REMOVAL LEMMAS AND APPROXIMATE HOMOMORPHISMS

JACOB FOX AND YUFEI ZHAO

Abstract. We study quantitative relationships between the triangle removal lemma and several of its variants. One such variant, which we call the *triangle-free lemma*, states that for each $\epsilon > 0$ there exists M such that every triangle-free graph G has an ϵ-approximate homomorphism to a triangle-free graph F on at most M vertices (here an *ϵ-approximate homomorphism* is a map $V(G) \to V(F)$ where all but at most $\epsilon |V(G)|^2$ edges of G are mapped to edges of F). One consequence of our results is that the least possible M in the triangle-free lemma grows faster than exponential in any polynomial in ϵ^{-1}. We also prove more general results for arbitrary graphs, as well as arithmetic analogues over finite fields, where the bounds are close to optimal.

1. Introduction

1.1. Graph removal and related results.

The triangle removal lemma of Ruzsa and Szemerédi [27] is a fundamental tool in extremal combinatorics.

Theorem 1.1 (Triangle removal lemma)**.** *For every $\epsilon > 0$, there exists $\delta > 0$ such that every n-vertex graph with fewer than δn^3 triangles can be made triangle-free by deleting at most ϵn^2 edges.*

Definition 1.2. Let $\delta_{TRL}(\epsilon)$ denote the largest possible constant δ in Theorem 1.1.

The standard proof of the triangle removal lemma, which uses Szemerédi's regularity lemma [30], gives an upper bound on $\delta_{TRL}(\epsilon)^{-1}$ which is a tower of 2's of height $\epsilon^{-O(1)}$. The tower height was improved to $O(\log(1/\epsilon))$ by Fox [8]. On the other hand, only a slightly superpolynomial lower bound $1/\delta_{TRL}(\epsilon) \geq (1/\epsilon)^{c \log(1/\epsilon)}$ is known [27], coming from the Behrend construction of large sets without 3-term arithmetic progressions [3].

The standard regularity proof of the triangle removal lemma actually shows that edges can be removed in a bounded complexity way.

Theorem 1.3 (Triangle removal lemma with bounded complexity)**.** *For every $\epsilon > 0$, there exist $\delta > 0$ and M such that for every n-vertex graph G with fewer than δn^3 triangles, there is a vertex partition $V(G) = V_1 \cup \ldots \cup V_M$, and a triangle-free graph G' on $V(G)$ which is complete or empty between each pair (V_i, V_j) and satisfying $|E(G) \setminus E(G')| \leq \epsilon n^2$.*

The above formulation of the removal lemma was highlighted by Tao [33], who gave a proof of the hypergraph removal lemma with similar bounded complexity features (the hypergraph removal lemma was independently proved by Gowers [14] and Rödl and Schacht [26]) and then used it to establish a removal lemma for sparse hypergraphs, which then led to the Gaussian integer analogue of the Green–Tao theorem [32] (also see [5] for an improvement and simplification).

We introduce the notion of an approximate graph homomorphism, which allows us to give a succinct restatement of the above result.

Definition 1.4 (Approximate homomorphisms)**.** Given graphs G and F, a map $\phi\colon V(G) \to V(F)$ is an *ϵ-approximate homomorphism* if at most $\epsilon |V(G)|^2$ edges of G do not map to edges of F under ϕ.

Fox was supported by a Packard Fellowship and by NSF award DMS-1855635. Zhao was supported by NSF Award DMS-1764176, a Sloan Research Fellowship, and the MIT Solomon Buchsbaum Fund.

The usual notion of a graph homomorphism corresponds to $\epsilon = 0$. With this notion, Theorem 1.3 is equivalent to the following statement.

Theorem 1.5 (Triangle removal lemma with bounded complexity, rephrased)**.** *For every $\epsilon > 0$, there exist $\delta > 0$ and M such that every n-vertex graph G with fewer than δn^3 triangles has an ϵ-approximate homomorphism into some triangle-free graph with at most M vertices.*

The following special case of Theorem 1.5 for triangle-free graphs G is already interesting.

Theorem 1.6 (Triangle-free lemma)**.** *For every $\epsilon > 0$, there exists M such that every triangle-free graph has an ϵ-approximate homomorphism to a triangle-free graph on at most M vertices.*

Definition 1.7. Let $M_{TFL}(\epsilon)$ denote the smallest possible M in Theorem 1.6.

Note that the triangle removal lemma (Theorem 1.1) and triangle-free lemma (Theorem 1.6) together imply Theorems 1.3 and 1.5. Indeed, starting with an n-vertex graph with fewer than $\delta_{TRL}(\epsilon/2)n^3$ triangles, first delete $(\epsilon/2)n^2$ edges to get rid of all triangles, and then find an $\epsilon/2$-approximate homomorphism into a triangle-free graph on $M_{TRL}(\epsilon/2)$ vertices.

Motivated by graph property testing, Hoppen, Kohayakawa, Lang, Lefmann, and Stagni [18] showed that one can deduce Theorems 1.3, 1.5 and 1.6 using the triangle removal lemma (Theorem 1.1) combined with the Frieze–Kannan weak regularity lemma [12]. In particular, the deduction does not need the full Szemerédi graph regularity lemma. This implies that

$$M_{TFL}(\epsilon) \le e^{O(\delta_{TRL}(\epsilon/C)^{-2})}, \tag{1.1}$$

which is already better than the usual bound of $M_{TFL} \le \text{tower}(\epsilon^{-O(1)})$ obtained from the standard regularity proof (here $\text{tower}(m)$ denotes an exponential tower of 2's of height m). Indeed, (1.1) is superior since $1/\delta_{TRL}(\epsilon) \le \text{tower}(O(\log(1/\epsilon)))$ [8], and potentially $1/\delta_{TRL}(\epsilon)$ could be much smaller. We include a proof sketch of (1.1) in Section 5.

We provide a complementary lower bound to $M_{TFL}(\epsilon)$ in terms of the following close cousin of the triangle removal lemma.

Theorem 1.8 (Diamond-free lemma)**.** *For every $\epsilon > 0$, there exists some N such that for every $n \ge N$, every n-vertex graph where each edge lies in a unique triangle has at most ϵn^2 edges.*

Definition 1.9. Let $N_{DFL}(\epsilon)$ denote the smallest constant N so that Theorem 1.8 holds.

The diamond-free lemma is a direct corollary of the triangle removal lemma, yielding $N_{DFL}(\epsilon) \le 1/\delta_{TRL}(\epsilon/3)$. Indeed, suppose we have a graph on $n \ge 1/\delta_{TRL}(\epsilon/3)$ vertices and each edge lies in a unique triangle. Then the number of triangles is at most a third times the number of edges, which is at most $n^2 \le \delta_{TRL}(\epsilon/3)n^3$. So by the triangle removal lemma, one can remove at most $(\epsilon/3)n^2$ edges to make this graph triangle-free. Since the graph was made up of edge-disjoint triangles, it has at most ϵn^2 edges.

A notable application of the diamond-free lemma is the graph theoretic proof of Roth's theorem on 3-term arithmetic progressions by Ruzsa and Szemerédi [26]. In fact, this application was one of the original motivations for the triangle removal lemma. Solymosi [29] also used the diamond-free lemma to give a short proof of the corners theorem of Ajtai and Szemerédi [1]. The best known lower bound on $N_{DFL}(\epsilon)$ has the form $(1/\epsilon)^{c\log(1/\epsilon)}$, which arises from the Behrend construction of large sets without 3-term arithmetic progressions (for recent improvements on the constant c coming from improved lower bound constructions related to the corners theorem, see [22, 16]).

Here is a representative case of our main result. It gives an exponential lower bound for the triangle-free lemma in terms of the bounds in the diamond-free lemma.

Theorem 1.10. *There exists a constant $C > 0$ such that, for every $\epsilon > 0$,*

$$M_{TFL}(\epsilon) \ge e^{\epsilon N_{DFL}(C\epsilon)/C}.$$

Using the best known lower bound on $N_{DFL}(\epsilon)$, we deduce the following superexponential lower bound on $M_{TFL}(\epsilon)$ in terms of $1/\epsilon$.

Corollary 1.11. *There exists a constant $c > 0$ such that for all $0 < \epsilon < 1/2$,*

$$M_{TFL}(\epsilon) \geq e^{(1/\epsilon)^{c \log(1/\epsilon)}}.$$

We suspect that $N_{DFL}(\epsilon)$ and $1/\delta_{TRL}(\epsilon)$ have similar growth. The next result provides evidence for this suspicion. We show that if $N_{DFL}(\epsilon)$ grows subexponentially in ϵ^{-1}, then $1/\delta_{TRL}(\epsilon)$ does as well. The proof of the theorem is based on a similar proof in the arithmetic setting by Fox and Lovász [9] but uses vertex subset sampling instead of subspace sampling.

