sigma < 1$ and C such that (2) is true.

We will prove Theorem 2 by induction on n; in what follows we will determine an integer n_0 such that $1/n_0$ is smaller than any other constant we use, and large enough to satisfy all of the following lemmas. From that, we will choose C large enough that (2) holds for every $n \le n_0$; for $n > n_0$ we will use induction.

First, we observe some facts about $\mathcal{Q}(n,H)$.

Lemma 2.1. $|\mathcal{Q}(n,H)| \ge |\mathcal{Q}(n-1,H)|2^{(1-\frac{1}{k})(n-1)}$.

Proof. Consider a graph $G \in \mathcal{Q}(n-1,H)$; it can be covered by s cliques and t independent sets for some witnessing pair (s,t). Form a new graph $G' \in \mathcal{Q}(n,H)$ by adding a new vertex, labelled 'n', to the smallest clique or independent set (break ties by choosing whichever has the lowest-indexed vertex). Then n must be adjacent to every vertex in its clique, or to no vertex in its independent set. It may or may not be adjacent to each other vertex in the graph; since we added n to the smallest clique or independent set there are at least $(1-\frac{1}{k})(n-1)$ vertices in other cliques and independent sets. Consequently, there are at least $2^{(1-\frac{1}{k})(n-1)}$ possible ways to form a new graph G' from each graph $G \in \mathcal{Q}(n-1, H)$. $\square$

Iterating Lemma 2.1 results in the following.

Corollary 2.2.

$$|\mathcal{Q}(n, H)| \geq |\mathcal{Q}(n-i, H)| 2^{(1-\frac{1}{k})in - \binom{i+1}{2}(1-\frac{1}{k})}.$$

Corollary 2.3.

$$2^{(1-\frac{1}{k})\binom{n}{2}} \leq |\mathcal{Q}(n, H)| < 2^k k^n 2^{(1-\frac{1}{k})\binom{n}{2}}.$$

Proof. For the lower bound, we use Corollary 2.2 with $i = n$ to obtain

$$|\mathcal{Q}(n, H)| \geq 2^{(1-\frac{1}{k})n^2 - \binom{n+1}{2}(1-\frac{1}{k})} \geq 2^{(1-\frac{1}{k})\binom{n}{2}},$$

as $n^2 - \frac{n^2+n}{2} = \binom{n}{2}$.

For the upper bound, there are fewer than k^n ways to partition n vertices into k classes, $\binom{k}{s}$ ways to let s of the classes be cliques, and $2^{(1-\frac{1}{k})\binom{n}{2}}$ ways to select edges between vertices in different classes. Since s may range from 0 to k, this gives

$$|\mathcal{Q}(n, H)| < \sum_{s=0}^{k} \binom{k}{s} k^n 2^{(1-\frac{1}{k})\binom{n}{2}} = 2^k k^n 2^{(1-\frac{1}{k})\binom{n}{2}}.$$

$\square$

2.2 Sparse and Dense classes

We will show that for almost every $G \in \mathcal{P}(n, H)$, G can be covered by s almost-cliques and t almost-independent sets, for some witnessing pair (s, t). We will define 'almost' later. To this end, we will define an auxiliary graph $A_\mu(G)$, called a cluster graph, and consider a partition of *its* vertices.

Definition 2.4. Given a graph G on n vertices and vertex sets $X, Y \subseteq V(G)$ with $X \cap Y = \emptyset$, the **edge density**, denoted $d(X, Y)$, is $\frac{e(X,Y)}{|X||Y|}$. We say a pair (X, Y) is **γ-regular** if for all $X' \subset X$ and $Y' \subset Y$ with $|X'| > \gamma|X|$ and $|Y'| > \gamma|Y|$, $|d(X', Y') - d(X, Y)| < \gamma$. An **equipartition** is a partition whose parts differ from each other in size by at most one. We say an equipartition $\mathcal{B} = (U_1, U_2, \ldots, U_m)$ is **γ-regular** if all but at most γm^2 of the pairs (U_i, U_j) are γ-regular.

If $\mu, \gamma > 0$ and $\mathcal{B} = (U_1, U_2, \ldots, U_m)$ is an equipartition of a subset of $V(G)$ such that every pair (X_i, X_j) is γ-regular, the **cluster graph** $A_\mu(G)$ has vertex set $[m]$ with ij an edge if and only if $\mu < d(U_i, U_j) < 1 - \mu$.

Given a graph $G \in \mathcal{P}(n, H)$ and a partition Π of $V(G)$, call a pair of vertices a **bad edge** if it appears as an edge in a class that has more non-edges than edges, or if it is a non-edge inside a class of G that has more edges than non-edges.

Definition 2.5. For $\vartheta > 0$, let $\mathcal{P}(n, H, \vartheta)$ be the set of graphs $G \in \mathcal{P}(n, H)$ that admit a partition of $V(G)$ into k parts which results in at most ϑn^2 bad edges.

The main lemma in this section is the following.

Lemma 2.6. *For n large enough there exist $\vartheta > 0$ and $\epsilon > 0$ such that*

$$|\mathcal{P}(n, H) - \mathcal{P}(n, H, \vartheta)| < 2^{(1-\frac{1}{k}-\frac{\epsilon}{5})\frac{n^2}{2}}.$$

Our main tools to prove Lemma 2.6 are the stability theorem of Simonovits [20] and a variant of the Regularity Lemma by Alon and Stav [3], and a corresponding embedding lemma.

Theorem 2.7. *[Stability Theorem] For any $\lambda > 0$ and any integers k, R, there exists an $\omega = \omega(k, \lambda, R) > 0$ such that if H is a graph on m vertices, for m sufficiently large, that does not contain $K_{k+1}(R)$ and*

$$e(H) > e(T_{m,k}) - \omega m^2,$$

then we can change H into $T_{m,k}$ by adding and/or removing at most λm^2 edges.

The following embedding lemma is from [3]:

Lemma 2.8. *For every real $0 < \mu < 1$ and integer $f \geq 1$ there exists $\gamma = \gamma_{2.8}(\mu, f)$ with the following property. Suppose that F is a graph on f vertices $v_1, \ldots, v_f$ and that $U_1, \ldots, U_f$ is an f-tuple of disjoint vertex sets of a graph G such that for every $1 \leq i < j \leq f$ the pair (U_i, U_j) is γ-regular. Moreover, suppose that whenever $(v_i, v_j) \in E(F)$ we have $d(U_i, U_j) \geq \mu$, and whenever $(v_i, v_j) \notin E(F)$ we have $d(U_i, U_j) \leq 1 - \mu$. Then, some f-tuple $u_1 \in U_1, \ldots, u_f \in U_f$ spans an **induced** copy of F, where each u_i plays the role of v_i.*

The following lemma, also from [3], is a variant of Szemerédi's Regularity Lemma [21], but says we can find an equipartition for which *every* pair is regular if we restrict our attention to an induced subgraph.

