TESTING LINEAR-INVARIANT PROPERTIES

JONATHAN TIDOR AND YUFEI ZHAO

Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract. Fix a prime p and a positive integer R. We study the property testing of functions $\mathbb{F}_p^n \to [R]$. We say that a property is testable if there exists an oblivious tester for this property with one-sided error and constant query complexity. Furthermore, a property is proximity oblivious-testable (PO-testable) if the test is also independent of the proximity parameter ϵ. It is known that a number of natural properties such as linearity and being a low degree polynomial are PO-testable. These properties are examples of linear-invariant properties, meaning that they are preserved under linear automorphisms of the domain. Following work of Kaufman and Sudan, the study of linear-invariant properties has been an important problem in arithmetic property testing.

A central conjecture in this field, proposed by Bhattacharyya, Grigorescu, and Shapira, is that a linear-invariant property is testable if and only if it is semi subspace-hereditary. We prove two results, the first resolves this conjecture and the second classifies PO-testable properties.

(1) A linear-invariant property is testable if and only if it is semi subspace-hereditary.
(2) A linear-invariant property is PO-testable if and only if it is locally characterized.

Our innovations are two-fold. We give a more powerful version of the compactness argument first introduced by Alon and Shapira. This relies on a new strong arithmetic regularity lemma in which one mixes different levels of Gowers uniformity. This allows us to extend the work of Bhattacharyya, Fischer, Hatami, Hatami, and Lovett by removing the bounded complexity restriction in their work. Our second innovation is a novel recoloring technique called patching. This Ramsey-theoretic technique is critical for working in the linear-invariant setting and allows us to remove the translation-invariant restriction present in previous work.

1. Introduction

In property testing, the aim is to find randomized algorithms that distinguish objects that have some given property from those that are far from satisfying the property by querying the given large object at a small number of locations. Property testing emerged from the linearity test of Blum, Luby, and Rubinfeld [10], and was formally defined and systematically studied by Rubinfeld and Sudan [29] and Goldreich, Goldwasser, and Ron [14]. There have been important developments especially in the following two settings: graph property testing and arithmetic property testing.

Two representative problems are: (1) given a large graph, test whether the graph is triangle-free or ϵ-far from triangle-free (an n-vertex graph is *ϵ-far* from a graph property if one needs to add and/or remove more than ϵn^2 edges in order to satisfy the property), and (2) given a function $f\colon \mathbb{F}_p^n \to \mathbb{F}_p$, test whether f is linear or ϵ-far from linear (for an arithmetic property, being *ϵ-far* means that one needs to change the value of the function on more than an ϵ-fraction of the domain in order to satisfy the property). In both cases, it is known that one can achieve the desired goal by sampling a fixed number of entries repeatedly $C(\epsilon)$ times. For testing whether a graph is triangle-free [30], one samples a uniformly random triple of vertices and checks whether they form a triangle, and for testing linearity [10], one samples $x, y \in \mathbb{F}_p^n$ uniformly and checks if $f(x)+f(y) = f(x+y)$.

E-mail address: {jtidor,yufeiz}@mit.edu.

Tidor was supported by NSF Graduate Research Fellowship Program DGE-1122374. Zhao was supported by NSF Award DMS-1764176, the MIT Solomon Buchsbaum Fund, and a Sloan Research Fellowship.

In this paper we give a property testing algorithm for a very general class of arithmetic properties. The goal is to determine whether a function $f\colon \mathbb{F}_p^n \to [R] := \{1, \dots, R\}$ (with fixed prime p and positive integer R) satisfies some given property or is ϵ-far from satisfying the property. All the properties we consider are *linear-invariant* in the sense that they are invariant under automorphisms of the vector space $\mathbb{F}_p^n$. Linear-invariant properties form an important general class of arithmetic properties, e.g., the work of Kaufman and Sudan [26] "highlights linear-invariance as a central theme in algebraic property testing."

We say that a property $\mathcal{P}$ is *testable* if there exists an oblivious tester with one-sided error (and constant query complexity) for the property. A tester for $\mathcal{P}$ produces a positive integer $d = d(\epsilon)$ and an oracle provides the tester with the restriction $f|_U$ where U is a uniform random d-dimensional linear subspace of the domain (if the domain is large enough that such a subspace exists; if the domain has dimension strictly less than d, the oracle provides the tester with all of f). We require our tester accepts functions f satisfying $\mathcal{P}$ with probability 1 and reject functions that are ϵ-far from satisfying $\mathcal{P}$ with probability at least $\delta = \delta(\epsilon)$ for some function $\delta\colon (0,1) \to (0,1)$. Furthermore, we say that $\mathcal{P}$ is *proximity oblivious-testable* (PO-testable) if $d = d(\epsilon)$ is a constant independent of ϵ. The idea of PO-testability was introduced by Goldreich and Ron [15] who, among other results, classified the PO-testable graph properties.

One surprising feature of property testing is that many natural properties, such as linearity, are testable and even PO-testable. A key feature of linearity is that it is subspace-hereditary meaning that if $f\colon \mathbb{F}_p^n \to \mathbb{F}_p$ is linear, then the same is true for $f|_U$ for every linear subspace $U \le \mathbb{F}_p^n$. To be precise, we say that a linear-invariant property $\mathcal{P}$ is *subspace-hereditary* if for every $f\colon \mathbb{F}_p^n \to [R]$ satisfying $\mathcal{P}$ and every linear subspace $U \le \mathbb{F}_p^n$, the restriction $f|_U$ also satisfies $\mathcal{P}$.

A central conjecture in this field, by Bhattacharyya, Grigorescu, and Shapira, is that all linear-invariant, subspace-hereditary properties are testable [9, Conjecture 4]. In fact, they conjecture that the slightly larger class of *semi subspace-hereditary* properties are testable and prove that no other properties can be tested.

Definition 1.1. A linear-invariant property $\mathcal{P}$ is *semi subspace-hereditary* if there exists a subspace-hereditary property $\mathcal{Q}$ such that

(i) every function satisfying $\mathcal{P}$ also satisfies $\mathcal{Q}$;
(ii) for all $\epsilon > 0$, there exists $N(\epsilon)$ such that if $f\colon \mathbb{F}_p^n \to [R]$ satisfies $\mathcal{Q}$ and is ϵ-far from satisfying $\mathcal{P}$, then $n < N(\epsilon)$.

It is known that there are subspace-hereditary properties where the dimension d sampled must grow as the proximity parameter ϵ approaches 0. To be PO-testable, a property must satisfy the following more restrictive condition.

Definition 1.2. A linear-invariant property $\mathcal{P}$ is *locally characterized* if there exists some d such that the following holds. For every $f\colon \mathbb{F}_p^n \to [R]$ with $n \ge d$, the function f satisfies $\mathcal{P}$ if and only if $f|_U$ satisfies $\mathcal{P}$ for every $U \le \mathbb{F}_p^n$ of dimension d.

Our first result is a resolution of the conjecture of Bhattacharyya, Grigorescu, and Shapira, classifying the testable linear-invariant properties. Our second result is a classification of the PO-testable linear-invariant properties.

Theorem 1.3. *A linear-invariant property is testable if and only if it is semi subspace-hereditary.*

Theorem 1.4. *A linear-invariant property is PO-testable if and only if it is locally characterized.*

Remark 1.5. Note that under our definition, the tester does not know the dimension of the domain. This rules out some "unnatural" properties such as those properties that behave differently depending on whether the dimension of the domain is even or odd.

Previous work in arithmetic property testing has focused on a number of special cases including monotone properties [27, 31], "complexity 1" properties over $\mathbb{F}_2$ [9], and bounded complexity translation-invariant properties [7].

We note that very little was previously known about general linear-invariant patterns. One simple way to define a class of linear-invariant patterns can be done by, for example, choosing an arbitrary subset of "allowable" maps $\mathbb{F}_p^2 \to [R]$ and defining a property of functions $\mathbb{F}_p^n \to [R]$ to consist of those whose restriction to every 2-dimensional linear subspace is allowable. Even this class of 2-dimensionally-defined patterns was not known to be testable in general prior to this work.

Our innovations are two-fold. We prove a strong arithmetic regularity lemma which, unlike previous arithmetic regularity lemmas, mixes different levels of Gowers uniformity. This allows us to give a more powerful version of the *compactness argument* first introduced by Alon and Shapira [4]. With this tool we can remove the *bounded complexity* restriction that was present in all previous work.

Our second innovation is a novel recoloring technique we call *patching*. This technique is critical for working in the linear-invariant setting and allows us to handle an important obstacle encountered by previous works. Roughly speaking, this obstacle is the inability of regularity methods to regularize functions in a neighborhood of the origin.

In the rest of this section we give a summary of the proof of the main theorem and its relation to previous work.

1.1. **Graph removal lemmas and property testing.** We begin with an overview of graph removal lemmas and their proof techniques (see also the survey [11]).

The triangle removal lemma of Ruzsa and Szemerédi [30] states that for all $\epsilon > 0$ there exists $\delta > 0$ such that any n-vertex graph with at most δn^3 triangles can be made triangle-free by removing ϵn^2 edges. This was generalized to the graph removal lemma, first stated explicitly by Alon, Duke, Lefmann, Rödl, and Yuster [1] and by Füredi [13].

A key tool for proving the graph removal lemma is a regularity lemma, namely Szemerédi's graph regularity lemma. Roughly speaking, the proof proceeds by using this regularity lemma to partition the input graph G into a small number of structured components. Then we "clean up" G by removing at most ϵn^2 edges. This is done in such a way that either the resulting graph is H-free or the original graph G contains many copies of H.

An important extension of the graph removal lemma is the induced graph removal lemma, proved by Alon, Fischer, Krivelevich, and Szegedy [2]. The induced graph removal lemma states that for every graph H (or finite collection $\mathcal{H}$ of graphs), for all $\epsilon > 0$ there exists $\delta > 0$ such that every n-vertex graph with at most $\delta n^{v(H)}$ induced copies of H can be made induced H-free by adding and/or removing at most ϵn^2 edges (here *induced H-free* means not containing any induced subgraph isomorphic to H).

The original proof of the induced removal lemma relies on an extension of Szemerédi's graph regularity lemma known as the "strong regularity lemma." Using such a regularity lemma combined with a random sampling argument, one can produce a "regular model", that is, a large induced subgraph $X := G[U]$ (on a constant fraction of the vertices of G) that is very regular and approximates the original graph well in a certain sense. Then we "clean up" G by adding and/or removing at most ϵn^2 edges in such a way that if the resulting graph is not induced H-free then X (in the original graph) must contain many induced copies of H.

Alon and Shapira [4] extended the induced graph removal lemma to an infinite collection of graphs. Namely they prove that for a (possibly infinite) set $\mathcal{H}$ of graphs and for $\epsilon > 0$ there exist $\delta > 0$ and k such that the following holds: if G is an n-vertex graph with at most $\delta n^{v(H)}$ copies of H for all $H \in \mathcal{H}$ with k or fewer vertices, then G can be made induced $\mathcal{H}$-free by adding and/or removing at most ϵn^2 edges (meaning the modified graph has no induced subgraph isomorphic to

H for every $H \in \mathcal{H}$). This theorem immediately implies (and is equivalent to) the fact that every hereditary graph property is testable with constant query-complexity and one-sided error.

This series of works, in addition to being important results in their own right, gives a framework for proving constant query-complexity property testing algorithms in other settings, given an appropriate regularity lemma. In particular, the hypergraph regularity lemma, proved by Gowers [17] and independently by Rödl et al. [28] can be used with the above techniques to prove an infinite induced hypergraph removal lemma [29]. Consequently, every hereditary hypergraph property is testable with constant query-complexity and one-sided error.

1.2. **Arithmetic analogs.** The problem of property testing for functions $f\colon \mathbb{F}_p^n \to [R]$ has been intensely studied, starting with the the classic work of Blum, Luby, and Rubinfeld [10] on linearity testing. Much of the work focuses on testing whether some function $f\colon \mathbb{F}_p^n \to \mathbb{F}_p$ has certain algebraic properties (e.g., a polynomial of some given type) [3, 26]. There is also much interest in testing properties that do not arrive from algebraic characterizations. Below we give an overview of the developments related to property testing in $\mathbb{F}_p^n$ from a perspective that is parallel to the graph regularity method developments discussed earlier.

The first arithmetic regularity lemma was proved by Green [19] using Fourier-analytic techniques, and it laid the groundwork for further developments of the regularity method in the arithmetic setting. These regularity lemmas has since found many applications in additive combinatorics and related fields. In particular, combined with the graph removal framework described above, Green's regularity lemma is suitable for proving an arithmetic removal lemma for "complexity 1" systems of linear forms (see Section 3 for the definition of complexity); e.g., see [6].

Král', Sera, and Vena [27] and independently Shapira [31] bypass the need for an arithmetic regularity lemma and prove the full arithmetic removal lemma by a direct reduction from the hypergraph removal lemma. Their results imply that all linear-invariant, subspace-hereditary *monotone* properties are testable with constant query-complexity and one-sided error. (A property of functions $\mathbb{F}_p^n \to \{0,1\}$ is *monotone* if changing 1's to 0's preserves the property.)

Note that the above result is an arithmetic removal lemma and not an *induced* arithmetic removal lemma (hence the restriction to monotone properties). Due to the nature of the reduction, the techniques do not seem to be capable of deducing the induced arithmetic removal lemma from the induced hypergraph removal lemma.

An alternative approach is to apply the strong graph regularity approach [2] of proving the induced graph removal lemma to Green's arithmetic regularity lemma. However there is also a major obstacle to the approach, related to the fact that the origin plays a special role in a vector space while there is no corresponding feature of graphs. It turns out that it is not always possible to regularize the space in a neighborhood of the origin [20].

Bhattacharyya, Grigorescu, and Shapira [9] managed to overcome this obstacle in the special case of vector spaces over $\mathbb{F}_2$. They follow the above strategy, implementing the strong regularity idea [2] in the style of Green's arithmetic regularity [19] along with one additional tool, namely a Ramsey-theoretic result, to prove an infinite induced arithmetic removal lemma for "complexity 1" patterns over $\mathbb{F}_2$. Unfortunately, it is known [20] that this Ramsey-theoretic result fails over all finite fields other than $\mathbb{F}_2$.

Bhattacharyya, Fischer, and Lovett [8] managed to overcome this obstacle in a different special case, namely for translation-invariant patterns. When all patterns considered are translation-invariant, the origin no longer plays a special role and one can essentially ignore it while carrying out the strong regularity framework. In addition, [8] allows one to handle higher complexity patterns, which requires developing and applying tools from higher-order Fourier analysis.

Higher-order Fourier analysis plays a central role in modern additive combinatorics. These techniques were initiated by Gowers [16] in his celebrated new proof of Szemerédi's theorem, and further developed in a sequence of works by Green, Tao, and Ziegler [21, 22, 23] settling classical

conjectures on the asymptotics of prime numbers patterns. A parallel theory of higher-order Fourier analysis was developed in finite field vector spaces by Bergelson, Tao, and Ziegler [5, 32, 33], leading to an inverse Gowers theorem over finite fields vector spaces.

For applications to property testing, this line of work culminated in the work of Bhattacharyya, Fischer, Hatami, Hatami, and Lovett [7] (extending [8]), who applied the inverse Gowers theorem over finite fields and developed further equidistribution tools to prove an infinite induced arithmetic removal lemma for all linear-invariant, subspace-hereditary properties that are also translation-invariant and bounded-complexity. Their work follows the strong regularity framework of [2, 4]. Our results improve upon this work by removing the translation-invariant and bounded-complexity restrictions.

In addition to their property testing algorithm, Bhattacharyya, Fischer, Hatami, Hatami, and Lovett [7] proved that a large class of somewhat algebraically structured properties are indeed affine-invariant, subspace-hereditary, and locally characterized. These are the so-called "degree-structural properties". A simple extension of their result [25, Theorem 16.3] implies that the larger class of "homogeneous degree-structural properties" are linear-invariant, linear subspace-hereditary (but not affine-invariant and not subspace-hereditary), and locally characterized, and thus these properties are testable by our main theorem. As an example, one can test whether a function $\mathbb{F}_p^n \to \mathbb{F}_p$ can be written as $A^2 + B^2$ where both A and B are homogeneous polynomials of some given degree d.

1.3. **Our contributions.**

1.3.1. *Patching.* In this paper, building on the authors' earlier work with Fox [12] for complexity 1 patterns, we develop a new technique called "patching" that allows us to overcome the obstacle faced by earlier approaches, namely that a neighborhood of the origin cannot be regularized and fails certain Ramsey properties (unless working over $\mathbb{F}_2$). In essence, the patching result states that if there exists some map $f\colon \mathbb{F}_p^n \to [R]$ that has low density of some colored patterns $\mathcal{H}$ for n large enough, then for all m there must exist some map $g\colon \mathbb{F}_p^m \to [R]$ that has no $\mathcal{H}$-instances.

Theorem 1.6 (Informal patching result)**.** *For every set of colored pattern $\mathcal{H}$, there exist $\epsilon_0 > 0$ and n_0 such that the following holds. Either:*

- *for every n, there exists a function $f\colon \mathbb{F}_p^n \to [R]$ that is $\mathcal{H}$-free; or*
- *for every function $f\colon \mathbb{F}_p^n \to [R]$ with $n \geq n_0$, the H-density in f is at least ϵ_0 for some $H \in \mathcal{H}$.*

Our proof proceeds in two steps. First, as in [7], following the strong regularity framework of [2] for proving induced graph removal lemmas, we apply a strong arithmetic regularity lemma, which produces a partition $\mathfrak{B}$ of $\mathbb{F}_p^n$ and a "regular model" $X \subseteq \mathbb{F}_p^n$ made up of a randomly sampled set of atoms from $\mathfrak{B}$. Unlike in the graph setting, we cannot ensure that the map $f\colon \mathbb{F}_p^n \to [R]$ is very regular on every atom of $\mathfrak{B}|_X$. In particular, it may be impossible to guarantee that f is regular on the atom containing the origin. Instead we only ensure that almost every atom of X is very regular. Unlike earlier proofs of removal lemmas, our "recoloring algorithm" has two components: for the regular atoms we "clean up" f as usual, while for the irregular atoms we apply our patching result. Our patching result implies that there is some new global coloring $g\colon \mathbb{F}_p^n \to [R]$ that avoids some appropriate set of colored patterns. To complete the proof we "patch" f by replacing it by g on all of the irregular atoms. If f has low density of some set of colored pattern, then our argument shows that these pattern cannot appear in the recoloring, thereby completing the proof of the induced arithmetic removal lemma.

Our proof does not give effective bounds on the rejection probability function $\delta(\epsilon)$ guaranteed by Theorem 1.3. The ineffectiveness is due to the fact that the current best-known bounds on the inverse theorem for non-classical polynomials are ineffective (the same occurs in [7]).

1.3.2. *Unbounded complexity.* The technique used to handle infinite removal lemmas is a compactness argument first introduced by Alon and Shapira [4] in the graph setting. A key ingredient of their proof is a strong regularity lemma.

Bhattacharyya, Fischer, Hatami, Hatami, and Lovett [7] prove that all linear-invariant subspace-hereditary properties that are also translation-invariant and bounded-complexity are testable. Their result follows from an infinite removal lemma for arithmetic patterns of bounded complexity. The proof of this result involves a strong arithmetic regularity lemma and a compactness argument in the spirit of Alon and Shapira.

To remove the bounded complexity assumption from [7], we prove a new strong arithmetic regularity lemma obtained by iterating a weaker arithmetic regularity lemma. The key innovation here is the level of Gowers uniformity used in each iteration is allowed to increase at each step of the process.

2. Colored patterns and removal lemmas

Theorem 1.3 and Theorem 1.4 both follow from an arithmetic removal lemma for colored linear patterns. In this section we define these objects and state the main removal lemma.

Definition 2.1. A **linear form** over $\mathbb{F}_p$ in ℓ variables is an expression L of the form

$$L(x_1, \dots, x_\ell) = \sum_{i=1}^{\ell} c_i x_i$$

with $c_i \in \mathbb{F}_p$. For any $\mathbb{F}_p$-vector space V, the linear form L gives rise to a function $L\colon V^\ell \to V$ that is linear in each variable.

Definition 2.2. For a prime p and a finite set $\mathcal{S}$, an **$\mathcal{S}$-colored pattern** over $\mathbb{F}_p$ consisting of m linear forms in ℓ variables is a pair $(\boldsymbol{L}, \psi)$ given by a system $\boldsymbol{L} = (L_1, \dots, L_m)$ of m linear forms in ℓ variables and a coloring $\psi\colon [m] \to \mathcal{S}$. Given a finite-dimensional $\mathbb{F}_p$-vector space V and a function $f\colon V \to \mathcal{S}$, an $(\boldsymbol{L}, \psi)$**-instance** in f is some $\boldsymbol{x} \in V^\ell$ such that $f(L_i(\boldsymbol{x})) = \psi(i)$ for all $i \in [m]$. An instance is called **generic** if $x_1, \dots, x_\ell$ are linearly independent. We say that $(\boldsymbol{L}, \psi)$ is **translation-invariant** if the coefficient of x_1 is 1 in each of $L_1, \dots, L_m$.

Given a finite-dimensional $\mathbb{F}_p$-vector space V and functions $f_1, \dots, f_m\colon V \to [-1, 1]$, we write

$$\Lambda_{\boldsymbol{L}}(f_1, \dots, f_m) := \mathbb{E}_{\boldsymbol{x} \in V^k}[f_1(L_1(\boldsymbol{x})) \cdots f_m(L_m(\boldsymbol{x}))].$$

Definition 2.3. For an $\mathcal{S}$-colored pattern over $\mathbb{F}_p$ consisting of m linear forms in k variables $(\boldsymbol{L}, \psi)$, a finite dimensional $\mathbb{F}_p$-vector space V, and a function $f\colon V \to \mathcal{S}$, define the $(\boldsymbol{L}, \psi)$**-density** in f to be $\Lambda_{\boldsymbol{L}}(f_1, \dots, f_m)$ where $f_i := 1_{f^{-1}(\psi(i))}$ for each $i \in [m]$.

Our main removal lemma is the following result.

Theorem 2.4 (Main removal lemma)**.** *Fix a prime p and a finite set $\mathcal{S}$. Let $\mathcal{H}$ be a (possibly infinite) set of $\mathcal{S}$-colored patterns over $\mathbb{F}_p$. For every $\epsilon > 0$, there exists a finite set $\mathcal{H}_\epsilon \subseteq \mathcal{H}$ and $\delta = \delta(\epsilon, \mathcal{H}) > 0$ such that the following holds. Let V be a finite-dimensional $\mathbb{F}_p$-vector space. If $f\colon V \to \mathcal{S}$ has H-density at most δ for every $H \in \mathcal{H}_\epsilon$, then there exists a recoloring $g\colon V \to \mathcal{S}$ that agrees with f on all but an at most ϵ-fraction of V such that g has no generic H-instances for every $H \in \mathcal{H}$.*

There are several difficulties in the proof of the main removal lemma. The first is that individual patterns $H \in \mathcal{H}$ may have "infinite complexity". Second, the set of patterns $\mathcal{H}$ may be infinite. Complicating this, even if all patterns in $\mathcal{H}$ have finite complexity, these complexities can be unbounded. Finally, there are major difficulties related to the fact that the patterns in $\mathcal{H}$ are not necessarily translation-invariant.

We use a trick called "projectivization" to reduce to the case where all patterns have finite complexity. To do this, we need a slightly modified version of the main removal lemma that we call the projective removal lemma (Theorem 2.8).

A "compactness argument" due to Alon and Shapira [4] reduces the problem of an infinite collection of patterns to a finite one at the expense of requiring a stronger arithmetic regularity lemma. If the collection of patterns is all complexity at most d, we only require a strong U^{d+1}-regularity lemma with rapidly decreasing error parameter. In the most general case when the collection of patterns has unbounded complexity we require an even stronger regularity lemma where the error parameter rapidly decreases and the degree of the uniformity norm rapidly increases.

Unless we restrict to the special case where all patterns in $\mathcal{H}$ are translation-invariant, the origin of the vector space plays a special role. This is unfortunate because it is impossible to regularize a function in the neighborhood of the origin. Since regularity methods are useless here, we turn to a new technique called patching, originally introduced by the authors and Fox [12], to deal with the portions of the vector space that cannot be regularized.

Definition 2.5. Let $\mathcal{S}$ be a finite set equipped with a group action of $\mathbb{F}_p^\times$ that we denote $c \cdot s$ for $c \in \mathbb{F}_p^\times$ and $s \in \mathcal{S}$. Given a finite-dimensional $\mathbb{F}_p$-vector space V, a function $f\colon V \to \mathcal{S}$ is **projective** if it preserves the action of $\mathbb{F}_p^\times$, i.e., $f(cx) = c \cdot f(x)$ for all $c \in \mathbb{F}_p^\times$ and all $x \in V$.

Definition 2.6. A list of linear forms $\boldsymbol{L} = (L_1, \dots, L_m)$ is **finite complexity** if no form is identically equal to zero, i.e., $L_i \not\equiv 0$ for all $i \in [m]$, and no two forms are linearly dependent, i.e., $L_i \not\equiv cL_j$ for all $i \neq j$ and $c \in \mathbb{F}_p$.

