The Minimum Degree Removal Lemma Thresholds

Lior Gishboliner* Zhihan Jin* Benny Sudakov*

Abstract

The graph removal lemma is a fundamental result in extremal graph theory which says that for every fixed graph H and $\varepsilon > 0$, if an n-vertex graph G contains εn^2 edge-disjoint copies of H then G contains $\delta n^{v(H)}$ copies of H for some $\delta = \delta(\varepsilon, H) > 0$. The current proofs of the removal lemma give only very weak bounds on $\delta(\varepsilon, H)$, and it is also known that $\delta(\varepsilon, H)$ is not polynomial in ε unless H is bipartite. Recently, Fox and Wigderson initiated the study of minimum degree conditions guaranteeing that $\delta(\varepsilon, H)$ depends polynomially or linearly on ε. In this paper we answer several questions of Fox and Wigderson on this topic.

1 Introduction

The graph removal lemma, first proved by Ruzsa and Szemerédi [23], is a fundamental result in extremal graph theory. It also have important applications to additive combinatorics and property testing. The lemma states that for every fixed graph H and $\varepsilon > 0$, if an n-vertex graph G contains εn^2 edge-disjoint copies of H then G it contains $\delta n^{v(H)}$ copies of H, where $\delta = \delta(\varepsilon, H) > 0$. Unfortunately, the current proofs of the graph removal lemma give only very weak bounds on $\delta = \delta(\varepsilon, H)$ and it is a very important problem to understand the dependence of δ on ε. The best known result, due to Fox [11], proves that $1/\delta$ is at most a tower of exponents of height logarithmic in $1/\varepsilon$. Ideally, one would like to have better bounds on $1/\delta$, where an optimal bound would be that δ is polynomial in ε. However, it is known [2] that $\delta(\varepsilon, H)$ is only polynomial in ε if H is bipartite. This situation led Fox and Wigderson [12] to initiate the study of minimum degree conditions which guarantee that $\delta(\varepsilon, H)$ depends polynomially or linearly on ε. Formally, let $\delta(\varepsilon, H; \gamma)$ be the maximum $\delta \in [0, 1]$ such that if G is an n-vertex graph with minimum degree at least γn and with εn^2 edge-disjoint copies of H, then G contains $\delta n^{v(H)}$ copies of H.

Definition 1.1. *Let H be a graph.*

1. *The* linear removal threshold *of H, denoted $\delta_{\text{lin-rem}}(H)$, is the infimum γ such that $\delta(\varepsilon, H; \gamma)$ depends linearly on ε, i.e. $\delta(\varepsilon, H; \gamma) \geq \mu\varepsilon$ for some $\mu = \mu(\gamma) > 0$ and all $\varepsilon > 0$.*
2. *The* polynomial removal threshold *of H, denoted $\delta_{\text{poly-rem}}(H)$, is the infimum γ such that $\delta(\varepsilon, H; \gamma)$ depends polynomially on ε, i.e. $\delta(\varepsilon, H; \gamma) \geq \mu\varepsilon^{1/\mu}$ for some $\mu = \mu(\gamma) > 0$ and all $\varepsilon > 0$.*

Trivially, $\delta_{\text{lin-rem}}(H) \geq \delta_{\text{poly-rem}}(H)$. Fox and Wigderson [12] initiated the study of $\delta_{\text{lin-rem}}(H)$ and $\delta_{\text{poly-rem}}(H)$, and proved that $\delta_{\text{lin-rem}}(K_r) = \delta_{\text{poly-rem}}(K_r) = \frac{2r-5}{2r-3}$ for every $r \geq 3$, where K_r is the clique on r vertices. They further asked to determine the removal lemma thresholds of odd cycles. Here we completely resolve this question. The following theorem handles the polynomial removal threshold.

Theorem 1.2. $\delta_{\text{poly-rem}}(C_{2k+1}) = \frac{1}{2k+1}$.

Theorem 1.2 also answers another question of Fox and Wigderson [12], of whether $\delta_{\text{lin-rem}}(H)$ and $\delta_{\text{poly-rem}}(H)$ can only obtain finitely many values on r-chromatic graphs H for a given $r \geq 3$. Theorem 1.2 shows that $\delta_{\text{poly-rem}}(H)$ obtains infinitely many values for 3-chromatic graphs. In contrast, $\delta_{\text{lin-rem}}(H)$ obtains only three possible values for 3-chromatic graphs. Indeed, the following theorem determines $\delta_{\text{lin-rem}}(H)$ for every 3-chromatic H. An edge xy of H is called *critical* if $\chi(H - xy) < \chi(H)$.

*Department of Mathematics, ETH, Zürich, Switzerland. Research supported in part by SNSF grant 200021_196965. Email: **{lior.gishboliner, zhihan.jin, benjamin.sudakov}@math.ethz.ch**.

Theorem 1.3. *For a graph H with $\chi(H) = 3$, it holds that*

$$\delta_{\textit{lin-rem}}(H) = \begin{cases} \frac{1}{2} & H \textit{ has no critical edge,} \\ \frac{1}{3} & H \textit{ has a critical edge and contains a triangle,} \\ \frac{1}{4} & H \textit{ has a critical edge and } \text{odd-girth}(H) \geq 5. \end{cases}$$

Theorems 1.2 and 1.3 show a separation between the polynomial and linear removal thresholds, giving a sequence of graphs (i.e. $C_5, C_7, \dots$) where the polynomial threshold tends to 0 while the linear threshold is constant $\frac{1}{4}$.

The parameters $\delta_{\text{poly-rem}}$ and $\delta_{\text{lin-rem}}$ are related to two other well-studied minimum degree thresholds: the chromatic threshold and the homomorphism threshold. The chromatic threshold of a graph H is the infimum γ such that every n-vertex H-free graph G with $\delta(G) \geq \gamma n$ has bounded cromatic number, i.e., there exists $C = C(\gamma)$ such that $\chi(G) \leq C$. The study of the chromatic threshold originates in the work of Erdős and Simonovits [10] from the '70s. Following multiple works [4, 15, 16, 7, 5, 25, 26, 19, 6, 14, 20], the chromatic threshold of every graph was determined by Allen et al. [1].

Moving on to the homomorphism threshold, we define it more generally for families of graphs. The *homomorphism threshold* of a graph-family $\mathcal{H}$, denoted $\delta_{\text{hom}}(\mathcal{H})$, is the infimum γ for which there exists an $\mathcal{H}$-free graph $F = F(\gamma)$ such that every n-vertex $\mathcal{H}$-free graph G with $\delta(G) \geq \gamma n$ is homomorphic to F. When $\mathcal{H} = \{H\}$, we write $\delta_{\text{hom}}(H)$. This parameter was widely studied in recent years [18, 22, 17, 8, 24]. It turns out that δ_{hom} is closely related to $\delta_{\text{poly-rem}}(H)$, as the following theorem shows. For a graph H, let $\mathcal{I}_H$ denote the set of all minimal (with respect to inclusion) graphs H' such that H is homomorphic to H'.

Theorem 1.4. *For every graph H, $\delta_{\textit{poly-rem}}(H) \leq \delta_{\textit{hom}}(\mathcal{I}_H)$.*

Note that $\mathcal{I}_{C_{2k+1}} = \{C_3, \dots, C_{2k+1}\}$. Using this, the upper bound in Theorem 1.2 follows immediately by combining Theorem 1.4 with the result of Ebsen and Schacht [8] that $\delta_{\text{hom}}(\{C_3, \dots, C_{2k+1}\}) = \frac{1}{2k+1}$. The lower bound in Theorem 1.2 was established in [12]; for completeness, we sketch the proof in Section 3.

The rest of this short paper is organized as follows. Section 2 contains some preliminary lemmas. In Section 3 we prove the lower bounds in Theorems 1.2 and 1.3. Section 4 gives the proof of Theorem 1.4, and Section 5 gives the proof of the upper bounds in Theorem 1.3. In the last section we discuss further related problems.

2 Preliminaries

Throughout this paper, we always consider *labeled copies of some fixed graph H* and write *copy of H* for simplicity. We use $\delta(G)$ for the minimum degree of G, and write $H \to F$ to denote that there is a homomorphism from H to F. For a graph H on $[h]$ and integers $s_1, s_2, \dots, s_h > 0$, we denote by $H[s_1, \dots, s_h]$ the blow-up of H where each vertex $i \in V(H)$ is replaced by a set S_i of size s_i (and edges are replaced with complete bipartite graphs). The following lemma is standard.

Lemma 2.1. *Let H be a fixed graph on vertex set $[h]$ and let $s_1, s_2, \dots, s_h \in \mathbb{N}$. There exists a constant $c = c(H, s_1, \dots, s_h) > 0$ such that the following holds. Let G be an n-vertex graph and $V_1, \dots, V_h \subseteq V(G)$. Suppose that G contains at least ρn^h copies of H mapping i to V_i for all $i \in [h]$. Then G contains at least $c\rho^{\frac{1}{c}} \cdot n^{s_1 + \dots + s_h}$ copies of $H[s_1, \dots, s_h]$ mapping S_i to V_i for all $i \in [h]$.*

Note that the sets $V_1, \dots, V_h$ in Lemma 2.1 do not have to be disjoint. The proof of Lemma 2.1 works by defining an auxiliary h-uniform hypergraph $\mathcal{G}$ whose hyperedges correspond to the copies of H in which vertex i is mapped to V_i. By assumption, $\mathcal{G}$ has at least ρn^h edges. By the hypergraph generalization of the Kovári-Sós-Turán theorem, see [9], $\mathcal{G}$ contains $\text{poly}(\rho) n^{s_1 + \dots + s_h}$ copies of $K^{(h)}_{s_1, \dots, s_h}$, the complete h-partite hypergraph with parts of size $s_1, \dots, s_h$. Each copy of $K^{(h)}_{s_1, \dots, s_h}$ gives a copy of $H[s_1, \dots, s_h]$ mapping S_i to V_i.