Lemma 2.9. *For every integer m_0 and every $\gamma > 0$ there exists an integer $M = M(m_0, \gamma)$ that satisfies the following. Any graph G on $n \geq M$ vertices has an equipartition $\mathcal{A} = \{V_i | 1 \leq i \leq m\}$ of $V(G)$ and an induced subgraph U of G, with an equipartition $\mathcal{B} = \{U_i | 1 \leq i \leq m\}$ of the vertices of U, that satisfy:*

1. *$m_0 \leq m \leq M$.*

2. *$U_i \subset V_i$ for all $i \geq 1$, and $|U_i| \geq n/M$.*

*3. In the equipartition $\mathcal{B}$, **all** pairs are γ-regular.*

4. All but at most $\gamma\binom{m}{2}$ of the pairs $1 \le i < j \le m$ are such that $|d(V_i, V_j) - d(U_i, U_j)| < \gamma$.

We will also use the Erdős-Stone Theorem [11].

Theorem 2.10. *For any integers k and t and real number $\epsilon > 0$ there exists $m_0(k, t, \epsilon)$ such that for $m > m_0$ any graph on m vertices with at least*

$$\left(1 - \frac{1}{k} + \epsilon\right)\binom{m}{2}$$

edges contains a copy of $K_{k+1}(t)$.

In Section 1.4 we fixed λ; notice that λ can be chosen to be arbitrarily small and still satisfy the conditions in 1.4.2. From Theorem 2.7 we obtain $\omega = \omega(k, \lambda, R(h))$, where $R(h)$ is the Ramsey number. Throughout what follows, we will have to make sure our constants remain small enough relative to ω that the cluster graph we consider has at least $e(T_{m,k}) - \omega m^2$ edges. We also fixed ϵ and μ in Section 1.4 and obtained $\gamma_{2.8}(\mu, k)$ from Lemma 2.8, then fixed $\gamma < \gamma_{2.8}$. Notice that if (U_i, U_j) is a γ-regular pair then it is also a $\gamma_{2.8}(\mu, k)$-regular pair. For m_0 defined by 1.4.7, consider the partition $\mathcal{B}$ guaranteed by Lemma 2.9 and let A be the cluster graph defined for that partition. Notice that by the definition of m in 1.4.7, $|V(A)| = m$ and $\frac{12}{\epsilon} < m_0 < m < M(\gamma, m_0)$.

We call a pair **sparse** if $d(U_i, U_j) \le \mu$, **average** if $\mu < d(U_i, U_j) < 1 - \mu$, and **dense** if $d(U_i, U_j) \ge 1 - \mu$. When it will not cause confusion, we will often refer to a pair of vertices (i, j) in the cluster graph A as "sparse", "average", or "dense" if the corresponding pair (U_i, U_j) is respectively sparse, average, or dense.

Lemma 2.11. *For sufficiently large n, all but at most $2^{(1-\frac{1}{k}-\frac{\epsilon}{6})\frac{n^2}{2}}$ graphs $G \in \mathcal{P}(n, H)$ have cluster graph A on m vertices that satisfies*

$$\left(1 - \frac{1}{k} - \epsilon\right)\binom{m}{2} < e(A) < \left(1 - \frac{1}{k} + \epsilon\right)\binom{m}{2}. \tag{4}$$

Proof. Let $r = R(h)$. If $K_{k+1}(r)$ is a subgraph of the cluster graph A it follows by Lemma 2.8 that G contains an induced copy of H; that is because each class of $K_{k+1}(r)$ contains either h mutually sparse pairs or h mutually dense pairs, which correspond to, respectively, an independent set or a clique in G. The edges in A are average pairs, and so may correspond to an edge or non-edge of G as needed. Since $\chi_c(H) = k + 1$, any collection of $k + 1$ cliques and independent sets cover H, and so by Lemma 2.8 G contains an induced copy of H. By the Erdős-Stone theorem, because $|V(A)| = m$, if $e(F) \ge (1 - \frac{1}{k} + \epsilon)\binom{m}{2}$ then $K_{k+1}(r)$ is a subgraph of A. Consequently, $e(A) < (1 - \frac{1}{k} + \epsilon)\binom{m}{2}$.

Now we will count the number of graphs whose cluster graph has fewer than $(1-\frac{1}{k}-\epsilon)\binom{m}{2}$ edges. Partition n vertices into m classes to determine a partition $\mathcal{A} = \{V_1, V_2, \ldots, V_m\}$, then choose a partition $\mathcal{B} = \{U_1, U_2, \ldots, U_m\}$ satisfying $U_i \subset V_i$ for all i. There are at most $m^{2n} = 2^{2\log(m)n}$ ways to do this. There are $3^{\binom{m}{2}} \le 2^{\log(3)m^2/2}$ ways to three-color the edges of K_m, which corresponds to determining which pairs of classes (U_i, U_j) are sparse, average, or dense.

There are at most $2^{\binom{n/m}{2}}$ ways to place edges inside one of the m classes $V_1, \ldots, V_m$, so there are at most $2^{m\binom{n/m}{2}} \leq 2^{n^2/2m}$ ways to place edges inside all m classes. If (U_i, U_j) is sparse or dense and $|d(V_i, V_j) - d(U_i, U_j)| < \gamma$, then there are at most $\sum_{i \leq 2\mu \frac{n^2}{m^2}} \binom{n^2/m^2}{i}$ ways to place edges between V_i and V_j (recall from 1.4.6 that $\gamma < \mu$, so $\mu + \gamma < 2\mu$). There are at most $\binom{m}{2}$ such pairs, so there are at most

$$\left[2 \sum_{i \leq 2\mu \frac{n^2}{m^2}} \binom{n^2/m^2}{i} \right]^{\binom{m}{2}} \leq 2^{m^2 + H(2\mu)\frac{n^2}{2}}$$

ways to place edges between such pairs (V_i, V_j).

There are at most $(1 - \frac{1}{k} - \epsilon)\binom{m}{2}$ average pairs (U_i, U_j) with $|d(V_i, V_j) - d(U_i, U_j)| < \gamma$, so there are at most $2^{n^2/m^2 (1-\frac{1}{k}-\epsilon)\binom{m}{2}} \leq 2^{(1-\frac{1}{k}-\epsilon)\frac{n^2}{2}}$ ways to place edges between such pairs (V_i, V_j). Finally, there are at most $\gamma\binom{m}{2}$ pairs (V_i, V_j) for which $|d(V_i, V_j) - d(U_i, U_j)| \geq \gamma$, so there are at most $2^{\frac{n^2}{m^2}\gamma\binom{m}{2}} \leq 2^{\gamma\frac{n^2}{2}}$ ways to place edges between such pairs.