Theorem 1.12. *Fix $0 < c < 1$. If $N_{DFL}(\epsilon) \leq 2^{\epsilon^{-c+o(1)}}$ as $\epsilon \to 0$, then $\delta_{TRL}(\epsilon) \geq 2^{-\epsilon^{-c/(1-c)+o(1)}}$ as $\epsilon \to 0$.*

If $N_{DFL}(\epsilon)$ and $1/\delta_{TRL}(\epsilon)$ have similar growth (as is the case if $N_{DFL}(\epsilon)$ grows subexponentially by Theorem 1.12), then Theorem 1.10 and the inequality (1.1) would give comparable lower and upper bounds on $M_{TFL}(\epsilon)$. Below we also discuss the arithmetic analogue, in which case the best lower and upper bounds indeed match.

Here is the proof strategy for Theorem 1.10. We start with a graph satisfying the hypotheses of the diamond-free lemma, namely that every edge lies in a unique triangle. We blow up this graph and then carefully construct a triangle-free subgraph. By the triangle-free lemma, this final graph we constructed must have an ϵ/C-approximate homomorphism to a triangle-free graph on $M_{TFL}(\epsilon/C)$ vertices, which then implies, by a novel entropy argument, that the original graph has at most $C\epsilon^{-1} \log M_{TFL}(\epsilon/C)$ triangles.

We state below extensions of the triangle removal lemma, the triangle-free lemma, and the diamond-free lemma from a triangle to an arbitrary graph H. These results are standard in the area, and their proofs use the same techniques as the triangle case.

Although some of these results are commonly stated in terms of H-free graphs (with caveats), it will be more natural and relevant for us to discuss them using the following formulations with H-homomorphism-free graphs. We say that a graph G is *H-homomorphism-free* if there is no graph homomorphism from H to G. A *homomorphic copy* of H in G is a subgraph of G that is the image of a homomorphism from H. The *core* of a graph H, denoted $\mathrm{core}(H)$, is defined to be the smallest subgraph of H that can arise as the image of a homomorphism of H (see [17]). The core of H is well-defined, i.e., it is unique up to graph isomorphism. Indeed, suppose $\phi, \psi \colon H \to H$ are both homomorphisms with images $\phi(H)$ and $\psi(H)$, then ψ gives a homomorphism from $\phi(H)$ to $\psi(H)$, and vice-versa with ϕ, so that the two images cannot both be minimal homomorphic copies of H unless they are isomorphic. For example, if H is a clique or an odd cycle, then $\mathrm{core}(H) = H$. Also, the core of H consists of a single edge if and only if H is bipartite and has at least one edge.

Theorem 1.13. *Let H be a graph. Let $\epsilon > 0$.*

(a) *There exists $\delta > 0$ such that every n-vertex graph with fewer than $\delta n^{|V(H)|}$ homomorphic copies of H can be made H-homomorphism-free by removing at most ϵn^2 edges.*
(b) *There exists some M such that every H-homomorphism-free graph has an ϵ-approximate homomorphism to an H-homomorphism-free graph on at most M vertices.*
(c) *Further suppose that H is connected and non-bipartite. There exists some N such that for every n-vertex graph G with $n \geq N$, if every edge of G lies in a unique homomorphic copy of $\mathrm{core}(H)$, then G has at most ϵn^2 edges.*

Definition 1.14. Let $\delta_H(\epsilon)$, $M_H(\epsilon)$, and $N_H(\epsilon)$ denote the optimal constants δ, M, and N, respectively, in Theorem 1.13.

Now we state our results comparing the bounds in Theorem 1.13, extending the earlier inequality (1.1) and Theorem 1.10 from triangles to general H. The lower bound is new. The upper bound below was already proved in [18], though we sketch a proof in Section 5.

Theorem 1.15 (Main theorem for graphs)**.** *For every connected non-bipartite graph H, there is some constant $C = C_H > 0$ such that, for every $0 < \epsilon < 1$,*

$$e^{\epsilon N_H(C\epsilon)/C} \leq M_H(\epsilon) \leq e^{C\delta_H(\epsilon/C)^{-2}}.$$

1.2. **Arithmetic analogue.** Green [15] developed an arithmetic analogue of Szemerédi's graph regularity lemma and used it to prove the following arithmetic analogue of the triangle removal lemma.

Let G be an abelian group. Given $X, Y, Z \subseteq G$, a *triangle* in $X \times Y \times Z$ is a triple $(x, y, z) \in X \times Y \times Z$ with $x + y + z = 0$.

Theorem 1.16 (Arithmetic triangle removal lemma)**.** *For every $\epsilon > 0$, there exists $\delta > 0$ such that for every finite abelian group G, and subsets $X, Y, Z \subseteq G$ with fewer than $\delta |G|^2$ triangles in $X \times Y \times Z$, we can remove all triangles by deleting at most $\epsilon |G|$ elements from each of X, Y, Z.*

Green's proof was Fourier analytic. It was later shown by Král, Serra, and Vena [20] that the arithmetic triangle removal lemma actually follows from the triangle removal lemma for graphs and even extends to all groups.

Here is the arithmetic analogue of the diamond-free lemma. It is a corollary of the arithmetic triangle-free lemma.

Theorem 1.17 (Arithmetic diamond-free lemma)**.** *For every $\epsilon > 0$, there exists N such that for every finite abelian group G with $|G| \geq N$, and $x_1, \ldots, x_l$, $y_1, \ldots, y_l$, $z_1, \ldots, z_l \in G$ satisfying $x_i + y_j + z_k = 0$ if and only if $i = j = k$, one has $l \leq \epsilon |G|$.*

The sets $\{x_1, \ldots, x_l\}$, $\{y_1, \ldots, y_l\}$, $\{z_1, \ldots, z_l\}$ in Theorem 1.17 are commonly known as "tricolor sum-free sets."

From now on, we restrict to the setting of $G = \mathbb{F}_p^n$ for a fixed p.

Definition 1.18. Let $\delta_p(\epsilon)$ denote the largest possible constant δ in Theorem 1.16 when restricted to groups of the form $G = \mathbb{F}_p^n$ for fixed prime p.

Definition 1.19. Let $N_p(\epsilon)$ denote the smallest positive integer so that Theorem 1.17 holds when restricted to groups of the form $G = \mathbb{F}_p^n$ with $p^n \geq N_p(\epsilon)$ and fixed prime p.

In this setting, Green's arithmetic regularity proof of Theorem 1.16 also gives us the following stronger statement, analogous of Theorems 1.3 and 1.5.

Theorem 1.20 (Arithmetic triangle removal lemma with bounded complexity)**.** *For every $\epsilon > 0$ and prime p, there exist $\delta > 0$ and a positive integer m such that if $X, Y, Z \subseteq \mathbb{F}_p^n$ are such that $X \times Y \times Z$ has fewer than δp^{2n} triangles, then there exist $X', Y', Z' \subseteq \mathbb{F}_p^m$ with $X' \times Y' \times Z'$ being triangle-free, and a linear map $\phi\colon \mathbb{F}_p^n \to \mathbb{F}_p^m$ such that at most ϵp^n elements from each of X, Y, Z do not get mapped to X', Y', Z' respectively.*

A special case is the following analogue of the triangle-free lemma (Theorem 1.6).

Theorem 1.21 (Arithmetic triangle-free lemma)**.** *For every $\epsilon > 0$ and prime p, there exists a positive integer m such that if $X, Y, Z \subseteq \mathbb{F}_p^n$ are such that $X \times Y \times Z$ is triangle-free, then there exist $X', Y', Z' \subseteq \mathbb{F}_p^m$ with $X' \times Y' \times Z'$ being triangle-free, and a linear map $\phi\colon \mathbb{F}_p^n \to \mathbb{F}_p^m$ such that at most ϵp^n elements from each of X, Y, Z do not get mapped to X', Y', Z' respectively.*

Definition 1.22. Let $m_p(\epsilon)$ denote the smallest m in Theorem 1.21. Let $M_p(\epsilon) = p^{m_p(\epsilon)}$.

Following a breakthrough of Croot, Lev, and Pach [6] and Ellenberg and Gijswijt [7] on the cap set problem, a number of developments together led to the following tight bound on $N_p(\epsilon)$. The upper bound on $N_p(\epsilon)$ was shown by Blasiak, Church, Cohn, Grochow, Naslund, Sawin, and Umans [4] and independently Alon (unpublished). The lower bound was first established by Kleinberg and Fu [13] for $p = 2$, and then in general by Kleinberg, Sawin, and Speyer [19] conditional on a conjecture later proved independently by Norin [24] and Pebody [25].