Definition 2.7. Fix a prime p and a positive integer ℓ. We consider two particular systems of linear forms. For $\boldsymbol{i} = (i_1, \dots, i_\ell) \in \mathbb{F}_p^\ell$, define

$$L_{\boldsymbol{i}}^\ell(x_1, \dots, x_\ell) := i_1 x_1 + \cdots + i_\ell x_\ell.$$

Then define $\boldsymbol{L}^\ell := (L_{\boldsymbol{i}}^\ell)_{\boldsymbol{i} \in \mathbb{F}_p^\ell}$, the system of p^ℓ linear forms in ℓ variables that defines an ℓ-dimensional subspace.

Let $E_\ell \subset \mathbb{F}_p^\ell$ be the set of non-zero vectors whose first non-zero coordinate is 1. Then define $\overline{\boldsymbol{L}}^\ell := (L_{\boldsymbol{i}}^\ell)_{\boldsymbol{i} \in E_\ell}$, a system of $(p^\ell - 1)/(p-1)$ linear forms in ℓ variables.

Note that unlike $\boldsymbol{L}^\ell$, the system $\overline{\boldsymbol{L}}^\ell$ has finite complexity. For technical reasons, it will be convenient to reduce the removal lemma for general patterns to the case where all patterns are defined by a system of the form $\boldsymbol{L}^\ell$. Then we reduce this to the following projective removal lemma where all patterns are defined by a system of the form $\overline{\boldsymbol{L}}^\ell$.

Theorem 2.8 (Projective removal lemma)**.** *Fix a prime p and a finite set $\mathcal{S}$ equipped with an $\mathbb{F}_p^\times$-action. Let $\mathcal{H}$ be a (possibly infinite) set consisting of $\mathcal{S}$-colored patterns over $\mathbb{F}_p$ of the form $(\overline{\boldsymbol{L}}^\ell, \psi)$ where ℓ is some positive integer and $\psi\colon E_\ell \to \mathcal{S}$ is some map (see Definition 2.7 for the definition of $\overline{\boldsymbol{L}}^\ell$ and E_ℓ). For every $\epsilon > 0$, there exists a finite subset $\mathcal{H}_\epsilon \subseteq \mathcal{H}$ and $\delta = \delta(\epsilon, \mathcal{H}) > 0$ such that the following holds. Let V be a finite-dimensional $\mathbb{F}_p$-vector space. If $f\colon V \to \mathcal{S}$ is a projective function with H-density at most δ for every $H \in \mathcal{H}_\epsilon$, then there exists a projective recoloring $g\colon V \to \mathcal{S}$ that agrees with f on all but an at most ϵ-fraction of V such that g has no generic H-instances for every $H \in \mathcal{H}$.*

3. Preliminaries on higher-order Fourier analysis

3.1. Gowers norms and complexity.

Definition 3.1. Fix a prime p, a finite-dimensional $\mathbb{F}_p$-vector space V, and an abelian group G. Given a function $f\colon V \to G$ and a shift $h \in V$, define the **additive derivative** $D_h f\colon V \to G$ by

$$(D_h f)(x) := f(x+h) - f(x).$$

Given a function $f\colon V \to \mathbb{C}$ and a shift $h \in V$, define the **multiplicative derivative** $\Delta_h f\colon V \to \mathbb{C}$ by

$$(\Delta_h f)(x) := f(x+h)\overline{f(x)}.$$

Definition 3.2. Fix a prime p and a finite-dimensional $\mathbb{F}_p$-vector space V. Given a function $f\colon V \to \mathbb{C}$ and $d \geq 1$, the **Gowers uniformity norm** $\|f\|_{U^d}$ is defined by

$$\|f\|_{U^d} := |\mathbb{E}_{x,h_1,\dots,h_d \in V}(\Delta_{h_1} \cdots \Delta_{h_d} f)(x)|^{1/2^d}.$$

Definition 3.3. A system $\boldsymbol{L} = (L_1, \dots, L_m)$ of m linear forms in ℓ variables is **complexity at most** d if for all $\epsilon > 0$ there exists $\delta > 0$ such that for all $f_1, \dots, f_\ell\colon V \to [-1,1]$ it holds that

$$|\Lambda_{\boldsymbol{L}}(f_1, \dots, f_\ell)| \leq \epsilon \qquad \text{whenever} \qquad \min_{1 \leq i \leq \ell} \|f_i\|_{U^{d+1}} \leq \delta. \tag{3.1}$$

The **complexity** of $\boldsymbol{L}$ is the smallest d such that the above holds, and infinite otherwise.

Remark 3.4. The above definition is sometimes known as true complexity. It is known that a pattern $(L_1, \dots, L_m)$ is complexity at most d if and only if $L_1^{d+1}, \dots, L_m^{d+1}$ are linearly independent as $(d+1)$th order tensors [18, 24].

Let $(L_1, \dots, L_m)$ be any pattern such that no form is identically zero and no two forms are linearly dependent. It is known (for example, because Cauchy-Schwarz complexity is an upper bound for true complexity [22]) that $(L_1, \dots, L_m)$ has complexity at most $m-2$.

It follows from the above discussion that the definition of complexity given in Definition 3.3 agrees with the definition of finite complexity given in Definition 2.6.

3.2. **Non-classical polynomials and homogeneity.** For ease of notation we write

$$\mathbb{U}_k := \tfrac{1}{p^k}\mathbb{Z}/\mathbb{Z} \subset \mathbb{R}/\mathbb{Z} \tag{3.2}$$

through the paper.

Definition 3.5. Fix a prime p, and a non-negative integer $d \geq 0$. Let V be a finite-dimensional $\mathbb{F}_p$-vector space. A **non-classical polynomial** of degree at most d is a map $P\colon V \to \mathbb{R}/\mathbb{Z}$ that satisfies

$$(D_{h_1} \cdots D_{h_{d+1}} P)(x) = 0$$

for all $h_1, \dots, h_{d+1}, x \in V$. The **degree** of P is the smallest $d > 0$ such that the above holds. The **depth** of P is the smallest $k \geq 0$ such that P takes values in a coset of $\mathbb{U}_{k+1}$.

See [33, Lemma 1.7] for some basic facts about non-classical polynomials. We record one such fact here.

Lemma 3.6 ([33, Lemma 1.7(iii)])**.** *Fix a prime p, and a finite-dimensional $\mathbb{F}_p$-vector space $V \simeq \mathbb{F}_p^n$. Then $P\colon V \to \mathbb{R}/\mathbb{Z}$ is a non-classical polynomial of degree at most d if and only if it can be expressed in the form*

$$P(x_1, \dots, x_n) = \alpha + \sum_{\substack{0 \leq i_1, \dots, i_n < p,\ j \geq 0: \\ 0 < i_1 + \cdots + i_n \leq d - k(p-1)}} \frac{c_{i_1,\dots,i_n,k} |x_1|^{i_1} \cdots |x_n|^{i_n}}{p^{k+1}} \pmod 1,$$

for some $\alpha \in \mathbb{R}/\mathbb{Z}$ and coefficients $c_{i_1,\dots,i_n,k} \in \{0, \dots, p-1\}$ and where $|\cdot|$ is the standard map $\mathbb{F}_p \to \{0, \dots, p-1\}$. Furthermore, this representation is unique.

As a corollary we see that in characteristic p, every non-classical polynomial of degree at most d has depth at most $\lfloor (d-1)/(p-1) \rfloor$.

Definition 3.7. A **homogeneous non-classical polynomial** is a non-classical polynomial $P\colon V \to \mathbb{R}/\mathbb{Z}$ that also satisfies the following. For all $b \in \mathbb{F}_p$ there exists $\sigma_b^{(P)} \in \mathbb{Z}/p^{k+1}\mathbb{Z}$ such that $P(bx) = \sigma_b^{(P)} P(x)$ for all $x \in V$.

Lemma 3.8 ([24, Lemma 3.3])**.** *Fix a prime p and integers $d > 0$ and $k \geq 0$ satisfying $k \leq \lfloor (d-1)/(p-1) \rfloor$. For each $b \in \mathbb{F}_p$ there exists $\sigma_b^{(d,k)} \in \mathbb{Z}/p^{k+1}\mathbb{Z}$ such that $\sigma_b^{(P)} = \sigma_b^{(d,k)}$ for all homogeneous non-classical polynomials P of degree d and depth k. Furthermore, for $b \neq 0$, $\sigma_b^{(d,k)}$ is uniquely determined by the following two properties:*

(i) $\sigma_b^{(d,k)} \equiv b^d \pmod{p}$

(ii) $\left(\sigma_b^{(d,k)}\right)^{p-1} = 1$ *(in $\mathbb{Z}/p^{k+1}\mathbb{Z}$).*

Theorem 3.9 ([24, Theorem 3.4])**.** *Let P be a non-classical polynomial of degree d and depth k. Then P can be written as the sum of homogeneous non-classical polynomials of degree at most d and depth at most k.*

3.3. **Polynomial factors.**

Definition 3.10. Fix a prime p. Define

$$D_p := \{(d,k) \in \mathbb{Z}_{>0} \times \mathbb{Z}_{\geqslant 0} : k \leq \lfloor (d-1)/(p-1) \rfloor\},$$

and

$$\mathcal{I}_p := \left\{ I \in \mathbb{Z}_{\geqslant 0}^{D_p} : \sum_{(d,k)\in D_p} I_{d,k} < \infty \right\}.$$

We call $I \in \mathcal{I}_p$ a **parameter list**. For $I \in \mathcal{I}_p$, we write $\|I\| := p^{\sum_{d,k}(k+1)I_{d,k}}$ and $\deg I$ for the largest d such that $I_{d,k} \neq 0$ for some k. We add and subtract parameter list coordinatewise. For $I, I' \in \mathcal{I}_p$, we write $I \leq I'$ if $I_{d,k} \leq I'_{d,k}$ for all $(d,k) \in D_p$.

Definition 3.11. For p a prime and $I \in \mathcal{I}_p$, define the **atom-indexing set** of I to be

$$A_I := \prod_{(d,k)\in D_p} \left(\tfrac{1}{p^{k+1}}\mathbb{Z}/\mathbb{Z}\right)^{I_{d,k}}. \tag{3.3}$$

(Note that $|A_I| = \|I\|$.)

For $I, I' \in \mathcal{I}_p$ with $I \leq I'$, write $\pi \colon A_{I'} \to A_I$ for the standard projection map defined by

$$\pi\left(\left(a_{d,k}^i\right)_{\substack{(d,k)\in D_p \\ i\in[I'_{d,k}]}}\right) \mapsto \left(a_{d,k}^i\right)_{\substack{(d,k)\in D_p \\ i\in[I_{d,k}]}}. \tag{3.4}$$

A_I is equipped with the following $\mathbb{F}_p^\times$-action:

$$c \cdot \left(a_{d,k}^i\right)_{\substack{(d,k)\in D_p \\ i\in[I'_{d,k}]}} := \left(\sigma_c^{(d,k)} a_{d,k}^i\right)_{\substack{(d,k)\in D_p \\ i\in[I'_{d,k}]}}, \tag{3.5}$$

where $\sigma_c^{(d,k)}$ is defined in Lemma 3.8.

Definition 3.12. Fix a prime p. Let V be a finite-dimensional $\mathbb{F}_p$-vector space and let $I \in \mathcal{I}_p$ be a parameter list. A **polynomial factor** on V with parameters I, denoted $\mathfrak{B}$, is a collection

$$\left(P_{d,k}^i\right)_{\substack{(d,k)\in D_p \\ i\in[I_{d,k}]}}$$

where $P_{d,k}^i$ is a homogeneous non-classical polynomial of degree d and depth k. We also use $\mathfrak{B}$ to denote the map $\mathfrak{B} \colon V \to A_I$ defined by evaluation of the polynomials. We also associate to $\mathfrak{B}$ the partition of V given by the fibers of this map. The atoms of this partition are called the atoms of $\mathfrak{B}$. We write $\|\mathfrak{B}\| := \|I\|$ and $\deg \mathfrak{B} := \deg I$.

Note that if $\mathfrak{B}$ is a polynomial factor on V with parameters I, then $\mathfrak{B}(cx) = c \cdot \mathfrak{B}(x)$ for all $c \in \mathbb{F}_p^\times$ and $x \in V$ where the $\mathbb{F}_p^\times$-action on A_I is defined in Eq. (3.5).

Definition 3.13. Fix a prime p. Let V be a finite-dimensional $\mathbb{F}_p$-vector space and let $I, I' \in \mathcal{I}_p$ be two parameter lists. Let $\mathfrak{B}$ and $\mathfrak{B}'$ be two polynomial factors on V with parameters I and I'. We say that $\mathfrak{B}'$ is a **refinement** of $\mathfrak{B}$ if $I \le I'$ and the lists of polynomials defining $\mathfrak{B}'$ are extensions of the lists of polynomials defining $\mathfrak{B}$. We say that $\mathfrak{B}'$ is a **weak refinement** of $\mathfrak{B}$ if the partition of V associated to $\mathfrak{B}'$ is a refinement of the partition associated to $\mathfrak{B}$.

Note that if $\mathfrak{B}'$ is a refinement of $\mathfrak{B}$, then $\mathfrak{B} = \pi \circ \mathfrak{B}'$ where $\pi \colon A_{I'} \to A_I$ is the projection defined in Eq. (3.4).

Definition 3.14. Fix a prime p and integer $d \ge 0$. Let V be a finite-dimensional $\mathbb{F}_p$-vector space. For a non-classical polynomial $P \colon V \to \mathbb{R}/\mathbb{Z}$, define the **$d$-rank of P**, denoted $\operatorname{rank}_d P$, to be the smallest integer r such that there exists non-classical polynomials $Q_1, \dots, Q_r \colon V \to \mathbb{R}/\mathbb{Z}$ of degree at most $d-1$ and a function $\Gamma \colon (\mathbb{R}/\mathbb{Z})^r \to \mathbb{R}/\mathbb{Z}$ such that $P(x) = \Gamma(Q_1(x), \dots, Q_r(x))$ for all $x \in V$.

For a polynomial factor $\mathfrak{B}$ on V with parameters $I \in I_p$, defined by a collection $(P^i_{d,k})_{(d,k)\in D_p, i \in [I_{d,k}]}$ where $P^i_{d,k}$ is a homogeneous non-classical polynomial of degree d and depth k, we define the **rank of $\mathfrak{B}$**, denoted $\operatorname{rank} \mathfrak{B}$, to be

$$\min_{\boldsymbol{\lambda} \in \prod_{(d,k)\in D_p} (\mathbb{Z}/p^{k+1}\mathbb{Z})^{I_{d,k}}} \operatorname{rank}_{d'} \left(\sum_{(d,k)\in D_p} \sum_{i=1}^{I_{d,k}} \lambda^i_{d,k} P^i_{d,k} \right)$$

where

$$d' := \min_{(d,k)\in D_p, i \in [I_{d,k}]} \deg \left(\lambda^i_{d,k} P^i_{d,k} \right).$$

Lemma 3.15. *Fix a prime p, a parameter list $I \in \mathcal{I}_p$, and a positive integer r. There exists a constant $n_{high-rank}(p, I, r)$ such that for every $n \ge n_{high-rank}(p, I, r)$, there exists a polynomial factor $\mathfrak{B}$ on $\mathbb{F}_p^n$ with parameters I and satisfying* $\operatorname{rank} \mathfrak{B} \ge r$.

Proof. For $(d,k) \in D_p$, write $d = (k+a)(p-1) + b$ where $a \ge 0$ and $b \in \{1, \dots, p-1\}$. By Lemma 3.6 there exists a non-classical polynomial of degree d and depth k. For example, consider

$$\frac{|x_1|^{p-1}|x_2|^{p-1} \cdots |x_a|^b}{p^{k+1}} \pmod 1.$$

By Theorem 3.9, we can decompose this non-classical polynomial as the sum of homogeneous non-classical polynomials of degree at most d and depth at most k. Let $P \colon \mathbb{F}_p^a \to \mathbb{U}_{k+1}$ be the homogeneous part of degree d and depth k.

Next define $Q \colon \mathbb{F}_p^{a \oplus N} \to \mathbb{U}_{k+1}$ by

$$Q(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) := P(\boldsymbol{x}_1) + \cdots + P(\boldsymbol{x}_N).$$

This is clearly a homogeneous non-classical polynomial of degree d and depth k. We claim that for N large enough, we have $\operatorname{rank}_d Q \ge r$. The proof of this fact uses several basic results from [33] that are not used elsewhere in this paper.

First, since P has degree exactly d, we have that $(D_{h_1} \cdots D_{h_d} P)(x)$ is a constant, independent of x, but is not identically zero. This implies that

$$c := \mathbb{E}_{h_1, \dots, h_d \in \mathbb{F}_p^a} e^{2\pi i (D_{h_1} \cdots D_{h_d} P)} < 1.$$

The quantity $-\log_p c$ is known as the *analytic rank* of P. Now a simple calculation shows that

$$\mathbb{E}_{\boldsymbol{h}_1, \dots, \boldsymbol{h}_d \in (\mathbb{F}_p^a)^{\oplus N}} e^{2\pi i (D_{\boldsymbol{h}_1} \cdots D_{\boldsymbol{h}_d} Q)} = c^N.$$

To conclude we use [33, Lemma 1.15(iii)] which implies that $\operatorname{rank}_d Q \ge -c_d \log_p(c^N)$ for some constant $c_d > 0$ only depending on d. Since $c < 1$, taking N large enough gives $\operatorname{rank}_d Q \ge r$, as desired.

Thus there exist homogeneous non-classical polynomials $Q_{d,k}\colon V_{d,k} \to \mathbb{U}_{k+1}$ of degree d and depth k that satisfy $\operatorname{rank}_d Q_{d,k} \geq r$ for each $(d,k) \in D_p$. Define the vector space

$$V := \bigoplus_{(d,k)\in D_p} V_{d,k}^{\oplus I_{d,k}}.$$

Define the homogeneous non-classical polynomials $\overline{Q}^i_{d,k}\colon V \to \mathbb{R}/\mathbb{Z}$ for each $(d,k) \in D_p$ and $i \in [I_{d,k}]$ such that $\overline{Q}^i_{d,k}$ is equal to $Q_{d,k}$ evaluated on the ith copy of $V_{d,k}$ and does not depend on the other coordinates. In particular, we define

$$\overline{Q}^i_{d,k}\left((x^{i'}_{d',k'})_{\substack{(d',k')\in D_p \\ i'\in[I_{d',k'}]}}\right) := Q_{d,k}(x^i_{d,k}).$$

These polynomials define a polynomial factor $\mathfrak{B}$ on V with parameters I such that each of the homogeneous non-classical polynomials defining $\mathfrak{B}$ has rank at least r. Furthermore, since the polynomials defining $\mathfrak{B}$ depend on disjoint sets of variables, it follows that all non-trivial linear combinations of the polynomials defining $\mathfrak{B}$ also have high rank. Setting $n_{high-rank}(p,I,R) := \dim V$, we have constructed the desired polynomial factor on $\mathbb{F}_p^{n_{high-rank}(p,I,r)}$. To extend this construction to $\mathbb{F}_p^n$ with $n \geq n_{high-rank}(p,I,r)$ one can simply add on extra variables that none of the polynomials depend on. □

Lemma 3.16 ([7, Lemma 2.13])**.** *Fix a prime p and a positive integers d, r. Let V be a finite-dimensional $\mathbb{F}_p$-vector space and let $P\colon V \to \mathbb{R}/\mathbb{Z}$ be a non-classical polynomial of degree d such that $\operatorname{rank}_d(P) \geq r + p$. Let $U \leq V$ be a codimension-1 hyperplane. Then $\operatorname{rank}_d(P|_U) \geq r$ unless $d = 1$ and $P|_U$ is identically zero.*

As a consequence, let $\mathfrak{B}$ be a polynomial factor on V and let $P\colon V \to \mathbb{R}/\mathbb{Z}$ be a linear polynomial. Write $\mathfrak{B}'$ for the common refinement of $\mathfrak{B}$ and $\{P\}$. If $\operatorname{rank}\mathfrak{B}' \geq r + p$, then $\operatorname{rank}\mathfrak{B}|_U \geq r$ where U is the codimension-1 hyperplane where P vanishes.

3.4. **Equidistribution and consistency sets.**

Definition 3.17. Fix a prime p, integers $d > 0$ and $k \geq 0$ satisfying $k \leq \lfloor (d-1)/(p-1) \rfloor$, and a system $\boldsymbol{L} = (L_1, \ldots, L_m)$ of m linear forms in ℓ variables. Define the (d,k)**-consistency set of** $\boldsymbol{L}$, denoted $\Phi_{d,k}(\boldsymbol{L})$, to be the subset of $\mathbb{U}_{k+1}^m$ consisting of the tuples $\boldsymbol{a} = (a_1, \ldots, a_m)$ such that there exists a finite-dimensional $\mathbb{F}_p$-vector space V, a homogeneous non-classical polynomial $P\colon V \to \mathbb{U}_{k+1}$ of degree d and depth k, and a tuple $\boldsymbol{x} \in V^\ell$ such that $a_i = P(L_i(\boldsymbol{x}))$ for all $i \in [m]$.

For a parameter list $I \in \mathcal{I}_p$, define the I**-consistency set of** $\boldsymbol{L}$ to be the set of tuples $\boldsymbol{a} = (a_1, \ldots, a_m) \in A_I^m$ such that for each $(d,k) \in D_p$ and $j \in [I_{d,k}]$ the tuple $\left((a_1)^j_{d,k}, \ldots, (a_m)^j_{d,k}\right)$ lies in $\Phi_{d,k}(\boldsymbol{L})$.

Lemma 3.18. *Fix a prime p, integers $d > 0$ and $k \geq 0$ satisfying $k \leq \lfloor (d-1)/(p-1) \rfloor$, and a system $\boldsymbol{L} = (L_1, \ldots, L_m)$ of m linear forms. The (d,k)-consistency set of $\boldsymbol{L}$ is a subgroup of $\mathbb{U}_{k+1}^m$.*

Proof. Suppose $\boldsymbol{a}, \boldsymbol{b} \in \Phi_{d,k}(\boldsymbol{L})$. We wish to show that $-\boldsymbol{a}$ and $\boldsymbol{a} + \boldsymbol{b}$ both lie in this set. By definition, there exist finite-dimensional $\mathbb{F}_p$-vector spaces V, W, homogeneous non-classical polynomials $P\colon V \to \mathbb{U}_{k+1}$ and $Q\colon W \to \mathbb{U}_{k+1}$ of degree d and depth k, and tuples $\boldsymbol{x} \in V^k$ and $\boldsymbol{y} \in W^k$ such that $a_i = P(L_i(\boldsymbol{x}))$ and $b_i = Q(L_i(\boldsymbol{y}))$ for all $i \in [m]$.

Note that $-\boldsymbol{a} \in \Phi_{d,k}(\boldsymbol{L})$ since $-P\colon V \to \mathbb{U}_{k+1}$ is a homogeneous non-classical polynomial of degree d and depth k that satisfies $(-P)(L_i(\boldsymbol{x})) = -a_i$.

Now define $P \oplus Q \colon V \oplus W \to \mathbb{U}_{k+1}$ by $(P \oplus Q)(v \oplus w) := P(v) + Q(w)$. One can easily check that $P \oplus Q$ is a homogeneous non-classical polynomial of degree d and depth k. Finally note that $(P \oplus Q)(L_i(\boldsymbol{x} \oplus \boldsymbol{y})) = a_i + b_i$, as desired.[1] □

Theorem 3.19 (Equidistribution [24, Theorem 3.10])**.** *Fix a prime p, a positive integer $d > 0$, and a parameter $\epsilon > 0$. There exists $r_{equi}(p, d, \epsilon)$ such that the following holds. Let V be a finite-dimensional $\mathbb{F}_p$-vector space and let $\mathfrak{B}$ be a polynomial factor on V with parameters I such that $\deg \mathfrak{B} \le d$ and $\operatorname{rank}(\mathfrak{B}) \ge r_{equi}(p, d, \epsilon)$. Then for a system of linear forms $\boldsymbol{L} = (L_1, \dots, L_m)$ consisting of m forms in ℓ variables, and a tuple of atoms $\boldsymbol{a} = (a_1, \dots, a_m) \in \Phi_I(\boldsymbol{L})$,*

$$\left| \Pr_{\boldsymbol{x} \in V^\ell} (\mathfrak{B}(L_i(\boldsymbol{x})) = a_i \text{ for all } i \in [m]) - \frac{1}{|\Phi_I(\boldsymbol{L})|} \right| \le \epsilon.$$

Remark 3.20. To be completely correct, the statement given above follows by combining [24, Theorem 3.10] and [24, Corollary 2.13].

Note that the probability above is 0 if $\boldsymbol{a} \notin \Phi_I(\boldsymbol{L})$. We typically apply the above theorem with ϵ that decreases rapidly with $\|I\|$, for example, taking $\epsilon = 2\|I\|^m$ and using the fact that $|\Phi_I(\boldsymbol{L})| \le \|\mathfrak{B}\|^m$, we see that in this case the probability above is at least $1/(2|\Phi_I(\boldsymbol{L})|)$.

Consistency sets are often hard to compute exactly. The next two lemmas give exact relations on the sizes of consistency sets in two special cases that occur in this paper.