Consequently, the number of graphs $G \in \mathcal{P}(n, H)$ whose cluster graph A on m vertices has fewer than $(1 - \frac{1}{k} - \epsilon)\binom{m}{2}$ edges is at most

$$2^{2\log(m)n + \log(3)\frac{m^2}{2} + m^2} 2^{\frac{n^2}{2m} + H(2\mu)\frac{n^2}{2} + (1-\frac{1}{k}-\epsilon)\frac{n^2}{2} + \gamma\frac{n^2}{2}}.$$

For large enough n, this is at most

$$2^{(1-\frac{1}{k}+\frac{1}{m}+H(2\mu)+\gamma-\frac{\epsilon}{2})\frac{n^2}{2}},$$

and n is chosen in Section 1.4 to be large enough. Recall that $H(2\mu) < \frac{\epsilon}{6}$ (from 1.4.5), $\frac{1}{m} < \frac{\epsilon}{12}$ (from 1.4.7), and $\gamma < \frac{\epsilon}{12}$ (from 1.4.6), and so $\frac{1}{m} + H(2\mu) + \gamma - \frac{\epsilon}{2} < -\frac{\epsilon}{6}$. Consequently, at most $2^{(1-\frac{1}{k}-\frac{\epsilon}{6})\frac{n^2}{2}}$ graphs $G \in \mathcal{P}(n, H)$ have a cluster graph with fewer than $(1 - \frac{1}{k} - \epsilon)\binom{m}{2}$ edges.

□

Notice that

$$e(A) > \left(1 - \frac{1}{k} - \epsilon\right)\binom{m}{2} = e(T_{m,k}) - \left[\left(1 - \frac{1}{k}\right)\frac{m}{2} + \epsilon\binom{m}{2}\right] > e(T_{m,k}) - \omega m^2,$$

and so by the Stability Theorem (Theorem 2.7) A differs from $T_{m,k}$ in at most λm^2 edges.

Given a partition of the vertices of a graph, we will call an edge **interior** if it connects vertices in the same part, and a **cross-edge** otherwise. Fix some k-partition $\Pi = (P_1, P_2, \ldots, P_k)$ of the vertices of A that maximizes the number of cross-edges. By Theorem 2.7, the number of missing cross-edges in A is at most λm^2 and the number of present interior edges is at most λm^2.

Lemma 2.12. *Let q be a prime and k any integer. Then we may pack q^2 edge-disjoint copies of K_k into $K_k(q)$.*

Proof. Order the k classes and label the vertices in each class with $\{0, 1, \ldots, q-1\}$. For any $a \in \{0, 1, \ldots, q-1\}$ and any $b \in \{0, 1, \ldots, q-1\}$ let v_i be the vertex in the ith class with label $a+(i-1)b$, modulo q. Let $F_{a,b}$ be the resulting induced copy of K_k; there are q^2 such copies. We claim that if $(a, b) \neq (a', b')$ then $F_{a,b}$ and $F_{a',b'}$ are edge-disjoint. Notice that because q is prime, $\mathbb{Z}_q$ is a field. Suppose the edge $\{a + (i-1)b, a + (j-1)b\}$ appears in $F_{a',b'}$. Then $a+(i-1)b = a'+(i-1)b'$ and $a+(j-1)b = a'+(j-1)b'$. Consequently, $(i-j)b = (i-j)b'$; because $\mathbb{Z}_q$ is a field, this implies $b = b'$. Therefore, $a + (i-1)b = a' + (i-1)b$ and so $a = a'$. $\square$

Definition 2.13. Call a subset $S \subset V(A)$ **sparse-homogeneous** if for each $ij \in E(A[S])$ the pair (U_i, U_j) is sparse or average. Call S **dense-homogeneous** if for each $ij \in E(A[S])$ the pair (U_i, U_j) is dense or average. Call S **non-homogeneous** otherwise.

Notice that by Lemma 2.8 if $S = \{v_1, \ldots, v_\ell\}$ is a sparse-homogeneous set in $V(A)$ then there is an independent set in G with exactly one vertex from each set U_{v_i}. Similarly, if S is a dense-homogeneous set in $V(A)$ there is a clique in G with exactly one vertex from each set U_{v_i}.

The following two lemmas will help to prove that each class of $V(A)$ under the partition Π either contains many sparse pairs or many dense pairs. This in turn will help us to find a partition of the original graph G whose classes again contain either many or very few edges. Let α be defined as in 1.4.8.

Lemma 2.14. *For large enough m and G satisfying the restrictions in Lemma 2.11, no class of the partition of $V(A)$, $\Pi = (P_1, P_2, \ldots, P_k)$, contains more than αm vertex-disjoint non-homogeneous sets of size h.*

Proof. Suppose the statement is false, and without loss of generality assume that P_1 is a contradictory class. Let $r = R(h)$ (the Ramsey number), and suppose that α is small enough that each class of the k-partition Π of $V(A)$ contains at least $\alpha m r$ vertices (since A is close to the Turán graph $T_{m,k}$, no class in the partition will be small). Then each class other than P_1 can be divided into αm vertex-disjoint subsets, each of which contains either an h-vertex sparse-homogeneous set or an h-vertex dense-homogeneous set. Call these αm subsets the *relevant* subsets.

On the other hand, P_1 contains αm non-homogeneous subsets of size h. For each such non-homogeneous set there is an induced subgraph in G with h vertices that is neither a clique nor an independent set, obtained by applying Lemma 2.8. Since H is critical, none of these αm subgraphs are in $\mathcal{P}(m/k, \mathcal{F}(H, s, t))$ for any $s + t = k - 1$. Consequently, for any $s + t = k - 1$, every one of the subgraphs contains some graph from $\mathcal{F}(H, s, t)$, and so any choice of one relevant subset from each P_i and one non-homogeneous set from P_1 results in a potential induced copy of H in G, if we can apply Lemma 2.8.

Bertrand's postulate states that there is a prime number between j and $2j$ for any integer j, so fix the smallest prime number q such that $\frac{2\alpha m}{4} \geq q \geq \lfloor \frac{\alpha m}{4} \rfloor$. Now by Lemma 2.12 and

Lemma 2.8, if every cross-edge is present then we can embed q^2 edge-disjoint copies of H into A in such a way that each induces H in G. Since G is H-free, at least one cross-edge must be missing from each such copy of H; because they are edge-disjoint, there are at least q^2 missing cross-edges in A. Consequently, there are $q^2 \geq \frac{\alpha^2 m^2}{100}$ cross-edges missing from A. By 1.4.8, $\frac{\alpha^2 m^2}{100} > \lambda m^2$; since there are at most λm^2 missing cross-edges, this is a contradiction. $\square$

Now, after removing vertex-disjoint non-homogeneous sets from a class of Π what remains of the class is either a sparse-homogeneous set or a dense-homogeneous set. We will call a class of a partition of the cluster graph "sparse" if what remains is sparse-homogeneous, and "dense" if what remains is dense-homogeneous. Call a pair (i, j) "bad" if $d(U_i, U_j) < \mu$ in a dense class or if $d(U_i, U_j) > 1 - \mu$ in a sparse class.