Theorem 1.23 (Optimal bounds in arithmetic diamond-free lemma for $\mathbb{F}_p^n$)**.** *For fixed prime p, as $\epsilon \to 0$, one has*

$$N_p(\epsilon) = \epsilon^{-1/c_p + o(1)}$$

with constant $0 < c_p < 1$ given by

$$p^{1-c_p} = \inf_{0<t<1} t^{-(p-1)/3}(1 + t + t^2 + \cdots + t^{p-1}). \tag{1.2}$$

Fox and Lovász [9] proved a polynomial dependence of parameters for the arithmetic triangle removal lemma over $\mathbb{F}_p^n$, and in fact determined the optimal exponent.

Theorem 1.24 (Optimal bounds in arithmetic triangle removal lemma for $\mathbb{F}_p^n$)**.** *For fixed prime p, as $\epsilon \to 0$, one has*

$$\delta_p(\epsilon) = \epsilon^{1+1/c_p + o(1)}$$

where $c_p > 0$ is the same constant defined in Theorem 1.23.

We prove the following analogue of Theorem 1.15.

Theorem 1.25 (Main theorem, arithmetic analogue)**.** *For any $0 < \epsilon < 1$ and prime p,*

$$p^{\epsilon N_p(5\epsilon)/p} \le M_p(\epsilon) \le p^{27\delta_p(\epsilon/4)^{-2}}.$$

Corollary 1.26. *For any fixed prime p, as $\epsilon \to 0$,*

$$\epsilon^{-1/c_p + 1 + o(1)} \le \log_p M_p(\epsilon) \le \epsilon^{-2/c_p - 2 + o(1)}.$$

One can check that $c_p = (0.172\cdots + o(1))/\log p$ as $p \to \infty$. Indeed, by writing $t = 1 - x/p$ we can deduce that $\lim_{p\to\infty}(\text{RHS of (1.2)})/p = \inf_{x>0} e^{x/3}(1 - e^{-x})/x = e^{-0.172\cdots}$. In particular, $c_p = \Theta(1/\log p)$. So we obtain the following bound.

Corollary 1.27. *There exists a universal constants $C > 0$ so that for all $0 < \epsilon < 1/2$ and prime p,*

$$\epsilon^{-(\log p)/C} \le \log_p M_p(\epsilon) \le \epsilon^{-C \log p}$$

For generalizations from triangles to longer cycles in $\mathbb{F}_p^n$, Lovász and Sauermann [23] extended the arithmetic diamond-free lemma with an optimal exponent, and Fox, Lovász, and Sauermann [10] extended the arithmetic removal lemma with a polynomial dependence but left open the optimal exponent.

It is possible to extend the above results from triangles to many other arithmetic patterns (including cycles), though we do not pursue this direction here so as not to further complicate matters. See [21, 28] for how to deduce removal lemmas for systems of linear equations over $\mathbb{F}_p$ from graph and hypergraph removal lemmas.

Organization. In Section 2 we prove the lower bound in Theorem 1.15, showing that the triangle-free lemma implies the diamond-free lemma with good bounds, as well as for general H. In Section 3, we prove Theorem 1.12, which shows that if the diamond-free lemma holds with subexponential bounds, then so does the triangle removal lemma. In Section 4 we prove the arithmetic analogue of the above, namely the lower bound in Theorem 1.25, which is based on similar ideas but has a somewhat cleaner execution. In Section 5 we prove the upper bounds in Theorems 1.15 and 1.25 by showing that, both for the graph version and the arithmetic analogue, the triangle removal lemma and the weak regularity lemma imply the diamond-free lemma with good bounds.

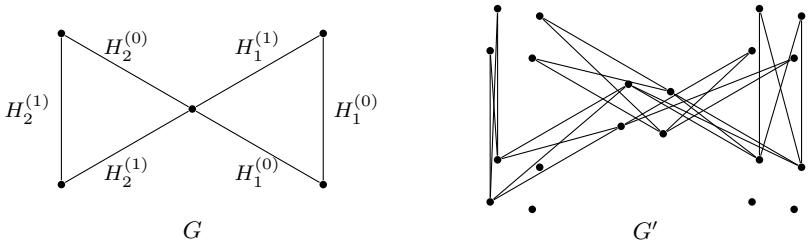

FIGURE 1. Illustration of the partial binary blow-up, Construction 2.1, for $H = K_3$.

2. Diamond-free versus triangle-free: graphs

Now we prove the lower bound $e^{\epsilon N_H(C\epsilon)/C} \le M_H(\epsilon)$ in Theorem 1.15. Note that being H-homomorphism-free is equivalent to being $\mathrm{core}(H)$-homomorphism-free. So it suffices to consider $H = \mathrm{core}(H)$, which will be the case for the rest of this section.

Construction 2.1 (Partial binary blow-up)**.** Suppose $H = \mathrm{core}(H)$ is connected and has more than one edge.

Let G be an n-vertex graph where every edge is contained in a unique homomorphic copy of H. Suppose there are exactly m homomorphic copies of H in G, and we enumerate them by $H_1, \dots, H_m$. We arbitrarily partition the edge-set of each H_i into two non-empty sets, resulting in $H_i = H_i^{(0)} \cup H_i^{(1)}$.

Let G' be a subgraph of the 2^m-blow-up of G constructed as follows. The vertices of G' are indexed by $V(G) \times \{0,1\}^m$. For each $i \in [m]$, $s \in \{0,1\}$, and $uv \in E(H_i^{(s)})$, the two vertices $(u, x_1, \dots, x_m)$ and $(v, y_1, \dots, y_m)$ in G' are adjacent if $x_i = y_i = s$. These are the only edges in G'.

See Figure 1 for an example of the construction.

Lemma 2.2. *The graph G' obtained in Construction 2.1 is H-homomorphism-free.*

Proof. Suppose we have a homomorphism $\phi\colon H \to G'$. We obtain a homomorphism $\psi\colon H \to G$ by composing ϕ with the homomorphism $G' \to G$ obtained by projection on the first coordinate of $V(G') = V(G) \times \{0,1\}^m$. Since every edge of G lies on a unique homomorphic copy of H, ψ must map H to some H_i (notated as in Construction 2.1). Consider the i-th binary coordinate of $\phi(v)$ for $v \in V(H)$. This coordinate must equal to 0 whenever $\psi(v)$ is an endpoint of an edge of $H_i^{(0)}$, and equal to 1 whenever $\psi(v)$ is an endpoint of an edge of $H_i^{(1)}$. This is impossible to satisfy simultaneously since H is connected. □

Next, we show that the G constructed above has no ϵ-approximate homomorphism to an H-homomorphism-free graph on a small number of vertices.

Proposition 2.3. *Suppose $H = \mathrm{core}(H)$ and $|E(H)| > 1$. Let G be an n-vertex graph where every edge is contained in a unique homomorphic copy of H. Let m be the number of homomorphic copies of H in G. Let G' be as in Construction 2.1.*

If $\epsilon \le m/(32n^2)$, then there is no ϵ-approximate homomorphism from G' to an H-homomorphism-free graph on at most $\exp(c_H m/n)$ vertices, where $c_H > 0$ is some constant that depends only on H.

We first give some intuition for the proof. Suppose $\phi\colon V(G') \to V(F)$ is an ϵ-approximate homomorphism and F is H-homomorphism-free. Consider the vertices and edges of G' corresponding to the vertices of some H_i, which is a homomorphic copy of H in G. Consider the bipartition $\mathcal{P}_i$ of $V(H_i) \times \{0,1\}^m \subseteq V(G')$ into two parts separated by the value of the i-th binary coordinate. If ϕ is nearly orthogonal to $\mathcal{P}_i$ on $V(H_i) \times \{0,1\}^m$ (in the sense that the two associated random

variables are nearly independent, as quantified by their mutual information), then the behavior of ϕ on $V(H_i) \times \{0,1\}^m$ would be similar to if the construction giving G' had instead used a full 2^m-blowup of H_i (without taking a subgraph, but with edge weights $1/4$ for normalization). It would then follow that many edges of G' inside $V(H_i) \times \{0,1\}^m$ cannot map to F, since F is H-homomorphism-free.

So ϕ cannot be nearly orthogonal to too many different $\mathcal{P}_i$'s. We then show that this would force its image $V(F)$ to be large. To illustrate this argument in an extreme scenario, consider a typical vertex of G that lies in cm/n homomorphic copies of H, each of which corresponds to some bipartition $\mathcal{P}_i$. If ϕ were to refine cm/n such $\mathcal{P}_i$'s, then the image of ϕ has size at least $2^{cm/n}$. We use entropy to give an approximate version of this argument.

Given joint discrete random variables X and Y, let $H(X)$ denote the (natural base) entropy of X, $H(X|Y) = H(X,Y) - H(Y)$ the conditional entropy, and $I(X;Y) = H(X) - H(X|Y)$ their mutual information.

Definition 2.4. Let P_0 and P_1 be two finite disjoint sets of equal size. We say that a non-empty subset $Q \subseteq P_0 \cup P_1$ is *η-nearly bisected by $\{P_0, P_1\}$* if the entropy of Bernoulli$(|Q \cap P_0| \,/\, |Q|)$ is at least $\log 2 - \eta^2$.