Definition 3.21. Fix a prime p. A system $\boldsymbol{L}$ of m linear forms in ℓ variables over $\mathbb{F}_p$ is **full dimensional** if $|\Phi_{d,k}(\boldsymbol{L})| = |\Phi_{d,k}(\boldsymbol{L}^\ell)|$ for all $(d, k) \in D_p$ (recall the system $\boldsymbol{L}^\ell$ defined in Definition 2.7 defines an ℓ-dimensional subspace).

Lemma 3.22. *Fix a prime p and a positive integer ℓ. Let $J \subseteq \mathbb{F}_p^\ell$ be a set that contains at least one vector in each direction (i.e., for each $\boldsymbol{i} \in \mathbb{F}_p^\ell$, there exists $\boldsymbol{j} \in J$ and $\in \mathbb{F}_p^\times$ such that $\boldsymbol{i} = b\boldsymbol{j}$). Consider the system $\boldsymbol{L}_J := (L^\ell_{\boldsymbol{i}})_{\boldsymbol{i} \in J}$ of $|J|$ linear forms in ℓ variables (recall the linear form $L^\ell_{\boldsymbol{i}}$ defined in Definition 2.7)). Then $\boldsymbol{L}_J$ is full dimensional.*

As a special case of this result we see that the system $\overline{\boldsymbol{L}}^\ell$, defined in Definition 2.7, is full rank.

Proof. Note that the system $\boldsymbol{L}_J$ is a subsystem of $\boldsymbol{L}^\ell$. This immediately implies that $|\Phi_{d,k}(\boldsymbol{L}_J)| \le |\Phi_{d,k}(\boldsymbol{L}^\ell)|$, since if $(a_{\boldsymbol{i}})_{\boldsymbol{i} \in \mathbb{F}_p^\ell} \in \Phi_{d,k}(\boldsymbol{L}^\ell)$, then $(a_{\boldsymbol{i}})_{\boldsymbol{i} \in J} \in \Phi_{d,k}(\boldsymbol{L}_J)$.

To go the other direction, we use the homogeneity of our polynomials. Suppose $(a_{\boldsymbol{j}})_{\boldsymbol{j} \in J} \in \Phi_{d,k}(\boldsymbol{L}_J)$. Thus there exists a finite-dimensional $\mathbb{F}_p$-vector space V and a homogeneous non-classical polynomial $P \colon V \to \mathbb{U}_{k+1}$ of degree d and depth k, and a vector $\boldsymbol{x} \in V^\ell$ such that $P(L^\ell_{\boldsymbol{j}}(\boldsymbol{x})) = a_{\boldsymbol{j}}$ for all $\boldsymbol{j} \in J$. Now by assumption, for $\boldsymbol{i} \in \mathbb{F}_p^\ell$, there exists $\boldsymbol{j} \in J$ and $b \in \mathbb{F}_p$ such that $\boldsymbol{i} = b\boldsymbol{j}$. Define $a_{\boldsymbol{i}} := \sigma_b^{(d,k)} a_{\boldsymbol{j}}$ (recall the definition of σ_b from Lemma 3.8). We claim that $(a_{\boldsymbol{i}})_{\boldsymbol{i} \in \mathbb{F}_p^\ell} \in \Phi_{d,k}(\boldsymbol{L}^\ell)$. This is true simply because

$$P(L^\ell_{\boldsymbol{i}}(x)) = P(bL^\ell_{\boldsymbol{j}}(\boldsymbol{x})) = \sigma_b^{(d,k)} P(L^\ell_{\boldsymbol{j}}(x)) = \sigma_b^{(d,k)} a_{\boldsymbol{j}} = a_{\boldsymbol{i}}$$

where $\boldsymbol{j} \in J$ and $b \in \mathbb{F}_p$ are defined as above. Thus $|\Phi_{d,k}(\boldsymbol{L}_J)| \ge |\Phi_{d,k}(\boldsymbol{L}^\ell)|$, as desired. □

Lemma 3.23. *Fix a prime p. Let $\boldsymbol{L}$ be a system of m linear forms in ℓ variables over $\mathbb{F}_p$. Say $\boldsymbol{L}$ is defined by M, an $m \times \ell$ matrix. (By this we mean that $L_i(x_1, \dots, x_\ell) = M_{i,1}x_1 + \cdots + M_{i,\ell}x_\ell$ for*

[1] To be completely correct, we also need to show that $\Phi_{d,k}(\boldsymbol{L})$ is non-empty, which follows from, for example Lemma 3.6 and Theorem 3.9, which together show the existence of homogeneous non-classical polynomial of degree d and depth k for every $(d, k) \in D_p$.

all $i \in [m]$.) Let $\boldsymbol{L}'$ be a system of $m(n+1)$ linear forms in $\ell + \ell'$ variables, defined by a matrix of the form

$$\left(\begin{array}{c|c} M & 0 \\ \hline c_1 M & \\ \vdots & N \\ c_n M & \end{array}\right)$$

where $c_1, \dots, c_n \in \mathbb{F}_p$ and N is an $mn \times \ell'$ matrix. Let $\boldsymbol{L}''$ be the system of $m(2n+1)$ linear forms in $\ell + 2\ell'$ variables, defined by the matrix

$$\left(\begin{array}{c|c|c} M & 0 & 0 \\ \hline c_1 M & & \\ \vdots & N & 0 \\ c_n M & & \\ \hline c_1 M & & \\ \vdots & 0 & N \\ c_n M & & \end{array}\right).$$

Then for all $(d,k) \in D_p$, we have

$$|\Phi_{d,k}(\boldsymbol{L})| \cdot \left|\Phi_{d,k}(\boldsymbol{L}'')\right| = \left|\Phi_{d,k}(\boldsymbol{L}')\right|^2.$$

Proof. We construct injections between $\Phi_{d,k}(\boldsymbol{L}) \times \Phi_{d,k}(\boldsymbol{L}'')$ and $\Phi_{d,k}(\boldsymbol{L}') \times \Phi_{d,k}(\boldsymbol{L}')$ in both directions. Write $\sigma_i := \sigma_{c_i}^{d,k}$ for $i \in [n]$ for the rest of the proof (see Lemma 3.8 for the definition).

Consider $\boldsymbol{a} = (a_1, \dots, a_m) \in \Phi_{d,k}(\boldsymbol{L})$ and $\boldsymbol{b} = (b_1, \dots, b_{m(2n+1)}) \in \Phi_{d,k}(\boldsymbol{L}'')$. By definition, there exists a finite-dimensional $\mathbb{F}_p$-vector space V, a homogeneous non-classical polynomial $P\colon V \to \mathbb{U}_{k+1}$ of degree d and depth k and a vector $\boldsymbol{x} \in V^\ell$ such that $P(L_i(\boldsymbol{x})) = a_i$ for all $i \in [m]$. Also by definition, there exists a finite-dimensional $\mathbb{F}_p$-vector space W, a homogeneous non-classical polynomial $Q\colon W \to \mathbb{U}_{k+1}$ of degree d and depth k and a vector $(\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{y}') \in W^\ell \times W^{\ell'} \times W^{\ell'}$ such that $Q(L_i''(\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{y}')) = b_i$ for all $i \in [m(2n+1)]$.

Now we map $(\boldsymbol{a}, \boldsymbol{b})$ to the pair $(\boldsymbol{a}', \boldsymbol{b}')$ where $b_i' = b_i$ for $i \in [m(n+1)]$ and $a_i' = a_i + b_i$ for $i \in [m]$ and $a'_{tm+i} = \sigma_t a_i + b_{m(n+t)+i}$ for $i \in [m]$ and $t \in [n]$. We can easily check that no two pairs $(\boldsymbol{a}, \boldsymbol{b})$ map to the same pair $(\boldsymbol{a}', \boldsymbol{b}')$. All that remains is to check that $\boldsymbol{a}', \boldsymbol{b}' \in \Phi_{d,k}(\boldsymbol{L}')$.

Define $P \oplus Q\colon V \oplus W \to \mathbb{U}_{k+1}$ by $(P \oplus Q)(x \oplus y) := P(x) + Q(y)$. This is clearly a homogeneous non-classical polynomial of degree d and depth k. Note that $\boldsymbol{z} := (\boldsymbol{x} \oplus \boldsymbol{x}', \boldsymbol{0} \oplus \boldsymbol{y}') \in (V \oplus W)^{\ell+\ell'}$ satisfies $(P \oplus Q)(L_i'(\boldsymbol{z})) = a_i'$ for all $i \in [m(n+1)]$. Similarly note that $\boldsymbol{z}' := (\boldsymbol{0} \oplus \boldsymbol{x}', \boldsymbol{0} \oplus \boldsymbol{y}) \in (V \oplus W)^{\ell+\ell'}$ satisfies $(P \oplus Q)(L_i'(\boldsymbol{z}')) = b_i'$ for all $i \in [m(n+1)]$. This demonstrates the first injection.

Now consider $\boldsymbol{a} = (a_1, \dots, a_{m(n+1)}) \in \Phi_{d,k}(\boldsymbol{L}')$ and $\boldsymbol{b} = (b_1, \dots, b_{m(n+1)}) \in \Phi_{d,k}(\boldsymbol{L}')$. By definition, there exists a finite-dimensional $\mathbb{F}_p$-vector space V, a homogeneous non-classical polynomial $P\colon V \to \mathbb{U}_{k+1}$ of degree d and depth k and a vector $(\boldsymbol{x}, \boldsymbol{y}) \in V^\ell \times V^{\ell'}$ such that $P(L_i'(\boldsymbol{x}, \boldsymbol{y})) = a_i$ for $i \in [m(n+1)]$. Also by definition, there exists a finite-dimensional $\mathbb{F}_p$-vector space W, a homogeneous non-classical polynomial $Q\colon W \to \mathbb{U}_{k+1}$ of degree d and depth k and a vector $(\boldsymbol{x}', \boldsymbol{y}') \in W^\ell \times W^{\ell'}$ such that $Q(L_i'(\boldsymbol{x}', \boldsymbol{y}')) = b_i$ for $i \in [m(n+1)]$.

We map $(\boldsymbol{a}, \boldsymbol{b})$ to the pair $(\boldsymbol{a}', \boldsymbol{b}')$ where $a_i' = a_i$ for $i \in [m]$ and $b_i' = a_i + b_i$ for $i \in [m]$ and $b'_{m+i} = a_{m+i}$ for $i \in [mn]$ and $b'_{m(n+1)+i} = b_{m+i}$ for $i \in [mn]$. We can easily check that no two pairs $(\boldsymbol{a}, \boldsymbol{b})$ map to the same pair $(\boldsymbol{a}', \boldsymbol{b}')$. All that remains is to check that $\boldsymbol{a}' \in \Phi_{d,k}$ and $\boldsymbol{b}' \in \Phi_{d,k}(\boldsymbol{L}'')$.

As above, define the homogeneous non-classical polynomial $P \oplus Q\colon V \oplus W \to \mathbb{U}_{k+1}$ of degree d and depth k by $(P \oplus Q)(x \oplus y) := P(x) + Q(y)$. Note that $\boldsymbol{z} := \boldsymbol{x} \oplus \boldsymbol{0} \in (V \oplus W)^\ell$ satisfies $(P \oplus Q)(\boldsymbol{L}_i(\boldsymbol{z})) = a_i'$ for $i \in [m]$ and $\boldsymbol{z}' := (\boldsymbol{x} \oplus \boldsymbol{x}', \boldsymbol{y} \oplus \boldsymbol{0}, \boldsymbol{0} \oplus \boldsymbol{y}') \in (V \oplus W)^\ell \times (V \oplus W)^{\ell'} \times (V \oplus W)^{\ell'}$ satisfies $(P \oplus Q)(L_i''(\boldsymbol{z}')) = b_i'$ for $i \in [m(2n+1)]$, as desired. ∎

3.5. **Subatom selection functions.** A situation that often occurs is the following. We have a polynomial factor $\mathfrak{B}$ with parameters I and a refinement $\mathfrak{B}'$ with parameters I'. We use the word atom to refer to the atoms of the partition induced by $\mathfrak{B}$; these atoms are indexed by A_I. We use the word subatom to refer to the atoms of the partition induced by $\mathfrak{B}'$; these atoms are indexed by $A_{I'}$. The projection map $\pi\colon A_{I'} \to A_I$, defined in Eq. (3.4), maps a subatom to the atom that it is contained in.

We wish to designate one subatom inside each atom as special. This choice is given by a map $s\colon A_I \to A_{I'}$ that is a right inverse for π. In this paper we define a certain class of these maps that we call subatom selection functions that have several desirable properties.

First we define certain polynomials $P_{d,k}\colon \mathbb{F}_p \to \mathbb{U}_{k+1}$ for each $(d,k) \in D_p$. For each $i \in \{1,\dots,p-1\}$, there exists a homogeneous non-classical polynomial $\mathbb{F}_p \to \mathbb{U}_{k+1}$ of degree $k(p-1)+i$ and depth k in one variable. (This follows from Lemma 3.6 and Theorem 3.9.) Let $P_{k(p-1)+i,k}$ be one such a polynomial. Finally, define

$$P_{(k+s)(p-1)+i,k+s} := P_{k(p-1)+i,k}$$

for all $i \in \{0,\dots,p-1\}$ and $s \ge 0$. This defines $P_{d,k}$ for each $(d,k) \in D_p$.

Definition 3.24. Fix a prime p and parameter lists $I, I' \in \mathcal{I}_p$ satisfying $I \le I'$. Let $c_{d,k}^{i,j} \in \mathbb{Z}/p^{k+1}\mathbb{Z}$ be arbitrary elements for $(d,k) \in D_p$ and $i \in [I_{1,0}]$ and $I_{d,k} < j \le I'_{d,k}$. A **subatom selection function** is a map of the form $s_{\boldsymbol{c}}\colon A_I \to A_{I'}$, defined by

$$[s_{\boldsymbol{c}}(a)]_{d,k}^i = \begin{cases} a_{d,k}^i & \text{if } i \le I_{d,k} \\ \sum_{j=1}^{I_{1,0}} c_{d,k}^{j,i} P_{d,k}(|a_{1,0}^j|) & \text{otherwise,} \end{cases}$$

where the maps $P_{d,k}\colon \mathbb{F}_p \to \mathbb{U}_{k+1}$ were defined in the preceding paragraph and $|\cdot|$ is the standard map $\mathbb{U}_1 \to \mathbb{F}_p$.

Lemma 3.25. *Fix a prime p, parameter lists $I, I' \in \mathcal{I}_p$ satisfying $I \le I'$, and a subatom selection function $s_{\boldsymbol{c}}\colon A_I \to A_{I'}$. The following hold:*

(i) $\pi \circ s = \mathrm{Id}$ (where $\pi\colon A_{I'} \to A_I$ is defined in Eq. (3.4));

(ii) for $a \in A_I$ and $b \in \mathbb{F}_p^\times$, we have $b \cdot s_{\boldsymbol{c}}(a) = s_{\boldsymbol{c}}(b \cdot a)$ (where the action of $\mathbb{F}_p^\times$ on A_I and $A_{I'}$ is defined in Eq. (3.5));

(iii) for every system $\boldsymbol{L}$ of m linear forms and every consistent tuple of atoms $(a_1,\dots,a_m) \in \Phi_I(\boldsymbol{L})$, we have

$$(s_{\boldsymbol{c}}(a_1),\dots,s_{\boldsymbol{c}}(a_m)) \in \Phi_{I'}(\boldsymbol{L})$$

(see Definition 3.17 for the definition of the consistency sets $\Phi_I(\boldsymbol{L})$ and $\Phi_{I'}(\boldsymbol{L})$).

Proof. Property (i) is immediate.

For the property (ii), by definition, we have

$$[b \cdot s_{\boldsymbol{c}}(a)]_{d,k}^i = \sigma_b^{(d,k)} [s_{\boldsymbol{c}}(a)]_{d,k}^i,$$

where $\sigma_b^{(d,k)}$ is defined in Lemma 3.8. Now $s_{\boldsymbol{c}}(b \cdot a)_{d,k}^i = \sigma_b^{(d,k)} a_{d,k}^i$ if $i \le I_{d,k}$, so we are done in this case.

Assume otherwise. Define d' such that $d' \equiv d \pmod{p-1}$ and $d = k(p-1)+i$ for some $i \in \{1,\dots,p-1\}$. Remember that $P_{d,k}$ is a homogeneous non-classical polynomial of degree d' and depth k. Note that $\sigma_b^{(1,0)} = b \in \mathbb{Z}/p\mathbb{Z}$. Then we have

$$[s_{\boldsymbol{c}}(b \cdot a)]_{d,k}^i = \sum_{j=1}^{I_{1,0}} c_{d,k}^{j,i} P_{d,k}(b|a_{1,0}^j|) = \sigma_b^{(d',k)} \sum_{j=1}^{I_{1,0}} c_{d,k}^{j,i} P_{d,k}(|a_{1,0}^j|).$$

To complete the proof of (ii) we need to show that $\sigma_b^{d,k} = \sigma_b^{d',k}$ whenever $d \equiv d' \pmod{p-1}$. This follows from Lemma 3.8 which implies that $\sigma_b^{d,k}$ is uniquely determined by the facts that $\sigma_b^{d,k} \equiv b^d$

(mod p) and $\left(\sigma_b^{(d,k)}\right)^{p-1} = 1$ in $\mathbb{Z}/p^{k+1}\mathbb{Z}$. The first property does not change when d changes by a multiple of $p-1$ (by Fermat's little theorem) and the second property does not depend on d at all. Thus we conclude the desired result.

Now we prove (iii). We know, by Eq. (3.1), that the consistency set $\Phi_{d,k}(\boldsymbol{L})$ is a subgroups of $\mathbb{U}_{k+1}^m$. Thus it suffices to prove that for $(a_1, \dots, a_m) \in \Phi_{1,0}(\boldsymbol{L}) \subseteq \mathbb{U}_1^m$ we have $(P_{d,k}(|a_1|), \dots, P_{d,k}(|a_m|)) \in \Phi_{d,k}(\boldsymbol{L})$. Given that $(a_1, \dots, a_m) \in \Phi_{1,0}$ we know that there exists a finite-dimensional $\mathbb{F}_p$-vector space V, a linear function $P\colon V \to \mathbb{U}_1$, and vectors $\boldsymbol{x} = (x_1, \dots, x_\ell) \in V^\ell$ such that $P(L_i(\boldsymbol{x})) = a_i$ for $i \in [m]$. Since P and L_i are linear, they commute, and thus $L_i(\boldsymbol{y}) = |a_i|$ for all $i \in [m]$ where $\boldsymbol{y} \in \mathbb{F}_p^\ell$ is defined by $y_i = |P(x_i)|$. Finally, since $P_{d,k}$ is a homogeneous non-classical polynomial of degree d' and depth k, we have $(P_{d,k}(|a_1|), \dots, P_{d,k}(|a_m|)) = (P_{d,k}(L_1(\boldsymbol{y})), \dots, P_{d,k}(L_m(\boldsymbol{y})) \in \Phi_{d',k}(\boldsymbol{L})$.

To complete the proof, let $Q\colon \mathbb{F}_p^n \to \mathbb{U}_{k+1}$ be a homogeneous non-classical polynomial of degree d and depth k. This exists by Lemma 3.6 and Theorem 3.9 as long as $n \geq \lceil (d-1)/(p-1) \rceil - k$. Then consider the map $P_{d,k} \oplus Q\colon \mathbb{F}_p \oplus \mathbb{F}_p^n \to \mathbb{U}_{k+1}$ defined as usual by $(P_{d,k} \oplus Q)(x \oplus y) := P_{d,k}(x) + Q(y)$. This is clearly a non-classical polynomial of degree d and depth k. Furthermore,

$$(P_{d,k} \oplus Q)(b(x \oplus y)) = \sigma_b^{(d',k)} P_{d,k}(x) + \sigma_b^{(d,k)} Q(y) = \sigma_b^{(d,k)} (P_{d,k} \oplus Q)(x \oplus y),$$

since $d' \equiv d \pmod{p-1}$. Considering $\boldsymbol{y} \oplus \boldsymbol{0} \in (\mathbb{F}_p \oplus \mathbb{F}_p^n)^\ell$ shows that $(P_{d,k}(|a_1|), \dots, P_{d,k}(|a_m|)) \in \Phi_{d,k}(\boldsymbol{L})$, as desired.[2] □

4. Arithmetic regularity and subatom selection

This section follows a fairly standard formula in the theory of regularity lemmas. We start with an inverse theorem, due to Tao and Ziegler [33]. Iterating the inverse theorem produces a weak regularity lemma (Lemma 4.3), iterating the weak regularity lemma produces a regularity lemma (Lemma 4.4), and iterating the regularity lemma gives a strong regularity lemma (Lemma 4.5). Finally we use the probabilistic method applied to the output of the strong regularity lemma to produce the desired "subatom selection" result (Theorem 4.6).

Lemma 4.3 and Lemma 4.4 are very similar to results in [8, 7], differing only in some technical details. The main innovation in this section is that Lemma 4.5 is much stronger than previous results. To accomplish this, we iterate Lemma 4.4 with the complexity parameter (i.e., degree of the non-classical polynomials) increasing at each step of the iteration. To our knowledge, this idea has not appeared previously in the literature.

Notation and conventions: Recall that a polynomial factor $\mathfrak{B}$ on a vector space V with parameters I gives rise to a partition (or σ-algebra) on V whose atoms are the fibers of the map $\mathfrak{B}\colon V \to A_I$. For a function $f\colon V \to \mathbb{C}$, we write $\mathbb{E}[f|\mathfrak{B}]\colon V \to \mathbb{C}$ for the projection of f onto the σ-algebra generated by $\mathfrak{B}$. Concretely, $\mathbb{E}[f|\mathfrak{B}](x)$ is defined to be the average of f over the atom of $\mathfrak{B}$ which contains x.

In this section we have to deal with many growth functions. Without loss of generality we always assume that these growth functions are monotone in all their parameters.

Theorem 4.1 (Inverse theorem [33, Theorem 1.10])**.** *Fix a prime p, a positive integer d, and a parameter $\delta > 0$. There exists $\epsilon_{inv}(p, d, \delta) > 0$ such that the following holds. Let V be a finite-dimensional $\mathbb{F}_p$-vector space. Given a function $f\colon V \to \mathbb{C}$ satisfying $\|f\|_\infty \leq 1$ and $\|f\|_{U^{d+1}} > \delta$, there exists a non-classical polynomial $P\colon V \to \mathbb{R}/\mathbb{Z}$ of degree at most d such that*

$$\left| \mathbb{E}_{x \in V} f(x) e^{-2\pi i P(x)} \right| \geq \epsilon_{inv}(p, d, \delta).$$

[2]This argument also shows that for any pattern $\boldsymbol{L}$, the consistency sets $\Phi_{d,k}(\boldsymbol{L})$ are nested as d increases by multiples of $p-1$, though this is the only time we will need that fact in this paper.

The next lemma is important for making factors high rank and its second claim will be critical in proving the stronger regularity lemma where we need to produce a refinement (instead of a weak refinement).

Lemma 4.2 (Making factors high rank [24], c.f. [7, Theorem 2.19])**.** *Fix a prime p, positive integer d, C_0, and a non-decreasing function $r\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$. There exist constants $C_{rank}(p, d, C_0, r)$ and $r_{rank}(p, d, C_0, r)$ such that the following holds. Let V be a finite-dimensional $\mathbb{F}_p$-vector space. Suppose that $\mathfrak{B}$ and $\mathfrak{B}'$ are polynomial factors on V with degree at most d such that $\mathfrak{B}'$ refines $\mathfrak{B}$ and $\|\mathfrak{B}'\| \le C_0$ and*

$$\operatorname{rank} \mathfrak{B} \ge r_{rank}(p, d, C_0, r).$$

Then there is a polynomial factor $\mathfrak{B}''$ on V that weakly refines $\mathfrak{B}'$, refines $\mathfrak{B}$, and satisfies $\|\mathfrak{B}''\| \le C_{rank}(p, d, C_0, r)$ and $\deg \mathfrak{B}'' \le d$ and $\operatorname{rank} \mathfrak{B}'' \ge r(\|\mathfrak{B}''\|)$.

Lemma 4.3 (Weak arithmetic regularity)**.** *Fix a prime p, positive integers d, R, C_0, a parameter $\eta > 0$, and a non-decreasing function $r\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$. There exist constants $C_{reg'}(p, d, R, C_0, \eta, r)$ and $r_{reg'}(p, d, R, C_0, \eta, r)$ such that the following holds. Let V be a finite-dimensional $\mathbb{F}_p$-vector space and let $\mathfrak{B}_0$ be a polynomial factor on V satisfying $\|\mathfrak{B}_0\| \le C_0$ and $\deg \mathfrak{B}_0 \le d$ and $\operatorname{rank} \mathfrak{B}_0 \ge r_{reg'}(p, d, R, C_0, \eta, r)$. Given functions $f^{(1)}, \dots, f^{(R)}\colon V \to [0,1]$, there exists a polynomial factor $\mathfrak{B}$ on V that refines $\mathfrak{B}_0$ with the following properties. There exists a decomposition*

$$f^{(\ell)} = f^{(\ell)}_{str} + f^{(\ell)}_{psr}$$

for each $\ell \in [R]$ such that:

(i) $f^{(\ell)}_{str} = \mathbb{E}[f^{(\ell)}|\mathfrak{B}]$ for each $\ell \in [R]$;
(ii) $\|f^{(\ell)}_{psr}\|_{U^{d+1}} < \eta$ for each $\ell \in [R]$;
(iii) $f^{(\ell)}_{str}$ has range $[0,1]$ and $f^{(\ell)}_{psr}$ has range $[-1,1]$ for each $\ell \in [R]$;
(iv) $\operatorname{rank} \mathfrak{B} \ge r(\|\mathfrak{B}\|)$;
(v) $\|\mathfrak{B}\| \le C_{reg'}(p, d, R, C_0, \eta, r)$ and $\deg \mathfrak{B} \le d$.