Lemma 2.15. *For a graph $G \in \mathcal{P}(n, H)$ with a cluster graph A satisfying (4) and $\Pi = (P_1, P_2, \ldots, P_k)$ a k-partition of $V(A)$ maximizing the number of cross-edges, A contains at most $\alpha m^2 h$ bad pairs.*

Proof. Each class P_i contains at most $\binom{\alpha mh}{2} + (|P_i| - \alpha mh)(\alpha mh)$ bad pairs. This can be rewritten as $\alpha mh|P_i| - \binom{\alpha mh+1}{2}$. Consequently, the total number of bad edges in A is at most $\sum_{i=1}^{k} (\alpha mh|P_i|) - \binom{\alpha mh+1}{2} \leq \sum_{i=1}^{k} (\alpha mh|P_i|) = \alpha m^2 h$. $\square$

We are now ready to prove Lemma 2.6.

Proof of Lemma 2.6. There are fewer than n ways to choose $m < n$. We have found that for all but at most $n2^{(1-\frac{1}{k}-\frac{\epsilon}{6})\frac{n^2}{2}}$ graphs $G \in \mathcal{P}(n, H)$, the cluster graph A of G has a partition $\Pi = (P_1, \ldots, P_k)$ with no more than αhm^2 bad pairs in each class. From Section 1.4, $n2^{(1-\frac{1}{k}-\frac{\epsilon}{6})\frac{n^2}{2}} < 2^{(1-\frac{1}{k}-\frac{\epsilon}{5})\frac{n^2}{2}}$. Form a partition of $V(G)$ by letting V_i be in class j if $i \in P_j$. In this partition of $V(G)$, call a class "sparse" if it contains more missing edges than edges and "dense" otherwise. Call an edge "bad" if it appears in a sparse class or is missing from a dense class.

Recall that $|V_i| = \frac{n}{m}$ for all i and that for all but at most $\gamma\binom{m}{2}$ pairs, $|d(V_i, V_j) - d(U_i, U_j)| < \gamma$. There are therefore at most $\gamma\binom{m}{2}\frac{n^2}{m^2} \leq \frac{\gamma n^2}{2}$ bad edges between all such pairs (V_i, V_j). Recall also that A contains at most λm^2 interior edges, which correspond to average pairs. There are therefore at most $\lambda(1 - \mu + \gamma)\frac{n^2}{m^2} \leq \frac{\lambda n^2}{m^2}$ bad edges between all such pairs (recall that $\gamma < \mu$, by 1.4.6). If (i, j) was a bad pair in A it results in at most $\frac{n^2}{m^2}$ bad edges in G, so there are at most $\alpha hm^2\frac{n^2}{m^2} = \alpha hn^2$ such bad edges. Edges within each V_i may be bad; there are at most $m\binom{n/m}{2} \leq \frac{n^2}{2m}$ such edges. In total, that is at most

$$\left(\frac{\gamma}{2} + \alpha h + \frac{\lambda}{m^2} + \frac{1}{2m}\right) n^2$$

bad edges. In other words, $G \in \mathcal{P}(n, H, \vartheta)$, for $\vartheta = \frac{\gamma}{2} + \alpha h + \frac{\lambda}{m^2} + \frac{1}{2m}$.

$\square$

2.3 Some more properties of $\mathcal{P}(n, H, \vartheta)$

Recall that for some $G \in \mathcal{P}(n, H, \vartheta)$ and a partition $\Pi = (P_1, P_2, \ldots, P_k)$ of the vertices of G, we call a pair of vertices a "bad edge" if it is missing from a dense class or if it is present in a sparse class. We call such a partition **optimal** if it minimizes the number of bad edges; recall that if $G \in \mathcal{P}(n, H, \vartheta)$ then an optimal partition of G contains at most ϑn^2 bad edges.

Lemmas 2.17, 2.18, and 2.20 are from [5], although Lemma 2.17 was updated in [4]. Note that although those papers deal with non-induced subgraphs the proofs of these lemmas can be adapted to our needs.

Definition 2.16. Let $\mathcal{P}(n, H, \vartheta, \delta)$ be the set of all $G \in \mathcal{P}(n, H, \vartheta)$ such that if $\Pi = (P_1, \ldots, P_k)$ is an optimal partition of $V(G)$ then

1. for any $A \subset P_i$ and $B \subset V(G_n) - P_i$ with $|A| = |B| = \lfloor \delta n \rfloor$, we have $\frac{1}{4}|A||B| < e(A, B) < \frac{3}{4}|A||B|$, and
2. $||P_i| - n/k| < (\sqrt{\vartheta} \log 1/\vartheta) n$.

In other words, $\mathcal{P}(n, H, \vartheta, \delta)$ consists of those graphs having an optimal partition Π that is somewhat close to being an equipartition and whose cross-edges behave somewhat randomly.

Lemma 2.17 (Lemma 6.1 in [5])**.** *For any $\delta \geq 2H(\vartheta)$ there is a positive constant $\rho_0 = \rho_0(\vartheta) > 0$ such that for sufficiently large n all but at most $2^{(1-\frac{1}{k})\frac{n^2}{2} - \rho_0 n^2}$ graphs in $\mathcal{P}(n, H, \vartheta)$ satisfy condition 1 of Definition 2.16.*

Lemma 2.18 (Lemma 6.6 in [5])**.** *Given ϑ sufficiently small, all but at most $2^{(1-\frac{1}{k})\frac{n^2}{2} - 2\vartheta n^2}$ graphs $G \in \mathcal{P}(n, H, \vartheta)$, satisfy condition 2 of Definition 2.16.*

Corollary 2.19. *For any $\delta \geq 2H(\vartheta)$ there exists $\rho = \rho(\vartheta)$ such that*

$$|\mathcal{P}(n, H, \vartheta) - \mathcal{P}(n, H, \vartheta, \delta)| < 2^{(1-\frac{1}{k})\frac{n^2}{2} - \rho n^2}.$$

Proof. By applying Lemmas 2.17 and 2.18 we obtain

$$|\mathcal{P}(n, H, \vartheta) - \mathcal{P}(n, H, \vartheta, \delta)| < 2^{(1-\frac{1}{k})\frac{n^2}{2} - \rho_0 n^2} + 2^{(1-\frac{1}{k})\frac{n^2}{2} - 2\vartheta n^2}.$$

We fixed $\rho' < \min\{\rho_0, 2\vartheta\}$ in 1.4.11. Then $|\mathcal{P}(n, H, \vartheta) - \mathcal{P}(n, H, \vartheta, \delta)| < 2^{(1-\frac{1}{k})\frac{n^2}{2} - \rho' n^2 + 1}$. In Section 1.4, ρ was chosen sufficiently smaller than ρ' to remove the 1 from the exponent. □

The following lemma is a variant of Lemma 6.10 in [5], but has an identical proof; we omit details.