Every Bernoulli random variable W satisfies (as can be verified by direct calculation or an application of Pinsker's inequality, e.g., see [31])

$$\left| \mathbb{P}(W = 0) - \frac{1}{2} \right| \leq \sqrt{\frac{\log 2 - H(W)}{2}}.$$

Thus, every Q that is η-nearly bisected by $\{P_0, P_1\}$ satisfies

$$\left| \frac{|Q \cap P_0|}{|Q|} - \frac{1}{2} \right| \leq \frac{\eta}{\sqrt{2}}. \tag{2.1}$$

The next technical lemma says that, if $P_0 \cup P_1$ is a partition with $|P_0| = |P_1|$, and $\mathcal{Q}$ is another nearly orthogonal partition of the same ground set, then the following two random processes are roughly equivalent: (i) choosing uniform random vertex of P_0 and (ii) first choosing a nearly bisected part Q of $\mathcal{Q}$ with probability proportional to $|Q|$, and then picking a uniform element of $P_0 \cap Q$.

Lemma 2.5. *Let $P_0 \cup P_1$ and $Q_1 \cup \cdots \cup Q_k$ be two partitions of some finite set U. Suppose $|P_0| = |P_1|$.*

Let u be a uniform random element of U, and define random variables $X \in \{0,1\}$ and $Y \in [k]$ so that $u \in P_X \cap Q_Y$. Let $\eta < 1/5$. Suppose $I(X;Y) \leq \eta^3$.

Let $J_{\mathrm{nb}} = \{j \in [k] : Q_j$ is η-nearly bisected by $\{P_0, P_1\}\}$. Let $U_{\mathrm{nb}} = \bigcup_{j \in J_{\mathrm{nb}}} Q_j$. Then $|U_{\mathrm{nb}}| \geq (1-\eta)\,|U|$.

Choose a random $j \in J_{\mathrm{nb}}$ where each $j \in J_{\mathrm{nb}}$ is chosen with probability proportional to $|Q_j|$. And then choose an element of $P_0 \cap Q_j$ uniformly at random. Let μ be the distribution of this random element. Then the total variation distance between μ and the uniform distribution on P_0 is at most 8η.

Proof. We have

$$I(X;Y) = H(X) - H(X|Y) = \log 2 - H(X|Y) = \sum_{j=1}^{k} \mathbb{P}(u \in Q_j)(\log 2 - H(X|u \in Q_j))$$

Since $H(X|u \in Q_j) < \log 2 - \eta^2$ for every part Q_j which is not η-nearly bisected by $\{P_0, P_1\}$, the above inequality combined with $I(X;Y) \leq \eta^3$ implies

$$|U_{\mathrm{nb}}| \geq (1-\eta)\,|U|\,. \tag{2.2}$$

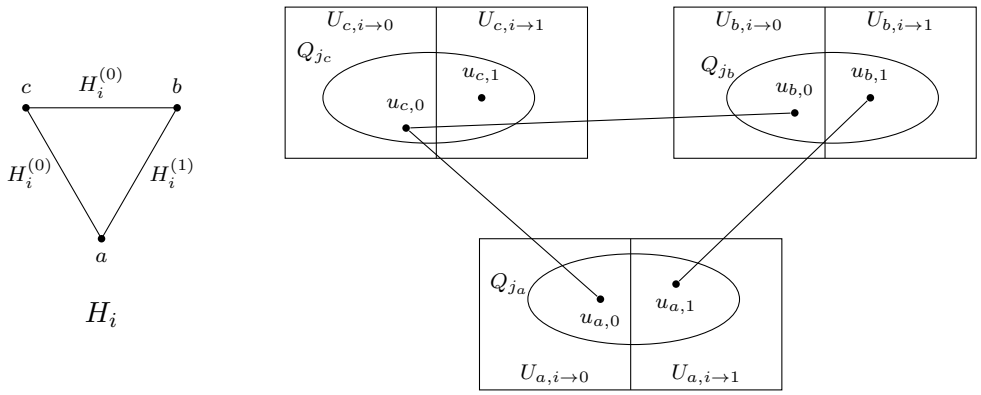

FIGURE 2. Illustration for Claim (†) in the proof of Proposition 2.3 with $H = K_3$. The vertices in Q_{j_a} all map to $j_a \in V(F)$ under ϕ, and likewise with Q_{j_b} and Q_{j_c}.

Then, for any $E \subseteq P_0$,

$$\begin{aligned}
\mu(E) &= \sum_{j \in J_{\mathrm{nb}}} \frac{|Q_j|}{|U_{\mathrm{nb}}|} \frac{|E \cap Q_j|}{|P_0 \cap Q_j|} \\
&= (2 \pm 4\eta) \sum_{j \in J_{\mathrm{nb}}} \frac{|E \cap Q_j|}{|U_{\mathrm{nb}}|} && \text{[by (2.1)]} \\
&= (2 + \eta \pm 4\eta) \sum_{j \in J_{\mathrm{nb}}} \frac{|E \cap Q_j|}{|U|}. && \text{[by (2.2)]}
\end{aligned}$$

If the final sum had been taken over all j (not just $j \in J_{\mathrm{nb}}$), then it would sum to exactly $|E| / |U|$. On the other hand, the j's not in J_{nb} contribute at most η to the sum due to (2.2). Thus this sum is at least $|E| / |U| - \eta$. Therefore, $\mu(E)$ differs from $2 |E| / |U| = |U| / |P_0|$ by at most 8η, which gives the claimed upper bound on total variance distance. □

Proof of Proposition 2.3. Let $\epsilon \leq m/(16n^2)$ and $\phi \colon G' \to F$ be an ϵ-approximate homomorphism where F is H-homomorphism-free.

For $v \in V(G)$, let U_v denote the set of vertices in G' of the form $(v, x_1, \ldots, x_m)$ for some $x_1, \ldots, x_m \in \{0, 1\}$. Let $U_{v,i\to 0} \subset U_v$ be those vertices with $x_i = 0$, and $U_{v,i\to 1} \subset U_v$ those vertices with $x_i = 1$. Then for each $i \in [m]$, there is a partition $U_v = U_{v,i\to 0} \cup U_{v,i\to 1}$.

For $i \in [m]$ and $v \in V(G)$, write

$$I_{i,v} := I(X; Y)$$

where X is the i-th binary coordinate of a uniform random vertex $u \in U_v$ and $Y \in V(F)$ is the image of the same u under ϕ.

Let $\eta = 1/(32\,|E(H)|)$.

(†) Claim: For a fixed i, if $I_{i,v} \leq \eta^3$ for all $v \in V(H_i)$, then at least 2^{2m-3} edges of G' in $\bigcup_{ab \in E(H_i)} U_a \times U_b$ do not map to an edge of F under ϕ.

The reader may find Figure 2 helpful when following the proof of this claim. The idea is that for each $a \in V(H_i)$ we are going to select a pair of vertices $(u_{a,0}, u_{a,1}) \in U_{a,i\to 0} \times U_{a,i\to 1}$ that agree on ϕ. Then for each $ab \in E(H_i)$, one of $u_{a,0}u_{b,0}$ and $u_{a,1}u_{b,1}$ must be an edge of G (which one depends on whether $ab \in E(H_i^{(0)})$ or $ab \in E(H_i^{(1)})$). If all these edges map to edges of F under ϕ, then we would obtain a homomorphic copy of H in F, which is impossible. So one of these edges does not get mapped to an edge of F, which then implies the claim by an averaging argument. The

averaging argument uses that each $u_{a,0}$ (and $u_{a,1}$) is nearly uniformly distributed on its domain by Lemma 2.5.

Now we proceed with the actual proof. Independently for each $a \in V(H_i)$, consider the following process for choosing a pair of vertices $u_{a,0}, u_{a,1} \in U_a$. Recall the partition of U_a into $U_{a,i\to 0} \cup U_{a,i\to 1}$ according to the value of the coordinate x_i. Also partition U_a into Q_j's according to fibers of ϕ, i.e., set $Q_j = \phi^{-1}(j) \cap U_a$ for each $j \in V(F)$. As in Lemma 2.5, we choose a random part Q_{j_a} that is η-nearly bisected by $\{U_{a,i\to 0}, U_{a,i\to 1}\}$, where each Q_{j_a} is chosen with probability proportional to $|Q_{j_a}|$. We choose a random vertex $u_{a,0} \in U_{a,i\to 0} \cap Q_{j_a}$ uniformly at random. Independently, we choose another random vertex $u_{a,1} \in U_{a,i\to 1} \cap Q_{j_a}$ uniformly at random.

