Minimum degree and the graph removal lemma

Jacob Fox* Yuval Wigderson†

Abstract

The clique removal lemma says that for every $r \geq 3$ and $\varepsilon > 0$, there exists some $\delta > 0$ so that every n-vertex graph G with fewer than δn^r copies of K_r can be made K_r-free by removing at most εn^2 edges. The dependence of δ on ε in this result is notoriously difficult to determine: it is known that δ^{-1} must be at least super-polynomial in ε^{-1}, and that it is at most of tower type in $\log \varepsilon^{-1}$.

We prove that if one imposes an appropriate minimum degree condition on G, then one can actually take δ to be a linear function of ε in the clique removal lemma. Moreover, we determine the threshold for such a minimum degree requirement, showing that above this threshold we have linear bounds, whereas below the threshold the bounds are once again super-polynomial, as in the unrestricted removal lemma.

We also investigate this question for other graphs besides cliques, and prove some general results about how minimum degree conditions affect the bounds in the graph removal lemma.

1 Introduction

One of the deepest results in extremal graph theory is the triangle removal lemma of Ruzsa and Szemerédi [33], as well as its extension to the graph removal lemma, proved independently by Alon–Duke–Lefmann–Rödl–Yuster [3] and Füredi [22]. Loosely, this result says that if a large graph G contains "few" copies of a fixed graph H, then it can be made H-free by deleting "few" edges. The formal statement is as follows.

Theorem 1.1. *Let H be a graph on h vertices. For every $\varepsilon > 0$, there exists a $\delta > 0$ such that the following holds. If G is an n-vertex graph with fewer than δn^h copies of H, then one can remove at most εn^2 edges from G to make it H-free.*

Despite its simple statement, the graph removal lemma is a deep result, with many applications in number theory, computer science, and graph theory. For more on the removal lemma and its history, we refer to the survey [6].

*Department of Mathematics, Stanford University, Stanford, CA 94305, USA. Email: `jacobfox@stanford.edu`. Research supported by a Packard Fellowship and by NSF award DMS-185563.

†Department of Mathematics, Stanford University, Stanford, CA 94305, USA. Email: `yuvalwig@stanford.edu`. Research supported by NSF GRFP Grant DGE-1656518.

Many important questions surrounding the graph removal lemma remain open. Most notably, the correct bound for δ in terms of ε is unknown. Formally, let $\delta(\varepsilon, H)$ denote the maximum δ such that every n-vertex graph with fewer than δn^h copies of H can be made H-free by removing at most εn^2 edges. The best lower bound on $\delta(\varepsilon, H)$, due to Fox [15], shows that $\delta(\varepsilon, H) \geq 1/T(O_h(\log \frac{1}{\varepsilon}))$, where T is the tower function, recursively defined by $T(0) = 1$ and $T(x) = 2^{T(x-1)}$ for $x \geq 1$. For the upper bound, Alon [2] (extending [33] and [11]) showed that if H is not bipartite[1], then $\delta(\varepsilon, H) \leq \varepsilon^{\Omega_H(\log \frac{1}{\varepsilon})}$ as $\varepsilon \to 0$. In particular, if H is not bipartite, then $1/\delta(\varepsilon, H)$ must be at least super-polynomial in $1/\varepsilon$. Even in the first non-trivial case, of $H = K_3$, these bounds remain the best known results. It is also a simple exercise to show that if H has h vertices, then the K_h removal lemma implies the H removal lemma, with similar bounds. Because of this, much of the study of the removal lemma is focused on cliques.

Another important class of results in extremal graph theory concerns structural results implied by minimum degree conditions. Notable examples include Dirac's theorem [8] on the existence of Hamiltonian cycles, its extension by Komlós, Sárközy, and Szemerédi [25] to powers of Hamiltonian cycles, and the Andrásfai–Erdős–Sós theorem [5] on when a K_r-free graph is $(r-1)$-partite.

In this paper, we study a natural minimum-degree version of the graph removal lemma. Formally, let us define $\delta(\varepsilon, H; \gamma)$ to be the maximum $\delta \in [0, 1]$ such that every n-vertex graph with fewer than δn^h copies of H and minimum degree at least γn can be made H-free by deleting at most εn^2 edges. Here, $\gamma \in [0, 1]$ is some constant which we think of as fixed, and we are interested in the behavior of $\delta(\varepsilon, H; \gamma)$ as $\varepsilon \to 0$. We remark that this function is non-decreasing in γ, and that setting $\gamma = 0$ recovers the earlier definition of $\delta(\varepsilon, H)$, that is, $\delta(\varepsilon, H; 0) = \delta(\varepsilon, H)$.

Our main results show that $\delta(\varepsilon, K_r; \gamma)$ is linear in ε if $\gamma > \frac{2r-5}{2r-3}$, but that it is super-polynomial in ε if $\gamma < \frac{2r-5}{2r-3}$. Formally, we first prove the following theorem, which asserts that $\delta(\varepsilon, K_r; \gamma)$ is linear in ε for $\gamma > \frac{2r-5}{2r-3}$.

Theorem 1.2. *For every $r \geq 3$, there exists $\mu_r > 0$ such that for all $\alpha, \varepsilon > 0$,*

$$\delta\left(\varepsilon, K_r; \frac{2r-5}{2r-3} + \alpha\right) \geq \mu_r \alpha \varepsilon,$$

meaning that $\delta(\varepsilon, K_r; \gamma)$ is linear in ε for all $\gamma > \frac{2r-5}{2r-3}$.

Our next result implies that below the threshold $\frac{2r-5}{2r-3}$, the K_r removal lemma must have super-polynomial bounds. In fact, we are able to relate the behavior of the restricted removal function $\delta(\varepsilon, K_r; \gamma)$ to that of the unrestricted *triangle* removal function $\delta(\varepsilon, K_3)$; since this function is known to have super-polynomial bounds, we conclude the same for $\delta(\varepsilon, K_r; \gamma)$. Formally, we prove the following result.

[1]If H is bipartite, then $\delta(\varepsilon, H) = \varepsilon^{\Theta(1)}$, i.e. the removal lemma has polynomial bounds [2]. The removal lemma is somewhat degenerate in case H is bipartite, since in this case the entire problem reduces to counting copies of bipartite graphs in dense graphs, which can be done with the method of Kővari–Sós–Turán [26]. This is closely related to a famous conjecture of Erdős–Simonovits and Sidorenko, see e.g. [7] for details.

Theorem 1.3. *For every integer $r \geq 3$ and every $\alpha > 0$, there exists some $C = C(r, \alpha) > 0$ such that for every $\varepsilon > 0$,*

$$\delta\left(\varepsilon, K_r; \frac{2r-5}{2r-3} - \alpha\right) \leq \delta(C\varepsilon, K_3).$$

In particular, $\delta(\varepsilon, K_r; \gamma)^{-1}$ is super-polynomial in ε^{-1} for fixed $\gamma < \frac{2r-5}{2r-3}$.

Somewhat surprisingly, our technique does not enable us to upper-bound $\delta(\varepsilon, K_r; \gamma)$ in terms of $\delta(\varepsilon, K_r)$ for $\gamma < \frac{2r-5}{2r-3}$. That is, to prove super-polynomial bounds on the restricted K_r removal function, we must use that such bounds are known for the unrestricted triangle removal function.

The results in Theorems 1.2 and 1.3 tell us that $\frac{2r-5}{2r-3}$ is a minimum degree threshold for the K_r removal lemma: below this threshold, the removal lemma has super-polynomial bounds, whereas above it we have linear bounds. We can formalize this notion of threshold as follows.