Lemma 2.20 (Lemma 6.10 and 6.11 in [5])**.** *For $\frac{8k^2}{3}\vartheta < \delta < \frac{1}{k^2}$, n sufficiently large (as defined by Lemma 2.17), $t \leq h$, and $G \in \mathcal{P}(n, H, \vartheta, \delta)$, let G' be a graph obtained from G by deleting t vertices. Then there are at most $2^{H(2\delta)nk}$ partitions of the vertices of G' into k classes that have at most ϑn^2 bad edges.*

2.4 Bad vertices of a graph

For $\beta > 0$, we say that a vertex v in a graph is β**-bad** with respect to an optimal partition if there are at least βn bad edges between v and other vertices in its class. We shall show that for β chosen in accordance with 1.4.12 almost every $G \in \mathcal{P}(n, H, \vartheta, \delta)$ contains no bad vertex.

Definition 2.21. Let $\mathcal{P}(n, H, \vartheta, \delta, \beta)$ be the set of graphs G in $\mathcal{P}(n, H, \vartheta, \delta)$ containing no β-bad vertex in any of the optimal partitions of $V(G)$.

If a graph $G \in \mathcal{P}(n, H, \vartheta, \delta)$ with an optimal k-partition $\Pi = (P_1, \ldots, P_k)$ contains a β-bad vertex v, then by the optimality of the partition Π we know that $|N(v) \cap P_i| \geq \beta n$ for every sparse class P_i and $|\overline{N(v)} \cap P_j| \geq \beta n$ for every dense class P_j. The following lemmas say that if both $N(v) \cap P_i$ and $\overline{N(v)} \cap P_i$ are large for every $1 \leq i \leq k$, we can find an induced copy of H in G. We need another embedding lemma first, Theorem 6.3 from [5]:

Lemma 2.22. *For every m there exists n_0, δ_0 such that if $n > n_0$ and $\delta < \delta_0 \ll \frac{\beta}{m}$ then for any graph G on n vertices and graph F on m vertices the following holds. Suppose $X_1, X_2, \ldots, X_m$ are disjoint sets in $V(G)$ with $|X_i| = \frac{\beta n}{m}$ such that for any pair $U_i \subset X_i$ and $U_j \subset X_j$ with $|U_i|, |U_j| \geq \delta n$, $e(U_i, U_j) \geq \frac{1}{4}|U_i||U_j|$ whenever $ij \in E(F)$ and $e(U_i, U_j) \leq \frac{3}{4}|U_i||U_j|$ whenever $ij \notin E(F)$. Then there is some choice of $x_i \in X_i$ such that $\{x_1, x_2, \ldots, x_m\}$ induces F in G.*

Lemma 2.23. *Given $G \in \mathcal{P}(n, H, \vartheta, \delta)$ with an optimal partition $\Pi = (P_1, P_2, \ldots, P_k)$, then G contains no bad vertex v such that $|N(v) \cap P_i| > \beta n$ and $|\overline{N(v)} \cap P_i| > \beta n$ for all P_i except perhaps the class containing v.*

Proof. Suppose there is such a vertex v and Π consists of s dense classes and t sparse classes. At least one of s, t is nonzero so without loss of generality suppose that v is in P_1 and P_1 is dense (a similar argument can be made if P_1 is sparse). Because H is critical and (s, t) is a witnessing pair with $s \geq 1$ by Lemma 1.8 there is some graph $F \in \mathcal{F}(s-1, t, H)$ such that F is the disjoint union of a clique and a single vertex. Given a copy of F, and $s-1$ cliques and t independent sets, each on h vertices, there is some way to place edges between them so as to induce a copy of H. Call such an embedding "proper". Notice that $|V(F)| \leq h$.

Now, since $\overline{N(v)} \cap P_1$ is large we can divide it into h sets, each of size at least $\frac{\beta n}{h}$. For $1 < i \leq k$, $N(v) \cap P_i$ and $\overline{N(v)} \cap P_i$ are all large, so we can divide each of those $2k-2$ sets into h sets of size at least $\frac{\beta n}{h}$. Let these sets together with $P_1 \cap \overline{N(v)}$ be called $X_1, X_2, \ldots, X_{2hk-h}$. We can now use Lemma 2.22 with $m = h-1$ and $F = H[V(H) - \{v\}]$ to find an induced copy of H in G. That is, we will find a clique in $P_1 \cap \overline{N(v)}$, cliques in the remaining dense classes, and independent sets in the remaining sparse classes such that everything respects the "proper" embedding (and exactly one vertex is chosen from each X_i). Notice that we will not be using every set X_i, only those that correspond to vertices in the covering of H by $s-1$ cliques, t independent sets, and F.

To apply Lemma 2.22 we require that $\delta \ll \frac{\beta}{h}$. Because Lemma 2.20 requires that $\frac{8k^2}{3}\vartheta < \delta$ this implies $\beta \gg \frac{8hk^2}{3}\vartheta$, which 1.4.12 guarantees. Recall from 1.4.10 that $\frac{3}{4}\delta^2 > \vartheta$. We suppose X_i and X_j are in the same dense class and $U_i \subset X_i$, $U_j \subset X_j$ with $|U_i|, |U_j| = \delta n$. Then $e(U_i, U_j) \geq \frac{1}{4}|U_i||U_j| = \frac{1}{4}\delta^2 n^2$, otherwise there are more than ϑn^2 edges missing from that dense class, which is a contradiction. Similarly, if X_i and X_j are in the same sparse class then $e(U_i, U_j) \leq \frac{3}{4}|U_i||U_j|$. Finally, if X_i and X_j are in different classes then by the definition of $\mathcal{P}(n, H, \vartheta, \delta)$, $\frac{1}{4}|U_i||U_j| \leq e(U_i, U_j) \leq \frac{3}{4}|U_i||U_j|$. Lemma 2.22 now provides the correct embedding, which is of course a contradiction. □

Lemma 2.24. *If $G \in \mathcal{P}(n, H, \vartheta, \delta, \beta)$ and $\Pi = (P_1, P_2, \ldots, P_k)$ is an optimal partition of $V(G)$, then G contains no bad vertex v such that both $N(v) \cap P_i$ and $\overline{N(v)} \cap P_i$ have sizes at least βn for all but one i.*

Proof. Suppose there is such a vertex, v. Without loss of generality, suppose that v is in the dense class P_1. Call P_i *evenly partitioned* if $N(v) \cap P_i$ and $\overline{N(v)} \cap P_i$ both have size at least βn. If P_1 is the class that is not evenly partitioned, then Lemma 2.23 still applies. Suppose that $N(v) \cap P_i$ and $\overline{N(v)} \cap P_i$ are both large for $1 \leq i < k$; that is, P_k is the class that is not evenly partitioned. If P_k is a dense class then $|\overline{N(v)} \cap P_k| > \beta n$, otherwise Π was not an optimal partition. Since both $\overline{N(v)} \cap P_i$ and $N(v) \cap P_i$ are large for $i \neq k$ and v is a bad vertex if $v \in P_k$, we can temporarily move v into P_k and apply the same proof as in Lemma 2.23. A similar argument applies if P_k is a sparse class. □

We need therefore only consider those graphs G for which an optimal partition Π results in a bad vertex whose neighbourhood leaves at least two classes of Π unevenly divided. The following lemma says that almost all $G \in \mathcal{P}(n, H, \vartheta, \delta)$ do *not* meet that description.