For each $s \in \{0,1\}$ and each $ab \in E(H_i^{(s)})$, consider the edge $u_{a,s}u_{b,s}$ of G' formed by the random vertices chosen earlier (both $u_{a,s}$ and $u_{b,s}$ have their i-th binary coordinate equal to s, so $u_{a,s}u_{b,s}$ is indeed an edge of G' by Construction 2.1). At least one of these $|E(H)|$ edges of G' cannot be mapped to F under ϕ, or else they would give a homomorphic copy of H in F. It follows that

$$\sum_{s\in\{0,1\}} \sum_{ab\in E(H_i^{(s)})} \mathbb{P}\left(\phi(u_{a,s})\phi(u_{b,s}) \notin E(F)\right) \geq 1.$$

Now choose $u'_{a,0} \in U_{a,i\to 0}$ and $u'_{a,1} \in U_{a,i\to 1}$ independently and uniformly at random for each $a \in V(H_i)$. By Lemma 2.5, the total variation distance between these random variables satisfies (using the triangle inequality and independence of random variables)

$$\begin{aligned} d_{\mathrm{TV}}(u_{a,0}u_{b,0}, u'_{a,0}u'_{b,0}) &\leq d_{\mathrm{TV}}(u_{a,0}u_{b,0}, u'_{a,0}u_{b,0}) + d_{\mathrm{TV}}(u'_{a,0}u_{b,0}, u'_{a,0}u'_{b,0}) \\ &= d_{\mathrm{TV}}(u_{a,0}, u'_{a,0}) + d_{\mathrm{TV}}(u_{b,0}, u'_{b,0}) \leq 16\eta. \end{aligned}$$

Thus, combining the above two displayed inequalities,

$$\begin{aligned} \sum_{s\in\{0,1\}} \sum_{ab\in E(H_i^{(s)})} \mathbb{P}\left(\phi(u'_{a,s})\phi(u'_{b,s}) \notin E(F)\right) &\geq \sum_{s\in\{0,1\}} \sum_{ab\in E(H_i^{(s)})} \left(\mathbb{P}\left(\phi(u_{a,s})\phi(u_{b,s}) \notin E(F)\right) - 16\eta\right) \\ &\geq 1 - 16\left|E(H)\right|\eta \geq \frac{1}{2}. \end{aligned}$$

The left-hand side, multiplied by 2^{2m-2}, equals the number of edges in $\bigcup_{uv\in E(H_i)} U_u \times U_v$ that do not map to F under ϕ. This implies the Claim (†).

For a fixed $v \in V(G)$, choose $X_1, \ldots, X_m \in \{0,1\}$ independently and uniformly at random. Let Y be the image under ϕ of the vertex $(v, X_1, \ldots, X_m)$. We have

$$\begin{aligned} \sum_{i=1}^m I_{i,v} = \sum_{i=1}^m I(X_i;Y) &= \sum_{i=1}^m (H(X_i) - H(X_i|Y)) \\ &= m\log 2 - \sum_{i=1}^m H(X_i|Y) \\ &\leq m\log 2 - H(X_1,\ldots,X_m|Y) \\ &= H(Y) \leq \log\left|V(F)\right|. \end{aligned}$$

Summing over $v \in V(G)$ we obtain

$$\sum_{v\in V(G)} \sum_{i=1}^m I_{i,v} \leq n\log\left|V(F)\right|.$$

Since ϕ is an ϵ-approximate homomorphism, at most $\epsilon n^2 2^{2m}$ edges of G' do not map to an edge of F. Thus the hypothesis of Claim (†) is satisfied for at most $8\epsilon n^2$ different $i \in [m]$. For all other

i, one has $I_{i,v} > \eta^3$ for some $v \in V(H_i)$, and thus $\sum_{v\in V(G)} I_{i,v} \geq \eta^3$. Summing over all i, we obtain

$$\sum_{v\in V(G)} \sum_{i=1}^{m} I_{i,v} \geq (m - 8\epsilon n^2)\eta = \frac{m - 8\epsilon n^2}{(32\,|E(H)|)^3} \geq \frac{m}{2(32\,|E(H)|)^3}.$$

Comparing the above two displayed inequalities, we obtain $\log |V(F)| \geq c_H m/n$, as claimed. □

Proof of the lower bound in Theorem 1.15. Let H be connected and non-bipartite and $0 < \epsilon < 1$. We would like to show that if $e^{\epsilon n/C} \geq M_H(\epsilon/C)$, where $C = C_H > 0$ is a sufficiently large constant, then any n-vertex graph G where every edge lies in a unique homomorphic copy of core(H) has at most ϵn^2 edges.

Since being H-homomorphism-free is equivalent to being core(H)-homomorphism free, we can replace H by its core, and assume from now on that $H = \text{core}(H)$, which has more than one edge since H was originally not bipartite. Suppose for contradiction that the number of homomorphic copies of H in G is $m > \epsilon n^2/|E(H)|$. Obtain G' using Construction 2.1. Then by Lemma 2.2, G' is H-homomorphism-free. Hence by Theorem 1.13(b), there exists an ϵ/C-approximate homomorphism from G' to an H-homomorphism-free graph on at most $M_H(\epsilon/C) \leq e^{\epsilon n/C}$ vertices. On the other hand, by Proposition 2.3, making sure that C is large enough so that $\epsilon/C \leq m/(32n^2)$, there is no ϵ-approximate homomorphism from G' to an H-homomorphism-free graph on fewer than $e^{c_H m/n}$ vertices, which contradicts the previous sentence if C is large enough. □

3. Diamond-free versus triangle removal: graphs

In this section we prove Theorem 1.12, following the techniques in [9]. Assuming that $N_{DFL}(\epsilon)$ grows subexponentially in ϵ^{-1}, it shows that $N_{DFL}(\epsilon)$ and $\delta_{TRL}(\epsilon)$ have similar growth.

Let $g : (0, 1] \longrightarrow \mathbb{R}^+$ satisfy that $g(\beta)$ increases as β decreases, $g(\beta)\beta$ decreases as β decreases, and $\sum_{i=1}^{\infty} 1/g(2^{-i}) < 1/2$. For example, we may take $g(x) = 100\log(100/x)(\log\log(100/x))^2$.

Lemma 3.1. *Suppose G is a graph on n vertices with δn^3 triangles and at least ϵn^2 edges need to be deleted to make G triangle-free. Then G has a subgraph with αn^3 triangles for some $0 < \alpha \leq \delta$ and no edge is in more than $g(\alpha/\delta)\alpha n/\epsilon$ triangles.*

Proof. We repeatedly delete edges from G one at a time in the most triangles until we arrive at the desired subgraph. Suppose that after removing a certain number of edges, the current remaining subgraph G' has βn^3 triangles with $\beta \leq \delta$. If no edge is in more than $g(\beta/\delta)\beta n/\epsilon$ triangles in G', then we will see that G' is the desired subgraph as less than ϵn^2 edges are deleted in total so we have $\beta > 0$. Otherwise, we delete the edge in G' in the most triangles.

To go from βn^3 triangles to at most $\beta n^3/2$ triangles, we remove at least $g(\beta/(2\delta))(\beta/2)n/\epsilon$ triangles for each edge deleted, so in total we delete at most

$$\frac{\beta n^3}{g(\frac{\beta}{2\delta})\frac{\beta n}{2\epsilon}} = \frac{2\epsilon n^2}{g(\frac{\beta}{2\delta})}$$

edges in halving the total number of triangles from βn^3 to at most $\beta n^3/2$. In total, we delete at most $\sum_{i=1}^{\infty} 2\epsilon n^2/g(1/2^i) < \epsilon n^2$ edges in this process. As the original graph G we assumed required at least ϵn^2 edges to be deleted to make triangle-free, the remaining subgraph when the process terminates still has at least one triangle and satisfies the desired properties. □

Lemma 3.2. *Suppose G is a graph on n vertices with αn^3 triangles and each edge is in at most $t \leq n/100$ triangles. There is a subgraph of G with $N = n/(9t)$ vertices and more than αN^3 edges in which every edge is in exactly one triangle.*

Proof. Pick a random subset S of $N = n/(9t)$ vertices. Call a triangle T of G *good* if it is a subset of S but no edge of T is in another triangle in S. The probability T is a subset of S is $\binom{n/(9t)}{3}/\binom{n}{3} \geq 1/(1000t^3)$. For each triangle T, there are at most $3t - 3$ other vertices that together

with an edge of T make a triangle in G. Conditioned on T being a subset of S, the probability that another particular vertex is in S is at most $\frac{n/(9t)-3}{n-3} \le 1/(9t)$. Thus, conditioning on T is in S, the probability that T is good is at least $1-(3t-3)/(9t) > 2/3$. Hence, the expected number of good triangles in S is at least $\frac{1}{1000t^3} \cdot \frac{2}{3} \cdot \alpha n^3 = 2\alpha(n/t)^3/3000$. The edges in the good triangles form a subgraph of G with $N = n/(9t)$ vertices in which each edge is in exactly one triangle and there are at least $\alpha(n/t)^3/500 > \alpha N^3$ edges. □

Now we prove Theorem 1.12, which, as a reminder, says that for fixed $0 < c < 1$, if $N_{DFL}(\epsilon) \le 2^{\epsilon^{-c+o(1)}}$ as $\epsilon \to 0$, then $\delta_{TRL}(\epsilon) \ge 2^{-\epsilon^{-c/(1-c)+o(1)}}$ as $\epsilon \to 0$.