Definition 1.4. Let H be a graph. We define the *linear removal threshold* of H to be

$$\delta_{\text{lin-rem}}(H) = \inf\{\gamma \in [0, 1] : \text{there exists } \mu > 0 \text{ so that } \delta(\varepsilon, H; \gamma) \geq \mu\varepsilon \text{ for all } \varepsilon \in (0, 1)\}.$$

Similarly, we define the *polynomial removal threshold* of H to be

$$\delta_{\text{poly-rem}}(H) = \inf\{\gamma \in [0, 1] : \text{there exists } \mu > 0 \text{ so that } \delta(\varepsilon, H; \gamma) \geq \mu\varepsilon^{1/\mu} \text{ for all } \varepsilon \in (0, 1)\}.$$

These thresholds measure the weakest possible minimum degree condition one can impose in order to have, respectively, linear and polynomial bounds in the graph removal lemma for H. In this language, Theorems 1.2 and 1.3 can be rephrased as saying that

$$\delta_{\text{lin-rem}}(K_r) = \delta_{\text{poly-rem}}(K_r) = \frac{2r-5}{2r-3}.$$

In addition to determining the linear and polynomial removal thresholds for K_r, we also prove some results about $\delta_{\text{lin-rem}}(H)$ and $\delta_{\text{poly-rem}}(H)$ for more general classes of graphs, and make a number of conjectures about the relationship between these thresholds and other well-known thresholds in extremal graph theory. We refer to Section 4 for more details.

We remark that other versions of the graph removal lemma have been studied under certain minimum degree–like assumptions, such as in [17, 18]. More generally, there is a long line of work on how the numerical dependencies in the removal lemma (and in Szemerédi's regularity lemma) can be improved under certain assumptions about the host graph, e.g. [4, 16, 19, 20, 21, 28, 30].

The rest of the paper is organized as follows. In the next section, we prove Theorem 1.2. In Section 3, we prove Theorem 1.3 by exhibiting a specific graph of high minimum degree and poor K_r removal properties. We end with some concluding remarks, where we generalize these results and study $\delta_{\text{lin-rem}}(H)$ and $\delta_{\text{poly-rem}}(H)$ for general graphs H, and discuss the connections this problem has to the chromatic and homomorphism thresholds of graphs. For the sake of clarity of presentation, we omit all floor and ceiling signs whenever they are not crucial.

2 Above the threshold: the proof of Theorem 1.2

In this section, we prove Theorem 1.2. Unlike all known proofs of the full graph removal lemma, our proof uses only simple averaging arguments to find a small set of edges, each of which lies in many copies of K_r. We then show that removing all these edges deletes all copies of K_r in G. Crucially, these averaging arguments only work because of our minimum degree assumption; as shown by Theorem 1.3, they cannot possibly work if the minimum degree is below $(\frac{2r-5}{2r-3} - \alpha)n$ for any fixed $\alpha > 0$.

Here is a restatement of Theorem 1.2, restated to indicate what exactly we will prove in this section.

Theorem 2.1. *For every $r \geq 3$, there exists $\mu_r > 0$ such that the following holds for all $\alpha, \varepsilon > 0$. Let G be an n-vertex graph with minimum degree at least $(\frac{2r-5}{2r-3} + \alpha)n$, and suppose that G contains at most $(\mu_r \alpha \varepsilon) n^r$ copies of K_r. Then G can be made K_r-free by deleting at most εn^2 edges.*

We will need the following simple fact from calculus (or basic algebra).

Lemma 2.2. *For any $x \geq 4$, we have that*

$$x \frac{2x-5}{2x-3} \geq x - 2 + \frac{2}{5}.$$

Proof. Differentiating shows that the function $f(x) = x\frac{2x-5}{2x-3} - (x-2)$ is monotonically increasing, so its value for all $x \geq 4$ is lower-bounded by its value at $x = 4$, and $f(4) = \frac{2}{5}$. □

Our main technical result is the following lemma, which says that if G has minimum degree at least $(\frac{2r-5}{2r-3} + \alpha)n$, then every K_r in G contains a "popular" edge, namely an edge lying in $\Omega_r(\alpha n^{r-2})$ copies of K_r.

Lemma 2.3. *Let $r \geq 3$ and $\alpha > 0$. If G is an n-vertex graph with minimum degree at least $(\frac{2r-5}{2r-3} + \alpha)n$, then every K_r in G contains an edge which lies in at least $c_r \alpha n^{r-2}$ copies of K_r, for some constant $c_r > 0$ depending only on r.*

Proof. Fix a copy of K_r in G, and let its vertices be $v_1, \ldots, v_r$. For $i \in [r]$, let $V_i = N(v_i)$ denote the neighborhood of v_i. Note that by the minimum degree condition, we have that $|V_i| \geq (\frac{2r-5}{2r-3} + \alpha)n$ for each i. We will prove the following claim by induction.

Claim. *For each integer $0 \leq t \leq r-3$, there exists a set $S_t \subseteq [r]$ of size $|S_t| = r - t$ and at least $c_{r,t} n^t$ copies of K_t whose vertices lie in $\bigcap_{i \in S_t} V_i$, for some constant $c_{r,t} > 0$.*

Proof of claim. The base case $t = 0$ is trivial, since we simply take $S_0 = [r]$ and $c_{r,0} = 1$. Inductively, suppose we have found a set S_t with the desired properties, for $t \leq r-4$. Let Q be a copy of K_t with vertices in $\bigcap_{i \in S_t} V_i$. Since every vertex in Q has degree at least $(\frac{2r-5}{2r-3} + \alpha)n$, there are at most $(\frac{2}{2r-3} - \alpha)n < \frac{2}{2r-3}n$ vertices not adjacent to any given vertex in Q. Thus, the common neighborhood of Q has size at least

$$m \coloneqq n - t\left(\frac{2}{2r-3}n\right) = \frac{2r-3-2t}{2r-3}n.$$

Consider an auxiliary bipartite graph B_t, whose first part consists of S_t, whose second part consists of m arbitrary common neighbors of the vertices in Q, and where a vertex v in the second part is adjacent to a vertex i in the first part if $v \in V_i$. Each vertex in the first part of B_t has degree at least

$$\left(\frac{2r-5}{2r-3}+\alpha\right)n-(n-m) > \frac{2r-5-2t}{2r-3}n = \frac{2r-5-2t}{2r-3-2t}m.$$

The first part of B_t has $r-t$ vertices. Hence, the average degree in the second part of B_t is at least $(r-t)\frac{2r-5-2t}{2r-3-2t}$, which is at least $r-t-2+\frac{2}{5}$, by Lemma 2.2 applied to $x=r-t$ and using the fact that $t \le r-4$, which implies that $x \ge 4$. By Markov's inequality, at least $m/5$ vertices in the second part of B_t have degree at least $r-t-1$. Therefore, there are at least $(c_{r,t}n^t)(m/5)$ choices of a clique Q contained in $\bigcap_{i\in S_t} V_i$, and a common neighbor of Q that lies in at least $r-t-1$ of the sets V_i for $i \in S_t$. Hence, by the pigeonhole principle, for at least $c_{r,t}n^t m/(5(r-t))$ of these choices, the same subset of S_t of order $r-t-1$ is used. We let S_{t+1} be this subset, and let

$$c_{r,t+1} = \frac{c_{r,t}}{5(r-t)}\frac{m}{n} = \frac{(2r-3-2t)}{5(2r-3)(r-t)}c_{r,t},$$

so that there are at least $c_{r,t+1}n^{t+1}$ choices of a K_{t+1} whose vertices lie in $\bigcap_{i\in S_{t+1}} V_i$. This completes the proof of the claim. □