Lemma 2.25. *There exists $\hat{\rho} = \hat{\rho}(\delta, \beta, \vartheta) > 0$ such that for all but at most $(C\sigma^{n-1} + 1)2^{-\hat{\rho}n}|\mathcal{Q}(n, H)|$ graphs $G \in \mathcal{P}(n, H, \vartheta, \delta)$, optimal partition $\Pi = (P_1, P_2, \ldots, P_k)$, and bad vertex v, there exists at most one i such that $v \notin P_i$ and $|N(v) \cap P_i| < \beta n$ if P_i is sparse, or $|\overline{N(v)} \cap P_i| < \beta n$ if P_i is dense.*

Proof. We will count the number of graphs G_n for which there is a bad vertex having two or more classes satisfying the conclusion of the lemma. Notice that if G is such a graph, Π is such a partition, and v is such a bad vertex then $G[V - \{v\}]$ is in $\mathcal{P}(n-1, H)$. Moreover the partition Π when restricted to $V - \{v\}$ still contains at most ϑn^2 bad edges. In other words, each such graph G_n is formed from some graph $G' \in \mathcal{P}(n-1, H)$ with a partition containing at most ϑn^2 bad edges.

There are at most n ways to choose v and by Lemma 2.20 there are at most $2^{H(2\delta)k(n-1)} < 2^{H(2\delta)kn}$ ways to partition $[n-1]$ that result in a k-partition with at most ϑn^2 bad edges. There are at most k^2 ways to choose the two classes, say P_1 and P_2. There are at most $\binom{\frac{n}{k} + (\sqrt{\vartheta}\log(\frac{1}{\vartheta}))n}{\beta n}^2 \leq 2^{2H(\beta)n}$ ways to place edges between v and the classes P_1 and P_2.

The remaining $k-2$ classes have no restrictions on their connections to v, so there are at most $2^{(k-2)(\frac{n}{k} + \sqrt{\vartheta}\log(\frac{1}{\vartheta})n)}$ ways to connect v with them. Let $\psi = \sqrt{\vartheta}\log(\frac{1}{\vartheta})$. The following is an upper bound for the number of graphs we are considering.

$$\begin{aligned}
& nk^2|\mathcal{P}(n-1,H)|2^{H(2\delta)kn+2H(\beta k)n+(k-2)(\frac{n}{k}+\psi n)} \\
\leq\ & |\mathcal{P}(n-1,H)|2^{H(2\delta)kn+2H(\beta k)n+(k-2)(\frac{n}{k}+\psi n)+2\log n} \\
\leq\ & (C\sigma^{n-1}+1)|\mathcal{Q}(n,H)|2^{-(1-\frac{1}{k})n}2^{H(2\delta)kn+2H(\beta k)n+(k-2)(\frac{n}{k}+\psi n)+2\log(n)} \\
=\ & (C\sigma^{n-1}+1)|\mathcal{Q}(n,H)|2^{-\hat{\rho}n},
\end{aligned}$$

where $\hat{\rho} = \hat{\rho}(n) = \frac{1}{k} - H(2\delta)k - 2H(\beta k) - (k-2)\psi - \frac{2\log(n)}{n}$. We make use of (1) and the inductive hypothesis (2), as well as Lemma 2.1. Recall that $H(2\delta) < \frac{1}{4k^2}$ (from 1.4.10), $2H(\beta k) < \frac{1}{4k}$ (from 1.4.12), and $\psi < \frac{1}{4k(k-2)}$ (from 1.4.9), and n is large enough that $\frac{2\log(n)}{n} < \frac{1}{4k}$ (from 1.4.15), so $\hat{\rho}$ is positive.

□

The following is now immediate, after fixing $C, \sigma, \delta, \vartheta, \beta$ and n_0 from the preceding lemmas.

Corollary 2.26. *For all $n > n_0$,*

$$|\mathcal{P}(n,H,\vartheta,\delta) - \mathcal{P}(n,H,\vartheta,\delta,\beta)| < (C\sigma^{n-1}+1)|Q(n,H)|2^{-\hat{\rho}n}.$$

2.5 Bad edges

It remains to show that for almost all $G \in \mathcal{P}(n,H,\vartheta,\delta,\beta)$, there is some optimal k-partition of the vertex set of G that produces no bad edges. Notice that if Π is an optimal k-partition of $V(G)$ and F is a subgraph of some part, with $h = |V(H)|$ vertices, this induces a natural partition of each other part of Π into 2^h classes based on neighbourhoods of the vertices of F. If each of these neighbourhoods is large enough, then we can once again use Lemma 2.22 to find an induced copy of H in G.

Lemma 2.27. *Suppose $G \in \mathcal{P}(n,H,\vartheta,\delta,\beta)$, where ϑ, δ, β are chosen in Section 1.4 and n is large enough. Let $\Pi = (P_1, P_2, \ldots, P_k)$ be an optimal partition of $V(G)$ with s dense and t sparse classes, for some s, t. If $uv \in G[P_i]$ is a bad edge and P_i is a dense class, then for each induced subgraph F in P_i with h vertices, including u and v, that contains as an induced subgraph some graph from $\mathcal{F}(H, s-1, t)$ there must exist some $j \neq i$ such that the natural partition of P_j induced by the vertices of F contains at least one part with size smaller than $\frac{|P_j|}{2^{h+1}}$. If P_i is a sparse class then an analogous statement is true for an $F \in \mathcal{F}(H, s, t-1)$.*

Proof. Suppose the statement is false. Assume without loss of generality that $i = 1$ and P_1 is a dense class and contains a missing edge (the proof is easily adapted if P_1 is sparse and contains an edge).

We will again apply Lemma 2.22 to obtain an embedding of H, which is a contradiction. Notice that since H is critical and n is large, any non-trivial graph on $|P_1|$ vertices contains some graph in $\mathcal{F}(H, s-1, t)$ as an induced subgraph. Fix one, say $\tilde{F}$, and extend it to an induced subgraph of P_1 on h vertices, say F; because $\tilde{F} \in \mathcal{F}(H, s-1, t)$ it has at most h vertices, so this is possible. Now, there is some way to place edges between $s-1$ copies of

K_h, t copies of $\overline{K_h}$, and F so as to induce H; call such an embedding "proper". Consider the natural neighbourhood partition induced in each P_j by the vertices of F; if we can find $\overline{K_h}$ in each sparse class and K_h in each dense class such that the vertices of each are in whichever neighbourhood respects the proper embedding, then we have found an induced copy of H in G. That would of course be a contradiction, since G was H-free.