Proof. Set $M := \lceil R\epsilon_{inv}(p, d, \eta)^{-2} \rceil$. Define non-decreasing functions $r_i\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ for each $i = 0, \dots, M$ such that $r_0 = r$ and $r_{i+1}(N) \ge r_{rank}(p, d, p^{d^3} N, r_i)$ and such that $r_{i+1}(N) \ge r_i(N)$ for all $i = 0, \dots, M-1$ and all $N \in \mathbb{Z}_{>0}$. Define $r_{reg'}(p, d, R, C_0, \eta, r) := r_M(C_0)$.

We construct a list of polynomial factors $\mathfrak{B}_0, \mathfrak{B}_1, \dots, \mathfrak{B}_m$ on V such that

- $\mathfrak{B}_i$ refines $\mathfrak{B}_{i-1}$ for $i = 1, \dots, m$;
- $\operatorname{rank} \mathfrak{B}_i \ge r_{M-i}(\|\mathfrak{B}_i\|)$ for $i = 0, \dots, m$;
- $\|\mathfrak{B}_i\| \le C_{rank}(p, d, p^{d^3}\|\mathfrak{B}_{i-1}\|, r_{M-i})$ and $\deg \mathfrak{B}_i \le d$ for $i = 1, \dots, m$.

Suppose we have constructed polynomial factors $\mathfrak{B}_0, \dots, \mathfrak{B}_i$ with the above properties. If $\|f^{(\ell)} - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_i]\|_{U^{d+1}} < \eta$ for each $\ell \in [R]$ we halt the iteration. Otherwise there is some $\ell \in [R]$ such that, writing $g := f^{(\ell)} - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_i]$, we have

$$\|g\|_{U^{d+1}} \ge \eta.$$

By Theorem 4.1, there exists a non-classical polynomial $P\colon V \to \mathbb{R}/\mathbb{Z}$ of degree at most d such that

$$\left|\mathbb{E}_{x \in V} g(x) e^{-2\pi i P(x)}\right| \ge \epsilon_{inv}(p, d, \eta).$$

By Theorem 3.9, we can write $P = P_1 + \cdots + P_C$ as the sum of homogeneous non-classical polynomials. There are at most $\sum_{i=1}^{d} 1 + \lfloor (i-1)/(p-1) \rfloor \le d^2$ terms in this sum. Let $\mathfrak{B}'_i$ to be the polynomial factor defined by the polynomials defining $\mathfrak{B}$ as well as the polynomials $P_1, \dots, P_C$. Note that $\|\mathfrak{B}'_i\| \le p^{d^3}\|\mathfrak{B}_i\|$. Finally let $\mathfrak{B}_{i+1}$ be the polynomial factor produced by applying

Lemma 4.2 to $\mathfrak{B}_i$ and $\mathfrak{B}'_i$ with parameters p, d, r_{M-i-1}. In particular $\mathfrak{B}'_i$ refines $\mathfrak{B}_i$ and

$$\begin{aligned}\operatorname{rank}\mathfrak{B}_i &\geq r_{M-i}(\|\mathfrak{B}_i\|)\\ &\geq r_{rank}(p,d,p^{d^3}\|\mathfrak{B}_i\|,r_{M-i-1})\\ &\geq r_{rank}(p,d,\|\mathfrak{B}'_i\|,r_{M-i-1}),\end{aligned}$$

so the hypotheses of Lemma 4.2 are satisfied. Thus we have defined $\mathfrak{B}_{i+1}$ with all the desired properties.

Next we claim that this iteration must stop after at most M steps. We claim that

$$\sum_{\ell=1}^{R}\left\|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_i]\right\|_2^2$$

increases by at least $\epsilon_{inv}(p,d,\eta)^2$ each time i increases. Since this sum is clearly bounded between 0 and R, it suffices to prove this claim.

First note that by the Cauchy-Schwarz inequality, $\|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_{i+1}]\|_2^2 \geq \|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_i]\|_2^2$ holds for all ℓ. Now pick $\ell \in [R]$ as in the ith iteration, define $g := f^{(\ell)} - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_i]$, and let P be the non-classical polynomial defined in the ith iteration. Note in particular that $e^{-2\pi i P(x)}$ is in the σ-algebra defined by $\mathfrak{B}'_i$. Then we compute

$$\begin{aligned}\|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_{i+1}]\|_2^2 - \|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_i]\|_2^2 &\geq \|\mathbb{E}[f^{(\ell)}|\mathfrak{B}'_i]\|_2^2 - \|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_i]\|_2^2\\ &= \|\mathbb{E}[f^{(\ell)}|\mathfrak{B}'_i] - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_i]\|_2^2\\ &= \|\mathbb{E}[g|\mathfrak{B}'_i]\|_2^2\\ &\geq \left\langle \mathbb{E}[g|\mathfrak{B}'_i], e^{2\pi i P}\right\rangle^2\\ &= \left\langle g, e^{2\pi i P}\right\rangle^2\\ &\geq \epsilon_{inv}(p,d,\eta)^2.\end{aligned}$$

Thus we have produced $\mathfrak{B}_m$ with $m \leq M$ such that $\mathfrak{B}_m$ refines $\mathfrak{B}_0$ and $\operatorname{rank}\mathfrak{B}_m \geq r_{M-m}(\|\mathfrak{B}_m\|) \geq r(\|\mathfrak{B}_m\|)$ and $\|f^{(\ell)} - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_m]\|_{U^{d+1}} < \eta$ for each $\ell \in [R]$. Defining $f^{(\ell)}_{str} := \mathbb{E}[f^{(\ell)}|\mathfrak{B}_m]$ and $f^{(\ell)}_{psr} := f^{(\ell)} - f^{(\ell)}_{str}$, we immediately see that conclusions (i), (ii), (iii), and (iv) hold. Conclusion (v) holds by defining $C_{reg'}(p,d,R,C_0,\eta,r)$ to be the M-fold iteration of the function $N \mapsto C_{rank}(p,d,p^{d^3}N,r_M)$ applied to C_0. □

Lemma 4.4 (Arithmetic regularity)**.** *Fix a prime p, positive integers d, R, C_0, a parameter $\theta > 0$, a non-increasing function $\eta\colon \mathbb{Z}_{>0} \to (0,1)$, and a non-decreasing function $r\colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$. There exist constants $C_{reg''}(p,d,R,C_0,\theta,\eta,r)$ and $r_{reg''}(p,d,R,C_0,\theta,\eta,r)$ such that the following holds. Let V be a finite-dimensional $\mathbb{F}_p$-vector space and let $\mathfrak{B}_0$ be a polynomial factor on V satisfying $\|\mathfrak{B}_0\| \leq C_0$ and $\deg\mathfrak{B}_0 \leq d$ and $\operatorname{rank}\mathfrak{B}_0 \geq r_{reg''}(p,d,r,I_0,\theta,\eta,r)$. Given functions $f^{(1)},\dots,f^{(R)}\colon V \to [0,1]$, there exists a polynomial factor $\mathfrak{B}$ on V that refines $\mathfrak{B}_0$ with the following properties. There exists a decomposition*

$$f^{(\ell)} = f^{(\ell)}_{str} + f^{(\ell)}_{psr} + f^{(\ell)}_{sml}$$

for each $\ell \in [R]$ such that:

- *(i) $f^{(\ell)}_{str} = \mathbb{E}[f^{(\ell)}|\mathfrak{B}]$ for each $\ell \in [R]$;*
- *(ii) $\|f^{(\ell)}_{psr}\|_{U^{d+1}} < \eta(\|\mathfrak{B}\|)$ for each $\ell \in [R]$;*
- *(iii) $f^{(\ell)}_{str}$ and $f^{(\ell)}_{str} + f^{(\ell)}_{sml}$ have range $[0,1]$ and $f^{(\ell)}_{psr}$ and $f^{(\ell)}_{sml}$ have range $[-1,1]$ for each $\ell \in [R]$;*
- *(iv) $\operatorname{rank}\mathfrak{B} \geq r(\|\mathfrak{B}\|)$;*
- *(v) $\|f^{(\ell)}_{sml}\|_2 < \theta$ for each $\ell \in [R]$;*
- *(vi) $\|\mathfrak{B}\| \leq C_{reg''}(p,d,R,C_0,\theta,\eta,r)$ and $\deg\mathfrak{B} \leq d$.*

Proof. Set $M := \lceil R\theta^{-2} \rceil$. Define non-decreasing functions $r_i \colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ for each $i = 0, \dots, M$ such that $r_0 = r$ and $r_{i+1}(N) \geq r_{reg'}(p, d, R, N, \eta(N), r_i)$ and such that $r_{i+1}(N) \geq r_i(N)$ for all $i = 0, \dots, M-1$ and all $N \in \mathbb{Z}_{>0}$. Define $r_{reg''}(p, d, R, C_0, \theta, \eta, r) := r_M(C_0)$.

We construct a list of polynomial factors $\mathfrak{B}_0, \mathfrak{B}_1, \dots, \mathfrak{B}_m$ on V such that

- $\mathfrak{B}_i$ refines $\mathfrak{B}_{i-1}$ for $i = 1, \dots, m$;
- $\operatorname{rank} \mathfrak{B}_i \geq r_{M-i}(\|\mathfrak{B}_i\|)$ for $i = 0, \dots, m$;
- $\|\mathfrak{B}_i\| \leq C_{reg'}(p, d, R, \|\mathfrak{B}_{i-1}\|, \eta(\|\mathfrak{B}_{i-1}\|), r_{M-i})$ and $\deg \mathfrak{B}_i \leq d$ for $i = 1, \dots, m$.

Suppose we have constructed polynomial factors $\mathfrak{B}_0, \dots, \mathfrak{B}_i$ with the above properties. If $i \geq 1$ and $\|E[f^{(\ell)}|\mathfrak{B}_i]\|_2^2 - \|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_{i-1}]\|_2^2 < \theta^2$ for each $\ell \in [R]$ we halt the iteration. Otherwise let $\mathfrak{B}_{i+1}$ be the polynomial factor produced by applying Lemma 4.3 to $\mathfrak{B}_i$ with parameters $p, d, R, \|\mathfrak{B}_i\|, \eta(\|\mathfrak{B}_i\|), r_{M-i-1}$. Note that

$$\begin{aligned} \operatorname{rank} \mathfrak{B}_i &\geq r_{M-i}(\|\mathfrak{B}_i\|) \\ &\geq r_{reg'}(p, d, R, \|\mathfrak{B}_i\|, \eta(\|\mathfrak{B}_i\|), r_{M-i-1}), \end{aligned}$$

so the hypotheses of Lemma 4.3 are satisfied.

Next we claim that this iteration must stop after at most M steps. This is obvious since

$$\sum_{\ell=1}^{R} \left\| \mathbb{E}[f^{(\ell)}|\mathfrak{B}_i] \right\|_2^2 - \sum_{\ell=1}^{R} \left\| \mathbb{E}[f^{(\ell)}|\mathfrak{B}_{i-1}] \right\|_2^2 \geq \theta^2$$

for $i = 1, \dots, m$ and the sum is bounded between 0 and R.

Thus we have produced $\mathfrak{B}_{m-1}$ with $m \leq M$ such that $\mathfrak{B}_{m-1}$ refines $\mathfrak{B}_0$ and $\operatorname{rank} \mathfrak{B}_{m-1} \geq r_{M-m+1}(\|\mathfrak{B}_{m-1}\|) \geq r(\|\mathfrak{B}_{m-1}\|)$ and $\|f^{(\ell)} - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_m]\|_{U^{d+1}} < \eta(\|\mathfrak{B}_{m-1}\|)$ for each $\ell \in [R]$ and $\left\|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_m] - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_{m-1}]\right\|_2 < \theta$ for each $\ell \in [R]$. Defining $f_{str}^{(\ell)} := \mathbb{E}[f^{(\ell)}|\mathfrak{B}_{m-1}]$ and $f_{psr}^{(\ell)} := f^{(\ell)} - E[f^{(\ell)}|\mathfrak{B}_m]$ and $f_{sml}^{(\ell)} := \mathbb{E}[f^{(\ell)}|\mathfrak{B}_m] - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_{m-1}]$, we immediately see that conclusions (i), (ii), (iii), (iv), and (v) hold. Conclusion (vi) holds by defining $C_{reg''}(p, d, R, C_0, \theta, \eta, r)$ to be the M-fold iteration of the function $N \mapsto C_{reg'}(p, d, R, N, \eta(N), r_M)$ applied to C_0. □

Lemma 4.5 (Strong arithmetic regularity)**.** *Fix a prime p, positive integers R, C_0, d_0, a parameter $\zeta > 0$, non-increasing functions $\eta, \theta \colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to (0,1)$, and non-decreasing functions $d, r \colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$. There exist constants $C_{reg'''}(p, R, C_0, d_0, \zeta, \eta, \theta, d, r)$ and $D_{reg'''}(p, R, C_0, d_0, \zeta, \eta, \theta, d, r)$ and $r_{reg'''}(p, R, C_0, d_0, \zeta, \eta, \theta, d, r)$ such that the following holds. Let V be a finite-dimensional $\mathbb{F}_p$-vector space and let $\mathfrak{B}_0$ be a polynomial factor on V satisfying $\|\mathfrak{B}_0\| \leq C_0$ and $\deg \mathfrak{B}_0 \leq d_0$ and $\operatorname{rank} \mathfrak{B}_0 \geq r_{reg'''}(p, R, C_0, d_0, \zeta, \eta, \theta, d, r)$. Given functions $f^{(1)}, \dots, f^{(R)} \colon V \to [0,1]$, there exist a polynomial factor $\mathfrak{B}$ and a refinement $\mathfrak{B}'$ both on V with parameters I and I' with the following properties. There exists a decomposition*

$$f^{(\ell)} = f_{str}^{(\ell)} + f_{psr}^{(\ell)} + f_{sml}^{(\ell)}$$

for each $\ell \in [R]$ such that:

(i) $f_{str}^{(\ell)} = \mathbb{E}[f^{(\ell)}|\mathfrak{B}']$ *for each $\ell \in [R]$;*
(ii) $\|f_{psr}^{(\ell)}\|_{U^{d(\deg \mathfrak{B}, \|\mathfrak{B}\|)+1}} < \eta(\deg \mathfrak{B}', \|\mathfrak{B}'\|)$ *for each $\ell \in [R]$;*
(iii) $f_{str}^{(\ell)}$ *and* $f_{str}^{(\ell)} + f_{sml}^{(\ell)}$ *have range* $[0,1]$ *and* $f_{psr}^{(\ell)}$ *and* $f_{sml}^{(\ell)}$ *have range* $[-1,1]$ *for each $\ell \in [R]$;*
(iv) $\operatorname{rank} \mathfrak{B} \geq r(\deg \mathfrak{B}, \|\mathfrak{B}\|)$ *and* $\operatorname{rank} \mathfrak{B}' \geq r(\deg \mathfrak{B}', \|\mathfrak{B}'\|)$*;*
(v) $\|f_{sml}^{(\ell)}\|_2 < \theta(\deg \mathfrak{B}, \|\mathfrak{B}\|)$ *for each $\ell \in [R]$;*
(vi) for all but at most a ζ-fraction of $a \in A_{I'}$ it holds that

$$\left| \mathbb{E}_{x \in \mathfrak{B}'^{-1}(a)}[f^{(\ell)}(x)] - \mathbb{E}_{x \in \mathfrak{B}^{-1}(\pi(a))}[f^{(\ell)}(x)] \right| < \zeta$$

for each $\ell \in [R]$ (recall the definition of the atom indexing sets from Eq. (3.3) *and the projection map $\pi \colon A_{I'} \to A_I$ from Eq.* (3.4)*);*

(vii) $\|\mathfrak{B}'\| \le C_{reg'''}(p,R,C_0,d_0,\zeta,\eta,\theta,d,r)$ *and* $\deg \mathfrak{B}' \le D_{reg'''}(p,R,C_0,d_0,\zeta,\eta,\theta,d,r)$.

Proof. Set $M := \lceil R\zeta^{-3} \rceil$. Define non-decreasing functions $r_i \colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ for each $i = 0,\dots,M$ such that $r_0 = r$ and $r_{i+1}(D,N) \ge r_{reg''}(p, d(D,N), R, N, \theta(D,N), \eta(d(D,N),\cdot), r_i(d(D,N),\cdot))$ and such that $r_{i+1}(D,N) \ge r_i(D,N)$ for all $i = 0,\dots,M-1$ and all $D,N \in \mathbb{Z}_{>0}$. Define $r_{reg'''}(p,R,C_0,d_0,\zeta,\eta,\theta,d,r) := r_M(d_0,C_0)$.

We construct a list of polynomial factors $\mathfrak{B}_0,\mathfrak{B}_1,\dots,\mathfrak{B}_m$ on V such that

- $\mathfrak{B}_i$ refines $\mathfrak{B}_{i-1}$ for $i = 1,\dots,m$;
- $\operatorname{rank}\mathfrak{B}_i \ge r_{M-i}(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|)$ for $i = 0,\dots,m$;
- $\|\mathfrak{B}_i\| \le C_{reg''}(p, d(\deg\mathfrak{B}_{i-1}, \|\mathfrak{B}_{i-1}\|), R, \|\mathfrak{B}_{i-1}\|, \theta, \eta, r_{M-i}(d(\deg\mathfrak{B}_{i-1}, \|\mathfrak{B}_{i-1}\|),\cdot))$ and $\deg\mathfrak{B}_i \le d(\deg\mathfrak{B}_{i-1}, \|\mathfrak{B}_{i-1}\|)$ for $i = 1,\dots,m$.

Suppose we have constructed polynomial factors $\mathfrak{B}_0,\dots,\mathfrak{B}_i$ with the above properties. If $i \ge 1$ and $\|E[f^{(\ell)}|\mathfrak{B}_i]\|_2^2 - \|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_{i-1}]\|_2^2 < \zeta^3$ for each $\ell \in [R]$ we halt the iteration. Otherwise let $\mathfrak{B}_{i+1}$ be the polynomial factor produced by applying Lemma 4.4 to $\mathfrak{B}_i$ with parameters $p, d(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|), R, \|\mathfrak{B}_i\|, \theta(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|), \eta(d(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|),\cdot), r_{M-i-1}(d(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|),\cdot)$. Note that

$$\operatorname{rank}\mathfrak{B}_i \ge r_{M-i}(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|) \ge r_{reg''}(p, d(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|), R, \|\mathfrak{B}_i\|, \theta(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|), \eta(d(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|),\cdot), r_{M-i-1}(d(\deg\mathfrak{B}_i, \|\mathfrak{B}_i\|),\cdot)),$$

so the hypotheses of Lemma 4.4 are satisfied.

Next we claim that this iteration must stop after at most M steps. This is obvious since

$$\sum_{\ell=1}^{R} \left\|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_i]\right\|_2^2 - \sum_{\ell=1}^{R} \left\|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_{i-1}]\right\|_2^2 \ge \zeta^3$$

for $i = 1,\dots,m$ and the sum is bounded between 0 and R.

Thus we have produced $\mathfrak{B}_{m-1}$ and $\mathfrak{B}_m$ with $m \le M$ such that $\mathfrak{B}_{m-1}$ refines $\mathfrak{B}_0$ and $\mathfrak{B}_m$ refines $\mathfrak{B}_{m-1}$. Furthermore, $\operatorname{rank}\mathfrak{B}_{m-1} \ge r_{M-m+1}(\deg\mathfrak{B}_{m-1}, \|\mathfrak{B}_{m-1}\|) \ge r(\deg\mathfrak{B}_{m-1}, \|\mathfrak{B}_{m-1}\|)$ and $\operatorname{rank}\mathfrak{B}_m \ge r_{M-m}(\deg\mathfrak{B}_m, \|\mathfrak{B}_m\|) \ge r(\deg\mathfrak{B}_m, \|\mathfrak{B}_m\|)$. Also $\left\|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_m] - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_{m-1}]\right\|_2^2 < \zeta^3$ for each $\ell \in [R]$. Let $f^{(\ell)} = f^{(\ell)}_{str} + f^{(\ell)}_{psr} + f^{(\ell)}_{sml}$ be the decomposition produced by the last application of Lemma 4.4. This decomposition satisfies conclusions (i), (ii), (iii), and (v). Conclusion (iv) we already verified, and conclusion (vi) follows from Markov's inequality applied to the bound $\left\|\mathbb{E}[f^{(\ell)}|\mathfrak{B}_m] - \mathbb{E}[f^{(\ell)}|\mathfrak{B}_{m-1}]\right\|_2^2 < \zeta^3$. Finally conclusion (vii) holds where we define the pair $(D_{reg'''}(p,R,C_0,d_0,\zeta,\eta,\theta,d,r), C_{reg'''}(p,R,C_0,d_0,\zeta,\eta,\theta,d,r))$ to be the M-fold iteration of the function $(D,N) \mapsto (C_{reg''}(p,d(D,N),R,N,\theta,\eta,r_M(d(D,N),\cdot)), d(D,N))$ applied to (d_0,C_0). □

Recall that for a parameter list $I \in \mathcal{I}_p$ and an atom $a \in A_I$, we write $a_{d,k} \in \mathbb{U}_{k+1}^{I_{d,k}}$ to be the degree d, depth k part of a. In the next theorem we will choose a subatom selection function $s \colon A_I \to A_{I'}$ (recall Definition 3.24) such that $s(a)$ is regular for all $a \in A_I$ except those with $a_{1,0} = 0$.

Theorem 4.6 (Subatom selection)**.** *Fix a prime p, positive integers R, c_0, a parameter $\zeta > 0$, non-increasing functions $\eta,\theta \colon \mathbb{Z}_{>0}\times\mathbb{Z}_{>0} \to (0,1)$, and non-decreasing functions $d,r \colon \mathbb{Z}_{>0}\times\mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$. There exist constants $C_{reg}(p,R,c_0,\zeta,\eta,\theta,d,r)$ and $D_{reg}(p,R,c_0,\zeta,\eta,\theta,d,r)$ and $n_{reg}(p,c_0,\zeta)$ such that the following holds. Let V be a finite-dimensional $\mathbb{F}_p$-vector space satisfying $\dim V \ge n_{reg}(p,c_0,\zeta)$. Given functions $f^{(1)},\dots,f^{(R)} \colon V \to [0,1]$, there exist a polynomial factor $\mathfrak{B}$ and a refinement $\mathfrak{B}'$ both on V with parameters I and I' with the following properties. There exists a subatom selection function $s \colon A_I \to A_{I'}$ and a decomposition*

$$f^{(\ell)} = f^{(\ell)}_{str} + f^{(\ell)}_{psr} + f^{(\ell)}_{sml}$$

for each $\ell \in [R]$ such that:

(i) $f_{str}^{(\ell)} = \mathbb{E}[f^{(\ell)}|\mathfrak{B}']$ *for each* $\ell \in [R]$*;*
(ii) $\|f_{psr}^{(\ell)}\|_{U^{d(\deg \mathfrak{B}, \|\mathfrak{B}\|)+1}} < \eta(\deg \mathfrak{B}', \|\mathfrak{B}'\|)$ *for each* $\ell \in [R]$*;*
(iii) $f_{str}^{(\ell)}$ *and* $f_{str}^{(\ell)} + f_{sml}^{(\ell)}$ *have range* $[0,1]$ *and* $f_{psr}^{(\ell)}$ *and* $f_{sml}^{(\ell)}$ *have range* $[-1,1]$ *for each* $\ell \in [R]$*;*
(iv) $\operatorname{rank} \mathfrak{B} \ge r(\deg \mathfrak{B}, \|\mathfrak{B}\|)$ *and* $\operatorname{rank} \mathfrak{B}' \ge r(\deg \mathfrak{B}', \|\mathfrak{B}'\|)$*;*
(v) for each $a \in A_I$ *with* $a_{1,0} \neq 0$*, it holds that*

$$\|f_{sml}^{(\ell)} 1_{\mathfrak{B}'^{-1}(s(a))}\|_2 < \theta(\deg \mathfrak{B}, \|\mathfrak{B}\|)\|1_{\mathfrak{B}'^{-1}(s(a))}\|_2$$

for each $\ell \in [R]$*;*
(vi) for all but at most a ζ*-fraction of* $a \in A_I$ *it holds that*

$$\left|\mathbb{E}_{x \in \mathfrak{B}^{-1}(a)}[f^{(\ell)}(x)] - \mathbb{E}_{x \in \mathfrak{B}'^{-1}(s(a))}[f^{(\ell)}(x)]\right| < \zeta$$

for each $\ell \in [R]$*;*
(vii) $I_{1,0} \ge c_0$*;*
(viii) $\|\mathfrak{B}'\| \le C_{reg}(p, R, c_0, \zeta, \eta, \theta, d, r)$ *and* $\deg \mathfrak{B}' \le D_{reg}(p, R, c_0, \zeta, \eta, \theta, d, r)$*.*

Proof. Define $n_{reg}(p, c_0, \zeta) := \max\{c_0, \lceil \log_p(2/\zeta) \rceil\}$. Let $\mathfrak{B}_0$ be a polynomial factor on V defined by $n_{reg}(p, c_0, \zeta)$ linearly independent linear functions. Define $\theta' \colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ by $\theta'(D, N) := \theta(D, N)/(2\sqrt{R}N)$. We apply Lemma 4.5 to $\mathfrak{B}_0$ with parameters $p, R, \|\mathfrak{B}_0\|, 1, \zeta/4, \eta, \theta', d, r$. Let $\mathfrak{B}$ and $\mathfrak{B}'$ be the polynomial factors produced. We immediately have conclusions (i), (ii), (iii), and (iv). Conclusion (vii) follows since $\mathfrak{B}$ refines $\mathfrak{B}_0$ which is defined by at least c_0 linear forms. Conclusion (viii) holds by defining $C_{reg}(p, R, c_0, \zeta, \eta, \theta, d, r) = C_{reg'''}(p, R, n_{reg}(p, c_0, \zeta), 1, \zeta/4, \eta, \theta', d, r)$ and $D_{reg}(p, R, c_0, \zeta, \eta, \theta, d, r) = D_{reg'''}(p, R, p^{n_{reg}(p, c_0, \zeta)}, 1, \zeta/4, \eta, \theta', d, r)$.