To conclude, we actually run the same argument for $t=r-3$, except that we need to be more careful about keeping track of the parameter α. Let $S=S_{r-3}$ be the set given by the claim for $t=r-3$, and let $c=c_{r,r-3}$. Let Q be a K_{r-3} whose vertices lie in $\bigcap_{i\in S} V_i$. Let B be the bipartite graph whose first part consists of three vertices, labeled by the elements of S, and whose second part consists of m common neighbors of the vertices in Q, where

$$m = n-(r-3)\left(\frac{2}{2r-3}-\alpha\right)n = \left(\frac{3}{2r-3}+(r-3)\alpha\right)n.$$

By the same argument as above, each vertex in the first part of B has degree at least

$$\begin{aligned}\left(\frac{2r-5}{2r-3}+\alpha\right)n-(n-m) &= \left(\frac{1}{2r-3}+(r-2)\alpha\right)n\\ &= \frac{\frac{1}{2r-3}+(r-2)\alpha}{\frac{3}{2r-3}+(r-3)\alpha}m\\ &= \frac{1+(r-2)(2r-3)\alpha}{3+(r-3)(2r-3)\alpha}m\\ &\ge \left(\frac{1}{3}+c'\alpha\right)m,\end{aligned}$$

for some constant $c'>0$ depending only on r. Therefore, the average degree in the second part of B is at least $1+3c'\alpha$. By Markov's inequality, this implies that at least $\frac{3}{2}c'\alpha m$

vertices in this part have at least two neighbors in the first part. Hence, there are at least $cn^{r-3} \cdot \frac{3}{2}c'\alpha m$ choices of a K_{r-3} and a vertex in its common neighborhood which lies in at least two of the three sets V_i for $i \in S$. By the pigeonhole principle, there is some $\{i, j\} \subset S$ such that $V_i \cap V_j$ contains at least $c_r \alpha n^{r-2}$ copies of K_{r-2}, where

$$c_r = \frac{cc'}{2}\frac{m}{n} = \frac{cc'}{2}\left(\frac{3}{2r-3} + (r-3)\alpha\right) \geq \frac{3cc'}{2(2r-3)}.$$

Therefore, the edge $\{v_i, v_j\}$ in our original K_r lies in at least $c_r \alpha n^{r-2}$ copies of K_r. □

Using Lemma 2.3, we can prove Theorem 2.1, and thus Theorem 1.2.

Proof of Theorem 2.1. Let $\mu_r = c_r / \binom{r}{2}$, where c_r is the constant from Lemma 2.3, and let $\delta = \mu_r \alpha \varepsilon$. Let G be an n-vertex graph with minimum degree at least $(\frac{2r-5}{2r-3} + \alpha)n$ and with at most δn^r copies of K_r.

Let E^* denote the set of edges in G which lie in at least $c_r \alpha n^{r-2}$ copies of K_r. Then the number of K_r in G is at least $\binom{r}{2}^{-1} c_r \alpha n^{r-2} |E^*|$, since each edge in E^* contributes at least $c_r \alpha n^{r-2}$ copies, and we count each copy at most $\binom{r}{2}$ times (once for each edge). By assumption, G has at most δn^r copies of K_r, and combining these bounds, we find that

$$|E^*| \leq \frac{\binom{r}{2}}{c_r}\frac{\delta}{\alpha} n^2 = \varepsilon n^2,$$

by our choice of $\delta = \mu_r \alpha \varepsilon$.

Additionally, by Lemma 2.3, we know that every K_r in G contains at least one edge from E^*. Hence, if we delete the edges in E^*, we are left with a K_r-free graph. Since we deleted at most εn^2 edges, this completes the proof. □

3 Below the threshold: the proof of Theorem 1.3

In this section, we prove Theorem 1.3 by constructing a graph with high minimum degree, few copies of K_r, but such that many edges must be removed to make it K_r-free. Formally, we will prove the following result.

Theorem 3.1. *For every integer $r \geq 3$, parameters $\alpha > 0$ and $\varepsilon > 0$, and all sufficiently large n, there exists an n-vertex graph G with minimum degree at least $(\frac{2r-5}{2r-3} - \alpha)n$ and with at most $\frac{\alpha^3}{(r/3)^r}\delta(\frac{(2r-3)^2}{\alpha^2}\varepsilon, K_3)n^r$ copies of K_r, but at least εn^2 edges must be deleted from G to make it K_r-free. Therefore,*

$$\delta\left(\varepsilon, K_r; \frac{2r-5}{2r-3} - \alpha\right) \leq \frac{\alpha^3}{(r/3)^r}\delta\left(\frac{(2r-3)^2}{\alpha^2}\varepsilon, K_3\right).$$

Note that Theorem 3.1 is somewhat stronger than Theorem 1.3, because we discarded the factor $\frac{\alpha^3}{(r/3)^r}$ in the statement of Theorem 1.3. We will first prove Theorem 3.1 in the case $r = 3$, and then show how to extend this construction to prove Theorem 3.1 for all $r \geq 4$.

We will first need the following simple and well-known lemma, which says that one can convert any construction for the triangle removal lemma into a tripartite construction with similar parameter dependencies. We remark that this lemma is not fully optimized, and one could obtain better constants through a more careful argument.

Lemma 3.2. *Suppose that H_0 is an n_0-vertex graph with at most δn_0^3 triangles, but such that at least εn_0^2 edges must be deleted to make H_0 triangle-free. Then there exists a tripartite graph H on $n := 3n_0$ vertices with at most $\frac{2}{9}\delta n^3$ triangles, such that at least $\frac{1}{9}\varepsilon n^2$ edges must be deleted to make H triangle-free.*

Proof. Consider the tensor product $H = H_0 \times K_3$, which is the graph whose vertices are pairs $(v, x) \in V(H) \times [3]$, and where two vertices (v, x) and (w, y) are adjacent if and only if $x \neq y$ and $v \sim w$ in H_0. We claim that H has the desired properties.

Indeed, by definition, H is tripartite and has $n = 3n_0$ vertices. Moreover, each triangle in H_0 yields precisely six triangles in H, so H has at most $6\delta n_0^3 = \frac{2}{9}\delta n^3$ triangles. To conclude, suppose that $E^* \subseteq E(H)$ is a set of edges whose deletion makes H triangle-free. Let $E_0^* \subseteq E(H_0)$ denote the edges of H_0 obtained by deleting the second coordinate of every vertex in every edge of E^*; in particular, $|E_0^*| \leq |E^*|$. We claim that deleting the edges in E_0^* makes H_0 triangle-free. Indeed, if $\{v_1, v_2, v_3\} \subseteq V(H_0)$ form a triangle in H_0 after the deletion of E_0^*, then we see that no edge in E^* can be of the form $\{(v_i, x), (v_j, y)\}$ for any $i \neq j$ and $x \neq y$. In particular, we find that $\{(v_1, 1), (v_2, 2), (v_3, 3)\}$ is a triangle in H whose edges do not intersect E^*, contradicting the assumption that the deletion of E^* destroyed all triangles in H. Hence, by the defining property of H_0, we conclude that

$$|E^*| \geq |E_0^*| \geq \varepsilon n_0^2 = \frac{1}{9}\varepsilon n^2,$$

as claimed. □

The next lemma is simply a restatement of Theorem 3.1 in the case $r = 3$. We state it as a separate lemma because the $r = 3$ construction will be used as a black box in the construction for larger values of r.