If there is no $j \neq 1$ such that the neighbourhood partition of P_j contains a part with fewer than $|P_j|/2^{h+1}$ vertices, then we can find such an embedding. This is achieved through another application of Lemma 2.22 in which $X_1, X_2, \ldots, X_m$ are the neighbourhood partition's parts, equally divided into h sub-parts, where $m = kh2^h$, and $\delta < \delta_0 \ll \frac{\beta}{kh2^h}$. Because $\frac{8k^2}{3}\vartheta < \delta$, this last inequalities implies that $\frac{8k^3h2^h}{3}\vartheta \ll \beta$, which 1.4.12 guarantees. Notice that $|P_i| \geq \frac{n}{k} - \psi n$ for every i. By 1.4.12 and 1.4.9, $\frac{\beta}{2} \leq 1 - \frac{1}{4k-8} < 1 - k\psi$, and so $|X_i| \geq \frac{|P_i|}{h2^{h+1}} \geq \frac{\beta n}{kh2^h}$. Recall that $\delta^2 > \frac{4\vartheta}{3}$, by 1.4.10.

Consider a pair (U_i, U_j) such that $U_i \subset X_i$ and $U_j \subset X_j$ and $|U_i|, |U_j| > \delta n$. If X_i and X_j are in different classes of Π then $\frac{1}{4}|U_i||U_j| < e(U_i, U_j) < \frac{3}{4}|U_i||U_j|$. If X_i and X_j are in the same sparse class of Π then $e(U_i, U_j) < \frac{1}{4}|U_i||U_j| \leq \frac{1}{4}\delta^2 n^2$, otherwise there are more than $\frac{3}{4}\delta^2 n^2 > \vartheta n^2$ bad edges in that sparse class. Similarly, if X_i and X_j are in the same dense class of Π then $e(U_i, U_j) > \frac{3}{4}|U_i||U_j|$. Consequently, by Lemma 2.22, there is a copy of K_h in each dense class and a copy of $\overline{K_h}$ in each sparse class such that the vertices of each respect the proper embedding, and cross-edges can be chosen properly, which is a contradiction. $\square$

Lemma 2.28. *There exists $\tilde{\rho} > 0$ such that for all but at most $(1 + C\sigma^{n-h})|\mathcal{Q}(n, H)|2^{-\tilde{\rho}n}$ graphs $G \in \mathcal{P}(n, H, \vartheta, \delta, \beta)$, there exists an optimal partition $\Pi = (P_1, P_2, \ldots, P_k)$ of $V(G)$ containing no bad edge.*

Proof. Suppose some $G \in \mathcal{P}(n, H, \vartheta, \delta, \beta)$ has a bad edge for every optimal k-partition of $V(G)$. Fix some optimal k-partition Π; let s be the number of dense classes in Π and t be the number of sparse classes in Π. Without loss of generality, assume that P_1 is a dense class and that there is a bad edge inside of P_1. Since H is critical, there is some induced subgraph of P_1, F, on h vertices containing a graph $\tilde{F} \in \mathcal{F}(H, s, t)$ (every graph in $\mathcal{F}(H, s, t)$ has at most h vertices). Removing the h vertices of F results in a graph $G' \in \mathcal{P}(n-h, H)$; it remains to count how many graphs G that contain a bad edge can be built from such a graph G'.

By Lemma 2.20 there are at most $2^{H(2\delta)kn}$ partitions of $[n-h]$ that have at most ϑn^2 bad edges. There are less than n^h ways to choose the h vertices of F. It remains to count how many ways these h vertices can be added to G'; that is, in how many ways these h vertices can be connected to the other $n - h$.

There are fewer than βn bad edges incident with each of these vertices, so there are at most $\binom{n/k+\sqrt{\vartheta}\log 1/\vartheta n}{\beta n}^h \leq 2^{H(\beta)nh}$ ways to connect them with vertices in their own class.

By Lemma 2.27 there is some class of the partition Π, say P_i, for which the partition induced by the neighbourhoods of the h vertices contains a part with size less than $\frac{|P_i|}{2^{h+1}}$. Let N be the small class in P_i; by Chernoff's inequality,

$$Pr\left(|N| < \frac{|P_i|}{2^{h+1}}\right) < \exp\left(-\frac{|P_i|}{2^{h+3}}\right) < 2^{-|P_i|/2^{h+3}}.$$

Consequently, there are at most $2^{h|P_i|-\frac{|P_i|}{2^{h+3}}} = 2^{|P_i|(h-\frac{1}{2^{h+3}})}$ ways to connect the h vertices to P_i. There are at most $2^{(n-|P_1|-|P_i|)h}$ ways to connect the h new vertices with vertices outside of P_1 and P_i. Let $\psi = \sqrt{\vartheta}\log(1/\vartheta)$. In total, at most

$$2^{H(2\delta)kn+h\log(n)+H(\beta)nh+(n-\frac{2n}{k}+2\psi n)h+(\frac{n}{k}+\psi n)(h-\frac{1}{2^{h+3}})} \tag{5}$$

"bad" graphs can be made from each graph in $\mathcal{P}(n-h,H)$. Let $-\rho'' = H(2\delta)k + H(\beta)h + 4\psi h - \frac{1}{k2^{h+3}} - \frac{\psi}{2^{h+3}}$, then (5) is at most

$$2^{(1-\frac{1}{k})hn-\rho'' n}.$$

Because $H(2\delta) < \frac{1}{3k^2 2^{h+3}}$ (from 1.4.10), and $H(\beta) < \frac{1}{3hk2^{h+3}}$ (from 1.4.12), and $\psi < \frac{1}{12hk2^{h+3}}$ (from 1.4.9), and n is large enough that $h\log(n)/n < \psi h$ (1.4.15) we see that $\rho'' > 0$. Consequently, there are at most

$$|\mathcal{P}(n-h,H)|2^{(1-\frac{1}{k})nh-\rho'' n}$$

"bad" graphs in $\mathcal{P}(n,H,\vartheta,\delta,\beta)$. By induction and by Lemma 2.2, the number of "bad" graphs is therefore at most

$$\begin{aligned}(1+C\sigma^{n-h})|\mathcal{Q}(n-h,H)|2^{(1-\frac{1}{k})nh-\rho'' n} &\leq (1+C\sigma^{n-h})|\mathcal{Q}(n,H)|2^{(1-\frac{1}{k})\binom{h}{2}-\rho'' n} \\ &\leq (1+C\sigma^{n-h})|\mathcal{Q}(n,H)|2^{-\frac{9}{10}\rho'' n},\end{aligned}$$

for large enough n. Let $\tilde{\rho} = \frac{9}{10}\rho''$. □

We are now ready to complete the proof of the main theorem.