Proof of Theorem 1.12. Let $g(x) = 100\log(100/x)(\log\log(100/x))^2$. Let G be a graph on n vertices with δn^3 triangles such that at least ϵn^2 edges need to be removed to make G triangle-free. By Lemma 3.1, G has a subgraph G' with αn^3 triangles for some $0 < \alpha \le \delta$ and no edge is in more than $t := g(\alpha/\delta)\alpha n/\epsilon$ triangles. Let $g = g(\alpha/\delta)$. So $\alpha/\delta = 2^{g^{1-o(1)}}$ as $g \to \infty$. Also let $\epsilon_0 = \epsilon/(9g)$.

Applying Lemma 3.2 to G', there is a subgraph G'' of G' on $N = n/(9t) = \epsilon_0/\alpha$ vertices with more than $\alpha N^3 = \epsilon_0 N^2$ edges and each edge is in exactly one triangle.

The graph G'' shows that $N_{DFL}(\epsilon_0) \ge N = \epsilon_0/\alpha$. On the other hand, by assumption, $N_{DFL}(\epsilon_0) \le 2^{\epsilon_0^{-c+o(1)}}$ as $g \to \infty$. These two bounds on $N_{DFL}(\epsilon_0)$ together imply

$$\epsilon_0^{-1} 2^{\epsilon_0^{-c+o(1)}} \ge \alpha^{-1} = (\delta/\alpha)\delta^{-1} = 2^{g^{1-o(1)}}\delta^{-1}.$$

This bound gives

$$\delta^{-1} \le 2^{\epsilon_0^{-c+o(1)} - g^{1-o(1)}} = 2^{(9g/\epsilon)^{c-o(1)} - g^{1-o(1)}} \le 2^{-\epsilon^{-c/(1-c)+o(1)}}.$$

The middle term is maximized when $g = \epsilon^{-c/(1-c)+o_{\epsilon\to 0}(1)}$ and gives the last inequality. □

4. Diamond-free versus triangle-free in $\mathbb{F}_p^n$

In this section, we prove the lower bound in Theorem 1.25 showing that, in $\mathbb{F}_p^n$, the triangle-free lemma (Theorem 1.21) implies the diamond-free lemma (Theorem 1.17) with good quantitative bounds. The idea is to construct a blow-up similar to that done in Section 2 for graphs, though the proof is cleaner here since partitions into cosets are much more rigid than arbitrary partitions.

Construction 4.1. Suppose $x_1, \dots, x_l, y_1, \dots, y_l, z_1, \dots, z_l \in \mathbb{F}_p^n$ satisfy $x_i + y_j + z_k = 0$ if and only if $i = j = k$.

Let X_i' denote the set of all elements of $\mathbb{F}_p^{n+l}$ whose first n coordinates form x_i, and whose $(n+i)$-th coordinate lies in $\{0, \dots, \lfloor (p-2)/3 \rfloor\}$. Let $X' = \bigcup_{i=1}^l X_i'$.

Let Y_i' denote the set of all elements of $\mathbb{F}_p^{n+l}$ whose first n coordinates form y_i, and whose $(n+i)$-th coordinate lies in $\{0, \dots, \lfloor (p-2)/3 \rfloor\}$. Let $Y' = \bigcup_{i=1}^l Y_i'$.

Let Z_i' denote the set of all elements of $\mathbb{F}_p^{n+l}$ whose first n coordinates form z_i, and whose $(n+i)$-th coordinate lies in $\{1, \dots, \lfloor (p-2)/3 \rfloor + 1\}$. Let $Z' = \bigcup_{i=1}^l Z_i'$.

Note that $X' \times Y' \times Z'$ above is triangle-free. Indeed, the first n coordinates of any such triangle must form (x_i, y_i, z_i) for some i, but then the $(n+i)$-th coordinate cannot sum to zero.

Proposition 4.2. *Let $x_1, \dots, x_l, y_1, \dots, y_l, z_1, \dots, z_l \in \mathbb{F}_p^n$ and $X', Y', Z' \subset \mathbb{F}_p^{n+l}$ be as in Construction 4.1.*

Let $\phi\colon \mathbb{F}_p^{n+l} \to \mathbb{F}_p^m$ be any linear map. Let $X'', Y'', Z'' \subset \mathbb{F}_p^m$. Suppose $X'' \times Y'' \times Z''$ is triangle-free. Then

$$\left|X' \setminus \phi^{-1}(X'')\right| + \left|Y' \setminus \phi^{-1}(Y'')\right| + \left|Z' \setminus \phi^{-1}(Z'')\right| \ge (l-m)p^l/4.$$

Proof. Since the rank of ϕ is at most m, there is some $w = (w_1, \dots, w_{n+l}) \in \mathbb{F}_p^{n+l}$ with $\phi(w) = 0$ such that the first n coordinates of w are all zero and w has at least $l - m$ nonzero coordinates. Say that $i \in [m]$ is "good" if $w_{n+i} \neq 0$.

Fix a good i. Writing X_i', Y_i', Z_i' as in Construction 4.1, we claim that

$$\left|X_i' \setminus \phi^{-1}(X'')\right| + \left|Y_i' \setminus \phi^{-1}(Y'')\right| + \left|Z_i' \setminus \phi^{-1}(Z'')\right| \geq p^{l-1}(\lfloor (p-2)/3 \rfloor + 1) \geq p^l/4. \tag{4.1}$$

Summing over all good i yields the claim.

Let us prove (4.1). Say that $x' \in \mathbb{F}_p^{n+l}$ lies *above* $x \in \mathbb{F}_p^n$ if their first n coordinates agree. Choose $x', y', z' \in \mathbb{F}_p^{n+l}$ uniformly at random among triples with $x' + y' + z' = 0$ such that x', y', z' lie above x_i, y_i, z_i respectively and furthermore the $(n+i)$-th coordinates of x', y', z' are all zero.

Let a, b, c be independent uniform random elements from $\{0, \dots, \lfloor (p-2)/3 \rfloor\}$. Multiplying w by a scalar, we may assume that its $(n+i)$-th coordinate equals to 1. Let

$$x = x' + aw \in X_i', \quad y = y' + bw \in Y_i', \quad \text{and } z = z' + (c+1)w \in Z_i'.$$

Note that x is uniformly distributed in X_i', and likewise with y in Y_i' and z in Z_i'.

Since $x' + y' + z' = 0$ and $\phi(w) = 0$, we have $\phi(x) + \phi(y) + \phi(z) = 0$. Due to the hypothesis on X'', Y'', Z'', we cannot simultaneously have $\phi(x) \in X''$, $\phi(y) \in Y''$, $\phi(z) \in Z''$. Therefore,

$$\mathbb{P}(\phi(x) \notin X'') + \mathbb{P}(\phi(y) \notin Y'') + \mathbb{P}(\phi(z) \notin Z'') \geq 1.$$

Multiplying both sides by $p^{l-1}(\lfloor (p-2)/3 \rfloor + 1)$ establishes (4.1). □

Proof of the lower bound in Theorem 1.25. Suppose $x_1, \dots, x_l, y_1, \dots, y_l, z_1, \dots, z_l \in \mathbb{F}_p^n$ satisfy $x_i + y_j + z_k = 0$ if and only if $i = j = k$. Let $m = m_p(\epsilon/5)$. It suffices to show that if $p^n \geq 5m/\epsilon$ then $l \leq \epsilon p^n$. Indeed, this would imply $N_p(\epsilon)/p \leq 5m/\epsilon$ since $N_p(\epsilon)$ is defined to be the smallest possible p^n (with n being a positive integer, which is why we have $N_p(\epsilon)/p$ on the left-hand side) so that we can guarantee the conclusion $l \leq \epsilon p^n$.

Apply Construction 4.1 to obtain sets $X', Y', Z' \subseteq \mathbb{F}_p^{n+l}$. Since $X' \times Y' \times Z'$ is triangle-free, by Theorem 1.21 there exist $X'', Y'', Z'' \subseteq \mathbb{F}_p^m$ with $X'' \times Y'' \times Z''$ triangle-free and a linear map $\phi\colon \mathbb{F}_p^{n+l} \to \mathbb{F}_p^m$ such that at most $(\epsilon/5)p^{n+l}$ elements from each of X', Y', Z' do not get mapped to X'', Y'', Z'' respectively. On the other hand, Proposition 4.2 tells us that at least $(l-m)p^l/4$ elements in total from X', Y', Z' combined do not get mapped to X'', Y'', Z'' respectively. So $(l-m)p^l/4 \leq (\epsilon/5)p^{n+l}$, and hence $l \leq (4\epsilon/5)p^n + m \leq \epsilon p^n$. □

5. Triangle-free versus triangle removal

5.1. **Sketch of the argument for graphs.** Here we sketch the proof of upper bound $M_H(\epsilon) \leq e^{C\delta_H(\epsilon/C)^{-2}}$ in Theorem 1.15, which was proved in [18, Section 3.3]. In the next subsection, we give the details of the analogous argument in the arithmetic setting.