Lemma 3.3. *For all $\alpha > 0$ and $\varepsilon > 0$ and all sufficiently large n, there exists a tripartite n-vertex graph G_0 with minimum degree at least $(\frac{1}{3} - \alpha)n$ and with at most $\alpha^3\delta(9\varepsilon/\alpha^2, K_3)n^3$ triangles, but at least εn^2 edges must be deleted from G_0 to make it triangle-free.*

Proof. We first claim that for all sufficiently large n, there exists an n-vertex graph with at most $2\delta(\varepsilon, K_3)n^3$ triangles such that at least εn^2 edges must be deleted to make it triangle-free. Indeed, by the definition of $\delta(\varepsilon, K_3)$, there must exist a graph H_0 on some fixed number n_0 of vertices with fewer than $\frac{3}{2}\delta(\varepsilon, K_3)n_0^3$ triangles such that at least εn_0^2 edges must be deleted to make it triangle-free. It is simple to check that for any $s \geq 1$, the balanced blowup $H_0[s]$ will have the same properties. Therefore, for n sufficiently large relative to n_0, we may take the blowup $H_0[\lfloor n/n_0 \rfloor]$ and add to it $n - n_0\lfloor n/n_0 \rfloor$ isolated vertices to obtain the desired graph.

Therefore, by Lemma 3.2 applied with parameters $\varepsilon' = 9\varepsilon/\alpha^2$ and $n' = \alpha n/3$, there exists a tripartite graph H on αn vertices with fewer than $\frac{2}{9} \cdot 2\delta(\varepsilon', K_3)(n')^3 < \delta(\varepsilon', K_3)(n')^3$ triangles such that at least $\frac{1}{9}\varepsilon'(n')^2$ edges must be deleted to make H triangle-free. Let Γ be a balanced blowup of the path with two edges, blown up so that it has $(1-\alpha)n$ vertices. Let G_0 be the graph obtained by taking the disjoint union of H and Γ, and placing a complete bipartite graph between the ith part of H and the ith part of Γ, for $i \in [3]$. Then G_0 is tripartite by definition. The construction is shown in Figure 1.

Since every vertex in G_0 is adjacent to all vertices in at least one part of Γ, we see that every vertex in G_0 has degree at least $\frac{1}{3}(1-\alpha)n > (\frac{1}{3} - \alpha)n$. Moreover, we see that every triangle in G_0 is actually contained in H, so the number of triangles in G_0 is at most $\delta(\varepsilon', K_3)(n')^3 = \alpha^3\delta(9\varepsilon/\alpha^2, K_3)n^3$. Finally, if we delete some edges to make G_0 triangle-free, we must in particular make H triangle-free. Therefore, the number of edges needed to make G_0 triangle-free is at least $\frac{1}{9}\varepsilon'(n')^2 = \frac{1}{9} \cdot \frac{9\varepsilon}{\alpha^2}(\alpha n)^2 = \varepsilon n^2$. □

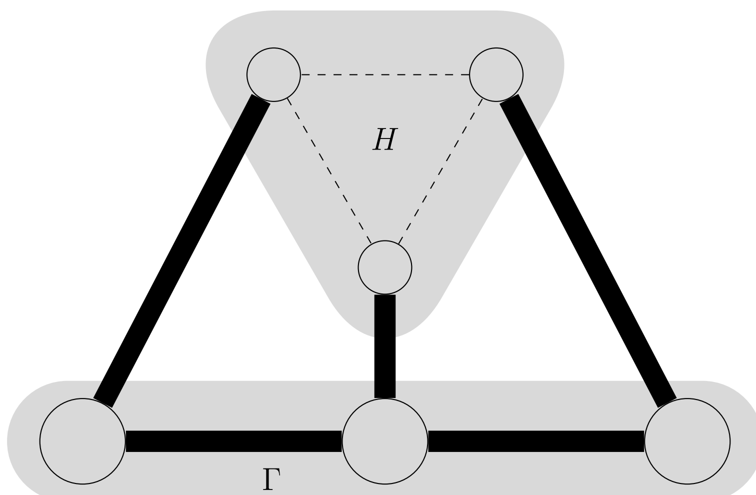

Figure 1: The construction in Lemma 3.3. Solid edges represent complete bipartite graphs, and dashed edges represent the edges of H as given by Lemma 3.2.

With this result, we are now ready to prove Theorem 3.1, and thus Theorem 1.3.

Proof of Theorem 3.1. If $r = 3$, then the result is precisely the statement of Lemma 3.3. So we henceforth assume that $r \geq 4$. Let G_0 be the graph from Lemma 3.3, applied with parameters $\alpha, \varepsilon' = \left(\frac{2r-3}{3}\right)^2 \varepsilon$, and $n' = \frac{3}{2r-3}n$. Let K be a complete $(r-3)$-partite graph where each part has $\frac{2}{2r-3}n$ vertices, and let G be the join of G_0 and K, i.e. the graph obtained by connecting every vertex in K to every vertex in G_0. Then G has $(r-3)\frac{2}{2r-3}n + \frac{3}{2r-3}n = n$ vertices. In G, every vertex coming from K has degree $(r-4)\frac{2}{2r-3}n + \frac{3}{2r-3}n = \frac{2r-5}{2r-3}n$, and every vertex coming from G_0 has degree at least

$$(r-3)\frac{2}{2r-3}n + \left(\frac{1}{3} - \alpha\right)\frac{3}{2r-3}n > \left(\frac{2r-5}{2r-3} - \alpha\right)n,$$

hence G has the desired minimum degree condition. Additionally, since G_0 is tripartite, it is K_t-free for all $t \geq 4$. Therefore, we see that every K_r in G must consist of a triangle in G and $r-3$ vertices from K. Hence, the number of K_r in G is at most

$$\alpha^3 \delta\left(\frac{9\varepsilon'}{\alpha^2}, K_3\right)(n')^3 \cdot \left(\frac{2n}{2r-3}\right)^{r-3} \leq \frac{\alpha^3}{(r/3)^r}\delta\left(\frac{(2r-3)^2}{\alpha^2}\varepsilon, K_3\right)n^r.$$

Moreover, if we delete some edges to make G be K_r-free, we must in particular make G_0 triangle-free. Thus, the number of edges that must be deleted is at least $\varepsilon'(n')^2 = \varepsilon n^2$. $\square$

4 Concluding remarks

4.1 The removal thresholds for other graphs

Recall the definition of the linear and polynomial removal thresholds from Definition 1.4. In this subsection, we make some remarks about the values of $\delta_{\text{lin-rem}}(H)$ and $\delta_{\text{poly-rem}}(H)$ for more general classes of graphs.

We begin by observing, directly from the definition, that $\delta_{\text{poly-rem}}(H) \leq \delta_{\text{lin-rem}}(H)$ for any H, since a linear bound on the removal lemma is in particular a polynomial bound. Moreover, if $\text{ex}(n, H)$ denotes the *extremal number* of H, that is the maximum number of edges in an H-free graph on n vertices, and if $\pi(H) := \lim_{n\to\infty} \binom{n}{2}^{-1} \text{ex}(n, H)$ is the *Turán density* of H, then we have that $\delta_{\text{lin-rem}}(H) \leq \pi(H)$. Indeed, this follows from the Erdős–Simonovits supersaturation theorem [13], which implies that for any $\alpha > 0$, there exists some $\delta_0 > 0$ such that every n-vertex graph G with minimum degree at least $(\pi(H) + \alpha)n$ has at least $\delta_0 n^{|V(H)|}$ copies of H.