Fox and Wigderson [12, Proposition 4.1] proved the following useful fact.

Lemma 2.2. *If $H \to F$ and F is a subgraph of H, then $\delta_{\textit{poly-rem}}(H) = \delta_{\textit{poly-rem}}(F)$.*

The following lemma is an asymmetric removal-type statement for odd cycles, which gives polynomial bounds. It may be of independent interest. A similar result has appeared very recently in [13].

Lemma 2.3. *For $1 \leq \ell < k$, there exists a constant $c = c(k) > 0$ such that if an n-vertex graph G has εn^2 edge-disjoint copies of $C_{2\ell+1}$, then it has at least $c\varepsilon^{1/c} n^{2k+1}$ copies of C_{2k+1}.*

Proof. Let $\mathcal{C}$ be a collection of εn^2 edge-disjoint copies of $C_{2\ell+1}$ in G. There exists a collection $\mathcal{C}' \subseteq \mathcal{C}$ such that $|\mathcal{C}'| \geq \varepsilon n^2/2$ and each vertex $v \in V(G)$ belongs to either 0 or at least $\varepsilon n/2$ of the cycles in $\mathcal{C}'$. Indeed, to obtain $\mathcal{C}'$, we repeatedly delete from $\mathcal{C}$ all cycles containing a vertex v which belongs to at least one but less than $\varepsilon n/2$ of the cycles in $\mathcal{C}$ (without changing the graph). The set of cycles left at the end is $\mathcal{C}'$. In this process, we delete at most $\varepsilon n^2/2$ cycles altogether (because the process lasts for at most n steps); hence $|\mathcal{C}'| \geq \varepsilon n^2/2$. Let V be the set of vertices contained in at least $\varepsilon n/2$ cycles from $\mathcal{C}'$, so $|V| \geq \varepsilon n/2$. With a slight abuse of notation, we may replace G with $G[V]$, $\mathcal{C}$ with $\mathcal{C}'$ and $\varepsilon/2$ with ε, and denote $|V|$ by n. Hence, from now on, we assume that each vertex $v \in V(G)$ is contained in at least εn of the cycles in $\mathcal{C}$. This implies that $|N(v)| \geq 2\varepsilon n$ for every $v \in V(G)$.

Fix any $v_0 \in V(G)$ and let $\mathcal{C}(v_0)$ be the set of cycles $C \in \mathcal{C}$ such that $C \cap N(v_0) \neq \emptyset$ and $v_0 \notin C$. The number of cycles $C \in \mathcal{C}$ intersecting $N(v_0)$ is at least $|N(v_0)| \cdot \varepsilon n/(2\ell+1) \geq 2\varepsilon^2 n^2/(2\ell+1)$, and the number of cycles containing v_0 is at most n. Hence, $|\mathcal{C}(v_0)| \geq 2\varepsilon^2 n^2/(2\ell+1) - n \geq \varepsilon^2 n^2/(\ell+1)$. Take a random partition $V_0, V_1, \ldots, V_\ell$ of $V(G) \setminus \{v_0\}$, where each vertex is put in one of the parts uniformly and independently. For a cycle $(x_1, \ldots, x_{2\ell+1}) \in \mathcal{C}(v_0)$ with $x_{\ell+1} \in N(v_0)$, say that $(x_1, \ldots, x_{2\ell+1})$ is good if $x_{\ell+1} \in V_0$ and $x_{\ell+1-i}, x_{\ell+1+i} \in V_i$ for $1 \leq i \leq \ell$ (so in particular $x_1, x_{2\ell+1} \in V_\ell$). The probability that $(x_1, \ldots, x_{2\ell+1})$ is good is $1/(\ell+1)^{2\ell+1}$, so there is a collection of good cycles $\mathcal{C}'(v_0) \subseteq \mathcal{C}_0$ of size $|\mathcal{C}'(v_0)| \geq |\mathcal{C}(v_0)|/(\ell+1)^{2\ell+1} \geq \varepsilon^2 n^2/(\ell+1)^{2\ell+2}$. Put $\gamma := \varepsilon^2/(\ell+1)^{2\ell+2}$. By the same argument as above, there is a collection $\mathcal{C}''(v_0) \subseteq \mathcal{C}'(v_0)$ with $|\mathcal{C}''(v_0)| \geq \gamma n^2/2$ such that each vertex is contained in either 0 or at least $\gamma n/2$ cycles from $\mathcal{C}''(v_0)$. Let W be the set of vertices contained in at least $\gamma n/2$ cycles from $\mathcal{C}''(v_0)$. Note that $W \cap V_0 \subseteq N(v_0)$ by definition. Also, each vertex in $W \cap V_\ell$ has at least $\gamma n/2$ neighbors in $W \cap V_\ell$, and for each $1 \leq i \leq \ell$, each vertex in $W \cap V_i$ has at least $\gamma n/2$ neighbors in $W \cap V_{i-1}$. It follows that $W \cap V_\ell$ contains at least $\frac{1}{2}|W \cap V_\ell| \cdot \prod_{i=0}^{2k-2\ell-2}(\gamma n/2 - i) = \text{poly}(\gamma) n^{2k-2\ell}$ paths of length $2k-2\ell-1$. We now construct a collection of copies of C_{2k+1} as follows. Choose a path $y_{\ell+1}, y_{\ell+2}, \ldots, y_{2k-\ell}$ of length $2k-2\ell-1$ in $W \cap V_\ell$. For each $i = \ell, \ldots, 1$, take a neighbor $y_i \in W \cap V_{i-1}$ of y_{i+1} and a neighbor $y_{2k-i+1} \in W \cap V_{i-1}$ of y_{2k-i}, such that the vertices $y_1, \ldots, y_{2k}$ are all different. Then $y_1, \ldots, y_{2k}$ is a path and $y_1, y_{2k} \in W \cap V_0 \subseteq N(v_0)$, so $v_0, y_1, \ldots, y_{2k}$ is a copy of $C_{2\ell+1}$. The number of choices for the path $y_{\ell+1}, y_{\ell+2}, \ldots, y_{2k-\ell}$ is $\text{poly}(\gamma) n^{2k-2\ell}$ and the number of choices for each vertex $y_i, y_{2k-i+1} \in V_{i-1}$ $(i = \ell, \ldots, 1)$ is at least $\gamma n/2$. Hence, the total number of choices for $y_1, \ldots, y_{2k}$ is $\text{poly}(\gamma) n^{2k}$. As there are n choices for v_0, we get a total of $\text{poly}(\gamma) n^{2k+1} = \text{poly}_k(\varepsilon) n^{2k+1}$ copies of C_{2k+1}, as required. ☐

3 Lower bounds

Here we prove the lower bounds in Theorems 1.2 and 1.3. The lower bound in Theorem 1.2 was proved in [12, Theorem 4.3]. For completeness, we include a sketch of the proof:

Lemma 3.1. $\delta_{\textit{poly-rem}}(C_{2k+1}) \geq \frac{1}{2k+1}$.

Proof. Fix an arbitrary $\alpha > 0$. In [2] it was proved that for every ε, there exists a $(2k+1)$-partite graph with parts $V_1, \ldots, V_{2k+1}$ of size $\alpha n/(2k+1)$ each, with εn^2 edge-disjoint copies of C_{2k+1}, but with only $\varepsilon^{\omega(1)} n^{2k+1}$ copies of C_{2k+1} in total (where the $\omega(1)$ term may depend on α). Add sets $U_1, \ldots, U_{2k+1}$ of size $(1-\alpha)n/(2k+1)$ each, and add the complete bipartite graphs (U_i, V_i), $1 \leq i \leq 2k+1$, and (U_i, U_{i+1}), $1 \leq i \leq 2k$. See Figure 1. It is easy to see that this graph has minimum degree $(1-\alpha)n/(2k+1)$, and every copy of C_{2k+1} is contained in $V_1 \cup \cdots \cup V_{2k+1}$. Letting $\alpha \to 0$, we get that $\delta_{\text{poly-rem}}(C_{2k+1}) \geq \frac{1}{2k+1}$. ☐

By combining the fact that $\delta_{\text{poly-rem}}(C_3) = \frac{1}{3}$ with Lemma 2.2 (with $F = C_3$), we get that $\delta_{\text{lin-rem}}(H) \geq \delta_{\text{poly-rem}}(H) = \frac{1}{3}$ for every 3-chromatic graph H containing a triangle. This proves the lower bound in the second case of Theorem 1.3. Now we prove the lower bounds in the other two cases. We prove a more general statement for r-chromatic graphs.

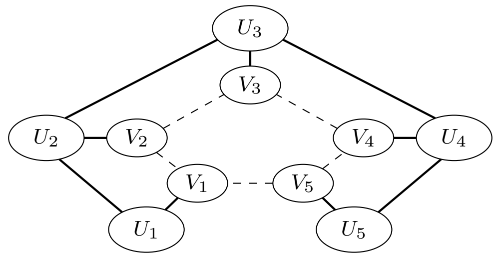

Figure 1: Proof of Lemma 3.1 for C_5. Heavy edges indicate complete bipartite graphs while dashed edges form the Ruzsa–Szemerédi construction for C_5 (see [2]).