Let $c_{d,k}^{i,j} \in \mathbb{Z}/p^{k+1}\mathbb{Z}$ be elements chosen independently and uniformly at random for each $(d, k) \in D_p$ and $i \in [I_{1,0}]$ and $I_{d,k} < j \le I'_{d,k}$. Consider the subatom selection function $s_{\boldsymbol{c}} \colon A_I \to A_{I'}$ defined in Definition 3.24. We claim that with positive probability, this $s_{\boldsymbol{c}}$ satisfies conclusions (v) and (vi).

Fix $a \in A_I$ with $a_{1,0} \neq 0$. We first claim that for this fixed a, as $\boldsymbol{c}$ varies, the subatom $s(a)$ is uniformly distributed over $\pi^{-1}(a) \subset A_{I'}$. To see this, first note that the univariate homogeneous non-classical polynomials $P_{d,k} \colon \mathbb{F}_p \to \mathbb{U}_{k+1}$ used in the definition of $s_{\boldsymbol{c}}$ satisfy $P_{d,k}(x) \notin \mathbb{U}_k$ for all $x \neq 0$. This follows by homogeneity: if $P_{d,k}(x) \in U_k$ for some $x \neq 0$, then $P_{d,k}$ always takes values in $\mathbb{U}_k$, contradicting the assumption that $P_{d,k}$ has depth exactly k. Thus for $a \in A_I$ with $a_{1,0} \neq 0$, we find that the vector $(P_{d,k}(|a_{1,0}^1|), \dots, P_{d,k}(|a_{1,0}^{I_{1,0}}|)) \in \mathbb{U}_{k+1}^{I_{1,0}}$ does not lie in $\mathbb{U}_k^{I_{1,0}}$.

Considering the definition of $s_{\boldsymbol{c}}$, we see that the vector $([s_{\boldsymbol{c}}]_{d,k}^i)_{I_{d,k} < i \le I'_{d,k}}$ is produced by the following matrix multiplication.

$$\begin{pmatrix} [s_{\boldsymbol{c}}]_{d,k}^{I_{d,k}+1} \\ [s_{\boldsymbol{c}}]_{d,k}^{I_{d,k}+2} \\ \vdots \\ [s_{\boldsymbol{c}}]_{d,k}^{I'_{d,k}} \end{pmatrix} = \begin{pmatrix} c_{d,k}^{I_{d,k}+1,1} & c_{d,k}^{I_{d,k}+1,2} & \cdots & c_{d,k}^{I_{d,k}+1,I_{d,k}} \\ c_{d,k}^{I_{d,k}+2,1} & c_{d,k}^{I_{d,k}+2,2} & \cdots & c_{d,k}^{I_{d,k}+2,I_{d,k}} \\ \vdots & \vdots & \ddots & \vdots \\ c_{d,k}^{I'_{d,k},1} & c_{d,k}^{I'_{d,k},2} & \cdots & c_{d,k}^{I'_{d,k},I_{d,k}} \end{pmatrix} \begin{pmatrix} P_{d,k}(|a_{1,0}^1|) \\ P_{d,k}(|a_{1,0}^2|) \\ \vdots \\ P_{d,k}(|a_{1,0}^{I_{d,k}}|) \end{pmatrix}$$

As stated above, the vector on the right lies in $\mathbb{U}_{k+1}^{I_{1,0}}$ but not in $\mathbb{U}_k^{I_{1,0}}$ while the matrix is uniform random with entries in $\mathbb{Z}/p^{k+1}\mathbb{Z}$. This implies that the vector on the left is uniformly distributed over $\mathbb{U}_{k+1}^{I'_{d,k}-I_{d,k}}$, as desired. This holds for each $(d, k) \in D_p$ and furthermore, since for each (d, k) the c matrices are chosen independently, one can see that the resulting vectors are also independent, proving the desired result.

Now note that by Lemma 4.5(vi), for each $\ell \in [R]$,

$$\sum_{a \in A_{I'}} \|f_{sml}^{(\ell)} 1_{\mathfrak{B}'^{-1}(a)}\|_2^2 \le \frac{\theta(\deg \mathfrak{B}, \|\mathfrak{B}\|)^2}{4R\|\mathfrak{B}\|^2} \sum_{a \in A_{I'}} \|1_{\mathfrak{B}'^{-1}(a)}\|_2^2.$$

Thus by Markov's inequality, for $a' \in A_{I'}$ chosen uniformly at random, with probability at least $1 - 1/(4R\|\mathfrak{B}\|^2)$ the following holds

$$\|f^{(\ell)}_{sml} 1_{\mathfrak{B}'^{-1}(a')}\|_2^2 \le \theta(\|\mathfrak{B}\|)^2 \|1_{\mathfrak{B}'^{-1}(a')}\|_2^2. \tag{4.1}$$

Thus for a fixed $a \in A_I$ with $a_{1,0} \neq 0$, we have that for $\boldsymbol{c}$ chosen at random, we have $s_{\boldsymbol{c}}(a)$ satisfies Eq. (4.1) with probability at least $1 - 1/(4R\|\mathfrak{B}\|)$. Taking a union bound over all choices of $a \in A_I$ and $\ell \in [R]$, we see that with probability at least $3/4$, conclusion (v) holds.

Now we deduce conclusion (vi). First note that the fraction of $a \in A_I$ which satisfy $a_{1,0} = 0$ is $p^{-I_{1,0}} \le \zeta/2$. For the other $a \in A_I$, the expected fraction of $a \in A_I$ that fail to satisfy the desired inequality is at most $\zeta/4$. Thus by Markov's inequality, with probability at least $1/2$, at most $\zeta/2$ fraction of $a \in A_I$ satisfy $a_{1,0} \neq 0$ and fail to satisfy the desired inequality. Thus with probability at least $1/2$, conclusion (vi) holds. □

5. Patching

To motivate the kind of results proved in this section consider the following result, which follows from an application of Ramsey's theorem.

> Let $\mathcal{H}$ be a finite set of red/blue edge-colored graphs. There exists an integer $n_0 = n_0(\mathcal{H})$ such that the following holds. Either:
> (a) either the all-red coloring of K_n or the all-blue coloring of K_n contains no subgraph from $\mathcal{H}$ for every n; or
> (b) every 2-edge-coloring of K_n with $n \ge n_0$ contains a subgraph from $\mathcal{H}$.

We call such a statement a dichotomy result. The first main result of this section, Theorem 5.6, is a dichotomy result for our setting. In our setting we consider colored labeled patterns instead of edge-colored subgraphs and instead of monochromatic colorings we have to consider so-called canonical colorings, defined below.

The second main result of this section, Theorem 5.9, is our patching result, which is a supersaturation version of the dichotomy result.

Definition 5.1. For a prime p, a finite set $\mathcal{S}$, and a parameter list $I \in \mathcal{I}_p$, an **$\mathcal{S}$-colored I-labeled pattern** over $\mathbb{F}_p$ consisting of m linear forms in ℓ variables is a triple $(\boldsymbol{L}, \psi, \phi)$ given by:

- a system $\boldsymbol{L} = (L_1, \dots, L_m)$ of m linear forms in ℓ variables,
- a coloring $\psi\colon [m] \to \mathcal{S}$, and
- a labeling $\phi\colon [m] \to A_I$ (recall the definition of the atom-indexing set A_I from Eq. (3.3)).

Given a finite-dimensional $\mathbb{F}_p$-vector space V, a function $f\colon V \to \mathcal{S}$, and a polynomial factor $\mathfrak{B}$ on V with parameters I, an **$(\boldsymbol{L}, \psi, \phi)$-instance** in $(f, \mathfrak{B})$ is some $\boldsymbol{x} \in V^\ell$ such that $f(L_i(\boldsymbol{x})) = \psi(i)$ for all $i \in [m]$ and $\mathfrak{B}(L_i(\boldsymbol{x})) = \phi(i)$ for all $i \in [m]$. An instance is called **generic** if $x_1, \dots, x_\ell$ are linearly independent.

Definition 5.2. For an $\mathcal{S}$-colored I-labeled pattern $(\boldsymbol{L}, \psi, \phi)$ consisting of m linear forms, a finite dimensional $\mathbb{F}_p$-vector space V, a function $f\colon V \to \mathcal{S}$, and a polynomial factor $\mathfrak{B}$ on V with parameters I, define the **$(\boldsymbol{L}, \psi, \phi)$-density** in $(f, \mathfrak{B})$ to be

$$\Lambda_{\boldsymbol{L}}(1_{f^{-1}(\psi(1)) \cap \mathfrak{B}^{-1}(\phi(1))}, \dots, 1_{f^{-1}(\psi(m)) \cap \mathfrak{B}^{-1}(\phi(m))}).$$

Given a set $X \subseteq V$, define the **relative density of $(\boldsymbol{L}, \psi, \phi)$ in X** to be

$$\frac{\Lambda_{\boldsymbol{L}}(f_1, \dots, f_m)}{\Lambda_{\boldsymbol{L}}(1_X, \dots, 1_X)}$$

where $f_i := 1_{X \cap f^{-1}(\psi(i)) \cap \mathfrak{B}^{-1}(\phi(i))}$.

Definition 5.3. Define the **first non-zero coordinate** function $\operatorname{fnz}\colon \mathbb{F}_p^n \to \mathbb{F}_p$ by $\operatorname{fnz}(0,\dots,0) := 0$ and $\operatorname{fnz}(x_1,\dots,x_n) := x_k$ where $x_1 = \cdots = x_{k-1} = 0$ and $x_k \neq 0$. Given a finite-dimensional $\mathbb{F}_p$-vector space V equipped with an isomorphism $\iota\colon V \xrightarrow{\sim} \mathbb{F}_p^n$, define the function $\operatorname{fnz}_\iota\colon V \to \mathbb{F}_p$ by $\operatorname{fnz}_\iota(x) := \operatorname{fnz}(\iota(x))$.

Definition 5.4. Fix a prime p, a finite set $\mathcal{S}$, a parameter list $I \in \mathcal{I}_p$, and a function $\xi\colon \mathbb{F}_p \times A_I \to \mathcal{S}$. For a finite-dimensional $\mathbb{F}_p$-vector space V equipped with an isomorphism $\iota\colon V \xrightarrow{\sim} \mathbb{F}_p^n$ and a polynomial factor $\mathfrak{B}$ on V with parameters I, define the **ξ-canonical coloring** $\Xi_{\xi,\iota,\mathfrak{B}}\colon V \to \mathcal{S}$ by $\Xi_{\xi,\iota,\mathfrak{B}}(x) := \xi(\operatorname{fnz}_\iota(x), \mathfrak{B}(x))$. Furthermore, if $\mathcal{S}$ is equipped with an $\mathbb{F}_p^\times$-action, say that ξ is **projective** if the same is true for every function $\Xi_{\xi,\iota,\mathfrak{B}}$. (Note that this property is equivalent to the condition that ξ preserves the action of $\mathbb{F}_p^\times$, i.e., $\xi(cx, c\cdot a) = c\cdot\xi(x,a)$ for all $c \in \mathbb{F}_p^\times$, all $x \in \mathbb{F}_p$, and all $a \in A_I$. Recall the action of $\mathbb{F}_p^\times$ on A_I defined in Eq. (3.5).)

Definition 5.5. Given a prime p, a finite set $\mathcal{S}$, a parameter list $I \in \mathcal{I}_p$, a function $\xi\colon \mathbb{F}_p \times A_I \to \mathcal{S}$, and a $\mathcal{S}$-colored I-labeled pattern $H = (\boldsymbol{L}, \psi, \phi)$, say that ξ **canonically induces** H if the following holds. There exists some $n \geq 0$ and a polynomial factor $\mathfrak{B}$ on $\mathbb{F}_p^n$ with parameters I such that there exists a generic H-instance in $(\Xi_{\xi,\mathrm{Id},\mathfrak{B}}, \mathfrak{B})$. For a finite set of $\mathcal{S}$-colored I-labeled patterns $\mathcal{H}$, say that ξ **canonically induces** $\mathcal{H}$ if ξ canonically induces some $H \in \mathcal{H}$.

It is not hard to show that if ξ canonically induces H, then there exists a generic H-instance in $(\Xi_{\xi,\iota,\mathfrak{B}}, \mathfrak{B})$ for every $V, \iota, \mathfrak{B}$ as long as $\dim V$ and $\operatorname{rank}\mathfrak{B}$ are large enough. Our first result is a strengthening of this: if every ξ canonically induces H, then there exists a generic H-instance in $(f, \mathfrak{B})$ for every $f\colon V \to \mathcal{S}$ and every $\mathfrak{B}$ as long as $\dim V$ and $\operatorname{rank}\mathfrak{B}$ are large enough.

Theorem 5.6 (Dichotomy)**.** *Fix a prime p, a finite set $\mathcal{S}$ with an $\mathbb{F}_p^\times$-action, a parameter list $I \in \mathcal{I}_p$, and a positive integer ℓ_0. There exist constants $n_{dich} = n_{dich}(p, |\mathcal{S}|, I, \ell_0)$ and $r_{dich} = r_{dich}(p, |\mathcal{S}|, I, \ell_0)$ such that the following holds. Let $\mathcal{H}$ be a finite set of $\mathcal{S}$-colored, I-labeled patterns each defined by a system of linear forms in at most ℓ_0 variables. Either:*

(a) there exists a projective $\xi\colon \mathbb{F}_p \times A_I \to \mathcal{S}$ that does not canonically induce $\mathcal{H}$; or
(b) for every finite-dimensional $\mathbb{F}_p$-vector space V satisfying $\dim V \geq n_{dich}$, every projective function $f\colon V \to \mathcal{S}$, and every polynomial factor $\mathfrak{B}$ on V with parameters I which has rank at least r_{dich}, there is a generic H-instance in $(f, \mathfrak{B})$ for some $H \in \mathcal{H}$.

Proof. Define m to be the smallest positive integer such that the following holds. Let H be a $\mathcal{S}$-colored, I-labeled patterns defined by a system of linear forms in at most ℓ_0 variables and let $\xi\colon \mathbb{F}_p \times A_I \to \mathcal{S}$ be a projective function. If ξ canonically induces H, then there exists some $n_H \leq m$ and a polynomial factor $\mathfrak{B}_H$ on $\mathbb{F}_p^{n_H}$ with parameters I such that there exists a generic H-instance in $(\Xi_{\xi,\mathrm{Id},\mathfrak{B}_H}, \mathfrak{B}_H)$. This is well defined since there are only a finite number of $\mathcal{S}$-colored, I-labeled patterns defined by a system of linear forms in at most ℓ_0 variables.

Lemma 5.7. *Fix a prime p, a finite set $\mathcal{S}$, a parameter list $I \in \mathcal{I}_p$, and positive integers m, r_0, n_0. There exist constants $n_{ramsey} = n_{ramsey}(p, |\mathcal{S}|, I, m, r_0, n_0)$ and $r_{ramsey} = r_{ramsey}(p, |\mathcal{S}|, I, m, r_0, n_0)$ such that the following holds. Let V be a finite dimensional $\mathbb{F}_p$-vector space satisfying $\dim V \geq n_{ramsey}$, let $\mathfrak{B}$ be a polynomial factor on $\mathbb{F}_p^n$ with parameters I such $\operatorname{rank}\mathfrak{B} \geq r_{ramsey}$, and let $f\colon V \to \mathcal{S}$ be a function. Then there exists a subspace $U \leq V$, and linear functions $P_1,\dots,P_m\colon V \to \mathbb{F}_p$, and a function $\xi\colon \mathbb{F}_p \times A_I \to \mathcal{S}$ such that the following holds:*

(i) $\xi(\operatorname{fnz}(P_1(x),\dots,P_m(x)), \mathfrak{B}(x)) = f(x)$ for all $x \in U$ that also satisfy $(P_1(x),\dots,P_m(x)) \neq (0,\dots,0)$;
(ii) $\xi(0,0) = f(0)$
(iii) the polynomial factor $\mathfrak{B}'$ on U defined by the homogeneous non-classical polynomials that define $\mathfrak{B}|_U$ in addition to the polynomials $P_1,\dots,P_m$ satisfies $\operatorname{rank}\mathfrak{B}' \geq r_0$;
(iv) $\dim U \geq n_0$.

Let us show how this lemma completes the proof. Define

$$r_0 := r_{equi}\left(p, \deg I, \frac{1}{2(p^m\|I\|)^{p^m}}\right) \qquad \text{and} \qquad n_0 := 2p^m(m + \lceil \log_p \|I\| \rceil).$$

Then define

$$n_{dich}(p, |\mathcal{S}|, I, \ell_0) := n_{ramsey}(p, |\mathcal{S}|, I, m, r_0, n_0),$$

and

$$r_{dich}(p, |\mathcal{S}|, I, \ell_0) := r_{ramsey}(p, |\mathcal{S}|, I, m, r_0, n_0).$$

Let $\mathcal{H}$ be a finite set of $\mathcal{S}$-colored, I-labeled patterns each defined by a system of linear forms in at most ℓ_0 variables. Suppose (a) does not hold. Thus for every projective $\xi\colon \mathbb{F}_p \times A_I \to \mathcal{S}$, there exists an $H \in \mathcal{H}$ such that ξ canonically induces H.

Now we apply Lemma 5.7 to $f\colon V \to \mathcal{S}$. This produces a subspace $U \leq V$, linear functions $P_1, \dots, P_m$, and a function $\xi\colon \mathbb{F}_p \times A_I \to \mathcal{S}$ with several desirable properties.

First note that since $f\colon V \to \mathcal{S}$ is projective, the same is true of $\xi\colon \mathbb{F}_p \times A_I \to \mathcal{S}$. Thus by assumption there exists an $H \in \mathcal{H}$ such that ξ canonically induces H. By the choices in the first paragraph, there exists a $n_H \leq m$ and a polynomial factor $\mathfrak{B}_H$ on $\mathbb{F}_p^{n_H}$ such that $(\Xi_{\xi,\mathrm{Id},\mathfrak{B}_H}, \mathfrak{B}_H)$ contains a generic H-instance.

To complete the proof, all we need to show is that there exists an injective linear map $\kappa\colon \mathbb{F}_p^{n_H} \to U$ such that $\mathfrak{B}_H(x) = \mathfrak{B}(\kappa(x))$ for all $x \in \mathbb{F}_p^{n_H}$ and $\mathrm{fnz}(x) = \mathrm{fnz}(P_1(\kappa(x)), \dots, P_m(\kappa(x)))$ for all $x \in \mathbb{F}_p^{n_H}$. This follows by an application of equidistribution (Theorem 3.19).

Recall the definition of $\boldsymbol{L}^{n_H}$ (Definition 2.7), the system of p^{n_H} linear forms in n_H variables that define an n_H-dimensional subspace.

Say that $\mathfrak{B}'$ has parameters I'. Thus the atom-indexing set of $\mathfrak{B}'$ can be written as $A_{I'} \simeq \mathbb{F}_p^m \times A_I$. We define the following tuple of atoms $\boldsymbol{a} = (a_{\boldsymbol{i}})_{\boldsymbol{i} \in \mathbb{F}_p^{n_H}}$ by $a_{\boldsymbol{i}} = ((i_1, \dots, i_{n_H}, 0, \dots, 0), \mathfrak{B}_H(\boldsymbol{i}))$ where there are $m - n_H$ 0's. We claim that $\boldsymbol{a} \in \Phi_{I'}(\boldsymbol{L}^{n_H})$. We can check this separately for the first and second coordinate; each is trivial.

Thus by Theorem 3.19 and the rank bound on $\mathfrak{B}'$, we have

$$\Pr_{x_1,\dots,x_{n_H} \in U}\left(\mathfrak{B}'(L_{\boldsymbol{i}}^{n_H}(x_1, \dots, x_{n_H})) = a_{\boldsymbol{i}} \text{ for all } \boldsymbol{i} \in \mathbb{F}_p^{n_H}\right) \geq \frac{1}{|\Phi_{I'}(\boldsymbol{L}^{n_H})|} - \frac{1}{2(p^m\|I\|)^{p^m}} \geq \frac{1}{2(p^m\|I\|)^{p^m}}.$$

We wish to find a single tuple $(x_1, \dots, x_{n_H}) \in V^{n_H}$ that satisfies the above condition and also has $x_1, \dots, x_{n_H}$ linearly independent. The number of linearly dependent tuples is small, so we calculate that the number of good tuples is at least

$$\frac{|U|^{n_H}}{2(p^m\|I\|)^{p^m}} - |U|^{n_H - 1}p^{n_H}.$$

This is positive by our assumption that $\dim U \geq n_0$. Thus there exists some good tuple $(x_1, \dots, x_{n_H}) \in V^{n_H}$. Defining $\kappa\colon \mathbb{F}_p^{n_H} \to U$ by $\kappa(\boldsymbol{i}) := L_{\boldsymbol{i}}^{n_H}(x_1, \dots, x_{n_H})$ has all the desired properties. Thus we have shown (b) assuming that (a) does not hold. □

Proof of Lemma 5.7. Define $M := m|\mathcal{S}|^{p\|I\|}$. Our strategy is to find a large subspace U_M and linear functions $P_1, \dots, P_M\colon V \to \mathbb{F}_p$ such that for $x \in U_M$, the value of $f(x)$ only depends on $\mathfrak{B}(x)$, $\mathrm{fnz}(P_1(x), \dots, P_M(x))$, and the index k such that $P_1(x) = \cdots = P_{k-1}(x) = 0$ and $P_k(x) \neq 0$. Once we have found such a configuration, we can complete the proof by a simple pigeonhole argument.

Define

$$\mathcal{L} := \{p^s P^i_{(p-1)s+1,s} : s \geq 0, i \in [I_{(p-1)s+1,s}]\},$$

where $P^i_{d,k}$ is the ith homogeneous non-classical polynomial of degree d and depth k defining $\mathfrak{B}$. Note that $\mathcal{L}$ is a finite set of linear functions (in particular, it is the set of all $p^s P$ that are classical linear polynomials where $s \geq 0$ is a non-negative integer and P is one of the homogeneous non-classical polynomials that define $\mathfrak{B}$.) It is immediate from the definition of rank that if $P_1, \dots, P_m$

are linear functions such that $\{P_1, \dots, P_m\} \cup \mathcal{L}$ are linearly independent, then $\operatorname{rank} \mathfrak{B}' = \operatorname{rank} \mathfrak{B}$. We will use this fact to guarantee conclusion (iii).

Our main tool is the following lemma which is a Van der Waerden-type result.

Lemma 5.8. *Fix a prime p, a finite set $\mathcal{S}$, a parameter list $I \in \mathcal{I}_p$, and positive integers n_0, r_0. There exist constants $n_{ramsey'} = n_{ramsey'}(p, |\mathcal{S}|, I, n_0, r_0)$ and $r_{ramsey'} = r_{ramsey'}(p, |\mathcal{S}|, I, n_0, r_0)$ such that the following holds. Let V be a finite dimensional $\mathbb{F}_p$-vector space satisfying $\dim V \ge n_{ramsey'}$ and let $P \colon V \to \mathbb{F}_p$ be a non-trivial linear function. Let $\mathfrak{B}$ be a polynomial factor on $\mathbb{F}_p^n$ with parameters I and let $\mathfrak{B}'$ be the common refinement of $\mathfrak{B}$ and $\{P\}$. Suppose that $\operatorname{rank} \mathfrak{B}' \ge r_{ramsey'}$. Let $f \colon V \to \mathcal{S}$ be a function. Then there exists a subspace $U \le V$ contained in the zero set of P, a vector $z \in V$ such that $P(z) = 1$, and a function $\chi \colon A_I \to \mathcal{S}$ such that the following holds:*

(i) $\chi(\mathfrak{B}(x)) = f(x)$ *for all* $x \in z + U$;
(ii) $\operatorname{rank} \mathfrak{B}|_U \ge r_0$;
(iii) $\dim U \ge n_0$.