Our first result in this section shows that $\delta_{\text{poly-rem}}$ is invariant under a certain natural relation on graphs. Recall that a graph homomorphism $H_2 \to H_1$ is a function $V(H_2) \to V(H_1)$ that maps every edge of H_2 to an edge of H_1.

Proposition 4.1. *If H_1 is a subgraph of H_2 and there exists a homomorphism $H_2 \to H_1$, then*

$$\delta_{\text{poly-rem}}(H_1) = \delta_{\text{poly-rem}}(H_2).$$

Proof. Let $h_1 = |V(H_1)|, h_2 = |V(H_2)|$, and fix a graph homomorphism $\varphi : H_2 \to H_1$. Let $s_1, \ldots, s_{h_1}$ be the sizes of the fibers of φ, i.e. $s_i = |\varphi^{-1}(v_i)|$ for $v_i \in V(H_1)$. Let $s = \max\{s_1, \ldots, s_{h_1}\}$, so that H_2 is a subgraph of the blowup $H_1[s]$.

We first prove that $\delta_{\text{poly-rem}}(H_1) \leq \delta_{\text{poly-rem}}(H_2)$. For this, let G be an n-vertex graph with minimum degree γn that has at most δn^{h_1} copies of H_1, such that at least εn^2 edges must be removed from G to make it H_1-free. By the same argument as in the first paragraph of Lemma 3.3, we may assume without loss of generality that n is sufficiently large. The number of non-injective homomorphisms $H_1 \to G$ is at most $\binom{h_1}{2} n^{h_1 - 1}$, which is at most δn^{h_1} for n sufficiently large. So the number of homomorphisms $H_1 \to G$ is at most $2\delta n^{h_1}$, which implies that the number of copies of H_1 in the blowup $G[s]$ is at most $2\delta(sn)^{h_1}$. Every copy of H_1 in $G[s]$ can be extended to a copy of H_2 in at most $(sn)^{h_2-h_1}$ ways, which implies that $G[s]$ contains at most $2\delta(sn)^{h_2}$ copies of H_2. An H_2-free subgraph $\Gamma \subseteq G[s]$ yields an H_1-free subgraph of G by keeping those edges of G all of whose lifts are present in Γ; any copy of H_1 in this subgraph would lift to a copy of $H_1[s] \supseteq H_2$ in Γ, a contradiction. This implies that at least $\varepsilon n^2 = \frac{\varepsilon}{s^2}(sn)^2$ edges must be deleted from $G[s]$ to make it H_2-free. Since $G[s]$ has minimum degree $\gamma(sn)$ and sn vertices, we conclude that

$$\delta(\varepsilon, H_1; \gamma) \geq \frac{1}{2}\delta\left(\frac{\varepsilon}{s^2}, H_2; \gamma\right),$$

and therefore that $\delta_{\text{poly-rem}}(H_1) \leq \delta_{\text{poly-rem}}(H_2)$.

For the reverse inequality, now let G be an n-vertex graph with minimum degree γn that has at most δn^{h_2} copies of H_2, such that at least εn^2 edges must be removed from G to make it H_2-free, and we again assume that n is sufficiently large. Since an H_1-free subgraph of G is also H_2-free, we see that at least εn^2 edges must be removed from G to make it H_1-free. Let $m = s_1 \cdots s_{h_1}$. We claim that G has at most $3h_1^{h_1}\delta^{1/m}n^{h_1}$ copies of H_1. For if not, then we can randomly partition $V(G)$ into h_1 sets to obtain an h_1-partite subgraph G' with at least $3\delta^{1/m}n^{h_1}$ canonical copies of H_1, where we say that a copy is *canonical* if the ith vertex of H_1 lies in the ith part of G' for all $i \in [h_1]$. Let $\mathcal{H}$ be the h_1-uniform hypergraph on $V(G)$ whose edges are the canonical copies of H_1 in G', so that $\mathcal{H}$ has edge density at least $\eta \coloneqq 3\delta^{1/m}$. An argument of Erdős [10] (see also [36]) implies that if $\mathcal{K}$ is a complete h_1-partite h_1-uniform hypergraph with m edges, then there at least $\eta^m n^{|V(\mathcal{K})|}$ homomorphisms $\mathcal{K} \to \mathcal{H}$. For n sufficiently large, and taking $\mathcal{K}$ to have parts of size $s_1, \ldots, s_{h_1}$, we conclude that $\mathcal{H}$ has at least $\frac{1}{2}\eta^m n^{h_2} > \delta n^{h_2}$ copies of $\mathcal{K}$. This implies that G' contains more than δn^{h_2} copies of H_2, a contradiction. We conclude that

$$\delta(\varepsilon, H_2; \gamma) \geq \frac{1}{3h_1^{h_1}}\delta(\varepsilon, H_1; \gamma)^m,$$

and therefore that $\delta_{\text{poly-rem}}(H_2) \leq \delta_{\text{poly-rem}}(H_1)$. $\square$

Using Proposition 4.1, we can determine the polynomial removal threshold of many graphs. For instance, if H is a non-bipartite graph whose clique number $\omega(H)$ equals its chromatic number $\chi(H)$, then

$$\delta_{\text{poly-rem}}(H) = \frac{2\omega(H) - 5}{2\omega(H) - 3},$$

since $K_{\omega(H)}$ is a subgraph of H and there is a homomorphism $H \to K_{\omega(H)}$. In particular, we are able to determine the polynomial removal threshold of all perfect graphs.

Additionally, Proposition 4.1 implies that $\delta_{\text{poly-rem}}(H) = \delta_{\text{poly-rem}}(H[s])$ for any graph H and any integer $s \geq 1$, that is that the polynomial removal threshold is invariant under blowups. More generally, we recall that the *core* of a graph H is defined as the inclusion-minimal subgraph K such that there exists a homomorphism $H \to K$; see e.g. [24] for more on this concept. Then by Proposition 4.1, we see that $\delta_{\text{poly-rem}}(H) = \delta_{\text{poly-rem}}(K)$ for any graph H and its core K. In other words, the polynomial removal threshold of a graph is completely determined by that of its core.

In contrast to the above results, the linear removal threshold does not satisfy such a nice invariance property. Indeed, our next result demonstrates that for any $r \geq 3$ and $s \geq 2$, the complete multipartite graph $K_r[s]$ has linear removal threshold $\frac{r-2}{r-1}$; this equals the Turán density $\pi(K_r[s])$, and is strictly larger than $\delta_{\text{lin-rem}}(K_r) = \frac{2r-5}{2r-3}$.

Proposition 4.2. *For any $r \geq 3$ and $s \geq 2$,*

$$\delta_{\text{lin-rem}}(K_r[s]) = \frac{r-2}{r-1}.$$

Proof. As remarked above, the inequality $\delta_{\text{lin-rem}}(K_r[s]) \leq \frac{r-2}{r-1}$ follows from the Erdős–Stone theorem [14], which says that $\pi(K_r[s]) = \frac{r-2}{r-1}$. For the reverse inequality, it suffices to construct an n-vertex graph G with minimum degree at least $\frac{r-2}{r-1}n$ and fewer than δn^{rs} copies of $K_r[s]$, but such that at least εn^2 edges must be deleted to make it $K_r[s]$-free, where ε cannot be taken to depend linearly on δ. We may assume throughout that n is sufficiently large.