Lemma 3.2. *Let H be a graph with $\chi(H) = r \geq 3$. Then, $\frac{3r-8}{3r-5} \leq \delta_{\text{lin-rem}}(H) \leq \frac{r-2}{r-1}$. Moreover, $\delta_{\text{lin-rem}}(H) = \frac{r-2}{r-1}$ if H contains no critical edge.*

Proof. Denote $h = |V(H)|$. The bound $\delta_{\text{lin-rem}}(H) \leq \frac{r-2}{r-1}$ holds for every r-chromatic graph H; this follows from the Erdős-Simonovits supersaturation theorem, see by [12, Section 4.1] for the details.

Suppose now that H contains no critical edge, and let us show that $\delta_{\text{lin-rem}}(H) \geq \frac{r-2}{r-1}$. To this end, we construct, for every small enough ε and infinitely many n, an n-vertex graph G with $\delta(G) \geq \frac{r-2}{r-1}n$, such that G has at most $\mathcal{O}(\varepsilon^2 n^h)$ copies of H, but $\Omega(\varepsilon n^2)$ edges must be deleted to turn G into an H-free graph. Let $T(n, r-1)$ be the Turán graph, i.e. the complete $(r-1)$-partite graph with balanced parts $V_1, \dots, V_{r-1}$. Add an εn-regular graph inside V_1 and let the resulting graph be G. We first claim that G contains $\mathcal{O}(\varepsilon^2 n^h)$ copies of H. As H contains no critical edge and $\chi(H) = r$, every copy of H in G contains two edges e and e' inside V_1. If e and e' are disjoint, then there are at most $n^2(\varepsilon n)^2 = \varepsilon^2 n^4$ choices for e and e' and then at most n^{h-4} choices for the other $h-4$ vertices of H. Therefore, there are at most $\varepsilon^2 n^h$ such H-copies. And if e and e' intersect, then there are at most $n(\varepsilon n)^2 = \varepsilon^2 n^3$ choices for e and e' and then at most n^{h-3} choices for the remaining vertices, again giving at most $\varepsilon^2 n^h$ such H-copies. So G indeed has $\mathcal{O}(\varepsilon^2 n^h)$ copies of H.

On the other hand, we claim that one must delete $\Omega(\varepsilon n^2)$ edges to destroy all H-copies in G. Observe that G has at least $\frac{1}{2}|V_1| \cdot \varepsilon n \cdot |V_2| \cdot \dots \cdot |V_{r-1}| = \Omega_r(\varepsilon n^r)$ copies of K_r, and every edge participates in at most n^{r-2} of these copies. Thus, deleting $c\varepsilon n^2$ edges can destroy at most $c\varepsilon n^r$ copies of K_r. If c is a small enough constant (depending on r), then after deleting any $c\varepsilon n^2$ edges, there are still $\Omega(\varepsilon n^r)$ copies of K_r. Then, by Lemma 2.1, the remaining graph contains $K_r[h]$, the h-blowup of K_r, and hence H. This completes the proof that $\delta_{\text{lin-rem}}(H) \geq \frac{r-2}{r-1}$.

We now prove that $\delta_{\text{lin-rem}}(H) \geq \frac{3r-8}{3r-5}$ for every r-chromatic graph H. It suffices to construct, for every small enough ε and infinitely many n, an n-vertex graph G with $\delta(G) \geq \frac{3r-8}{3r-5}n$, such that G has at most $\mathcal{O}(\varepsilon^2 n^h)$ copies of H but at least $\Omega(\varepsilon n^2)$ edges must be deleted to turn G into an H-free graph. The vertex set of G consists of $r+1$ disjoint sets $V_0, V_1, V_2, \dots, V_r$, where $|V_i| = \frac{n}{3r-5}$ for $i = 0, 1, 2, 3$ and $|V_i| = \frac{3n}{3r-5}$ for $i = 4, 5, \dots, r$. Put complete bipartite graphs between V_0 and V_1, between $V_0 \cup V_1$ and $V_4 \cup \dots \cup V_r$, and between V_i to V_j for all $2 \leq i < j \leq r$. Put εn-regular bipartite graphs between V_1 and V_2, and between V_1 and V_3. The resulting graph is G (see Figure 2). It is easy check that $\delta(G) \geq \frac{3r-8}{3r-5}n$. Indeed, let $0 \leq i \leq r$ and $v \in V_i$. If $4 \leq i \leq r$ then v is connected to all vertices except for V_i; if $i \in \{2, 3\}$ then v is connected to all vertices except $V_0 \cup V_1 \cup V_i$; and if $i \in \{0, 1\}$ then v is connected to all vertices except $V_2 \cup V_3 \cup V_i$. In any case, the neighborhood of v misses at most $\frac{3n}{3r-5}$ vertices.

We claim that G has at most $\mathcal{O}(\varepsilon^2 n^h)$ copies of H. Indeed, observe that if we delete all edges between V_1 and V_2 then the remaining graph is $(r-1)$-colorable with coloring $V_1 \cup V_2, V_0 \cup V_3, V_4, \dots, V_r$. Hence, every copy of H must contain an edge e between V_1 and V_2. Similarly, every copy of H must contain an edge e' between V_1 and V_3. If e, e' are disjoint then there are at most $n^2(\varepsilon n)^2 = \varepsilon^2 n^4$ ways to choose e, e' and then at most n^{h-4} ways to choose the remaining vertices of H. And if e and e' intersect then there are at most $n(\varepsilon n)^2 = \varepsilon^2 n^3$ ways to choose e, e' and at most n^{h-3} for the remaining $h-3$ vertices of H. In both cases, the number of H-copies is at most $\varepsilon^2 n^h$, as required.

Now we show that one must delete $\Omega(\varepsilon n^2)$ edges to destroy all copies of H in G. Observe that G has $|V_1| \cdot (\varepsilon n)^2 \cdot |V_4| \cdot \dots \cdot |V_r| = \Omega(\varepsilon^2 n^r)$ copies of K_r between the sets $V_1, \dots, V_r$. We claim that every edge f

Figure 2: Proof of Lemma 3.2, $r = 3$ (left) and $r = 4$ (right). Heavy edges indicate complete bipartite graphs while dashed edges indicate εn-regular bipartite graphs.

participates in at most εn^{r-2} of these r-cliques. Indeed, by the same argument as above, every copy of K_r containing f must contain an edge e from $E(V_1, V_2)$ and an edge e' from $E(V_1, V_3)$. Suppose without loss of generality that $e \neq f$ (the case $e' \neq f$ is symmetric). In the case $f \cap e = \emptyset$, there are at most $n \cdot \varepsilon n = \varepsilon n^2$ choices for e and at most n^{r-4} choices for the remaining vertices of K_r, giving at most εn^{r-2} copies of K_r containing f. And if f, e intersect, then there are at most εn choices for e and at most n^{r-3} for the remaining $r - 3$ vertices, giving again εn^{r-2}.

We see that deleting $c\varepsilon n^2$ edges of G can destroy at most $c\varepsilon^2 n^r$ copies of K_r. Hence, if c is a small enough constant, then after deleting any $c\varepsilon n^2$ edges there are still $\Omega(\varepsilon^2 n^r)$ copies of K_r left. By Lemma 2.1, the remaining graph contains a copy of $K_r[h]$ and hence H. This completes the proof. □

4 Polynomial removal thresholds: Proof of Theorem 1.4

We say that an n-vertex graph G is *ε-far* from a graph property $\mathcal{P}$ (e.g. being H-free for a given graph H, or being homomorphic to a given graph F) if one must delete at least εn^2 edges to make G satisfy $\mathcal{P}$. Trivially, if G has εn^2 edge-disjoint copies of H, then it is ε-far from being H-free. We need the following result from [21].

Theorem 4.1. *For every graph F on f vertices and for every $\varepsilon > 0$, there is $q = q_F(\varepsilon) = poly(f/\varepsilon)$, such that the following holds. If a graph G is ε-far from being homomorphic to F, then for a sample of q vertices $x_1, \dots, x_q \in V(G)$, taken uniformly with repetitions, it holds that $G[\{x_1, \dots, x_q\}]$ is not homomorphic to F with probability at least $\frac{2}{3}$.*

Theorem 4.1 is proved in Section 2 of [21]. In fact, [21] proves a more general result on property testing of the so-called 0/1-partition properties. Such a property is given by an integer f and a function $d : [f]^2 \to \{0, 1, \perp\}$, and a graph G satisfies the property if it has a partition $V(G) = V_1 \cup \dots \cup V_f$ such that for every $1 \leq i, j \leq f$ (possibly $i = j$), it holds that (V_i, V_j) is complete if $d(i, j) = 1$ and (V_i, V_j) is empty if $d(i, j) = 0$ (if $d(i, j) = \perp$ then there are no restrictions). One can express the property of having a homomorphism into F in this language, simply by setting $d(i, j) = 0$ for $i = j$ and $ij \notin E(F)$. In [21], the class of these partition properties is denoted $\mathcal{GPP}_{0,1}$, and every such property is shown to be testable by sampling poly(f/ε) vertices. This implies Theorem 4.1.