Define $r_1, \dots, r_M$ and $n_1, \dots, n_M$ by

$$n_i := \max\{n_{ramsey'}(p, |\mathcal{S}|^{p-1}, I, n_{i-1}, r_{i-1}), |\mathcal{L}| + 1\}$$

and

$$r_i := r_{ramsey'}(p, |\mathcal{S}|^{p-1}, I, n_{i-1}, r_{i-1}).$$

Then define

$$n_{ramsey}(p, |\mathcal{S}|, I, m, r_0, n_0) := n_M \qquad \text{and} \qquad r_{ramsey}(p, |\mathcal{S}|, I, m, r_0, n_0) := r_M.$$

We will find nested subspaces $V = U_0 \ge U_1 \ge \cdots \ge U_M$, linear functions $P_i \colon V \to \mathbb{F}_p$, and functions $\xi_i \colon (\mathbb{F}_p \setminus \{0\}) \times A_I \to \mathcal{S}$ such that the following holds for each $i \in [M]$:

- $\xi_i(P_i(x), \mathfrak{B}(x)) = f(x)$ for all $x \in U_i$ that satisfy $P_1(x) = \cdots = P_{i-1}(x) = 0$ and $P_i(x) \ne 0$;
- $\{P_1, \dots, P_i\} \cup \mathcal{L}$ are linearly independent;
- $\operatorname{rank} \mathfrak{B}|_{W_i} \ge r_{M-i}$ where $W_i := \{x \in U_i : P_1(x) = \cdots = P_i(x)\}$;
- $\dim W_i \ge n_{M-i}$ where $W_i := \{x \in U_i : P_1(x) = \cdots = P_i(x)\}$.

Suppose we have defined $V = U_0 \ge U_1 \ge \cdots \ge U_i$, linear functions $P_1, \dots, P_i \colon V \to \mathbb{F}_p$, and functions $\xi_1, \dots, \xi_i \colon (\mathbb{F}_p \setminus \{0\}) \times A_I \to \mathcal{S}$ with the above properties.

Define $W := \{x \in U_i : P_1(x) = \cdots = P_i(x) = 0\}$. We have $\dim W \ge m_i > |\mathcal{L}|$. Pick an arbitrary $y \in W$ such $y \ne 0$ but all of the linear functions $\mathcal{L}$ vanish on y. Let $P_{i+1} \colon V \to \mathbb{F}_p$ be an arbitrary linear function such that $P_{i+1}(y) = 1$. Note that automatically we have $\{P_1, \dots, P_i, P_{i+1}\} \cup \mathcal{L}$ are linearly independent.

Define $W' := \{x \in W : P_{i+1}(x) = 0\}$. Note that the subspace W is partitioned into hyperplanes as $W = W' \sqcup (y + W') \sqcup (2y + W') \sqcup \cdots$. Write $\overline{\mathcal{S}} := \mathcal{S}^{\mathbb{F}_p^\times}$. Then define $\overline{f} \colon W \to \overline{\mathcal{S}}$ by

$$\overline{f}(x + ty) := (f(bx + by))_{b \in \mathbb{F}_p^\times} \qquad \text{for } x \in W' \text{ and } t \in \mathbb{F}_p.$$

We apply Lemma 5.8 to $\overline{f}, \mathfrak{B}|_W, P_{i+1}$ with parameters n_{M-i-1}, r_{M-i-1} to produce a subspace $U_i' \le W'$, a vector $z \in y + U_i'$ and a function $\chi \colon A_I \to \overline{\mathcal{S}}$ with several desirable properties.

We have $\chi(\mathfrak{B}(x + z)) = \overline{f}(x + z)$ for all $x \in U_i'$. Looking at the bth coordinate of this equation for some $b \in \mathbb{F}_p^\times$ gives $\chi(\mathfrak{B}(x + z))_a = f(bx + bz)$. Finally using the homogeneity of $\mathfrak{B}$ (recall the action of $\mathbb{F}_p^\times$ on A_I defined in Eq. (3.5)) gives $\chi(b^{-1} \cdot \mathfrak{B}(bx + bz))_a = f(bx + bz)$ for all $x \in U_i'$ and all $b \in \mathbb{F}_p^\times$.

Define $U_{i+1} \le U_i$ to be a $(\dim W + i + 1)$-dimensional subspace of U_i that contains z and W and such that none of $P_1, \dots, P_i$ are identically 0 on U_{i+1}. Then define $\xi_{i+1} \colon (\mathbb{F}_p \setminus \{0\}) \times A_I \to \mathcal{S}$ by

$$\xi_{i+1}(b, a) := \chi(b^{-1} \cdot a)_b.$$

Note $\dim U_{i+1} \geq \dim W' \geq m_{i+1}$ and $\operatorname{rank} \mathfrak{B}|_{U_{i+1}} \geq \operatorname{rank} \mathfrak{B}|_{W'} \geq r_{i+1}$. Furthermore, for $x \in U_{i+1}$ such that $P_1(x) = \cdots = P_{i-1}(x) = 0$ and $P_i(x) = b \neq 0$, we can write $x = bx' + bz$ for some $x' \in W'$. Then $f(x) = \overline{f}(x' + z)_b = \chi(\mathfrak{B}(x' + z))_b = \chi(b^{-1} \cdot \mathfrak{B}(bx' + bz))_b = \xi_{i+1}(b, \mathfrak{B}(x))$, as desired.

Thus we have defined a sequence of nested subspaces $V = U_0 \leq \cdots \leq U_M$, linear functions $P_1, \ldots, P_m \colon V \to \mathbb{F}_p$ and functions $\xi_1, \ldots, \xi_M \colon (\mathbb{F}_p \setminus \{0\}) \times A_I \to \mathcal{S}$ with the above properties.

Finally note that the number of possible functions $(\mathbb{F}_p \setminus \{0\}) \times A_I \to \mathcal{S}$ is at most $|\mathcal{S}|^{p\|I\|}$. Thus by the pigeonhole principle, there exists $1 \leq i_1 < \cdots < i_m \leq M$ such that $\xi_1 = \cdots = \xi_m$. Define $\xi \colon \mathbb{F}_p \times A_I \to \mathcal{S}$ by $\xi(0, a) = f(0)$ for all $a \in A$ and $\xi(b, a) = \xi_1(b, a)$ for all $b \in (\mathbb{F}_p \setminus \{0\})$ and all $a \in A$. Define $W_M := \{x \in U_M : P_1(x) = \cdots = P_M(x) = 0\}$ and let U be a $(\dim W_M + m)$-dimensional subspace of V that contains W_M and such that none of $P_{i_1}, \ldots, P_{i_m}$ are identically 0 on U. Then U, $P_{i_1}, \ldots, P_{i_m}$, and ξ have all the desired properties. □

Proof of Lemma 5.8. Define constants

$$n_1 := \max\{n_0, n_{high-rank}(p, I, r_0 + p)\},$$

and

$$\theta := \frac{1}{8(2|\mathcal{S}|)^{2p^{n_1}}} \qquad \text{and} \qquad \theta' := \frac{\theta}{\|I\|\sqrt{2p|\mathcal{S}|}}$$

Define the non-increasing functions $\alpha \colon \mathbb{Z}_{>0} \to (0, 1)$ by

$$\alpha(N) := \frac{1}{2N^{2p^{n_1}}},$$

and $\eta \colon \mathbb{Z}_{>0} \to (0, 1)$ by

$$\eta(N) := \frac{1}{8(3|\mathcal{S}|N)^{p^{n_1}}},$$

and define the non-decreasing function $r \colon \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ by

$$r(N) := r_{equi}(p, p^{n_1}, \alpha(N)).$$

Define

$$n_{ramsey'}(p, |\mathcal{S}|, I, n_0, r_0) := 2p^{n_1} \left\lceil \log_p(16|\mathcal{S}|C_{reg''}(p, p^{n_1}, |\mathcal{S}|, p\|I\|, \theta', \eta, r)) \right\rceil$$

and

$$r_{ramsey'}(p, |\mathcal{S}|, I, n_0, r_0) := r_{reg''}(p, p^{n_1}, |\mathcal{S}|, p\|I\|, \theta', \eta, r)$$

For $c \in \mathcal{S}$, define $f^{(c)} \colon V \to [0, 1]$ by $f^{(c)} := 1_{f^{-1}(c)}$. We are now ready to proceed with the proof. We apply the arithmetic regularity lemma, Lemma 4.4, to the polynomial factor $\mathfrak{B}'$ on V and the functions $(f^{(c)})_{c \in \mathcal{S}}$ with parameters $p, p^{n_1}, |\mathcal{S}|, p\|I\|, \theta', \eta, r$. This produces a polynomial factor $\mathfrak{B}''$ refining $\mathfrak{B}'$ and decompositions $f^{(c)} = f^{(c)}_{str} + f^{(c)}_{psr} + f^{(c)}_{sml}$ with several desirable properties.

Say that $\mathfrak{B}'$ has parameters I' and $\mathfrak{B}''$ has parameters I'' (note that $\|I'\| = p\|I\|$). We consider the atom-indexing set of $\mathfrak{B}''$ (see Eq. (3.3) for the definition) as $A_{I''} \simeq \mathbb{F}_p \times A_I \times A_{I''-I'}$.

Say that an atom $a \in A_{I''}$ is *regular* if

$$\|f^{(c)}_{sml} 1_{\mathfrak{B}''^{-1}(a)}\|_2 \leq \theta \|1_{\mathfrak{B}''^{-1}(a)}\|_2 \qquad \text{for all } c \in \mathcal{S}.$$

Our first goal is to find $s \in A_{I''-I'}$ such that all atoms of the form $(1, a, s) \in A_{I''} \simeq \mathbb{F}_p \times A_I \times A_{I''-I'}$ are regular.

By Lemma 4.4(v), for each $c \in \mathcal{S}$,

$$\sum_{a \in A_{I''}} \|f^{(c)}_{sml} 1_{\mathfrak{B}''^{-1}(a)}\|_2^2 \leq \frac{\theta^2}{2p|\mathcal{S}|\|I\|^2} \sum_{a \in A_{I''}} \|1_{\mathfrak{B}''^{-1}(a)}\|_2^2.$$

Thus at least a $(1 - 1/(2p\|I\|^2))$-fraction of atoms are regular. For each $a \in A_I$, at least a $(1 - 1/(2\|I\|))$-fraction of the atoms of the form $(1, a, s)$ are regular, for $s \in A_{I''-I'}$. Thus by a union

bound there exists some $s \in A_{I''-I'}$ such that $(1, a, s)$ is regular for all $a \in A_I$. Fix this value of s for the rest of the proof.

Define $\chi\colon A_I \to \mathcal{S}$ such that $\chi(a)$ is a color that appears in the atom $\mathfrak{B}^{-1}(1, a, s)$ with density at least $1/|\mathcal{S}|$.

By the definition of n_1 and Lemma 3.15, there exists a polynomial factor $\mathfrak{B}_1$ on $\mathbb{F}_p^{n_1}$ with parameters I and satisfying $\operatorname{rank} \mathfrak{B}_1 \ge r_0 + p$. Our goal is to find vectors $x_0, x_1, \dots, x_{n_1} \in V$ such that

- $\mathfrak{B}''(x_0 + i_1x_1 + \cdots + i_{n_1}x_{n_1}) = (1, \mathfrak{B}_1(i_1, \dots, i_{n_1}), s)$ for all $(i_1, \dots, i_{n_1}) \in \mathbb{F}_p^{n_1}$;
- $f(x_0 + i_1x_1 + \cdots + i_{n_1}x_{n_1}) = \chi(\mathfrak{B}_1(i_1, \dots, i_{n_1}))$ for all $(i_1, \dots, i_{n_1}) \in \mathbb{F}_p^{n_1}$;
- $x_1, \dots, x_{n_1}$ are linearly independent.

We choose $x_0, x_1, \dots, x_{n_1} \in V$ independently and uniformly at random. Let p_1 be the probability that this choice of $\boldsymbol{x}$ satisfies all three conditions above. First note that the probability that $x_0, \dots, x_{n_1}$ are linearly dependent is at most $p^{n_1+1}/|V|$. Let p_2 be the probability that this choice of $\boldsymbol{x}$ satisfies the first two conditions above. We have shown that

$$p_1 \ge p_2 - p^{n_1+1}/|V|.$$

Let $\boldsymbol{L} = (L_{\boldsymbol{i}})_{\boldsymbol{i} \in \mathbb{F}_p^{n_1}}$ to be the system of p^{n_1} linear forms in $n_1 + 1$ variables defined by

$$L_{\boldsymbol{i}}(x_0, x_1, \dots, x_{n_1}) := x_0 + i_1x_1 + \cdots + i_{n_1}x_{n_1}.$$

This system defines an n_1-dimensional affine subspace. One can easily see that $\boldsymbol{L}$ is finite complexity and in fact its complexity is at most p^{n_1} (see Remark 3.4).

Define $g^{(\boldsymbol{i})} := f^{(\chi(\mathfrak{B}_1(\boldsymbol{i})))}$ and define $g^{(\boldsymbol{i})}_{str}, g^{(\boldsymbol{i})}_{sml}, g^{(\boldsymbol{i})}_{psr}$ similarly. Define $h^{(\boldsymbol{i})}\colon V \to [0,1]$ by $h^{(\boldsymbol{i})} := 1_{\mathfrak{B}''^{-1}(1, \mathfrak{B}_1(\boldsymbol{i}), s)}$.

We compute p_2 as

$$\begin{aligned}
p_2 &= \mathbb{E}_{\boldsymbol{x}} \left[\prod_{\boldsymbol{i} \in \mathbb{F}_p^{n_1}} g^{(\boldsymbol{i})}(L_{\boldsymbol{i}}(\boldsymbol{x})) h^{(\boldsymbol{i})}(L_{\boldsymbol{i}}(\boldsymbol{x})) \right] \\
&= \mathbb{E}_{\boldsymbol{x}} \left[\prod_{\boldsymbol{i}} \left(g^{(\boldsymbol{i})}_{str}(L_{\boldsymbol{i}}(\boldsymbol{x})) + g^{(\boldsymbol{i})}_{sml}(L_{\boldsymbol{i}}(\boldsymbol{x})) + g^{(\boldsymbol{i})}_{psr}(L_{\boldsymbol{i}}(\boldsymbol{x})) \right) h^{(\boldsymbol{i})}(L_{\boldsymbol{i}}(\boldsymbol{x})) \right] \\
&\ge \mathbb{E}_{\boldsymbol{x}} \left[\prod_{\boldsymbol{i}} \left(g^{(\boldsymbol{i})}_{str}(L_{\boldsymbol{i}}(\boldsymbol{x})) + g^{(\boldsymbol{i})}_{sml}(L_{\boldsymbol{i}}(\boldsymbol{x})) \right) h^{(\boldsymbol{i})}(L_{\boldsymbol{i}}(\boldsymbol{x})) \right] - 3^{p^{n_1}} \eta(\|\mathfrak{B}''\|).
\end{aligned}$$

The inequality follows from Lemma 4.4(ii) and the counting lemma, Eq. (3.1).

Write p_3 for the expectation in the last line above. We have $p_2 \ge p_3 - 3^{p^{n_1}} \eta(\|\mathfrak{B}''\|)$. Expanding the product, there are at most $2^{p^{n_1}}$ terms involving $g^{(\boldsymbol{j})}_{sml}$ for some $\boldsymbol{j} \in \mathbb{F}_p^{n_1}$. Each of these is bounded in magnitude by

$$\mathbb{E}_{\boldsymbol{x}} \left[\left| g^{(\boldsymbol{j})}_{sml}(L_{\boldsymbol{j}}(\boldsymbol{x})) \right| \prod_{\boldsymbol{i}} h^{(\boldsymbol{i})}(L_{\boldsymbol{i}}(\boldsymbol{x})) \right].$$

By applying a change of coordinates, we can transform to the case that $\boldsymbol{j} = 0$ (and $L_{\boldsymbol{0}}(\boldsymbol{x}) = x_0$). Then by the Cauchy-Schwarz inequality, the square of the above expression is bounded by

$$\mathbb{E}_{x_0} \left[\left| g^{(0)}_{sml}(x_0) \right|^2 h^{(0)}(x_0) \right] \mathbb{E}_{x_0} \left[h^{(0)}(x_0) \mathbb{E}_{x_1, \dots, x_{n_1}} \left[\prod_{\boldsymbol{i} \ne 0} h^{(\boldsymbol{i})}(L_{\boldsymbol{i}}(\boldsymbol{x})) \right]^2 \right].$$

The first term is at most $\theta(\|\mathfrak{B}\|)^2 \|h^{(0)}\|_2^2$ by the fact that $(1, \mathfrak{B}_1(\boldsymbol{i}), s)$ is a regular atom for all $\boldsymbol{i}$. The second term can be counted by equidistribution applied to the system $\boldsymbol{L}'$ of $2p^{n_1} - 1$ linear

forms in $2n_1 - 1$ variables defined as follows. Set

$$L'_0(x_0, x_1, \dots, x_{n_1}, x'_1, \dots, x'_{n_1}) := x_0,$$

and for $\boldsymbol{i} \in \mathbb{F}_p^{n_1} \setminus \{0\}$, define

$$L'_{\boldsymbol{i},1}(x_0, x_1, \dots, x_{n_1}, x'_1, \dots, x'_{n_1}) := x_0 + i_1 x_1 + \cdots + i_{n_1} x_{n_1},$$
$$L'_{\boldsymbol{i},2}(x_0, x_1, \dots, x_{n_1}, x'_1, \dots, x'_{n_1}) := x_0 + i_1 x'_1 + \cdots + i_{n_1} x'_{n_1}.$$

By Lemma 3.23, we know that $\|\mathfrak{B}''\| \cdot |\Phi_{I''}(\boldsymbol{L}')| = |\Phi_{I''}(\boldsymbol{L})|^2$ (see also [7, Lemma 5.13].)

Thus by equidistribution, Theorem 3.19, and the rank bound on $\mathfrak{B}''$, we have the second term is at most

$$\frac{1}{|\Phi_{I''}(\boldsymbol{L}')|} + \alpha(\|\mathfrak{B}''\|) = \frac{\|\mathfrak{B}''\|}{|\Phi_{I''}(\boldsymbol{L})|^2} + \alpha(\|\mathfrak{B}''\|) \le \frac{2\|\mathfrak{B}''\|}{|\Phi_{I''}(\boldsymbol{L})|^2}.$$

Applying equidistribution again we have that the first term is at most

$$\theta^2 \left(\frac{1}{\|\mathfrak{B}''\|} + \alpha(\|\mathfrak{B}''\|) \right) \le \frac{2\theta^2}{\|\mathfrak{B}''\|}.$$

Combining these bounds and summing over all terms that contain some $g^{(\boldsymbol{j})}_{sml}$, we see that

$$p_3 \ge \mathbb{E}_{\boldsymbol{x}} \left[\prod_{\boldsymbol{i}} g^{(\boldsymbol{i})}_{str}(L_{\boldsymbol{i}}(\boldsymbol{x})) h^{(\boldsymbol{i})}(L_{\boldsymbol{i}}(\boldsymbol{x})) \right] - 2^{p^{n_1}+1} \frac{\theta}{|\Phi_{I''}(\boldsymbol{L})|}.$$

Write p_4 for the expectation in the last line. The quantity $g^{(\boldsymbol{i})}_{str}(L_{\boldsymbol{i}}(\boldsymbol{x}))$ is the density of $\chi(\mathfrak{B}_1(\boldsymbol{i}))$ in the atom of $\mathfrak{B}''$ that $L_{\boldsymbol{i}}(\boldsymbol{x})$ lies in. When $\mathfrak{B}(L_{\boldsymbol{i}}(\boldsymbol{x})) = (1, \mathfrak{B}_1(\boldsymbol{i}), s)$, the choice of χ implies that this density is at least $1/|\mathcal{S}|$. Thus

$$p_4 \ge \frac{1}{|\mathcal{S}|^{p^{n_1}}} \mathbb{E}_{\boldsymbol{x}} \left[\prod_{\boldsymbol{i}} h^{(\boldsymbol{i})}(L_{\boldsymbol{i}}(\boldsymbol{x})) \right].$$

Write p_5 for the expectation in the last line. Unwrapping the definition of $h^{(\boldsymbol{i})}$, this can be written as

$$p_5 = \Pr_{\boldsymbol{x}} \left(\mathfrak{B}''(L_{\boldsymbol{i}}(\boldsymbol{x})) = (1, \mathfrak{B}_1(\boldsymbol{i}), s) \text{ for all } \boldsymbol{i} \in \mathbb{F}_p^{n_1} \right).$$

We claim that $((1, \mathfrak{B}_1(\boldsymbol{i}), s))_{\boldsymbol{i} \in \mathbb{F}_p^{n_1}}$ is an $\boldsymbol{L}$-consistent tuple of atoms. We check this coordinate by coordinate. Obviously $(\mathfrak{B}_1(\boldsymbol{i}))_{\boldsymbol{i} \in \mathbb{F}_p^{n_1}}$ is $\boldsymbol{L}$-consistent. Furthermore any constant tuple is also obviously $\boldsymbol{L}$-consistent (this follows from the fact that $\boldsymbol{L}$ is a translation-invariant pattern). Thus by another application of equidistribution and the rank bound on $\mathfrak{B}''$, we have

$$p_5 \ge \frac{1}{|\Phi_{I''}(\boldsymbol{L}')|} - \alpha(\|\mathfrak{B}'\|) \ge \frac{1}{2|\Phi_{I''}(\boldsymbol{L}')|}.$$

Combining all these inequalities, we see that

$$p_1 \ge \frac{1}{|\mathcal{S}|^{p^{n_1}}} \frac{1}{2|\Phi_{I''}(\boldsymbol{L}')|} - 2^{p^{n_1}+1} \frac{\theta}{|\Phi_{I''}(\boldsymbol{L})|} - 3^{p^{n_1}} \eta(\|\mathfrak{B}''\|) - \frac{p^{n_1+1}}{|V|}.$$

This expression is positive by the definition of θ, η and the assumption that $\dim V \ge n_{ramsey'}$.

Thus we have define a function $\chi \colon A_I \to \mathcal{S}$ and found linearly independent $x_0, x_1, \dots, x_{n_1} \in V$ with several desirable properties. Define $z := x_0$ and $U := \operatorname{span}\{x_1, \dots, x_{n_1}\}$. Since $P \colon V \to \mathbb{F}_p$ is a linear function and $P(x_0 + i_1 x_1 + \cdots + i_{n_1} x_{n_1}) = 1$ for all $\boldsymbol{i} \in \mathbb{F}_p^{n_1}$, we conclude that $P(z) = P(x_0) = 1$ and U is contained in the zero set of P. Furthermore, $\mathfrak{B}(x_0 + i_1 x_1 + \cdots + i_{n_1} x_{n_1}) = \mathfrak{B}_1(i_1, \dots, i_{n_1})$ for all $\boldsymbol{i} \in \mathbb{F}_p^{n_1}$. Since $\mathfrak{B}_1$ was chosen such that $\operatorname{rank} \mathfrak{B}_1 \ge r_0 + p$ and $n_1 \ge n_0$, we see that $\dim U \ge n_0$ and $\operatorname{rank} \mathfrak{B}|_{\operatorname{span}\{x_0, U\}} \ge r_0 + p$. By Lemma 3.16, we conclude that $\operatorname{rank} \mathfrak{B}|_U \ge r_0$. Finally, $f(x_0 + i_1 x_1 + \cdots + i_{n_1} x_{n_1}) = \chi(\mathfrak{B}_1(i_1, \dots, i_{n_1})) = \chi(\mathfrak{B}(x_0 + i_1 x_1 + \cdots + i_{n_1} x_{n_1}))$ for all $\boldsymbol{i} \in \mathbb{F}_p^{n_1}$, which proves the desired result. □

We boost this Ramsey dichotomy result to a density result using a Cauchy-Schwarz supersaturation argument. For technical reasons we need this result to hold inside "subvarieties" of a vector space, i.e., the zero sets of a sufficiently high-rank collection of non-classical polynomials. Again for technical reasons, this supersaturation argument only works for full dimensional patterns (recall Definition 3.21).

Theorem 5.9 (Patching). *Fix a prime p, a finite set $\mathcal{S}$ with an $\mathbb{F}_p^\times$-action, parameter lists $I, I' \in \mathcal{I}_p$ satisfying $I \le I'$, and a positive integer ℓ_0. There exist constants $n_{patch} = n_{patch}(p, |\mathcal{S}|, I', \ell_0)$ and $\beta_{patch} = \beta_{patch}(p, |\mathcal{S}|, I, \ell_0) > 0$ and a non-decreasing function $r_{patch} = r_{patch}(p, |\mathcal{S}|, I, \ell_0) \colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ such that the following holds. Let $\mathcal{H}$ be a finite set of $\mathcal{S}$-colored, I-labeled patterns such that each pattern is defined by a full dimension system of linear forms in at most ℓ_0 variables (recall Definition 3.21). Either:*

(a) there exists a projective $\xi \colon \mathbb{F}_p \times A_I \to \mathcal{S}$ that does not canonically induce $\mathcal{H}$; or

(b) for every finite-dimensional $\mathbb{F}_p$-vector space V satisfying $\dim V \ge n_{patch}$, every projective function $f \colon V \to \mathcal{S}$, every polynomial factor $\mathfrak{B}$ on V with parameters I that satisfies $\operatorname{rank} \mathfrak{B} \ge r_{patch}(\deg \mathfrak{B}, \|\mathfrak{B}\|)$, and every polynomial factor $\mathfrak{B}'$ on V with parameters I' that refines $\mathfrak{B}$ and satisfies $\operatorname{rank} \mathfrak{B}' \ge r_{patch}(\deg \mathfrak{B}', \|\mathfrak{B}'\|)$, there is a pattern $H \in \mathcal{H}$ such that in $(f, \mathfrak{B})$, the relative density of H in $\mathfrak{B}'^{-1}(A_I \times \{0\})$ is at least β_{patch}.