To do so, we let $T(n, r-1)$ denote the Turán graph, that is the complete $(r-1)$-partite graph with parts of size $|S_1| = \cdots = |S_{r-1}| = \frac{n}{r-1}$ (where we assume for simplicity that $2(r-1)$ divides n). Inside the part S_1 of $T(n, r-1)$, we place a random graph $G(\frac{n}{r-1}, p)$ for some fixed $p \in (0, \frac{1}{4s^2})$; in other words, we connect every pair in S_1 by an edge with probability p, independently over all these choices. We let G be the resulting graph. Then we immediately see that G has minimum degree at least $\frac{r-2}{r-1}n$, since that was the case in $T(n, r-1)$. For clarity, we now fix $G[S_1]$ to be a graph where every vertex has degree $(p + o(1))|S_1|$ and the $K_{s,s}$ density in $G[S_1]$ is $p^{s^2} + o(1)$. This is possible since both these properties hold in $G(\frac{n}{r-1}, p)$ with high probability. Note that every copy of $K_r[s]$ in G must contain a copy of $K_{s,s}$ in S_1. Therefore, G contains at most δn^{rs} copies of $K_r[s]$, where $\delta = p^{s^2}$.

Now, suppose that G' is a $K_r[s]$-free subgraph of G with the maximum possible number of edges. Recall that in G, every vertex of S_1 has degree at most $2p\frac{n}{r-1}$ inside S_1. Therefore, if any vertex $v \in S_1$ has more than $2p\frac{n}{r-1}$ non-neighbors in $S_2 \cup \cdots \cup S_{r-1}$, we can find a $K_r[s]$-free subgraph of G with more edges than G' by deleting all edges incident to v in S_1 and adding all missing edges to $S_2 \cup \cdots \cup S_{r-1}$.

We now show $G'[S_1]$ is $K_{s,s}$-free. Suppose not, and consider a copy of $K_{s,s}$ in $G'[S_1]$. Since every vertex of this $K_{s,s}$ has at most $2p\frac{n}{r-1}$ non-neighbors in $S_2 \cup \cdots \cup S_{r-1}$, we find that the vertices of the $K_{s,s}$ have in total at most $2s \cdot 2p\frac{n}{r-1}$ non-neighbors outside S_1. As $2s \cdot 2p\frac{n}{r-1} \leq \frac{1}{2}|S_i|$ since $p < 1/(4s^2) \leq 1/(8s)$, we see that there are subsets $S_i' \subseteq S_i$ for $2 \leq i \leq r-1$ such that $|S_i'| = \frac{1}{2}|S_i|$ and every vertex in S_i' is complete to the vertices in the $K_{s,s}$. If $r = 3$, we arrive at a contradiction as $|S_2'| \geq s$ for n sufficiently large, so the $K_{s,s}$ together with s vertices from S_2' yields a copy of $K_3[s]$. We now assume $r \geq 4$. As G' is $K_r[s]$-free, $G'[S_2' \cup \cdots \cup S_{r-1}']$ does not contain a copy of $K_{r-2}[s]$. When taking s random vertices from each S_i', at least one edge must be deleted in obtaining G' from G, which implies by averaging that there are at least $\frac{1}{s^2}|S_2'||S_3'| = \frac{1}{4s^2}(\frac{n}{r-1})^2$ edges deleted to obtain G' from G. However, this is more than the number of edges in $G[S_1]$, so the graph obtained from G by deleting all edges inside S_1 has more edges than G' and is also $K_r[s]$-free, a contradiction. Thus, $G'[S_1]$ is $K_{s,s}$-free.

Any $K_{s,s}$-free subgraph of $G[S_1]$ has at most $O(n^{2-1/s})$ edges by the Kővári–Sós–Turán theorem [26], so we conclude that at least εn^2 edges must have been deleted when going from G to G', where $\varepsilon = \frac{p}{4r^2}$. Since G has at most δn^{rs} copies of $K_r[s]$, where $\delta = p^{s^2}$, we see that the dependence between ε and δ cannot be linear. $\square$

To conclude this subsection, we turn our attention to cycles. Since an even cycle is bipartite, its Turán density is 0, and hence so are its linear and polynomial removal thresholds. For odd cycles, we are able to prove the following lower bound, though we do not know if it is tight. The construction is a natural and simple generalization of that in Lemma 3.3, which corresponds to the case $k = 1$ in the following theorem.

Theorem 4.3. *For every positive integer k, we have that $\delta_{\text{poly-rem}}(C_{2k+1}) \geq \frac{1}{2k+1}$.*

Proof. For every $k \geq 1$, every sufficiently small $\varepsilon_0 > 0$, and every sufficiently large m, Alon [2] constructed a graph H on m vertices with vertex set $V_0 \sqcup \cdots \sqcup V_{2k}$ with the following properties. Every edge in H goes between V_i and V_{i+1} for some i (with the indices taken modulo $2k+1$), at least $\varepsilon_0 m^2$ edges must be removed from H to make it C_{2k+1}-free, and H has at most δm^{2k+1} copies of C_{2k+1}, where $\delta = {\varepsilon_0}^{-c\log\varepsilon_0}$, for some constant $c > 0$ depending only on k.

We set $\varepsilon_0 = \varepsilon/\alpha^2$ and $m = \alpha n$ and adjoin to this graph H a balanced blowup Γ of the path on $2k+1$ vertices, blown up so that Γ has $(1-\alpha)n$ vertices in total. Finally, we place a complete bipartite graph between the ith part of Γ and the ith part of H. Then the resulting graph has minimum degree at least $(\frac{1}{2k+1} - \alpha)n$, and every C_{2k+1} in this graph is contained in H. Therefore, this resulting graph has at most $\delta m^{2k+1} \leq \delta n^{2k+1}$ copies of C_{2k+1}, but at least $\varepsilon_0 m^2 = \varepsilon n^2$ edges must be removed to make it C_{2k+1}-free. Since $1/\delta$ is super-polynomial in $1/\varepsilon$, this gives the theorem. $\square$

4.2 The popular edge threshold

Our proof of Theorem 1.2 used Lemma 2.3, which is a very natural way of proving linear bounds on the removal lemma. Recall that Lemma 2.3 says that if G has minimum degree at least $(\frac{2r-5}{2r-3} + \alpha)n$, then every copy of K_r in G has a "popular" edge, namely an edge that lies in $\Omega(n^{r-2})$ copies of K_r. Given this statement, the proof of Theorem 1.2 is simple, since we simply delete a popular edge from each copy of K_r, which necessarily yields linear bounds for the K_r removal lemma.

This discussion naturally leads to the following definition.

Definition 4.4. Let H be a graph. The *popular edge threshold* of H is defined as the infimum of all $\gamma \in [0, 1]$ for which the following holds. There exists $\beta = \beta(\gamma) > 0$ such that for every n-vertex graph G with minimum degree at least γn, every copy of H in G contains an edge which lies in at least $\beta n^{|V(H)|-2}$ copies of H.

From the same argument as before, we see that $\delta_{\text{lin-rem}}(H) \le \delta_{\text{pop-edge}}(H)$, and Lemma 2.3 shows that $\delta_{\text{pop-edge}}(K_r) \le \frac{2r-5}{2r-3}$, which is a tight bound by Theorem 1.3. However, in general, $\delta_{\text{pop-edge}}(H)$ can be strictly larger than $\delta_{\text{lin-rem}}(H)$, as shown in the following proposition.

Proposition 4.5. *Let H be a graph with no isolated vertices. Then $\delta_{\text{pop-edge}}(H) = 0$ if and only if H is bipartite and has minimum degree* 1.