Proof of Theorem 1.4. Recall that $\mathcal{I}_H$ is the set of minimal graphs H' (with respect to inclusion) such that H is homomorphic to H'. For convenience, put $\delta := \delta_{\text{hom}}(\mathcal{I}_H)$. Our goal is to show that $\delta_{\text{poly-rem}}(H) \leq \delta + \alpha$ for every $\alpha > 0$. So fix $\alpha > 0$ and let G be a graph with minimum degree $\delta(G) \geq (\delta + \alpha)n$ and with εn^2 edge-disjoint copies of H. By the definition of the homomorphism threshold, there is an $\mathcal{I}_H$-free graph F (depending only on $\mathcal{I}_H$ and α) such that if a graph G_0 is $\mathcal{I}_H$-free and has minimum degree at least $(\delta + \frac{\alpha}{2}) \cdot |V(G_0)|$, then G_0 is homomorphic to F. Observe that if a graph G_0 is homomorphic to F then G_0 is H-free, because F is free of any homomorphic image of H. It follows that G is ε-far from being homomorphic to F, because G is ε-far from being H-free. Now we apply Theorem 4.1. Let $q = q_F(\varepsilon)$ be given by Theorem 4.1. We assume that $q \gg \frac{\log(1/\alpha)}{\alpha^2}$ and $n \gg q^2$ without loss of generality. Sample q vertices $x_1, \dots, x_q \in V(G)$ with repetition and let $X = \{x_1, \dots, x_q\}$. By Theorem 4.1, $G[X]$ is not homomorphic to F with probability at least $2/3$. As $n \gg q^2$, the vertices $x_1, \dots, x_q$ are pairwise-distinct with probability at least 0.99. Also, for every $i \in [q]$, the number of indices $j \in [q] \setminus \{i\}$ with $x_i x_j \in E(G)$ dominates a binomial distribution $\mathrm{B}(q-1, \frac{\delta(G)}{n})$. By the Chernoff bound (see e.g. [3, Appendix A]) and as $\delta(G) \geq (\delta+\alpha)n$, the number of such indices is at least $(\delta + \frac{\alpha}{2})q$ with probability $1 - e^{-\Omega(q\alpha^2)}$. Taking the union bound over $i \in [q]$, we get that $\delta(G[X]) \geq (\delta + \frac{\alpha}{2})|X|$ with probability at least $1 - qe^{-\Omega(q\alpha^2)} \geq 0.9$, as $q \gg \frac{\log(1/\alpha)}{\alpha^2}$. Hence, with probability at least $\frac{1}{2}$ it holds that $\delta(G[X]) \geq (\delta + \frac{\alpha}{2})|X|$ and $G[X]$ is not homomorphic to F. If this happens, then $G[X]$ is not $\mathcal{I}_H$-free (by the choice of F), hence $G[X]$ contains a copy of some $H' \in \mathcal{I}_H$. By averaging, there is $H' \in \mathcal{I}_H$ such that $G[X]$ contains a copy of H' with probability at least $\frac{1}{2|\mathcal{I}_H|}$. Put $k = |V(H')|$ and let M be the number of copies of H' in G. The probability that $G[X]$ contains a copy of H' is at most $M(\frac{q}{n})^k$. Using the fact that $q = \mathrm{poly}_{H,\alpha}(\frac{1}{\varepsilon})$, we conclude that $M \geq \frac{1}{2|\mathcal{I}_H|} \cdot (\frac{n}{q})^k \geq \mathrm{poly}_{H,\alpha}(\varepsilon)n^k$. As $H \to H'$, there exists H'', a blow-up of H', such that H'' have the same number of vertices as H, and that $H \subset H''$. By Lemma 2.1 for H' with $V_i = V(G)$ for all i, there exist $\mathrm{poly}_{H,\alpha}(\varepsilon)n^{v(H'')}$ copies of H'' in G, and thus $\mathrm{poly}_{H,\alpha}(\varepsilon)n^{v(H)}$ copies of H. This completes the proof. □

5 Linear removal thresholds: Proof of Theorem 1.3

Here we prove the upper bounds in Theorem 1.3; the lower bounds were proved in Section 3. The first case of Theorem 1.3 follows from Lemma 3.2, so it remains to prove the other two cases. We begin with some preparation. For disjoint sets $A_1, \dots, A_m$, we write $\bigcup_{i\in[m]} A_i \times A_{i+1}$ to denote all pairs of vertices which have one endpoint in A_i and one in A_{i+1} for some $1 \leq i \leq m$, with subscripts always taken modulo m. So a graph G has a homomorphism to the cycle C_m if and only if there is a partition $V(G) = A_1 \cup \dots \cup A_m$ with $E(G) \subseteq \bigcup_{i\in[m]} A_i \times A_{i+1}$.

Lemma 5.1. *Suppose H is a graph such that $\chi(H) = 3$, H contains a critical edge xy, and* odd-girth$(H) \geq 2k+1$. *Then,*

- *There is a partition $V(H) = A_1 \uplus A_2 \uplus A_3 \uplus B$ such that $A_1 = \{x\}, A_2 = \{y\}$ and $E(H) \subseteq (A_3 \times B) \cup (\bigcup_{i\in[3]} A_i \times A_{i+1})$;*
- *if $k \geq 2$, there is a partition $V(H) = A_1 \uplus A_2 \uplus \dots \uplus A_{2k+1}$ such that $A_1 = \{x\}, A_2 = \{y\}$ and $E(H) \subseteq \bigcup_{i\in[2k+1]} A_i \times A_{i+1}$. In particular, H is homomorphic to C_{2k+1}.*

Proof. Write $H' = H - xy$, so H' is bipartite. Let $V(H) = V(H') = L \uplus R$ be a bipartition of H'. As $\chi(H) = 3$, x and y must both lie in the same side of the bipartition. Without loss of generality, assume that $x, y \in L$. For the first item, set $A_1 = \{x\}$, $A_2 = \{y\}$, $A_3 = R$ and $B = L\backslash\{x, y\}$. Then every edge of G goes between B and A_3 or between two of the sets A_1, A_2, A_3, as required.

Suppose now that $k \geq 2$, i.e. odd-girth$(H) = 2k+1 \geq 5$. For $1 \leq i \leq k$, let X_i be the set of vertices at distance $(i-1)$ from x in H', and let Y_i be the set of vertices at distance $(i-1)$ from y in H'. Note that $X_1 = \{x\}$ and $Y_1 = \{y\}$. Also, X_i, Y_i lie in L if i is odd and in R if i is even. Write

$$L' := L\backslash \bigcup_{i=1}^{k}(X_i \cup Y_i), \quad R' := R\backslash \bigcup_{i=1}^{k}(X_i \cup Y_i),$$

We first claim that $\{X_1, \dots, X_k, Y_1, \dots, Y_k, L', R'\}$ forms a partition of $V(H)$. The sets $X_1, \dots, X_k$ are clearly pairwise-disjoint, and so are $Y_1, \dots, Y_k$. Also, all of these sets are disjoint from L', R' by definition. So we only need to check X_i and Y_j are disjoint for every pair $1 \leq i, j \leq k$. Suppose for contradiction that there exists $u \in X_i \cap Y_j$ for some $1 \leq i, j \leq k$. Then $i \equiv j \pmod 2$, because otherwise X_i, Y_j are contained in different parts of the bipartition L, R. By the definition of X_i and Y_j, H' has a path $x = x_1, x_2, \dots, x_i = u$ and a path $y = y_1, y_2, \dots, y_j = u$. Then, $x = x_1, x_2, \dots, x_i = u = y_j, y_{j-1}, \dots, y_1, y, x$ forms a closed walk of length $i+j-1$, which is odd as $i \equiv j \pmod 2$. Hence, odd-girth$(H) \leq 2k-1$, contradicting our assumption.

By definition, there are no edges between X_i and X_j for $j - i \geq 2$, and similarly for Y_i, Y_j. Also, there are no edges between $L' \cup R'$ and $\bigcup_{i=1}^{k-1}(X_i \cup Y_i)$ because the vertices in $L' \cup R'$ are at distance more than k to x, y. Moreover, if k is even then there are no edges between $X_k \cup Y_k$ and R', and if k is odd then there are no edges between $X_k \cup Y_k$ and L'. Next, we show that there are no edges between X_i and Y_j for any $1 \leq i, j \leq k$ except $(i, j) = (1, 1)$. Indeed, if $i = j$ then $e(X_i, Y_j) = 0$ because X_i, Y_j are on the same side

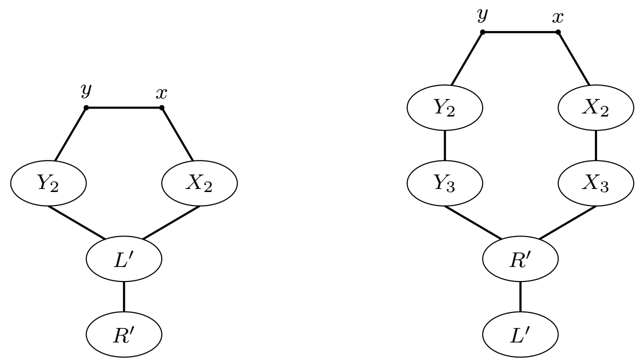

Figure 3: Proof of Lemma 5.1, $k = 2$ (left) and $k = 3$ (right). Edges indicate bipartite graphs where edges can be present.

of the bipartition L, R. So suppose that $i \neq j$, say $i < j$, and assume by contradiction that there is an edge uv with $u \in X_i, v \in Y_j$. Then v is at distance at most $i + 1 \leq k$ from x, implying that Y_j intersects $X_1 \cup \dots \cup X_{i+1}$, a contradiction.

Finally, we define the partition $A_1, \dots, A_{2k+1}$ that satisfies the assertion of the second item. If k is even then take $A_1, \dots, A_{2k+1}$ to be $X_1, Y_1, \dots, Y_{k-1}, Y_k \cup R', L', X_k, \dots, X_2$, and if k is odd then take $A_1, \dots, A_{2k+1}$ to be $X_1, Y_1, \dots, Y_{k-1}, Y_k \cup L', R', X_k, \dots, X_2$. See Figure 3 for an illustration. By the above, in both cases it holds that $E(H) \subseteq \bigcup_{i \in [2k+1]} A_i \times A_{i+1}$, as required. □

For vertex $u \in V(G)$, denote by $N_G(u)$ the neighborhood of u and let $\deg_G(u) = |N_G(u)|$. For vertices $u, v \in V(G)$, denote by $N_G(u, v)$ the common neighborhood of u, v and let $\deg_G(u, v) = |N_G(u, v)|$.