Note that n_{patch} may depend on I', but critically β_{patch} does not depend on I'.

Proof. First we define several parameters.

Write $r := r_{dich}(p, |\mathcal{S}|, I, \ell_0)$ for brevity. Define

$$n_1 := \max\{n_{dich}(p, |\mathcal{S}|, I, \ell_0), n_{high-rank}(p, I, r)\}$$

where n_0 is defined in Lemma 3.15.

Define the constants

$$n_{patch}(p, |\mathcal{S}|, I', \ell_0) := 2p^{n_1} \left\lceil \log_p(\|I'\|) \right\rceil,$$

and

$$\beta_{patch}(p, |\mathcal{S}|, I, \ell_0) := \frac{1}{6400 \ell_0^2 p^{2n_1 \cdot \ell} \|I\|^{2p^{n_1}}},$$

and define the non-increasing function $\alpha \colon \mathbb{Z}_{>0} \to (0, 1)$ by

$$\alpha(N) := \frac{1}{2N^{2p^{n_1}}},$$

and the non-decreasing function $r_{patch}(p, |\mathcal{S}|, I, \ell_0) \colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ by

$$r_{patch}(p, |\mathcal{S}|, I, \ell_0)(D, N) := r_{equi}(p, D, \alpha(N)).$$

We now proceed to the proof. Let $\mathcal{H}$ be a finite set of $\mathcal{S}$-colored, I-labeled patterns each defined by a full dimension system of linear forms in at most ℓ_0 variables. We apply Theorem 5.6 to $\mathcal{H}$. If Theorem 5.6(a) holds, then clearly conclusion (a) holds. Now assume that Theorem 5.6(b) holds. We wish to show conclusion (b).

Let V be a finite-dimensional $\mathbb{F}_p$-vector space satisfying $\dim V \ge n_{patch}$, let $f \colon V \to \mathcal{S}$ be a projective function, let $\mathfrak{B}$ be a polynomial factor on V with parameters I that satisfies $\operatorname{rank} \mathfrak{B} \ge r_{patch}(\|\mathfrak{B}\|)$, and let $\mathfrak{B}'$ be a polynomial factor on V with parameters I' that refines $\mathfrak{B}$ and satisfies $\operatorname{rank} \mathfrak{B}' \ge r_{patch}(\|\mathfrak{B}'\|)$.

Write $X := \mathfrak{B}'^{-1}(A_I \times \{0\})$ and $n := \dim V$. By assumption, $n \ge n_{patch}$. We wish to count $\mathcal{H}$ instances in $(f, \mathfrak{B})$ that are contained in X. Write $\mathcal{H} = \bigsqcup_{\ell=1}^{\ell_0} \mathcal{H}_\ell$ where $\mathcal{H}_\ell$ is defined to be the subset of $\mathcal{H}$ consisting of colored labeled patterns defined by a system of linear forms in exactly ℓ variables. We define sets $\mathfrak{U}_1, \dots, \mathfrak{U}_{\ell_0}$ as follows. $\mathfrak{U}_\ell$ is the set of ℓ-dimensional subspaces U of V which satisfy the following:

- $U \subseteq X$;

- there exists a colored labeled pattern $H \in \mathcal{H}_\ell$ in ℓ variables and a generic H-instance $x_1, \dots, x_\ell \in U$.

Note that the requirement that $x_1, \dots, x_\ell \in U$ are generic implies that $U = \operatorname{span}\{x_1, \dots, x_\ell\}$. Thus $\sum_{\ell=1}^{\ell_0} |\mathfrak{U}_\ell|$ is a lower bound on the number of $\mathcal{H}$-instances in X.

Define the following counting function $c\colon \mathfrak{U}_\ell \to \mathbb{Z}_{\geqslant 0}$ for each $\ell \in [\ell_0]$ as follows. For $U \in \mathfrak{U}_\ell$, let $c(U)$ be the number of n_1-dimensional subspaces that contain U and are contained in X. An application of the Cauchy-Schwarz inequality implies that

$$|\mathfrak{U}_\ell| \geq \frac{\left(\sum_{U \in \mathfrak{U}_\ell} c(U)\right)^2}{\sum_{U \in \mathfrak{U}_\ell} c(U)^2}. \tag{5.1}$$

Define S_1 to be the number of n_1-dimensional subspaces W of V such that $W \subseteq X$ and $\operatorname{rank}(\mathfrak{B}|_W) \geq r$. By Theorem 5.6(b), every such W contains a generic $\mathcal{H}$-instance. Thus

$$\sum_{\ell=1}^{\ell_0} \sum_{U \in \mathfrak{U}_\ell} c(U) \geq S_1.$$

By the pigeonhole principle, there exists some $\ell \in [\ell_0]$ such that

$$\sum_{U \in \mathfrak{U}_\ell} c(U) \geq \frac{S_1}{\ell_0}. \tag{5.2}$$

We fix such a value of $\ell \in [\ell_0]$ for the rest of the proof.

Define S_2 to be the number of ordered n_1-tuples $(x_1, \dots, x_{n_1}) \in V^{n_1}$ such that $x_1, \dots, x_{n_1}$ are linearly independent, $\operatorname{span}\{x_1, \dots, x_{n_1}\} \subseteq X$, and $\operatorname{rank}(\mathfrak{B}|_{\operatorname{span}\{x_1, \dots, x_{n_1}\}}) \geq r$. We can compute

$$S_2 = S_1 \prod_{i=0}^{n_1-1} (p^{n_1} - p^i) \leq p^{n_1^2} S_1. \tag{5.3}$$

Define S_3 to be the number of ordered n_1-tuples $(x_1, \dots, x_{n_1}) \in V^{n_1}$ such that $\operatorname{span}\{x_1, \dots, x_{n_1}\} \subseteq X$ and $\operatorname{rank}(\mathfrak{B}|_{\operatorname{span}\{x_1, \dots, x_{n_1}\}}) \geq r$. We can easily bound

$$S_2 \geq S_3 - p^{n \cdot n_1} p^{n_1 - n}. \tag{5.4}$$

By the definition of n_1 and Lemma 3.15, there exists a polynomial factor $\mathfrak{B}_1$ on $\mathbb{F}_p^{n_1}$ with parameters I and rank at least r. Define S_4 to be the number of ordered n_1-tuples $(x_1, \dots, x_{n_1}) \in V^{n_1}$ such that $\mathfrak{B}_1(i_1, \dots, i_{n_1}) = \mathfrak{B}(i_1 x_1 + \cdots + i_{n_1} x_{n_1})$ and $\operatorname{span}\{x_1, \dots, x_n\} \subseteq X$. Notice that $S_3 \geq S_4$.

Write $\boldsymbol{L}' := \boldsymbol{L}^{n_1}$, the system of p^{n_1} linear forms in n_1 variables that define an n_1-dimensional subspace (see Definition 2.7). By definition, $(P^i_{d,k}(i_1, \dots, i_{n_1}))_{\boldsymbol{i} \in \mathbb{F}_p^{n_1}} \in \Phi_{d,k}(\boldsymbol{L}')$ for every $(d,k) \in D_p$ and $i \in [I_{d,k}]$ where $P^i_{d,k}$ is the ith non-classical polynomial of degree d and depth k defining $\mathfrak{B}_1$. Also, $(0, \dots, 0) \in \Phi_{d,k}(\boldsymbol{L}')$ for all $(d,k) \in D_p$. Define $\boldsymbol{a} \in A_{I'}^{\mathbb{F}_p^{n_1}}$ by

$$(a_{\boldsymbol{i}})^i_{d,k} = \begin{cases} P^i_{d,k}(i_1, \dots, i_{n_1}) & \text{if } i \leq I_{d,k}, \\ 0 & \text{if } I_{d,k} < i \leq I'_{d,k}. \end{cases}$$

By the above discussion, $\boldsymbol{a}$ is $\boldsymbol{L}'$-consistent, so by Theorem 3.19 and the rank assumption of $\mathfrak{B}'$, we find

$$\begin{aligned} S_3 \geq S_4 &= \left|\{\boldsymbol{x} \in V^{n_1} : \mathfrak{B}'(L'_{\boldsymbol{i}}(\boldsymbol{x})) = a_{\boldsymbol{i}} \text{ for all } \boldsymbol{i} \in \mathbb{F}_p^{n_1}\}\right| \\ &\geq \left(\frac{1}{|\Phi_{I'}(\boldsymbol{L}')|} - \alpha(\|I'\|)\right) p^{n \cdot n_1} \geq \frac{p^{n \cdot n_1}}{2|\Phi_{I'}(\boldsymbol{L}')|}. \end{aligned} \tag{5.5}$$

Combining Eq. (5.2), Eq. (5.3), Eq. (5.4), and Eq. (5.5), we conclude

$$\sum_{U\in\mathfrak{U}_\ell} c(U) \ge \frac{1}{\ell_0 p^{n_1^2}} p^{n\cdot n_1}\left(\frac{1}{2|\Phi_{I'}(\boldsymbol{L}')|} - p^{n_1-n}\right) \ge \frac{1}{4\ell_0 p^{n_1^2}|\Phi_{I'}(\boldsymbol{L}')|} p^{n\cdot n_1}. \tag{5.6}$$

Next we find an upper bound on $\sum_{U\in\mathfrak{U}_\ell} c(U)^2$. Define T_1 to be the number of triples (U, W, W') where U is a ℓ-dimensional subspace of V and W, W' are both n_1-dimensional subspaces of V that contain U and are contained in X. First note that

$$\sum_{U\in\mathfrak{U}_\ell} c(U)^2 \le T_1. \tag{5.7}$$

Now define T_2 to be the number of ordered $(2n_1-\ell)$-tuples $(x_1,\dots,x_\ell,y_1,\dots,y_{n_1-\ell},z_1,\dots,z_{n_1-\ell}) \in V^{2n_1-\ell}$ such that $x_1,\dots,x_\ell,y_1,\dots,y_{n_1-\ell}$ are linearly independent, $x_1,\dots,x_\ell,z_1,\dots,z_{n_1-\ell}$ are linearly independent, $\operatorname{span}\{x_1,\dots,x_\ell,y_1,\dots,y_{n_1-\ell}\} \subseteq X$, and $\operatorname{span}\{x_1,\dots,x_\ell,z_1,\dots,z_{n_1-\ell}\} \subseteq X$. We compute

$$T_2 = \left(\prod_{i=0}^{\ell-1}(p^\ell - p^i)\right)\left(\prod_{i=\ell}^{n_1-1}(p^{n_1}-p^i)\right)^2 T_1 \ge \frac{p^{2n_1(n_1-\ell)}}{100} T_1. \tag{5.8}$$

Next define T_3 to be the number of ordered $(2n_1-\ell)$-tuples $(x_1,\dots,x_\ell,y_1,\dots,y_{n_1-\ell},z_1,\dots,z_{n_1-\ell}) \in V^{2n_1-\ell}$ such that $\operatorname{span}\{x_1,\dots,x_\ell,y_1,\dots,y_{n_1-\ell}\} \subseteq X$ and $\operatorname{span}\{x_1,\dots,x_\ell,z_1,\dots,z_{n_1-\ell}\} \subseteq X$. Clearly $T_2 \le T_3$.

Define $\boldsymbol{L}''$ to be the following system of $2p^{n_1} - p^\ell$ linearly forms in $2n_1 - n_\ell$ variables. For $\boldsymbol{i} \in \mathbb{F}_p^\ell$, define

$$L''_{\boldsymbol{i}}(x_1,\dots,x_\ell,y_1,\dots,y_{n_1-\ell},z_1,\dots,z_{n_1-\ell}) := i_1x_1 + \cdots + i_\ell x_\ell.$$

For $\boldsymbol{i} \in \mathbb{F}_p^{n_1} \setminus \left(\mathbb{F}_p^\ell \times \{0\}^{n_1-\ell}\right)$, define

$$\begin{aligned} L''_{\boldsymbol{i},1}(x_1,\dots,x_\ell,y_1,\dots,y_{n_1-\ell},z_1,\dots,z_{n_1-\ell}) &:= i_1x_1 + \cdots + i_\ell x_\ell + i_{\ell+1}y_1 + \cdots + i_{n_1}y_{n_1-\ell}, \\ L''_{\boldsymbol{i},2}(x_1,\dots,x_\ell,y_1,\dots,y_{n_1-\ell},z_1,\dots,z_{n_1-\ell}) &:= i_1x_1 + \cdots + i_\ell x_k + i_{\ell+1}z_1 + \cdots + i_{n_1}z_{n_1-\ell}. \end{aligned}$$

Now let $\mathfrak{B}''$ be the polynomial factor on V with parameters $I' - I$ defined by homogeneous non-classical polynomials

$$\left(P^i_{d,k}\right)_{\substack{(d,k)\in D_p \\ I_{d,k}<i\le I'_{d,k}}}$$

where $P^i_{d,k}$ is the ith non-classical polynomial of degree d and depth k that defines $\mathfrak{B}'$. The important property of this polynomial factor is that $X = \mathfrak{B}''^{-1}(\mathbf{0})$. Also note that $\operatorname{rank}\mathfrak{B}'' \ge \operatorname{rank}\mathfrak{B}' \ge r(\|\mathfrak{B}'\|)$. By Theorem 3.19 and the bounds on $\operatorname{rank}\mathfrak{B}''$, we find

$$\begin{aligned} T_2 \le T_3 &= \left|\{\boldsymbol{x} \in V^{2n_1-\ell} : \mathfrak{B}''(\boldsymbol{L}''(\boldsymbol{x})) = \mathbf{0}\}\right| \\ &\le \left(\frac{1}{|\Phi_{I'-I}(\boldsymbol{L}'')|} + \alpha(\|I'\|)\right) p^{n(2n_1-\ell)} \le \frac{2p^{n(2n_1-\ell)}}{|\Phi_{I'-I}(\boldsymbol{L}'')|}. \end{aligned} \tag{5.9}$$

Combining Eq. (5.7), Eq. (5.8), and Eq. (5.9), we conclude

$$\sum_{U\in\mathfrak{U}_\ell} c(U)^2 \le \frac{200p^{2n_1\cdot\ell}}{p^{2n_1^2}|\Phi_{I'-I}(\boldsymbol{L}'')|} p^{n(2n_1-\ell)}. \tag{5.10}$$

Finally, we combine Eq. (5.1), Eq. (5.6), and Eq. (5.10) to find

$$|\mathfrak{U}_\ell| \ge \frac{|\Phi_{I'-I}(\boldsymbol{L}'')|}{3200\ell_0^2 p^{2n_1\cdot\ell}|\Phi_{I'}(\boldsymbol{L}')|^2} p^{n\cdot\ell}.$$

Consider $\boldsymbol{L}^\ell$, the system of p^ℓ linear forms in ℓ variables that define an ℓ-dimension subspace (see Definition 2.7). By Lemma 3.23, we have $|\Phi_{I'-I}(\boldsymbol{L}^\ell)| \cdot |\Phi_{I'-I}(\boldsymbol{L}'')| = |\Phi_{I'-I}(\boldsymbol{L}')|^2$. Thus the above expression simplifies to

$$|\mathfrak{U}_\ell| \geq \frac{1}{3200\ell_0^2 p^{2n_1\cdot\ell}|\Phi_I(\boldsymbol{L}')|^2} \cdot \frac{p^{n\cdot\ell}}{|\Phi_{I'-I}(\boldsymbol{L}^\ell)|} \geq \frac{1}{3200\ell_0^2 p^{2n_1\cdot\ell}\|I\|^{2p^{n_1}}} \cdot \frac{p^{n\cdot\ell}}{|\Phi_{I'-I}(\boldsymbol{L}^\ell)|}.$$

Therefore there exists some colored labeled pattern $H = (\boldsymbol{L}, \psi, \phi) \in \mathcal{H}_\ell$ where $\boldsymbol{L}$ is a full dimension system of linear forms in ℓ variables and such that the number of generic H-instances in $(f, \mathfrak{B})$ which are contained in X is at least $1/|\mathcal{H}_\ell|$ times the right-hand side of the above equation. Note that by equidistribution, Theorem 3.19, and the rank bound on $\mathfrak{B}''$,

$$\Lambda_{\boldsymbol{L}}(1_X, \ldots, 1_X) \leq \frac{1}{|\Phi_{I'-I}(\boldsymbol{L})|} + \alpha(\|I'\|) \leq \frac{2}{|\Phi_{I'-I}(\boldsymbol{L})|}. \tag{5.11}$$

. Noting that since $\boldsymbol{L}$ is full dimensional, we have $|\Phi_{I'-I}(\boldsymbol{L})| = |\Phi_{I'-I}(\boldsymbol{L}^\ell)|$. Thus dividing the two above quantities, we find that the relative density of the above H in X is at least

$$\frac{1}{6400\ell_0^2 p^{2n_1\cdot\ell}\|I\|^{2p^{n_1}}} = \beta_{patch}(p, |\mathcal{S}|, I, \ell_0). \qquad \square$$

6. Proof of removal lemmas

As usual, for an atom $a \in A_I$ (defined in Eq. (3.3)), we use $a_{d,k} \in \mathbb{U}_{k+1}^{I_{d,k}}$ to denote the degree d, depth k part of a. We use the notation $\tilde{A}_I \subset A_I$ to denote the set $\tilde{A}_I := \{a \in A_I : a_{1,0} = 0\}$. This is the set of atoms that are regularized by Theorem 4.6. Also define $\tilde{I} \in \mathcal{I}_p$ by $\tilde{I}_{1,0} = 0$ and $\tilde{I}_{d,k} = I_{d,k}$ otherwise.

Definition 6.1. Fix a prime p, a finite set $\mathcal{S}$ equipped with an $\mathbb{F}_p^\times$-action, and a parameter list $I \in \mathcal{I}_p$. A **summary function** with parameters I is a pair (F, ξ) consisting of a function $F\colon (A_I \setminus \tilde{A}_I) \to 2^{\mathcal{S}} \setminus \{\emptyset\}$ and a projective function $\xi\colon \mathbb{F}_p \times A_{\tilde{I}} \to \mathcal{S}$.

Definition 6.2. For an $\mathcal{S}$-colored pattern $H = (\boldsymbol{L}, \psi)$ consisting of m linear forms and a summary function (F, ξ) with parameters I, say that (F, ξ) **partially induces** H if there exists a tuple of atoms $\boldsymbol{a} \in A_I^m$ such that the following holds:

(i) $\boldsymbol{a}$ is $\boldsymbol{L}$-consistent, i.e., $\boldsymbol{a} \in \Phi_I(\boldsymbol{L})$;
(ii) for each $i \in [m]$ such that $a_i \notin \tilde{A}_I$, we have $\psi(i) \in F(a_i)$;
(iii) defining $J := \{i \in [m] : a_i \in \tilde{A}_i\}$ and $H_J := ((L_i)_{i\in J}, \psi|_J, (a_i)_{i\in J}\}$, an $\mathcal{S}$-colored I-labeled pattern, we have ξ canonically induces H_J.

Proof of Theorem 2.8. We are given a parameter $\epsilon > 0$ and a possibly infinite set $\mathcal{H}$ of $\mathcal{S}$-colored patterns over $\mathbb{F}_p$ of the form $(\overline{\boldsymbol{L}}^\ell, \psi)$ where ℓ is some positive integer and $\psi\colon E_\ell \to \mathcal{S}$ is some map. (See Definition 2.7 for the definition of $\overline{\boldsymbol{L}}^\ell$ and E_ℓ.)

We begin with a "compactness argument" based on ideas of Alon and Shapira that allows us to reduce to the case when $\mathcal{H}$ is finite size.

For each parameter list $I \in \mathcal{I}_p$, we define a finite subset $\mathcal{H}_I \subseteq \mathcal{H}$ as follows. Consider the set of all summary functions (F, ξ) with parameters I. If there exists any $H \in \mathcal{H}$ such (F, ξ) partially induces H, include one such H in $\mathcal{H}_I$. Note that $|\mathcal{H}_I|$ is at most the number of summary functions with parameters I, which is finite.

Define the compactness functions $\Psi_{\mathcal{H}}\colon \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ as follows. Let $\Psi_{\mathcal{H}}(D, N)$ be the largest positive integer ℓ such that there exists a parameter list $I \in \mathcal{I}_p$ satisfying $\deg I \leq D$ and $\|I\| \leq N$ such that a pattern of the form $(\overline{\boldsymbol{L}}^\ell, \psi)$ exists in $\mathcal{H}_I$.

Now we set several parameters. Define non-increasing functions $\eta\colon \mathbb{Z}_{>0}\times\mathbb{Z}_{>0}\to\mathbb{Z}_{>0}$ by

$$\eta(D,N) := \frac{1}{40}\left(\frac{\epsilon}{12N|\mathcal{S}|}\right)^{p^{\Psi_{\mathcal{H}}(D,N)}},$$

$\beta\colon \mathbb{Z}_{>0}\times\mathbb{Z}_{>0}\to\mathbb{Z}_{>0}$ by

$$\beta(D,N) := \min_{I\in\mathcal{I}_p:\deg I\le D,\|I\|\le N} \beta_{patch}(p,|\mathcal{S}|,I,\Psi_{\mathcal{H}}(D,N)),$$

$\theta\colon \mathbb{Z}_{>0}\times\mathbb{Z}_{>0}\to\mathbb{Z}_{>0}$ by

$$\theta(D,N) := \frac{\beta(D,N)}{40}\left(\frac{\epsilon}{8|\mathcal{S}|}\right)^{p^{\Psi_{\mathcal{H}}(D,N)}},$$

and $\alpha\colon \mathbb{Z}_{>0}\times\mathbb{Z}_{>0}\to(0,1)$ by

$$\alpha(D,N) := \frac{\beta(D,N)}{10N^{2p^{\Psi_{\mathcal{H}}(D,N)}}}$$

Then define non-decreasing functions $r\colon \mathbb{Z}_{>0}\times\mathbb{Z}_{>0}\to\mathbb{Z}_{>0}$ by

$$r(D,N) := \max\{r_{equi}(p,D,\alpha(D,N)), \max_{I\in\mathcal{I}_p:\deg I\le D,\|I\|\le N} r_{patch}(p,|\mathcal{S}|,I,\Psi_{\mathcal{H}}(D,N))(D,N)+p\left\lceil\log_p N\right\rceil\}$$

and $d\colon \mathbb{Z}_{>0}\times\mathbb{Z}_{>0}\to\mathbb{Z}_{>0}$ by

$$d(D,N) := p^{\Psi_{\mathcal{H}}(D,N)}.$$

Define parameters

$$\zeta := \frac{\epsilon}{16|\mathcal{S}|} \qquad \text{and} \qquad c_0 := \left\lceil\log_p(2/\epsilon)\right\rceil.$$

Then define

$$C_{max} := C_{reg}(p,|\mathcal{S}|,c_0,\zeta,\eta,\theta,d,r),$$
$$D_{max} := D_{reg}(p,|\mathcal{S}|,c_0,\zeta,\eta,\theta,d,r),$$

$$n_{min} := \max\left\{n_{reg}(p,c_0,\zeta), \max_{I\in\mathcal{I}_p:\ \deg I\le D_{max},\|I\|\le C_{max}} n_{patch}(p,|\mathcal{S}|,I,\Psi_{\mathcal{H}}(D_{max},C_{max}))+\left\lceil\log_p C_{max}\right\rceil\right\}.$$

Finally, define

$$\delta(\epsilon,\mathcal{H}) := \min\left\{\frac{\beta(D_{max},C_{max})}{40}\left(\frac{\epsilon}{4C_{max}|\mathcal{S}|}\right)^{p^{\Psi_{\mathcal{H}}(D_{max},C_{max})}}, p^{-n_{min}\cdot\Psi_{\mathcal{H}}(D_{max},C_{max})}\right\}$$

and

$$\mathcal{H}_\epsilon := \bigcup_{I\in\mathcal{I}_p:\|I\|\le C_{max},\deg I\le D_{max}} \mathcal{H}_I.$$

Since the union is over a finite set of I, we have $\mathcal{H}_\epsilon$ is finite. We will show that this choice of δ and $\mathcal{H}_\epsilon$ satisfies the desired conclusion.

Let V be a finite-dimensional $\mathbb{F}_p$-vector space and $f\colon V\to\mathcal{S}$ be a projective function with H-density at most $\delta(\epsilon,\mathcal{H})$ for every $H\in\mathcal{H}_\epsilon$. Our goal is to produce a projective recoloring $g\colon V\to\mathcal{S}$ that agrees with f on all but an at most ϵ-fraction of V that has no generic H-instances for every $H\in\mathcal{H}$.

First note that if $\dim V < n_{min}$, the theorem is easily follows. This is because for the pattern $H=(\overline{\boldsymbol{L}}^\ell,\psi)$, if there exists an H-instance in f, then the H-density in f is at least $1/|V|^\ell$. Thus taking $g=f$ and noticing that we chose $\delta(\epsilon,\mathcal{H})\le p^{-n_{min}\cdot\Psi_{\mathcal{H}}(C_{max},D_{max})}$, the theorem holds in this case.