First one shows that the graph removal lemma Theorem 1.13(a) can be extended to allow edge-weights on the n-vertex graph with edge-weights in $[0, 1]$. When counting H in a weighted graph, we weigh each homomorphic copy of H by the product of the edge-weights.

Theorem 5.1 (Weighted graph removal lemma)**.** *For every H and $\epsilon > 0$, there exists $\delta > 0$ such that for every n-vertex edge-weighted graph G with edge-weights in $[0, 1]$, if the weighted number of homomorphisms from H to G is less than $\delta n^{|V(H)|}$, then G can be made H-homomorphism-free by removing edges with total weight at most ϵn^2.*

In [18], the weighted version of the removal lemma was derived from the unweighted version as follows. Starting with a weighted graph G, consider the unweighted graph G' consisting of all edges whose edge-weight is at least $\epsilon/2$. If G has H-homomorphism-density at most $\delta_H(\epsilon/2)(\epsilon/2)^{|E(H)|}$, then G' has H-homomorphism-density at most $\delta_H(\epsilon/2)$, so by the removal lemma, G' can be made H-homomorphism-free by removing at most $\epsilon n^2/2$ edges. Now we remove the same edges from G,

along with all edges with individual weight less than $\epsilon/2$, and then the resulting weighted graph is H-homomorphism-free.

The above argument shows that in Theorem 5.1, one can take $\delta = \delta_H(\epsilon/2)(\epsilon/2)^{|E(H)|}$. This is good for most purposes, though we sketch a different argument showing that one can take $\delta = \delta_H(\frac{\epsilon}{|E(H)|+1})$ in Theorem 5.1 (the latter bound is superior when $\delta_H(\epsilon) = \epsilon^{\Theta(1)}$, which is the case if and only if H is bipartite [2], but also in the arithmetic analogue below). See Theorem 5.4 below for the details of a completely analogous argument in the arithmetic setting.

Let G be a weighted n-vertex graph with H-homomorphism density less than $\delta = \delta_H(\frac{\epsilon}{|E(H)|+1})$. Randomly blow G up to an mn-vertex graph G'. This means replacing every edge $xy \in E(G)$ with edge-weight $w(x,y)$ by a random bipartite graph with m vertices in each part and random edges appearing independently with probability $w(x,y)$. We view G as fixed and consider $m \to \infty$. Then with probability $1-o(1)$, the H-homomorphism density in G is less than δ. So by the graph removal lemma (Theorem 1.13(a)), one can delete at most $\epsilon m^2 n^2/(|E(H)|+1)$ edges from G' to make it H-homomorphism-free. For each edge xy of G, delete it from G if more than $w(x,y)m^2/(|E(H)|+1)$ edges sitting above it were deleted from G'. This then deletes edges from G with total weight at most ϵn^2. Furthermore, with probability $1-o(1)$, no homomorphic copy of H remains. Indeed, suppose some homomorphic copy H_0 of H were to remain. Consider a random copy of H_0 in G' above H_0. A linearity of expectations argument (here we use that with high probability all edges of G' between the same pair of parts lie in roughly the same number of copies of H_0, as can be verified by a Chernoff bound argument) shows that with positive probability one of these copies of H_0 does not contain any deleted edges, which violates that we had deleted edges from G' and made it H-homomorphism-free.

Now we sketch the argument in [18] that derives Theorem 1.13(b) from Theorem 1.13(a) yielding the bound $M_H(\epsilon) \le e^{C\delta_H(\epsilon/C)^{-2}}$ in Theorem 1.15. Starting with an H-homomorphism-free graph G, we can apply the Frieze–Kannan weak regularity lemma to obtain a δ/C-weak-regular partition $\mathcal{P}$ of G with $M = e^{O(\delta^{-2})}$ parts. Let $G/\mathcal{P}$ be the corresponding reduced weighted graph whose vertices are the parts of the partition and weights being the edge densities between the corresponding pairs of parts. By the counting lemma, the H-homomorphism-densities in G and $G/\mathcal{P}$ differ by $O_H(\delta/C)$. We can choose the constant C so that $G/\mathcal{P}$ has H-homomorphism density at most δ. Then Theorem 1.13(a) allows us to make the reduced graph H-homomorphism-free by removing weighted edges in $G/\mathcal{P}$ corresponding to at most ϵn^2 edges in G. Then the map from $V(G)$ to $\mathcal{P}$ gives an ϵ-approximate homomorphism from G to an H-homomorphism-free graph with M parts.

5.2. **Arithmetic analogue.** Now we provide the arithmetic analogue of the argument sketched above, thereby showing the upper bound $M_p(\epsilon) \le p^{27\delta_p(\epsilon/4)^{-2}}$ in Theorem 1.25.

Given a function $f\colon \mathbb{F}_p^n \to \mathbb{R}$, and a subspace $H \le \mathbb{F}_p^n$, we write $f_H\colon \mathbb{F}_p^n \to \mathbb{R}$ for the function that is constant on every H-coset, so that on $x+H$ the value of f_H equals to the average of f on $x+H$. We say that H is *ϵ-weakly-regular* for f the L^∞ norm of the Fourier transform of $f - f_H$ is at most ϵ. We say that H is ϵ-weakly regular for a set $X \subseteq \mathbb{F}_p^n$ if it is so for its indicator function $f = 1_X$. Here the normalization of the Fourier transform is given by $\widehat{f}(y) = \mathbb{E}_x f(x) e^{-2\pi i(x\cdot y)/p}$. Also we write $\|f\|_1 = \mathbb{E}_x |f(x)|$.

We recall the weak regularity lemma and the associated counting lemma, both of which are standard (e.g., see [11, Section 2]). These are versions of Green's arithmetic regularity results [15] analogous to the Frieze–Kannan weak regularity lemma [12].

Lemma 5.2 (Weak arithmetic regularity lemma)**.** *Let p be a prime and $\epsilon > 0$. For every $X, Y, Z \subseteq \mathbb{F}_p^n$, there exists a subspace H of $\mathbb{F}_p^n$ of codimension at most $3\epsilon^{-2}$ that is ϵ-weakly-regular for each of X, Y, Z.*

A quick proof sketch: take H to be the subspace orthogonal to all non-trivial characters with Fourier transform magnitude at least ϵ for any of X, Y, Z. There are at most ϵ^{-2} such characters for X by Parseval, and likewise with Y and Z.

For $f, g, h\colon \mathbb{F}_p^n \to [0,1]$, let us denote their triangle density by

$$\Lambda(f,g,h) := \mathbb{E}_{x,y,z\in\mathbb{F}_p^n : x+y+z=0} f(x)g(y)h(z).$$

The following counting lemma is also standard (e.g., see [11, Lemma 4] for a proof).

Lemma 5.3 (Counting lemma)**.** *Let p be a prime and $\epsilon > 0$. For every $f, g, h\colon \mathbb{F}_p^n \to [0,1]$ and a subspace H of $\mathbb{F}_p^n$ that is ϵ-regular with respect to each of f, g, h, then*

$$|\Lambda(f,g,h) - \Lambda(f_H, g_H, h_H)| \le 3\epsilon.$$

We need the following weighted version of the arithmetic triangle removal lemma. The proof follows the second argument sketched in the previous subsection (the first argument sketched there, by considering edges with weight at least $\epsilon/2$, would be too lossy).

Theorem 5.4 (Weighted arithmetic triangle removal lemma)**.** *If $f, g, h\colon \mathbb{F}_p^n \to [0,1]$ are such that $\Lambda(f,g,h) < \delta_p(\epsilon/4)$, then there exist $f', g, h'\colon \mathbb{F}_p^n \to [0,1]$ such that $\Lambda(f', g', h') = 0$ and $\|f - f'\|_1$, $\|g - g'\|_1$, $\|h - h'\|_1 \le \epsilon$.*

Proof. In this proof, we fix $f, g, h\colon \mathbb{F}_p^n \to [0,1]$ with $\Lambda(f,g,h) < \delta_p(\epsilon/4)$. All the asymptotics are with respect to a new parameter $m \to \infty$.

We say that $y \in \mathbb{F}_p^{n+m}$ is *above* $x \in \mathbb{F}_p^n$ if the first n coordinates of y form x.

Let X be a random subset of $\mathbb{F}_p^{n+m}$ obtained by independently keeping each element above every $x \in \mathbb{F}_p^n$ with probability $f(x)$. Likewise define $Y, Z \subseteq \mathbb{F}_p^{n+m}$ from g, h respectively.