Lemma 5.2. *Let H be a graph on h vertices such that $\chi(H) = 3$ and H contains a critical edge xy. Let G be a graph on n vertices with $\delta(G) \geq \alpha n$. Let $ab \in E(G)$ such that $\deg_G(a, b) \geq \alpha n$. Then, there are at least poly$(\alpha)n^{h-2}$ copies of H in G mapping $xy \in E(H)$ to $ab \in E(G)$.*

Proof. By the first item of Lemma 5.1, there is a partition $V(H) = A_1 \cup A_2 \cup A_3 \cup B$ such that $A_1 = \{x\}, A_2 = \{y\}$ and $E(H) \subseteq (A_3 \times B) \cup \bigcup_{i \in [3]} A_i \times A_{i+1}$. Let $s = |A_3|$ and $t = |B|$. Each $u \in N_G(a, b)$ has at least $\alpha n - 2 \geq \frac{\alpha n}{2}$ neighbors not equal to a, b. Hence, there are at least $\frac{1}{2} \cdot |N_G(a, b)| \cdot \frac{\alpha n}{2} \geq \frac{\alpha^2 n^2}{4}$ edges uv with $u \in N_G(a, b)$ and $v \notin \{a, b\}$. Applying Lemma 2.1 with $H = K_2$, $V_1 = N_G(a, b)$ and $V_2 = V(G) \setminus \{a, b\}$, we see that there are poly$(\alpha)n^{s+t}$ pairs of disjoint sets (S, T) such that $|S| = s, |T| = t$, $S \subseteq N_G(a, b)$, $a, b \notin T$, and S, T form a complete bipartite graph in G. Given any such pair, it is safe to map x to a, y to b, A_3 to S and B to T to obtain an H-copy. Hence, G contains at least poly$(\alpha)n^{s+t} =$ poly$(\alpha)n^{h-2}$ copies of H mapping xy to ab. □

Lemma 5.3. *Let H be a graph on h vertices such that $\chi(H) = 3$, H contains a critical edge xy, and* odd-girth$(H) \geq 5$. *Let G be a graph on n vertices, let $ab \in E(G)$, and suppose that there exists $A \subset N_G(a)$ and $B \subset N_G(b)$ such that $|A|, |B| \geq \alpha n$ and $|N_G(a', b')| \geq \alpha n$ for all distinct $a' \in A$ and $b' \in B$. Then there are at least poly$(\alpha)n^{h-2}$ copies of H in G mapping $xy \in E(H)$ to $ab \in E(G)$.*

Proof. By Lemma 5.1 (using odd-girth$(H) \geq 5$), there exists a partition $V(H) = A_1 \cup \dots \cup A_5$ such that $A_1 = \{x\}, A_2 = \{y\}$, and $E(H) \subseteq \bigcup_{i \in [5]} A_i \times A_{i+1}$. Put $s_i = |A_i|$ for $i \in [5]$.

There at least $(|A||B| - |A|)/2 \geq \alpha^2 n^2/3$ pairs $\{a', b'\}$ of distinct vertices with $a' \in A, b' \in B$ (the factor of 2 is due to the fact that each pair in $A \cap B$ is counted twice). Each such pair a', b' has at least $\alpha n - 2 \geq \alpha n/2$ common neighbors $c' \notin \{a, b\}$, by assumption. Therefore, there are at least $\frac{\alpha^2 n^2}{3} \cdot \frac{\alpha n}{2} = \frac{\alpha^3 n^3}{6}$ triples (a', b', c') such that $a' \in A, b' \in B$, and $c' \neq a, b$ is a common neighbor of a', b'. By Lemma 2.1 with $H = K_{2,1}$ and $V_1 = A, V_2 = B, V_3 = V(G) \setminus \{a, b\}$, there are at least poly$(\alpha)n^{s_3+s_4+s_5}$ corresponding copies of $K_{2,1}[s_3, s_5, s_4]$, i.e., triples of disjoint sets (R, S, T) such that $R \subseteq A$, $S \subseteq B$, $a, b \notin T$, $|R| = s_5, |S| = s_3, |T| = s_4$, and (R, T) and (S, T) form complete bipartite graphs in G. Given any such triple, we can safely map $A_1 = \{x\}$ to a, $A_2 = \{y\}$ to b, A_5 to R, A_3 to S, and A_4 to T to obtain a copy of H. Thus, there are at least $\text{poly}(\alpha)n^{s_3+s_4+s_5} = \text{poly}(\alpha)n^{h-2}$ copies of H mapping xy to ab. $\square$

In the following theorem we prove the upper bound in the second case of Theorem 1.3.

Theorem 5.4. *Let H be a graph such that $\chi(H) = 3$, H has a critical edge xy, and H contains a triangle. Then, $\delta_{\text{lin-rem}}(H) \leq \frac{1}{3}$.*

Proof. Write $h = v(H)$. Fix an arbitrary $\alpha > 0$, and let G be an n-vertex graph with minimum degree $\delta(G) \geq (\frac{1}{3} + \alpha)n$ and with a collection $\mathcal{C} = \{H_1, \ldots, H_m\}$ of $m := \varepsilon n^2$ edge-disjoint copies of H. For each $i = 1, \ldots, m$, there exist $u, v, w \in V(H_i)$ forming a triangle (because H contains a triangle). As $\deg_G(u) + \deg_G(v) + \deg_G(w) \geq 3\delta(G) \geq (1 + 3\alpha)n$, two of u, v, w have at least αn common neighbors. We denote these two vertices by a_i and b_i. By Lemma 5.2, G has at least $\text{poly}(\alpha)n^{h-2}$ copies of H which map xy to $a_i b_i$. The edges $a_1b_1, \ldots, a_mb_m$ are distinct because $H_1, \ldots, H_m$ are edge-disjoint. Hence, summing over all $i = 1, \ldots, m$, we see that G contains at least $\varepsilon n^2 \cdot \text{poly}(\alpha)n^{h-2} = \text{poly}(\alpha)\varepsilon n^h$ copies of H. This proves that $\delta_{\text{lin-rem}}(H) \leq \frac{1}{3} + \alpha$, and taking $\alpha \to 0$ gives $\delta_{\text{lin-rem}}(H) \leq \frac{1}{3}$. $\square$

In what follows, we need the following very well-known observation, originating in the work of Andrásfai, Erdős and Sós, see [4, Remark 1.6].

Lemma 5.5. *If $\delta(G) > \frac{2}{2k+1}n$ and $\text{odd-girth}(G) \geq 2k + 1$ for $k \geq 2$, then G is bipartite.*

Proof. Suppose by contradiction that G is not bipartite and take a shortest odd cycle C in G, so $|C| \geq 2k+1$. As $\sum_{x \in C} \deg(x) \geq (2k+1)\delta(G) > 2n$, there exists a vertex $v \notin C$ with at least 3 neighbors on C. Then there are two neighbors $x, y \in C$ of v such that the distance of x, y along C is not equal to 2. Then by taking the odd path between x, y along C and adding the edges vx, vy, we get a shorter odd cycle, a contradiction. $\square$

We will also use the following result of Letzter and Snyder, see [17, Corollary 32].

Theorem 5.6 ([17])**.** *Let G be a $\{C_3, C_5\}$-free graph on n vertices with $\delta(G) > \frac{n}{4}$. Then G is homomorphic to C_7.*

We can now finally prove the upper bound in the last case of Theorem 1.3.

Theorem 5.7. *Let H be a graph such that $\chi(H) = 3$, H contains critical edge xy, and $\text{odd-girth}(H) \geq 5$. Then $\delta_{\text{lin-rem}}(H) \leq \frac{1}{4}$.*

Proof. Denote $h = |V(H)|$. Write $\text{odd-girth}(G) = 2k + 1 \geq 5$. By the second item of Lemma 5.1, there is a partition $V(H) = A_1 \cup A_2 \cup \cdots \cup A_{2k+1}$ such that $|A_1| = |A_2| = 1$, and $E(H) \subseteq \bigcup_{i \in [2k+1]} A_i \times A_{i+1}$. Denote $s_i = |A_i|$ for each $i \in [2k + 1]$, so H is a subgraph of the blow-up $C_{2k+1}[s_1, \ldots, s_{2k+1}]$ of C_{2k+1}. Let $c_1 = c_1(C_{2k+1}, s_1, \ldots, s_{2k+1}) > 0$ and $c_2 = c_2(k) > 0$ be the constants given by Lemma 2.1 and Lemma 2.3, respectively. According to Theorem 1.2, $\delta_{\text{poly-rem}}(C_{2k+1}) = \frac{1}{2k+1} < \frac{1}{4}$, and hence there exists a constant $c_3 = c_3(k) > 0$ such that if G is a graph on n vertices with $\delta(G) \geq \frac{n}{4}$ and at least εn^2 edge-disjoint C_{2k+1}-copies, then G contains at least $c_3 \varepsilon^{\frac{1}{c_3}} n^{2k+1}$ copies of C_{2k+1}. Set $c := c_1 \cdot \min(c_2, c_3)$.