Recall that if H is bipartite, then $\pi(H) = 0$ and therefore $\delta_{\text{lin-rem}}(H) = 0$ as well. Thus, every bipartite graph with minimum degree at least 2 is an example of a graph whose popular edge threshold is strictly larger than its linear removal threshold.

Proof of Proposition 4.5. Let $h = |V(H)|$. First suppose that H is not bipartite and has minimum degree at least 2. For any integer $s \ge 1$, consider the blowup $C_{h^2}[s]$, where we label the parts $0, 1, \ldots, h^2 - 1$. Form a graph G by adding to $C_{h^2}[s]$ a single copy of H, with one vertex in each of the parts labeled $0, h, 2h, \ldots, h^2 - h$. Then G has $n \coloneqq h^2 s$ vertices and minimum degree $2s = \frac{2}{h^2}n$. If H is not bipartite, then the only odd cycles of length at most h in G are in the added copy of H. Since any odd cycle can be extended to a copy of H in at most $O_H(n^{h-3})$ ways, we conclude that G has at most $O_H(n^{h-3})$ copies of H. For any fixed $\beta > 0$, by letting s (and thus n) be sufficiently large, this implies that G contains fewer than βn^{h-2} copies of H, and in particular the added copy of H has no popular edge. Similarly, if H is bipartite and has minimum degree at least 2, then every edge of H lies in a cycle of length at most h. For any edge in the added copy of H, the only cycles of length at most h it participates in are in the added copy, so any such edge appears in at most $O_H(n^{h-4})$ copies of H, again showing that this copy of H has no popular edge. This implies that $\delta_{\text{pop-edge}}(H) \ge \frac{2}{h^2} > 0$.

For the reverse implication, suppose that H is bipartite and has a vertex of degree 1. Let the bipartition of H be $V(H) = A \cup B$, where $|A| = a, |B| = b$, and $h = a + b$, and fix some vertex $u \in b$ of degree 1. Let $\gamma > 0$, and let G be a graph with n vertices and minimum degree at least γn. We claim that for any vertex $v \in V(G)$, there are at least $\Omega_{\gamma,a,b}(n^{h-2})$ copies of $K_{a,b-1}$ in G which contain v as one of the a vertices in the first part. This follows from the Kővári–Sós–Turán theorem [26] on the problem of Zarankiewicz. By this theorem, we obtain many copies of $K_{a-1,b-1}$ in the auxiliary bipartite graph with parts $V(G) \setminus \{v\}$ and $N(v)$, whose edges are given by adjacency in G.

This shows that every edge in G lies in at least $\Omega_{\gamma,a,b}(n^{h-2})$ copies of H, since H is a subgraph of the graph gotten by adding a pendant edge to $K_{a,b-1}$. In particular, in any copy of H in G, any edge is a popular edge, showing that $\delta_{\text{pop-edge}}(H) \le \gamma$. Letting $\gamma \to 0$ gives the desired result. $\square$

This example demonstrates that our approach to upper-bounding the linear removal threshold via the popular edge threshold will not give tight bounds in general. Nevertheless, we think that it is interesting to study when this approach will yield a tight bound, i.e. to understand when $\delta_{\text{lin-rem}}(H) = \delta_{\text{pop-edge}}(H)$.

Question 4.6. *For which graphs H does $\delta_{\text{lin-rem}}(H) = \delta_{\text{pop-edge}}(H)$?*

4.3 The chromatic and homomorphism thresholds

The number $\frac{2r-5}{2r-3}$, which emerges from Theorems 1.2 and 1.3 as the linear and polynomial removal threshold of K_r, is a well-known number in extremal graph theory. Indeed, it turns out that $\frac{2r-5}{2r-3}$ is also both the *chromatic threshold* and *homomorphism threshold* of K_r. These are defined as follows. For a family of graphs $\mathcal{F}$ and a parameter $\gamma \in [0, 1]$, let $\mathcal{G}(\mathcal{F}, \gamma)$ denote the set of $\mathcal{F}$-free graphs G with minimum degree at least $\gamma|V(G)|$.

Definition 4.7. Let $\mathcal{F}$ be a family of graphs. The *chromatic threshold* of $\mathcal{F}$ is the number

$$\delta_\chi(\mathcal{F}) = \inf\{\gamma \in [0, 1] : \text{there exists } M > 0 \text{ such that } \chi(G) \leq M \text{ for all } G \in \mathcal{G}(\mathcal{F}, \gamma)\}$$

and the *homomorphism threshold* of $\mathcal{F}$ is

$$\begin{aligned}\delta_{\text{hom}}(\mathcal{F}) = \inf\{\gamma \in [0, 1] : \text{there exists an } \mathcal{F}\text{-free graph } G_0 \text{ such that for all } G \in \mathcal{G}(\mathcal{F}, \gamma), \\ G \text{ has a homomorphism to } G_0\}.\end{aligned}$$

If $\mathcal{F} = \{F\}$ consists of a single graph, we denote these by $\delta_\chi(F)$ and $\delta_{\text{hom}}(F)$.

In other words, the chromatic threshold measures what minimum degree conditions force an $\mathcal{F}$-free graph to have a homomorphism to a graph of bounded order, and the homomorphism threshold further requires this bounded graph to itself be $\mathcal{F}$-free. Due to the efforts of many researchers [12, 23, 29, 31, 32, 34], it is now known that

$$\delta_\chi(K_r) = \delta_{\text{hom}}(K_r) = \frac{2r-5}{2r-3}.$$

Despite the fact that we get the same answer for all cliques, we were not able to find any *a priori* relationship between the two removal thresholds and the chromatic or homomorphism thresholds. Moreover, such a relationship does not hold in general. For instance, Thomassen [35] proved that $\delta_\chi(C_{2k+1}) = 0$ for all $k \geq 2$, while Theorem 4.3 shows that $\delta_{\text{lin-rem}}(C_{2k+1}) \geq \delta_{\text{poly-rem}}(C_{2k+1}) \geq \frac{1}{2k+1}$. Thus, the chromatic threshold of C_{2k+1} is different from both removal thresholds. The homomorphism threshold of C_{2k+1} is unknown for any $k \geq 2$, though Letzter–Snyder [27] and Ebsen–Schacht [9] proved that $\delta_{\text{hom}}(C_{2k+1}) \leq \frac{1}{2k+1}$, and that $\delta_{\text{hom}}(\{C_3, C_5, \ldots, C_{2k+1}\}) = \frac{1}{2k+1}$. It would be very interesting to determine if in fact $\delta_{\text{hom}}(C_{2k+1}) = \frac{1}{2k+1}$, as well as what the values of $\delta_{\text{poly-rem}}(C_{2k+1})$ and $\delta_{\text{lin-rem}}(C_{2k+1})$ are. As a first pass, we make the following conjecture.

Conjecture 4.8. $\delta_{\text{poly-rem}}(C_{2k+1}) > \delta_{\text{hom}}(C_{2k+1})$ *for all* $k \geq 2$.

Odd cycles provide an example of graphs where the polynomial removal threshold is strictly larger than the chromatic threshold. In the other direction, we see from Proposition 4.1 that $\delta_{\text{poly-rem}}(K_3[2]) = \delta_{\text{poly-rem}}(K_3) = \frac{1}{3}$, while it is well-known that $\delta_\chi(K_3[2]) = \frac{1}{2}$ (see [1], where this is stated as a folklore result). Thus, $K_3[2]$ (or more generally a non-trivial balanced blowup of a clique) is an example of a graph whose polynomial removal threshold is strictly smaller than its chromatic threshold. However, in this case, we have the curious situation that $\delta_{\text{lin-rem}}(K_3[2]) = \delta_\chi(K_3[2])$, by Proposition 4.2. We conjecture that this is a coincidence, and that in general, the four thresholds have nothing to do with one another.