With high probability (meaning probability $1 - o(1)$ as $m \to \infty$), $X \times Y \times Z$ has at most $\delta_p(\epsilon/4)p^{2(n+m)}$ triangles, so by the arithmetic triangle removal lemma (Theorem 1.16), we can remove all triangles by deleting at most $\epsilon p^{n+m}/4$ elements from each of X, Y, Z.

For each $x \in \mathbb{F}_p^n$, we set $f'(x) = 0$ if we deleted least $f(x)p^m/4$ elements of X above x, and set $f'(x) = f(x)$ otherwise. Then the number of elements deleted from X is at least $\sum_{x\in\mathbb{F}_p^n}(f - f')(x)p^m/4$. Thus $\|f - f'\|_1 \le \epsilon$. Similarly define g' and h'.

Finally, we claim with high probability, $\Lambda(f', g', h') = 0$. Suppose otherwise. Fix some $x+y+z = 0$ in $\mathbb{F}_p^n$ with $f'(x), g'(y), h'(z) > 0$. Among all triples $(x', y', z') \in X \times Y \times Z$ sitting above (x, y, z) and satisfying $x' + y' + z' = 0$, choose a triple uniformly at random. One of x', y', z' must be deleted to make X, Y, Z triangle-free, so

$$\mathbb{P}(x' \text{ is deleted}) + \mathbb{P}(y' \text{ is deleted}) + \mathbb{P}(z' \text{ is deleted}) \ge 1.$$

On the other hand, the total variation distance between x' and a uniform random element of X above x is $o(1)$ with high probability (e.g., a second moment argument shows that almost all x' lies in nearly the same number of such triples (x', y', z')). So if r_x is the fraction of elements above x that are deleted, then $r_x + r_y + r_z \ge 1 - o(1)$ with high probability, thereby contradicting $r_x, r_y, r_z \le 1/4$. □

Proof of the upper bound in Theorem 1.25. We want to show that the arithmetic triangle-free lemma (Theorem 1.21) holds with some $m \le 27\delta^{-2}$, where $\delta := \delta_p(\epsilon/4)$.

Let $X, Y, Z \le \mathbb{F}_p^n$ be such that $X \times Y \times Z$ is triangle-free. Applying the arithmetic weak regularity lemma (Lemma 5.2), we find a subspace H of $\mathbb{F}_p^n$ with codimension $m \le 27\delta^{-2}$ that is $\delta/3$-weakly-regular to each of X, Y, Z.

Let $\phi\colon \mathbb{F}_p^n \to \mathbb{F}_p^m$ be a linear map with kernel H. Define $f, g, h\colon \mathbb{F}_p^m \to [0,1]$ by setting, for each $x \in \mathbb{F}_p^m$, $f(x) = \left|\phi^{-1}(x) \cap X\right| / p^{n-m}$. In other words, $f(x)$ is the fraction of the coset $H + \phi^{-1}(x)$ that belongs to X. Likewise define g and h based on Y and Z.

Applying the counting lemma (Lemma 5.3),

$$\Lambda(f,g,h) = \Lambda((1_X)_H, (1_Y)_H, (1_Y)_H) \leq \Lambda(1_X, 1_Y, 1_Y) + \delta = \delta.$$

By the weighted arithmetic triangle removal lemma, Theorem 5.4, there are $f', g', h' \colon \mathbb{F}_p^m \to [0,1]$ so that $\Lambda(f', g', h') = 0$ and $\|f - f'\|_1$, $\|g - g'\|_1$, $\|h - h'\|_1 \leq \epsilon$. The conclusion of Theorem 1.21 follows then by taking X', Y', Z' to be the respective supports of f', g', h'. □

Acknowledgments. We thank Yuval Wigderson and the anonymous reviewer for careful readings and comments on the manuscript.

References

[1] M. Ajtai and E. Szemerédi, *Sets of lattice points that form no squares*, Studia Sci. Math. Hungar. **9** (1974), 9–11.
[2] Noga Alon, *Testing subgraphs in large graphs*, Random Structures Algorithms **21** (2002), 359–370.
[3] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331–332.
[4] Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans, *On cap sets and the group-theoretic approach to matrix multiplication*, Discrete Anal. (2017), Paper No. 3, 27.
[5] David Conlon, Jacob Fox, and Yufei Zhao, *A relative Szemerédi theorem*, Geom. Funct. Anal. **25** (2015), 733–762.
[6] Ernie Croot, Vsevolod F. Lev, and Péter Pál Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*, Ann. of Math. (2) **185** (2017), 331–337.
[7] Jordan S. Ellenberg and Dion Gijswijt, *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression*, Ann. of Math. (2) **185** (2017), 339–343.
[8] Jacob Fox, *A new proof of the graph removal lemma*, Ann. of Math. (2) **174** (2011), 561–579.
[9] Jacob Fox and László Miklós Lovász, *A tight bound for Green's arithmetic triangle removal lemma in vector spaces*, Adv. Math. **321** (2017), 287–297.
[10] Jacob Fox, László Miklós Lovász, and Lisa Sauermann, *A polynomial bound for the arithmetic k-cycle removal lemma in vector spaces*, J. Combin. Theory Ser. A **160** (2018), 186–201.
[11] Jacob Fox and Huy Tuan Pham, *Popular Progression Differences in Vector Spaces*, Int. Math. Res. Not. IMRN (2021), 5261–5289.
[12] Alan Frieze and Ravi Kannan, *Quick approximation to matrices and applications*, Combinatorica **19** (1999), 175–220.
[13] Hu Fu and Robert Kleinberg, *Improved lower bounds for testing triangle-freeness in Boolean functions via fast matrix multiplication*, Approximation, randomization, and combinatorial optimization, LIPIcs. Leibniz Int. Proc. Inform., vol. 28, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2014, pp. 669–676.
[14] W. T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. (2) **166** (2007), 897–946.
[15] B. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), 340–376.
[16] Ben Green, *Lower bounds for corner-free sets*, New Zealand J. Math. **51** (2021), 1–2.
[17] Pavol Hell and Jaroslav Nešetřil, *The core of a graph*, vol. 109, 1992, Algebraic graph theory (Leibnitz, 1989), pp. 117–126.
[18] Carlos Hoppen, Yoshiharu Kohayakawa, Richard Lang, Hanno Lefmann, and Henrique Stagni, *Estimating parameters associated with monotone properties*, Combin. Probab. Comput. **29** (2020), 616–632.
[19] Robert Kleinberg, David E. Speyer, and Will Sawin, *The growth of tri-colored sum-free sets*, Discrete Anal. (2018), Paper No. 12, 10.
[20] Daniel Král, Oriol Serra, and Lluís Vena, *A combinatorial proof of the removal lemma for groups*, J. Combin. Theory Ser. A **116** (2009), 971–978.
[21] Daniel Kráľ, Oriol Serra, and Lluís Vena, *A removal lemma for systems of linear equations over finite fields*, Israel J. Math. **187** (2012), 193–207.
[22] Nati Linial and Adi Shraibman, *Larger corner-free sets from better NOF exactly-N protocols*, Discrete Anal. 2021, Paper No. 19, 9.
[23] László Miklós Lovász and Lisa Sauermann, *A lower bound for the k-multicolored sum-free problem in $\mathbb{Z}_m^n$*, Proc. Lond. Math. Soc. (3) **119** (2019), 55–103.
[24] Sergey Norin, *A distribution on triples with maximum entropy marginal*, Forum Math. Sigma **7** (2019), e46, 12.
[25] Luke Pebody, *Proof of a conjecture of Kleinberg-Sawin-Speyer*, Discrete Anal. (2018), Paper No. 13, 7.
[26] Vojtěch Rödl and Mathias Schacht, *Generalizations of the removal lemma*, Combinatorica **29** (2009), 467–501.

[27] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, 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.
[28] Asaf Shapira, *A proof of Green's conjecture regarding the removal properties of sets of linear equations*, J. Lond. Math. Soc. (2) **81** (2010), 355–373.
[29] József Solymosi, *Note on a generalization of Roth's theorem*, Discrete and computational geometry, Algorithms Combin., vol. 25, Springer, Berlin, 2003, pp. 825–827.
[30] Endre Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401.
[31] Terence Tao, *Entropy and rare events*, blog post `https://terrytao.wordpress.com/2015/09/20/entropy-and-rare-events/`.
[32] Terence Tao, *The Gaussian primes contain arbitrarily shaped constellations*, J. Anal. Math. **99** (2006), 109–176.
[33] Terence Tao, *A variant of the hypergraph removal lemma*, J. Combin. Theory Ser. A **113** (2006), 1257–1280.

Department of Mathematics, Stanford University, Stanford, CA 94305, USA
Email address: `jacobfox@stanford.edu`

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
Email address: `yufeiz@mit.edu`