Proof of Theorem 1.9. Notice, from the triangle inequality, that

$$\begin{aligned}|\mathcal{P}(n,H)-\mathcal{Q}(n,H)| &\leq |\mathcal{P}(n,H)-\mathcal{P}(n,H,\vartheta)| + |\mathcal{P}(n,H,\vartheta)-\mathcal{P}(n,H,\vartheta,\delta)| + \\ &+ |\mathcal{P}(n,H,\vartheta,\delta)-\mathcal{P}(n,H,\vartheta,\delta,\beta)| + |\mathcal{P}(n,H,\vartheta,\delta,\beta)-\mathcal{Q}(n,H)|.\end{aligned}$$

We have, however, found bounds on the size of each of these sets:

$$\begin{aligned}|\mathcal{P}(n,H)-\mathcal{P}(n,H,\vartheta)| &< 2^{(1-\frac{1}{k})\frac{n^2}{2}-\frac{\epsilon}{10}n^2} && \text{(Lemma 2.6)},\\ |\mathcal{P}(n,H,\vartheta)-\mathcal{P}(n,H,\vartheta,\delta)| &< 2^{(1-\frac{1}{k})\frac{n^2}{2}-\rho n^2} && \text{(Corollary 2.19)},\\ |\mathcal{P}(n,H,\vartheta,\delta)-\mathcal{P}(n,H,\vartheta,\delta,\beta)| &< (C\sigma^{n-1}+1)|\mathcal{Q}(n,H)|2^{-\hat{\rho}n} && \text{(Corollary 2.26)},\\ |\mathcal{P}(n,H,\vartheta,\delta,\beta)-\mathcal{Q}(n,H)| &< (C\sigma^{n-h}+1)|\mathcal{Q}(n,H)|2^{-\tilde{\rho}n} && \text{(Lemma 2.28)}.\end{aligned}$$

It is obvious now that $|\mathcal{P}(n,H)-\mathcal{Q}(n,H)| < 4T$, where we define T to be the largest of the four bounds above. Recall that $|\mathcal{Q}(n,H)| \geq 2^{(1-\frac{1}{k})\binom{n}{2}}$. We may therefore fix n_0 large enough that whenever $n > n_0$ then

$$\min\{|\mathcal{Q}(n,H)|2^{-\hat{\rho}n}, |\mathcal{Q}(n,H)|2^{-\tilde{\rho}n}\} > \max\{2^{(1-\frac{1}{k})\frac{n^2}{2}-\frac{\epsilon}{10}n^2}, 2^{(1-\frac{1}{k})\frac{n^2}{2}-\rho n^2}\}.$$

In that case, the first two bounds above are smaller than the second two. It now suffices to show that

$$(C\sigma^{n-1}+1)2^{-\hat{\rho}n} + (C\sigma^{n-h}+1)2^{-\tilde{\rho}n} \leq \frac{1}{2}C\sigma^n. \tag{6}$$

We now fix $\sigma > \max\{2^{-\hat{\rho}}, 2^{-\tilde{\rho}}\}$ such that $\sigma < 1$, as in 1.4.14. This is possible, since both $2^{-\hat{\rho}}$ and $2^{-\tilde{\rho}}$ are smaller than 1. Since C is chosen in Section 1.4 to be large enough to make the statement of Theorem 1.9 true for $n \leq n_0$, we may assume that $C > 8$. Now, for $n > n_0$,

$$\begin{array}{rclcl} C\sigma^{n-1}2^{-\hat{\rho}n} & < & C\sigma^{n-1}\frac{1}{8}\sigma & = & \frac{1}{8}C\sigma^n \\ C\sigma^{n-h}2^{-\tilde{\rho}n} & < & C\sigma^{n-h}\frac{1}{8}\sigma & = & \frac{1}{8}C\sigma^n \\ 2^{-\hat{\rho}n} & < & \sigma^n & < & \frac{1}{8}C\sigma^n \\ 2^{-\tilde{\rho}n} & < & \sigma^n & < & \frac{1}{8}C\sigma^n, \end{array}$$

so (6) is true. This completes the proof of Theorem 1.9. □

References

[1] V. E. Alekseev. Range of values of entropy of hereditary classes of graphs. *Diskret. Mat.*, 4(2):148–157, 1992.

[2] N. Alon, J. Balogh, B. Bollobás, and R. Morris. The structure of almost all graphs in a hereditary property. to appear in Journal of Combinatorial Theory, Series B.

[3] N. Alon and U. Stav. The maximum edit distance from hereditary graph properties. *J. Combin. Theory Ser. B*, 98(4):672–697, 2008.

[4] J. Balogh, B. Bollobás, and M. Simonovits. The fine structure of octahedron-free graphs. to appear in Journal of Combinatorial Theory, Series B.

[5] J. Balogh, B. Bollobás, and M. Simonovits. The number of graphs without forbidden subgraphs. *J. Combin. Theory Ser. B*, 91(1):1–24, 2004.

[6] J. Balogh, B. Bollobás, and M. Simonovits. The typical structure of graphs without given excluded subgraphs. *Random Structures and Algorithms*, 34(3):305–318, 2009.

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

[8] B. Bollobás and A. Thomason. Projections of bodies and hereditary properties of hypergraphs. *Bull. London Math. Soc.*, 27(5):417–424, 1995.

[9] B. Bollobás and A. Thomason. Hereditary and monotone properties of graphs. *The mathematics of Paul Erdős, II*, 14:70–78, 1997.

[10] P. Erdős, D. Kleitman, and B. Rothschild. Asymptotic enumeration of K_n-free graphs. pages 19–27. Atti dei Convegni Lincei, No. 17, 1976.

[11] P. Erdős and A. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[12] P. Kolaitis, H. Prömel, and B. Rothschild. K_{l+1}-free graphs: asymptotic structure and a 0-1 law. *Trans. Amer. Math. Soc.*, 303(2):637–671, 1987.

[13] H. Prömel and A. Steger. Excluding induced subgraphs: quadrilaterals. *Random Structures and Algorithms*, 2(1):55–71, 1991.

[14] H. Prömel and A. Steger. Almost all Berge graphs are perfect. *Combin. Probab. Comput.*, 1(1):53–79, 1992.

[15] H. Prömel and A. Steger. The asymptotic number of graphs not containing a fixed color-critical subgraph. *Combinatorica*, 12(4):463–473, 1992.

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

[17] H. Prömel and A. Steger. Excluding induced subgraphs. II. Extremal graphs. *Discrete Appl. Math.*, 44(1-3):283–294, 1993.

[18] H. Prömel and A. Steger. Counting H-free graphs. *Discrete Math.*, 154(1-3):311–315, 1996.

[19] E. Scheinerman and J. Zito. On the size of hereditary classes of graphs. *J. Combin. Theory Ser. B*, 61(1):16–39, 1994.

[20] M. Simonovits. A method for solving extremal problems in graph theory, stability problems. *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319, 1968.

[21] E. Szemerédi. Regular partitions of graphs. *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, 260:399–401, 1978.