Conjecture 4.9. *There exists a graph H for which $\delta_\chi(H), \delta_{\text{hom}}(H), \delta_{\text{lin-rem}}(H), \delta_{\text{poly-rem}}(H)$ are all distinct.*

More generally, the numbers $\delta_\chi(H), \delta_{\text{hom}}(H), \delta_{\text{lin-rem}}(H), \delta_{\text{poly-rem}}(H)$ may appear in any order in $[0,1]$, subject to the constraints $\delta_\chi(H) \le \delta_{\text{hom}}(H)$ and $\delta_{\text{poly-rem}}(H) \le \delta_{\text{lin-rem}}(H)$.

To conclude, we remark that Allen, Böttcher, Griffiths, Kohayakawa, and Morris [1] determined $\delta_\chi(H)$ for all graphs H. Moreover, they showed that if $\chi(H) = r \ge 3$, then $\delta_\chi(H)$ can take only one of three values, namely

$$\delta_\chi(H) \in \left\{ \frac{r-3}{r-2}, \frac{2r-5}{2r-3}, \frac{r-2}{r-1} \right\}.$$

We are not bold enough to make any specific conjecture about the removal thresholds for arbitrary graphs. But we do leave the following open question, inspired by the Allen–Böttcher–Griffiths–Kohayakawa–Morris theorem.

Question 4.10. *Is it the case that for each $r \ge 3$, there exists a finite set $\Delta_r \subset [0,1]$ such that $\delta_{\text{poly-rem}}(H) \in \Delta_r$ for every graph H with $\chi(H) = r$? What if we replace $\delta_{\text{poly-rem}}(H)$ by $\delta_{\text{lin-rem}}(H)$, by $\delta_{\text{pop-edge}}(H)$, or by $\delta_{\text{hom}}(H)$?*

References

[1] P. Allen, J. Böttcher, S. Griffiths, Y. Kohayakawa, and R. Morris, The chromatic thresholds of graphs, *Adv. Math.* **235** (2013), 261–295.

[2] N. Alon, Testing subgraphs in large graphs, *Random Structures Algorithms* **21** (2002), 359–370.

[3] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, and R. Yuster, The algorithmic aspects of the regularity lemma, *J. Algorithms* **16** (1994), 80–109.

[4] N. Alon, E. Fischer, and I. Newman, Efficient testing of bipartite graphs for forbidden induced subgraphs, *SIAM J. Comput.* **37** (2007), 959–976.

[5] B. Andrásfai, P. Erdős, and V. T. Sós, On the connection between chromatic number, maximal clique and minimal degree of a graph, *Discrete Math.* **8** (1974), 205–218.

[6] D. Conlon and J. Fox, Graph removal lemmas, in *Surveys in combinatorics 2013*, *London Math. Soc. Lecture Note Ser.*, vol. 409, Cambridge Univ. Press, Cambridge, 2013, pp. 1–49.

[7] D. Conlon, J. Fox, and B. Sudakov, An approximate version of Sidorenko's conjecture, *Geom. Funct. Anal.* **20** (2010), 1354–1366.

[8] G. A. Dirac, Some theorems on abstract graphs, *Proc. London Math. Soc. (3)* **2** (1952), 69–81.

[9] O. Ebsen and M. Schacht, Homomorphism thresholds for odd cycles, *Combinatorica* **40** (2020), 39–62.

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

[11] 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.

[12] P. Erdős and M. Simonovits, On a valence problem in extremal graph theory, *Discrete Math.* **5** (1973), 323–334.

[13] P. Erdős and M. Simonovits, Supersaturated graphs and hypergraphs, *Combinatorica* **3** (1983), 181–192.

[14] P. Erdös and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087–1091.

[15] J. Fox, A new proof of the graph removal lemma, *Ann. of Math. (2)* **174** (2011), 561–579.

[16] J. Fox, L. Gishboliner, A. Shapira, and R. Yuster, The removal lemma for tournaments, *J. Combin. Theory Ser. B* **136** (2019), 110–134.

[17] J. Fox, H. Huang, and B. Sudakov, On graphs decomposable into induced matchings of linear sizes, *Bull. Lond. Math. Soc.* **49** (2017), 45–57.

[18] J. Fox and P.-S. Loh, On a problem of Erdős and Rothschild on edges in triangles, *Combinatorica* **32** (2012), 619–628.

[19] J. Fox, J. Pach, and A. Suk, A polynomial regularity lemma for semialgebraic hypergraphs and its applications in geometry and property testing, *SIAM J. Comput.* **45** (2016), 2199–2223.

[20] J. Fox, J. Pach, and A. Suk, Erdős-Hajnal conjecture for graphs with bounded VC-dimension, *Discrete Comput. Geom.* **61** (2019), 809–829.

[21] J. Fox and B. Sudakov, Unavoidable patterns, *J. Combin. Theory Ser. A* **115** (2008), 1561–1569.

[22] Z. Füredi, Extremal hypergraphs and combinatorial geometry, in *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, Birkhäuser, Basel, 1995, pp. 1343–1352.

[23] W. Goddard and J. Lyle, Dense graphs with small clique number, *J. Graph Theory* **66** (2011), 319–331.

[24] P. Hell and J. Nešetřil, The core of a graph, *Discrete Math.* **109** (1992), 117–126.

[25] J. Komlós, G. N. Sárközy, and E. Szemerédi, Proof of the Seymour conjecture for large graphs, *Ann. Comb.* **2** (1998), 43–60.

[26] T. Kövari, V. T. Sós, and P. Turán, On a problem of K. Zarankiewicz, *Colloq. Math.* **3** (1954), 50–57.

[27] S. Letzter and R. Snyder, The homomorphism threshold of $\{C_3, C_5\}$-free graphs, *J. Graph Theory* **90** (2019), 83–106.

[28] L. Lovász and B. Szegedy, Regularity partitions and the topology of graphons, in *An irregular mind*, *Bolyai Soc. Math. Stud.*, vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 415–446.

[29] T. Łuczak, On the structure of triangle-free graphs of large minimum degree, *Combinatorica* **26** (2006), 489–493.

[30] M. Malliaris and S. Shelah, Regularity lemmas for stable graphs, *Trans. Amer. Math. Soc.* **366** (2014), 1551–1585.

[31] V. Nikiforov, Chromatic number and minimum degree of K_r-free graphs, 2010. Preprint available at `https://arxiv.org/pdf/1001.2070.pdf`.

[32] H. Oberkampf and M. Schacht, On the structure of dense graphs with bounded clique number, *Combin. Probab. Comput.* **29** (2020), 641–649.

[33] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, in *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, *Colloq. Math. Soc. János Bolyai*, vol. 18, North-Holland, Amsterdam-New York, 1978, pp. 939–945.

[34] C. Thomassen, On the chromatic number of triangle-free graphs of large minimum degree, *Combinatorica* **22** (2002), 591–596.

[35] C. Thomassen, On the chromatic number of pentagon-free graphs of large minimum degree, *Combinatorica* **27** (2007), 241–243.

[36] Y. Wigderson, Complete r-partite r-graphs are Sidorenko: a brief exposition, 2021. Not intended for publication. Available at `http://web.stanford.edu/~yuvalwig/math/expository/HypergraphSidorenko.pdf`.