Let $\alpha > 0$ and ε be small enough; it suffices to assume that $\varepsilon < \left(\frac{\alpha^2}{200k(k+2)}\right)^{1/c}$. Let G be a graph on n vertices with $\delta(G) \geq (\frac{1}{4} + \alpha)n$ which contains at least εn^2 edge-disjoint copies of H. Our goal is to show that G contains $\Omega_{H,\alpha}(\varepsilon n^h)$ copies of H. Suppose first that G contains at least $\varepsilon^c n^2$ edge-disjoint copies of $C_{2\ell+1}$ for some $1 \leq \ell \leq k$. If $\ell < k$, then G contains $\Omega_k(\varepsilon^{c/c_2} n^{2k+1}) = \Omega_k(\varepsilon^{c_1} n^{2k+1})$ copies of C_{2k+1} by Lemma 2.3 and the choice of c_2. And if $\ell = k$, then G contains $\Omega_k(\varepsilon^{c/c_3} n^{2k+1}) = \Omega_k(\varepsilon^{c_1} n^{2k+1})$ copies of C_{2k+1} by Theorem 1.2 and the choice of c_3. In either case, G contains $\Omega_k(\varepsilon^{c_1} n^{2k+1})$ copies of C_{2k+1}. But then, by Lemma 2.1 (with $V_1 = \cdots = V_{2k+1} = V(G)$), G contains at least $\Omega_H(\varepsilon^{c_1/c_1} n^h) = \Omega_H(\varepsilon n^h)$ copies of $C_{2k+1}[s_1, \ldots, s_{2k+1}]$, and hence $\Omega_{H,\alpha}(\varepsilon n^h)$ copies of H. This concludes the proof of this case.

From now on, assume that G contains at most $\varepsilon^c n^2$ edge-disjoint $C_{2\ell+1}$-copies for every $\ell \in [k]$. Let $\mathcal{C}_\ell$ be a maximal collection of edge-disjoint $C_{2\ell+1}$-copies in G, so $|\mathcal{C}_\ell| \leq \varepsilon^c n^2$. Let E_c be the set of edges which are contained in one of the cycles in $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_k$. Let S be the set of vertices which are incident with at least $\frac{\alpha n}{10}$ edges from E_c. Then

$$|E_c| \le \sum_{\ell=1}^{k}(2\ell+1)\varepsilon^c n^2 = k(k+2)\varepsilon^c n^2 \text{ and } |S| \le \frac{2\,|E_c|}{\alpha n/10} \le \frac{20k(k+2)\varepsilon^c}{\alpha} n < \frac{\alpha n}{10}, \tag{1}$$

where the last inequality holds by our assumed bound on ε. Let G' be the subgraph of G obtained by deleting the edges in E_c and the vertices in S. Note that $G' \subseteq G - E_c$ is $\{C_3, C_5, \ldots, C_{2k+1}\}$-free because for every $1 \le \ell \le k$, we removed all edges from a maximal collection of edge-disjoint $C_{2\ell+1}$-copies.

Claim 5.8. *$|V(G')| > (1 - \frac{\alpha}{10})n$ and $\delta(G') > (\frac{1}{4} + \frac{4\alpha}{5})n$.*

Proof. The first inequality follows from (1) as $|V(G')| = n - |S|$. Each $v \in V(G)\backslash S$ has at most $\frac{\alpha n}{10}$ incident edges from E_c, and at most $|S| < \frac{\alpha n}{10}$ neighbors in S, thereby $\deg_{G'}(v) > \deg_G(v) - \frac{\alpha n}{5} \ge (\frac{1}{4} + \frac{4\alpha}{5})n$. Hence, $\delta(G') > (\frac{1}{4} + \frac{4\alpha}{5})n$. □

Claim 5.9. *G' is homomorphic to C_7. Moreover, G' is bipartite unless $k = 2$.*

Proof. Recall that G' is $\{C_3, C_5, \ldots, C_{2k+1}\}$-free. As $k \ge 2$, G' is $\{C_3, C_5\}$-free. Also, $\delta(G') > \frac{n}{4} \ge \frac{|V(G')|}{4}$ by Claim 5.8. So G' is homomorphic to C_7 by Theorem 5.6. If $k \ge 3$, i.e. odd-girth$(H) \ge 7$, then odd-girth$(G') \ge 2k + 3 \ge 9$. As $\delta(G') > \frac{n}{4}$, G' is bipartite by Lemma 5.5. □

The rest of the proof is divided into two cases based whether or not G' is bipartite. These cases are handled by Propositions 5.10 and 5.11, respectively.

Proposition 5.10. *Suppose that G' is bipartite. Then G has $\Omega_{H,\alpha}(\varepsilon n^h)$ copies of H.*

Proof. Let (L', R') be a bipartition of G', so $V(G) = L' \cup R' \cup S$. Let $L_1 \subseteq S$ (resp. $R_1 \subseteq S$) be the set of vertices of S having at most $\frac{\alpha n}{5}$ neighbors in L' (resp. R'). Let G'' be the bipartite subgraph of G induced by the bipartition $(L'', R'') := (L' \cup L_1, R' \cup R_1)$. Let $S'' = V(G)\backslash(L'' \cup R'')$, so $V(G) = L'' \cup R'' \cup S''$.

We claim that $\delta(G'') \ge (\frac{1}{4} + \frac{\alpha}{2})n$. First, as G' is a subgraph of G'', we have $\deg_{G''}(v) > (\frac{1}{4} + \frac{4\alpha}{5})n$ for each $v \in V(G') \subseteq V(G'')$ by Claim 5.8. Now we consider vertices in $V(G'') \setminus V(G') = L_1 \cup R_1$. Each $v \in L_1$ has at most $|S| \le \frac{\alpha n}{10}$ neighbors in S and at most $\frac{\alpha n}{5}$ neighbors in L', by the definition of L_1. Hence, v has at least $\deg_G(v) - \frac{3\alpha}{10}n \ge (\frac{1}{4} + \frac{\alpha}{2})n$ neighbors in $R' \subseteq V(G'')$. By the symmetric argument for vertices $v \in R_1$, we get that $\delta(G'') \ge (\frac{1}{4} + \frac{\alpha}{2})n$, as required.

For an edge $uv \in E(G)\backslash E(G'')$, we say uv is of type I if $u, v \in L''$ or $u, v \in R''$, and we say that uv is of type II if $u \in S''$ or $v \in S''$. Every edge in $E(G)\backslash E(G'')$ is of type I or II. Since $\chi(H) = 3$ and G'' is bipartite, each copy of H in G must contain an edge of type I or an edge of type II (or both). As G has εn^2 edge-disjoint H-copies, G contains at least $\frac{\varepsilon n^2}{2}$ edges of type I or at least $\frac{\varepsilon n^2}{2}$ edges of type II. We now consider these two cases separately. See Fig. 4 for an illustration. Recall that $xy \in E(H)$ denotes a critical edge of H.

Case 1: *G contains $\frac{\varepsilon n^2}{2}$ edges of type I.* Fix any edge $ab \in E(G)$ of type I. Without loss of generality, assume $a, b \in L''$ (the case $a, b \in R''$ is symmetric). We claim that G has poly$(\alpha)n^{h-2}$ copies of H mapping $xy \in E(H)$ to $ab \in E(G)$. If $\deg_G(a, b) \ge \frac{\alpha n}{2}$ then this holds by Lemma 5.2. Otherwise, $\deg_G(a, b) < \frac{\alpha n}{2}$, and thus

$$|R''| \ge |N_{G''}(a) \cup N_{G''}(b)| \ge \deg_{G''}(a) + \deg_{G''}(b) - \deg_G(a, b) > 2\delta(G'') - \frac{\alpha n}{2} > \frac{n}{2},$$

using that $\delta(G'') \ge (\frac{1}{4} + \frac{\alpha}{2})n$. Thus, $|L''| < \frac{n}{2}$. This implies that for all $a' \in N_{G''}(a), b' \in N_{G''}(b)$,

$$\deg_{G''}(a', b') \ge 2\delta(G'') - |L''| \ge \alpha n.$$

Now, by Lemma 5.3 (with $A = N_{G''}(a)$ and $B = N_{G''}(b)$), there are poly$(\alpha)n^{h-2}$ copies of H mapping xy to ab, as claimed. Summing over all edges ab of type I, we get $\frac{\varepsilon n^2}{2} \cdot \text{poly}(\alpha)n^{h-2} = \text{poly}(\alpha)\varepsilon n^h$ copies of H. This completes the proof in Case 1.

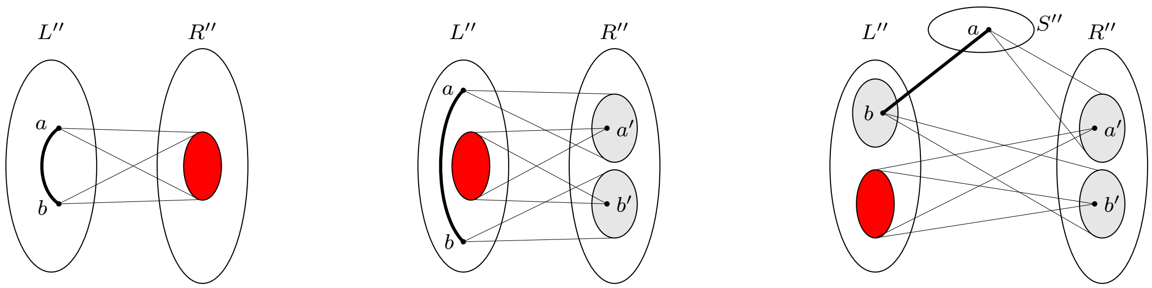

Figure 4: Proof of Proposition 5.10: Case 1 with $\deg_G(a,b) \geq \frac{\alpha n}{2}$ (left), Case 1 with $\deg_G(a,b) < \frac{\alpha n}{2}$ (middle) and Case 2 (right). The red part is the common neighborhood of a and b (or a' and b').

Case 2: *G contains $\frac{\varepsilon n^2}{2}$ edges of type II.* Note that the number of edges of type II is trivially at most $|S''|n$. Thus, $|S''| \geq \frac{\varepsilon n}{2}$. Fix some $a \in S''$. By the definition of L'', R'' and S'', v has at least $\frac{\alpha n}{5}$ neighbors in $L' \subseteq L''$ and at least $\frac{\alpha n}{5}$ neighbors in $R' \subseteq R''$. Without loss of generality, assume $|L''| \leq |R''|$, thereby $|L''| \leq \frac{n}{2}$. Now fix any $b \in L''$ adjacent to a; there are at least $\frac{\alpha n}{5}$ choices for b. We have $|N_G(a) \cap R''| \geq \frac{\alpha n}{5}$ and $|N_{G''}(b)| \geq \delta(G'') > \frac{n}{4}$, and for all $a' \in N_G(a) \cap R'', b' \in N_{G''}(b) \subseteq R''$ it holds that $\deg_{G''}(a',b') \geq 2\delta(G'') - |L''| \geq \alpha n$. Therefore, by Lemma 5.3, G has $\mathrm{poly}(\alpha)n^{h-2}$ copies of H mapping xy to ab. Enumerating over all $a \in S''$ and $b \in N_G(a) \cap L''$, we again get $\Omega_{H,\alpha}(\varepsilon n^h)$ copies of H in G. This completes the proof of Proposition 5.10. □

Proposition 5.11. *Suppose G' is non-bipartite but homomorphic to C_7. Then G has $\Omega_{H,\alpha}(\varepsilon n^h)$ copies of H.*

Proof. By Claim 5.9 we must have $k = 2$, so odd-girth$(H) = 5$. The proof is similar to that of Proposition 5.10, but instead of a bipartition of G', we use a partition corresponding to a homomorphism into C_7. Let $V(G)\backslash S = V(G') = V_1' \cup V_2' \cup \cdots \cup V_7'$ be a partition of $V(G')$ such that $E(G') \subseteq \bigcup_{i\in[7]} V_i' \times V_{i+1}'$. Here and later, all subscripts are modulo 7. We have $V_i' \neq \emptyset$ for all $i \in [7]$, because otherwise G' would be bipartite. For $i \in [7]$, let S_i be the set of vertices in S having at most $\frac{2\alpha n}{5}$ neighbors in $V(G') \setminus (V_{i-1}' \cup V_{i+1}')$. In case v lies in multiple S_i's, we put v arbitrarily in one of them. Set $V_i'' := V_i' \cup S_i$. Let G'' be the 7-partite subgraph of G with parts $V_1'', \dots, V_7''$ and with all edges of G between V_i'' and V_{i+1}'', $i = 1, \dots, 7$. By definition, G' is a subgraph of G'', and G'' is homomorphic to C_7 via the homomorphism $V_i'' \mapsto i$. Put $S'' := V(G)\backslash V(G'') = S \setminus \bigcup_{i=1}^7 S_i$. We now collect the following useful properties.

Claim 5.12. *The following holds:*

(i) $\delta(G'') \geq (\frac{1}{4} + \frac{\alpha}{2})n$.

(ii) For every $i \in [7]$ and for every $u, v \in V_i''$ or $u \in V_i'', v \in V_{i+2}''$, it holds that $\deg_{G''}(u,v) \geq \alpha n$.

(iii) For every $i \in [7]$, every $v \in V_i''$ has at least αn neighbors in V_{i-1}'' and at least αn neighbors in V_{i+1}''.

(iv) For every $a \in S''$, there are i, j with $j - i \equiv 1, 3 \pmod 7$ and $|N_G(a) \cap V_i''|, |N_G(a) \cap V_j''| > \frac{2\alpha n}{25}$.

Proof.

(i) Let $i \in [7]$ and $v \in V_i''$. If $v \in V(G')$, then $\deg_{G''}(v) \geq \deg_{G'}(v) \geq \delta(G') > (\frac{1}{4} + \frac{\alpha}{2})n$, using Claim 5.8. Otherwise, $v \in S_i$. By definition, v has at most $\frac{2\alpha n}{5}$ neighbours in $V(G') \setminus (V_{i-1}' \cup V_{i+1}')$. Also, v has at most $|S| \leq \frac{\alpha n}{10}$ neighbours in S. It follows that v has at least $\deg_G(v) - \frac{2\alpha n}{5} - \frac{\alpha n}{10} \geq (\frac{1}{4} + \frac{\alpha}{2})n$ neighbors in $V_{i-1}'' \cup V_{i+1}''$. Hence, $\deg_{G''}(v) > (\frac{1}{4} + \frac{\alpha}{2})n$.

(ii) First, observe that

$$|V_i''| + |V_{i+2}''| \geq \left(\frac{1}{4} + \frac{\alpha}{2}\right)n \tag{2}$$

for all $i \in [7]$. Indeed, V''_{i+1} is non-empty, and fixing any $v \in V''_{i+1}$, we have $|V''_i| + |V''_{i+2}| \geq \deg_{G''}(v) \geq \delta(G'') \geq (\frac{1}{4} + \frac{\alpha}{2})n$. By applying (2) to the pairs $(i+2, i+4)$ and $(i-2, i)$, we get

$$|V''_{i-1}| + |V''_{i+1}| + |V''_{i+3}| \leq n - (|V''_{i+2}| + |V''_{i+4}|) - (|V''_{i-2}| + |V''_i|) \leq n - 2\left(\frac{1}{4} + \frac{\alpha}{2}\right)n < \frac{n}{2}. \quad (3)$$

Now let $i \in [7]$. For $u, v \in V''_i$ we have $N_{G''}(u) \cup N_{G''}(v) \subseteq V''_{i-1} \cup V''_{i+1}$, and for $u \in V''_i, v \in V''_{i+2}$ we have $N_{G''}(u) \cup N_{G''}(v) \subseteq V''_{i-1} \cup V''_{i+1} \cup V''_{i+3}$. In both cases, $|N_{G''}(u) \cup N_{G''}(v)| < \frac{n}{2}$ by (3). As $\deg_{G''}(u) + \deg_{G''}(v) \geq 2\delta(G'') \geq (\frac{1}{2} + \alpha)n$, we have $\deg_{G''}(u, v) > \alpha n$, as required.

(iii) We first argue that $|V''_i| \leq (\frac{1}{4} - \frac{3\alpha}{2})n$ for each $i \in [7]$. Indeed, by applying (2) to the pairs $(i-1, i+1)$, $(i+2, i+4)$, $(i+3, i+5)$, we get

$$|V''_i| \leq n - (|V''_{i-1}| + |V''_{i+1}|) - (|V''_{i+2}| + |V''_{i+4}|) - (|V''_{i+3}| + |V''_{i+5}|) \leq n - 3\left(\frac{1}{4} + \frac{\alpha}{2}\right)n = \left(\frac{1}{4} - \frac{3\alpha}{2}\right)n.$$

Now, for every $v \in V''_i$, we have $N_{G''}(v) \subseteq V''_{i-1} \cup V''_{i+1}$ and $|V''_{i-1}|, |V''_{i+1}| < (\frac{1}{4} - \frac{3\alpha}{2})n$. Hence, v has at least $\deg_{G''}(v) - (\frac{1}{4} - \frac{3\alpha}{2})n \geq \alpha n$ neighbors in each of V''_{i-1}, V''_{i+1}.

(iv) Let I be the set of i with $|N_G(a) \cap V''_i| \geq \frac{2\alpha n}{25}$. If I is empty, then a has less than $5 \cdot \frac{2\alpha n}{25} = \frac{2\alpha n}{5}$ neighbors in every $V(G') \setminus (V'_{i-1} \cup V'_{i+1})$ and therefore can not be in S''. Suppose for contradiction that there exist no $i, j \in I$ with $j - i \equiv 1, 3 \pmod 7$. We claim that there is $j \in [7]$ such that $I \subseteq \{j, j+2\}$. Fix an arbitrary $i \in I$. Then, $i \pm 1, i \pm 3 \notin I$ by assumption. Also, at most one of $i+2, i-2$ is in I, because $(i-2) - (i+2) \equiv 3 \pmod 7$. So $I \subseteq \{i, i+2\}$ or $I \subseteq \{i-2, i\}$, proving our claim that $I \subseteq \{j, j+2\}$ for some j. By the definition of I, a has at most $5 \cdot \frac{2\alpha n}{25} = \frac{2\alpha n}{5}$ neighbors in $V(G') \setminus (V'_j \cup V'_{j+2})$. Hence, $a \in S_{j+1}$. This contradicts the fact that $a \in S''$, as $S'' \cap S_{i+1} = \emptyset$.

□

We continue with the proof of Proposition 5.11. Recall that the edges in $E(G) \setminus E(G'')$ are precisely the edges of G not belonging to $\bigcup_{i \in [7]} V''_i \times V''_{i+1}$. For an edge $ab \in E(G) \setminus E(G'')$, we say ab is of type I if $a, b \in V(G'')$, and of type II if $a \in S''$ or $b \in S''$. Clearly, every edge in $E(G) \setminus E(G'')$ is either of type I or of type II. Since odd-girth$(H) = 5$ and C_5 is not homomorphic to C_7, every H-copy in G must contain some edge of type I or of type II (or both). As G has εn^2 edge-disjoint H-copies, G must have at least $\frac{\varepsilon n^2}{2}$ edges of type I or at least $\frac{\varepsilon n^2}{2}$ edges of type II. We consider these two cases separately. See Fig. 5 for an illustration. Recall that $xy \in E(H)$ denotes a critical edge of H.

Case 1: *G contains $\frac{\varepsilon n^2}{2}$ edges of type I.* Fix any edge ab of type I, where $a \in V''_i$ and $b \in V''_j$ for $i, j \in [7]$. We now show that G has poly$(\alpha)n^{h-2}$ copies of H mapping $xy \in E(H)$ to ab. As $ab \notin E(G'')$, we have $i - j \equiv 0, \pm 2, \pm 3 \pmod 7$. When $j - i \equiv 0, \pm 2 \pmod 7$, we have $\deg_G(a, b) \geq \deg_{G''}(a, b) > \alpha n$ by Claim 5.12 (ii). Then, by Lemma 5.2, G has poly$(\alpha)n^{h-2}$ copies of H mapping xy to ab, as required. Now suppose that $j - i \equiv \pm 3 \pmod 7$, say $j \equiv i + 3 \pmod 7$. Denote $A := N_G(a) \cap V''_{i-1}$ and $B := N_G(b) \cap V''_{j+1} = N_G(b) \cap V''_{i-3}$. We have that $|A|, |B| \geq \alpha n$ by Claim 5.12 (iii), and $|N_G(a', b')| > \alpha n$ for all $a' \in A, b' \in B$ by Claim 5.12 (ii). Now, by Lemma 5.3, G has poly$(\alpha)n^{h-2}$ copies of H mapping xy to ab, proving our claim. Summing over all edges ab of type I, we get $\frac{\varepsilon n^2}{2} \cdot \text{poly}(\alpha)n^{h-2} = \Omega_{H,\alpha}(\varepsilon n^h)$ copies of H in G, finishing this case.

Case 2: *G contains $\frac{\varepsilon n^2}{2}$ edges of type II.* Notice that the number edges incident to S'' is at most $|S''|\, n$, meaning that $|S''| \geq \frac{\varepsilon n}{2}$. Fix any $a \in S''$. By Claim 5.12 (iv), there exist $i, j \in [7]$ with $j - i \equiv 1, 3 \pmod 7$ and $|N_G(a) \cap V''_i|, |N_G(a) \cap V''_j| > \frac{2\alpha n}{25}$. Fix any $b \in N_G(a) \cap V''_i$ (there are at least $\frac{2\alpha n}{25}$ choices for b). Take $A = N_G(a) \cap V''_j$ and $B = N_G(b) \cap V''_{i+1}$. We have that $|A| \geq \frac{2\alpha n}{25}$, and $|B| \geq \alpha n$ by Claim 5.12 (iii). Further, as $j - (i+1) \equiv 0, 2 \pmod 7$, Claim 5.12 (ii) implies that $|N_G(a', b')| > \alpha n$ for all $a' \in A, b' \in B$. Now, by Lemma 5.3, G has poly$(\alpha)n^{h-2}$ copies of H mapping xy to ab. Summing over all choices of $a \in S''$ and $b \in V''_i$, we acquire $|S''| \cdot \frac{2\alpha n}{25} \cdot \text{poly}(\alpha)n^{h-2} = \Omega_{H,\alpha}(\varepsilon n^h)$ copies of H in G. This completes the proof of Case 2, and hence the proposition. □

Propositions 5.10 and 5.11 imply the theorem. □

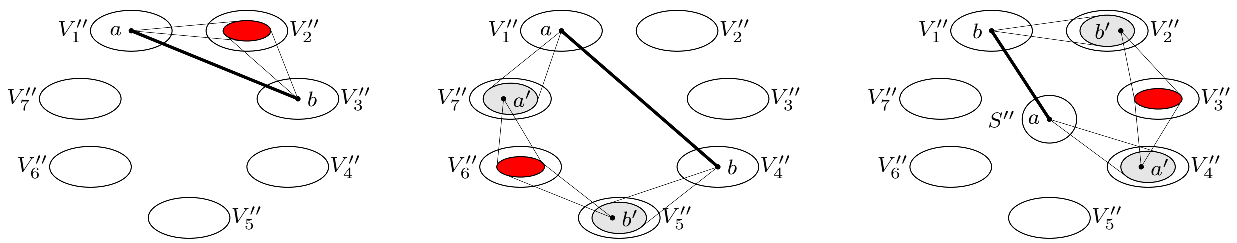

Figure 5: Proof of Proposition 5.11: Case 1 for $j = i + 2$ (left), Case 1 for $j = i + 3$ (middle) and Case 2 for $j = i + 3$ (right). The red part is the common neighborhood of a and b (or a' and b').

6 Concluding remarks and open questions

It would be interesting to determine the possible values of $\delta_{\text{poly-rem}}(H)$ for 3-chromatic graphs H. So far we know that $\frac{1}{2k+1}$ is a value for each $k \geq 1$. Is there a graph H with $\frac{1}{5} < \delta_{\text{poly-rem}}(H) < \frac{1}{3}$? Also, is it true that $\delta_{\text{poly-rem}}(H) > \frac{1}{5}$ if H is not homomorphic to C_5?

Another question is whether the inequality in Theorem 1.4 is always tight, i.e. is it always true that $\delta_{\text{poly-rem}}(H) = \delta_{\text{hom}}(\mathcal{I}_H)$?

Finally, we wonder whether the parameters $\delta_{\text{poly-rem}}(H)$ and $\delta_{\text{lin-rem}}(H)$ are monotone, in the sense that they do not increase when passing to a subgraph of H. We are not aware of a way of proving this without finding $\delta_{\text{poly-rem}}(H), \delta_{\text{lin-rem}}(H)$.

References

[1] P. Allen, J. Böttcher, S. Griffiths, Y. Kohayakawa, and R. Morris. The chromatic thresholds of graphs. *Advances in Mathematics*, 235:261–295, 2013. 1

[2] N. Alon. Testing subgraphs in large graphs. *Random Structures & Algorithms*, 21(3-4):359–370, 2002. 1, 3, 1

[3] N. Alon and J. H. Spencer. *The probabilistic method*. John Wiley & Sons, 2016. 4

[4] B. Andrásfai, P. Erdös, and V. T. Sós. On the connection between chromatic number, maximal clique and minimal degree of a graph. *Discrete Mathematics*, 8(3):205–218, 1974. 1, 5

[5] S. Brandt. On the structure of dense triangle-free graphs. *Combinatorics, Probability and Computing*, 8(3):237–245, 1999. 1

[6] S. Brandt and S. Thomassé. Dense triangle-free graphs are four-colorable: A solution to the Erdős-Simonovits problem. *preprint*, 2011. 1

[7] C.-C. Chen, G. P. Jin, and K. M. Koh. Triangle-free graphs with large degree. *Combinatorics, Probability and Computing*, 6(4):381–396, 1997. 1

[8] O. Ebsen and M. Schacht. Homomorphism thresholds for odd cycles. *Combinatorica*, 40(1):39–62, 2020. 1, 1

[9] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel Journal of Mathematics*, 2(3):183–190, 1964. 2

[10] P. Erdős and M. Simonovits. On a valence problem in extremal graph theory. *Discrete Mathematics*, 5(4):323–334, 1973. 1

[11] J. Fox. A new proof of the graph removal lemma. *Annals of Mathematics*, pages 561–579, 2011. 1

[12] J. Fox and Y. Wigderson. Minimum degree and the graph removal lemma. *Journal of Graph Theory*, 2021. 1, 1, 1, 1, 2, 3, 3

[13] L. Gishboliner, A. Shapira, and Y. Wigderson. An efficient asymmetric removal lemma and its limitations. *arXiv preprint arXiv:2301.07693*, 2023. 2

[14] W. Goddard and J. Lyle. Dense graphs with small clique number. *Journal of Graph Theory*, 66(4):319–331, 2011. 1

[15] R. Häggkvist. Odd cycles of specified length in non-bipartite graphs. In *North-Holland Mathematics Studies*, volume 62, pages 89–99. Elsevier, 1982. 1

[16] G. Jin. Triangle-free four-chromatic graphs. *Discrete Mathematics*, 145(1-3):151–170, 1995. 1

[17] S. Letzter and R. Snyder. The homomorphism threshold of $\{C_3, C_5\}$-free graphs. *Journal of Graph Theory*, 90(1):83–106, 2019. 1, 5, 5.6

[18] T. Łuczak. On the structure of triangle-free graphs of large minimum degree. *Combinatorica*, 26(4):489–493, 2006. 1

[19] T. Łuczak and S. Thomassé. Coloring dense graphs via VC-dimension. *arXiv preprint arXiv:1007.1670*, 2010. 1

[20] J. Lyle. On the chromatic number of H-free graphs of large minimum degree. *Graphs and Combinatorics*, 27(5):741–754, 2011. 1

[21] Y. Nakar and D. Ron. On the testability of graph partition properties. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018. 4, 4

[22] H. Oberkampf and M. Schacht. On the structure of dense graphs with bounded clique number. *Combinatorics, Probability and Computing*, 29(5):641–649, 2020. 1

[23] 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,*, volume 18, pages 939–945. North-Holland, Amsterdam-New York, 1978. 1

[24] M. Sankar. Homotopy and the homomorphism threshold of odd cycles. *arXiv preprint arXiv:2206.07525*, 2022. 1

[25] C. Thomassen. On the chromatic number of triangle-free graphs of large minimum degree. *Combinatorica*, 22(4):591–596, 2002. 1

[26] C. Thomassen. On the chromatic number of pentagon-free graphs of large minimum degree. *Combinatorica*, 27(2):241–243, 2007. 1