Now assume that $\dim V\ge n_{min}$. We apply Theorem 4.6 to the functions $\{1_{f^{-1}(c)}\}_{c\in\mathcal{S}}$ with parameters $p,|\mathcal{S}|,c_0,\zeta,\eta,\theta,d,r$. This produces a polynomial factor $\mathfrak{B}$ and a refinement $\mathfrak{B}'$ both on V with parameters I and I' and a subatom selection function $s\colon A_I\to A_{I'}$ satisfying several other desirable properties.

As above, define $\tilde{A}_I \subset A_I$ to be the set $\tilde{A}_I := \{a \in A_I : a_{1,0} = 0\}$. We call the atoms $a \in \tilde{A}_I$ *irregular* and the remaining atoms $a \in A_I \setminus \tilde{A}_I$ *regular*. Define $\tilde{V}$ to be the codimension-$I_{1,0}$ subspace of V that is the common zero set of all $I_{1,0}$ linear polynomials defining $\mathfrak{B}$. The irregular atoms of $\mathfrak{B}$ exactly consist of $\tilde{V}$, i.e., $\mathfrak{B}^{-1}(\tilde{A}_I) = \tilde{V}$.

Define $\tilde{\mathfrak{B}}$ to be the polynomial factor on $\tilde{V}$ defined by the restrictions of the homogeneous non-classical polynomials that define $\mathfrak{B}$ to $\tilde{V}$, except for the linear polynomials (which restrict to the zero function). Let $\tilde{I} \in \mathcal{I}_p$ be the parameter list of $\tilde{\mathfrak{B}}$ ($\tilde{I}_{1,0} = 0$ and $\tilde{I}_{d,k} = I_{d,k}$ otherwise). Also define $\tilde{\mathfrak{B}}'$ to be the polynomial factor on $\tilde{V}$ defined by the restrictions of the homogeneous non-classical polynomial that define $\mathfrak{B}'$ to $\tilde{V}$, except for the linear polynomials that also define $\mathfrak{B}$. Let $\tilde{I}' \in \mathcal{I}_p$ be the parameter list of $\tilde{\mathfrak{B}}'$ ($\tilde{I}'_{1,0} = I'_{1,0} - I_{1,0}$ and $\tilde{I}'_{d,k} = I'_{d,k}$ otherwise). Note that by Lemma 3.16 and our definition of r, we have

$$\operatorname{rank} \tilde{\mathfrak{B}} \geq r_{patch}(p, |\mathcal{S}|, \tilde{I}, \Psi_{\mathcal{H}}(\deg \mathfrak{B}, \|\mathfrak{B}\|))(\deg \mathfrak{B}, \|\mathfrak{B}\|),$$

$$\operatorname{rank} \tilde{\mathfrak{B}}' \geq r_{patch}(p, |\mathcal{S}|, \tilde{I}, \Psi_{\mathcal{H}}(\deg \mathfrak{B}, \|\mathfrak{B}\|))(\deg \mathfrak{B}', \|\mathfrak{B}'\|).$$

We will "clean up" the regular atoms by removing low-density colors in a projective manner. We will "patch" the irregular atoms by replacing the coloring by a new coloring $\Xi_{\xi,\iota,\tilde{\mathfrak{B}}}$ for some projective $\xi \colon \mathbb{F}_p \times A_{\tilde{I}} \to \mathcal{S}$ and $\iota \colon \tilde{V} \xrightarrow{\sim} \mathbb{F}_p^{\dim \tilde{V}}$.

Note that to check that the recoloring $g \colon V \to \mathcal{S}$ is projective, it suffices to check this fact separately on $\tilde{V}$ and on $V \setminus \tilde{V}$.

Clean up regular atoms: For each $a \in (A_I \setminus \tilde{A}_I)$, say that a color $c \in \mathcal{S}$ is *high-density in* a if it appears in $\mathfrak{B}'^{-1}(s(a))$ with density at least $\epsilon/(4|\mathcal{S}|)$. Say that a color is *low-density in* a otherwise.

First note that a basic property of subatom selection functions, Lemma 3.25(ii), is the following. For $a \in A_I$ and $b \in \mathbb{F}_p^\times$, we have $b \cdot s(a) = s(b \cdot a)$. Combined with the projectiveness of f, this implies that for a color $c \in \mathcal{S}$ and $b \in \mathbb{F}_p^\times$, the c-density in $\mathfrak{B}'^{-1}(s(a))$ is the same as the $(b \cdot c)$-density in $\mathfrak{B}^{-1}(s(b \cdot a))$. Thus a color c is high-density in a if and only if $b \cdot c$ is high density in $b \cdot a$.

We pick a single high-density color $c_a \in \mathcal{S}$ for each regular atom $a \in (A_I \setminus \tilde{A}_I)$ in a projective way, namely such that $b \cdot c_a = c_{b \cdot a}$ for all $a \in (A_I \setminus \tilde{A}_I)$ and $b \in \mathbb{F}_p^\times$. By the argument in the above paragraph, this is possible.

Now we define our recoloring of the regular atoms, $g \colon (V \setminus \tilde{V}) \to \mathcal{S}$ as follows. For each $a \in (A_I \setminus \tilde{A}_I)$ and $x \in \mathfrak{B}^{-1}(a)$, we define $g(x) := f(x)$ unless $f(x)$ is low-density in a, in which case we define $g(x) := c_a$. Note that g is a projective function. Furthermore we claim that g differs from f on at most an $(\epsilon/2)$-fraction of V.

By Theorem 4.6(vi), for all but at most a ζ-fraction of $a \in A_I$, the c-density in $\mathfrak{B}^{-1}(a)$ and the c-density in $\mathfrak{B}'^{-1}(s(a))$ differ by at most ζ for all $s \in \mathcal{S}$. Thus for most atoms a, each low-density color appears in $\mathfrak{B}^{-1}(a)$ with density at most $\epsilon/(4|\mathcal{S}|) + \zeta$, so f and g differ on at most an $(\epsilon/4 + \zeta|\mathcal{S}|)$-fraction of these atoms. The functions f and g may differ completely on the other atoms, but there are at most $\zeta\|\mathfrak{B}\|$ of these. It follows by equidistribution, Theorem 3.19, and the rank bound on $\mathfrak{B}$ that each atom of $\mathfrak{B}$ is at most an $(\|\mathfrak{B}\|^{-1} + \alpha(\deg \mathfrak{B}, \|\mathfrak{B}\|))$-fraction of V. Putting this all together, we see that g differs from f on at most the following fraction of V

$$\zeta\|\mathfrak{B}\| \left(\frac{1}{\|\mathfrak{B}\|} + \alpha(\deg \mathfrak{B}, \|\mathfrak{B}\|) \right) + \left(\frac{\epsilon}{4|\mathcal{S}|} + \zeta \right) |\mathcal{S}| < \frac{\epsilon}{2}.$$

Patch irregular atoms: We define $\tilde{\mathcal{H}}$ to be the set of all $\mathcal{S}$-colored I-labeled patterns that are defined by a full dimension system of linear forms in at most $\Psi_{\mathcal{H}}(\deg I, \|I\|)$ variables and whose relative density in $\tilde{\mathfrak{B}}'^{-1}(A_{\tilde{I}} \times \{0\})$ is less than $\beta_{patch}(p, |\mathcal{S}|, \tilde{I}, \Psi_{\mathcal{H}}(\deg I, \|I\|))$.

We apply our patching result, Theorem 5.9, to the set $\tilde{\mathcal{H}}$. Our definitions are exactly such that $f|_{\tilde{V}}$ with $\tilde{\mathfrak{B}}, \tilde{\mathfrak{B}}'$ demonstrate that Theorem 5.9(b) does not hold. In particular, we checked the

rank assumptions on $\tilde{\mathfrak{B}}$ and $\tilde{\mathfrak{B}}'$ above when they were defined. Furthermore, we assumed that $\dim V \geq n_{min}$, which implies that $\dim \tilde{V} \geq n_{patch}(p, |\mathcal{S}|, \tilde{I}', \Psi_{\mathcal{H}}(\deg I, \|I\|))$. Finally we defined $\tilde{\mathcal{H}}$ to be the set of patterns which appear with very low density in $\tilde{\mathfrak{B}}'^{-1}(A_{\tilde{I}} \times \{0\})$. Thus we conclude that Theorem 5.9(a) holds. This means that there exists a projective $\xi\colon \mathbb{F}_p \times A_{\tilde{I}} \to \mathcal{S}$ that does not canonically induce $\tilde{\mathcal{H}}$. In particular, this means that for any fixed isomorphism $\iota\colon \tilde{V} \xrightarrow{\sim} \mathbb{F}_p^{\dim \tilde{V}}$, there are no generic H-instances in $(\Xi_{\xi,\iota,\tilde{\mathfrak{B}}})$ for any $H \in \tilde{\mathcal{H}}$.

We complete our definition of $g\colon V \to \mathcal{S}$ by defining $g(x) := \Xi_{\xi,\iota,\tilde{\mathfrak{B}}}(x)$ for all $x \in \tilde{V}$. To conclude this portion of the proof, we make sure that $|\tilde{V}| \leq (\epsilon/2)|V|$. By assumption,

$$|\tilde{V}|/|V| = p^{-I_{1,0}} \leq p^{-c_0} = \epsilon/2,$$

as desired.

Proof of correctness: We claim that g has no generic H-instances for each $H \in \mathcal{H}$. Define $F\colon (A_I \setminus \tilde{A}_I) \to 2^{\mathcal{S}} \setminus \{0\}$ to map a to the set of high-density colors in a and recall the projective function $\xi\colon \mathbb{F}_p \times A_{\tilde{I}} \to \mathcal{S}$ defined in the "patch irregular atoms" section. Now suppose that the desired conclusion does not hold, i.e., that there is a generic H'-instance in g for some $H' \in \mathcal{H}$. By the construction of g, this means that (F, ξ) partially induces H'. By the definition of the $\mathcal{H}_I$, this means that there is some $H \in \mathcal{H}_I \subseteq \mathcal{H}_\epsilon$ so that (F, ξ) also partially induces H. We will reach a contradiction by showing that this implies that the H-density in f is larger than $\delta(\epsilon, \mathcal{H})$.

Say that $H = (\overline{\boldsymbol{L}}^\ell, \psi)$ (note that $\ell \leq \Psi_{\mathcal{H}}(\deg I, \|I\|)$). Since (F, ξ) partially induces H, this implies that there exists a tuple of atoms $\boldsymbol{a} \in \Phi_I(\overline{\boldsymbol{L}}^\ell) \subseteq A_I^{E_\ell}$ with several desirable properties. Define $J := \{i \in E_\ell : a_i \in \tilde{A}_I\}$. Recalling that $\tilde{A}_I$ is just the set of atoms whose linear part is zero, we can conclude that $J \subseteq E_\ell \subset \mathbb{F}_p^\ell$ is the intersection of E_ℓ with some linear subspace $U \leq \mathbb{F}_p^\ell$ of dimension $\ell' \leq \ell$. This means that the system $(L_i^\ell)_{i \in J}$ is equivalent to the system $(L_i^{\ell'})_{i \in J}$ where now we view $J \subset U \simeq \mathbb{F}_p^{\ell'}$. Define, $H_J := ((L_i^{\ell'})_{i \in E_\ell}, \psi|_J, (a_i)_{i \in J})$, an $\mathcal{S}$-colored I-labeled pattern. By Lemma 3.22, we see that H_j is a full dimension pattern.

The first property, that $\boldsymbol{a}$ is $\overline{\boldsymbol{L}}^\ell$-consistent, implies that $s(\boldsymbol{a})$ is also $\overline{\boldsymbol{L}}^\ell$-consistent, by Lemma 3.25(iii). The second property implies that for each $i \in (E_\ell \setminus J)$, the color $\psi(i)$ is high-density in a_i. The third property, together with our definition of ξ implies that in $(f, \mathfrak{B})$, the relative density of H_J in $\tilde{\mathfrak{B}}'^{-1}(A_{\tilde{I}} \times \{0\})$ is at least $\beta_{patch}(p, |\mathcal{S}|, \tilde{I}, \Psi_{\mathcal{H}}(\deg I, \|I\|) \geq \beta(\deg I, \|I\|))$.

Now we put everything together as follows. Write $f^{(i)}$ for $1_{f^{-1}(\psi(i))}$. There is a decomposition $f^{(i)} = f^{(i)}_{str} + f^{(i)}_{sml} + f^{(i)}_{psr}$ given by Theorem 4.6 for each $i \in E_\ell$. Let p_1 be the H-density in f. We lower bound p_1 as follows.

$$\begin{aligned}
p_1 &= \mathbb{E}_{\boldsymbol{x}} \left[\prod_{i \in E_\ell} f^{(i)}(L_i^\ell(\boldsymbol{x})) \right] \\
&= \mathbb{E}_{\boldsymbol{x}} \left[\prod_{i \in J} f^{(i)}(L_i^\ell(\boldsymbol{x})) \prod_{i \in E_\ell \setminus J} \left(f^{(i)}_{str}(L_i^\ell(\boldsymbol{x})) + f^{(i)}_{sml}(L_i^\ell(\boldsymbol{x})) + f^{(i)}_{psr}(L_i^\ell(\boldsymbol{x})) \right) \right] \\
&\geq \mathbb{E}_{\boldsymbol{x}} \left[\prod_{i \in J} f^{(i)}(L_i^\ell(\boldsymbol{x})) \prod_{i \in E_\ell \setminus J} \left(f^{(i)}_{str}(L_i^\ell(\boldsymbol{x})) + f^{(i)}_{sml}(L_i^\ell(\boldsymbol{x})) \right) \right] - 3^{|E_\ell|} \eta(\deg \mathfrak{B}', \|\mathfrak{B}'\|).
\end{aligned}$$

The inequality follows from Theorem 4.6(iii), the counting lemma (Eq. (3.1)), and the fact that the complexity of H is at most $d(\deg I, \|I\|) = p^{\Psi_{\mathcal{H}}(\deg I, \|I\|)}$.

Write p_2 for the expectation in the last line above. By Theorem 4.6(iv), the expression inside the expectation is non-negative so we can restrict the expectation to $\boldsymbol{x}$ such that $\overline{\boldsymbol{L}}^\ell(\boldsymbol{x}) = s(\boldsymbol{a})$.

Thus

$$p_2 \geq \mathbb{E}_{\boldsymbol{x}}\left[\prod_{i\in J} f^{(i)}(L_i^\ell(\boldsymbol{x}))1_{\mathfrak{B}'^{-1}(s(a_i))}(L_i^\ell(\boldsymbol{x}))\prod_{i\in E_\ell\setminus J}\left(f_{str}^{(i)}(L_i^\ell(\boldsymbol{x}))+f_{sml}^{(i)}(L_i^\ell(\boldsymbol{x}))\right)1_{\mathfrak{B}'^{-1}(s(a_i))}(L_i^\ell(\boldsymbol{x}))\right].$$

Write p_3 for the expectation in the last line above. Expanding the product, there are at most $2^{|E_\ell|}$ terms involving $f_{sml}^{(j)}$ for some $j \in E_\ell \setminus J$. Each of these is bounded in magnitude by

$$\mathbb{E}_{\boldsymbol{x}}\left[\left|f_{sml}^{(j)}(L_j(\boldsymbol{x}))\right|\prod_{i\in E_\ell}1_{\mathfrak{B}'^{-1}(s(a_i))}(L_i^\ell(\boldsymbol{x}))\right].$$

By applying a change of coordinates, we can assume that $L_j^\ell(\boldsymbol{x}) = x_1$. Then by the Cauchy-Schwarz inequality, the square of the above expression is bounded by

$$\mathbb{E}_{x_1}\left[\left|f_{sml}^{(j)}(x_1)\right|^2 1_{\mathfrak{B}'^{-1}(s(a_j))}(x_1)\right]\mathbb{E}_{x_1}\left[1_{\mathfrak{B}'^{-1}(s(a_j))}(x_1)\mathbb{E}_{x_2,\ldots,x_\ell}\left[\prod_{i\neq j}1_{\mathfrak{B}'^{-1}(s(a_i))}(L_i^\ell(\boldsymbol{x}))\right]^2\right].$$

The first term is at most $\theta(\deg\mathfrak{B},\|\mathfrak{B}\|)^2\|1_{\mathfrak{B}'^{-1}(s(a_j))}\|_2^2$ by Theorem 4.6(vi) and the fact that a_j is a regular atom for $j \in E_\ell \setminus J$. The second term can be counted by equidistribution applied to the system $\boldsymbol{L}'$ of $2|E_k|-1$ linear forms in $2\ell-1$ variables defined as follows. Set

$$L'_j(x_1,x_2,\ldots,x_\ell,x'_2,\ldots,x'_\ell) := x_1,$$

and for $i \in E_\ell \setminus \{j\}$, define

$$L'_{i,1}(x_1,x_2,\ldots,x_\ell,x'_2,\ldots,x'_\ell) := L_i^\ell(x_1,x_2\ldots,x_\ell),$$

$$L'_{i,2}(x_1,x_2,\ldots,x_\ell,x'_2,\ldots,x'_\ell) := L_i^\ell(x_1,x'_2,\ldots,x'_\ell).$$

By Lemma 3.23, we know that $\|\mathfrak{B}'\|\cdot|\Phi_{I'}(\boldsymbol{L}')| = |\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|^2$ (see also [7, Lemma 5.13].)

Thus by equidistribution (Theorem 3.19) and the rank bound on $\mathfrak{B}'$, we have the second term is at most

$$\frac{1}{|\Phi_{I'}(\boldsymbol{L}')|}+\alpha(\deg\mathfrak{B}',\|\mathfrak{B}'\|) = \frac{\|\mathfrak{B}'\|}{|\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|^2}+\alpha(\deg\mathfrak{B}',\|\mathfrak{B}'\|) \leq \frac{2\|\mathfrak{B}'\|}{|\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|^2}.$$

Applying equidistribution again we have that the first term is at most

$$\theta(\deg\mathfrak{B},\|\mathfrak{B}\|)^2\left(\frac{1}{\|\mathfrak{B}'\|}+\alpha(\deg\mathfrak{B}',\|\mathfrak{B}'\|)\right) \leq \frac{2\theta(\deg\mathfrak{B},\|\mathfrak{B}\|)^2}{\|\mathfrak{B}'\|}.$$

Combining these bounds and summing over all terms that contain some f_{sml}^j, we see that

$$p_3 \geq \mathbb{E}_{\boldsymbol{x}}\left[\prod_{i\in J} f^{(i)}(L_i^\ell(\boldsymbol{x}))1_{\mathfrak{B}'^{-1}(s(a_i))}(L_i^\ell(\boldsymbol{x}))\prod_{i\in E_\ell\setminus J} f_{str}^{(i)}(L_i^\ell(\boldsymbol{x}))1_{\mathfrak{B}'^{-1}(s(a_i))}(L_i^\ell(\boldsymbol{x}))\right]-2^{|E_\ell|+1}\frac{\theta(\deg\mathfrak{B},\|\mathfrak{B}\|)}{|\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|}.$$

Write p_4 for the expectation above. The quantity $f_{str}^{(i)}(L_i^\ell(\boldsymbol{x}))$ is the density of $\psi(i)$ in the atom of $\mathfrak{B}'$ that $L_i^\ell(\boldsymbol{x})$ lies in. When $\mathfrak{B}(L_i^\ell(\boldsymbol{x})) = s(a_i)$, the fact that $\psi(i)$ is high density in a_i for all $i \in E_\ell \setminus J$ implies that

$$p_4 \geq \left(\frac{\epsilon}{4|\mathcal{S}|}\right)^{|E_\ell|}\mathbb{E}_{\boldsymbol{x}}\left[\prod_{i\in J} f^{(i)}(L_i^\ell(\boldsymbol{x}))1_{\mathfrak{B}'^{-1}(s(a_i))}(L_i^\ell(\boldsymbol{x}))\prod_{i\in E_\ell\setminus J}1_{\mathfrak{B}'^{-1}(s(a_i))}(L_i^\ell(\boldsymbol{x}))\right].$$

Write p_5 for the expectation above. We write $\boldsymbol{L}_J := (L_i^\ell)_{i\in J}$. By assumption, we know that

$$\mathbb{E}_{\boldsymbol{x}}\left[\prod_{i\in J} f^{(i)}(L_i^\ell(\boldsymbol{x}))1_{\mathfrak{B}'^{-1}(s(a_i))}(L_i^\ell(\boldsymbol{x}))\right] \ge \frac{\beta(\deg I, \|I\|)}{|\Phi_{I'}(\boldsymbol{L}_J)|}. \tag{6.1}$$

We want to use this inequality to show that p_5 is at least on the order of $\beta(\deg I, \|I\|)/|\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|$. For simplicity, write $\beta := \beta(\deg I, \|I\|)$ in the rest of this argument.

By applying a change of coordinates, we can assume that $\boldsymbol{L}_J$ only depends on $x_1, \dots, x_{\ell'}$ and is independent of $x_{\ell'+1}, \dots, x_\ell$. To lower bound p_5, we want to show that each tuple $(x_1, \dots, x_{\ell'})$ that lies in certain atoms extends to a tuple $(x_1, \dots, x_\ell)$ that still lies in certain atoms in approximately the same number of ways. We do this by a Cauchy-Schwarz argument. Define $\boldsymbol{L}''$ to be the following system of $2|E_\ell| - |J|$ linear forms in $2\ell - \ell'$ variables. For $i \in J$, set

$$L_i''(x_1, \dots, x_\ell, x'_{\ell'+1}, \dots, x'_\ell) := L_i^{\ell'}(x_1, \dots, x_{\ell'}),$$

and for $i \in E_\ell \setminus J$, define

$$L''_{i,1}(x_1, \dots, x_\ell, x'_{\ell'+1}, \dots, x'_\ell) := L_i^\ell(x_1, \dots, x_\ell),$$
$$L''_{i,1}(x_1, \dots, x_\ell, x'_{\ell'+1}, \dots, x'_\ell) := L_i^\ell(x_1, \dots, x_{\ell'}, x_{\ell'+1}, \dots, x_\ell).$$

By Lemma 3.23, we know that $|\Phi_{I'}(\boldsymbol{L}_J)| \cdot |\Phi_{I'}(\boldsymbol{L}'')| = |\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|^2$.

Define $S \subseteq V^{\ell'}$ to be the set of tuples $\boldsymbol{x} = (x_1, \dots, x_{\ell'})$ such that $\mathfrak{B}'(L_i^\ell(\boldsymbol{x})) = s(a_i)$ for each $i \in J$. For $\boldsymbol{x} \in S$, let $c_{\boldsymbol{x}}$ be the number of tuples $\boldsymbol{x}' = (x_1, \dots, x_\ell)$ such that $\mathfrak{B}'(L_i^\ell(\boldsymbol{x})) = s(a_i)$ for each $i \in E_\ell$. By applying equidistribution (Theorem 3.19) to $\boldsymbol{L}_J$ and $\overline{\boldsymbol{L}}^\ell$ and $\boldsymbol{L}''$, we find that

$$|S| = \sum_{\boldsymbol{x}\in S} 1 \le (1 + \beta/10)\frac{|V|^{\ell'}}{|\Phi_{I'}(\boldsymbol{L}_J)|}, \tag{6.2}$$

$$\sum_{\boldsymbol{x}\in S} c_{\boldsymbol{x}} \ge (1 - \beta/10)\frac{|V|^{\ell}}{|\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|}, \tag{6.3}$$

$$\sum_{\boldsymbol{x}\in S} c_{\boldsymbol{x}}^2 \le (1 + \beta/10)\frac{|V|^{2\ell-\ell'}}{|\Phi_{I'}(\boldsymbol{L}'')|} = (1 + \beta/10)\frac{|V|^{2\ell-\ell'}|\Phi_{I'}(\boldsymbol{L}_J)|}{|\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|^2}. \tag{6.4}$$

Define $T \subseteq S \subseteq V^{\ell'}$ to be the set of tuples $\boldsymbol{x} = (x_1, \dots, x_{\ell'})$ such that $\mathfrak{B}'(L_i^\ell(\boldsymbol{x})) = s(a_i)$ and $f^{(i)}(L_i^\ell(\boldsymbol{x})) = 1$ for each $i \in J$. Eq. (6.1) implies that

$$|T| \ge \beta\frac{|V|^{\ell'}}{|\Phi_{I'}(\boldsymbol{L}_J)|}. \tag{6.5}$$

We express p_5 as follows.

$$p_5 = \frac{1}{|V|^\ell}\sum_{\boldsymbol{x}\in T} c_{\boldsymbol{x}} = \frac{1}{|V|^\ell}\left(\sum_{\boldsymbol{x}\in S} c_{\boldsymbol{x}} - \sum_{\boldsymbol{x}\in S\setminus T} c_{\boldsymbol{x}}\right) \ge (1 - \beta/10)\frac{1}{|\Phi_{I'}(\boldsymbol{L}_J)|} - \frac{1}{|V|^\ell}\sum_{\boldsymbol{x}\in S\setminus t} c_{\boldsymbol{x}}.$$

Then combining Eq. (6.2), Eq. (6.4), Eq. (6.5) with the Cauchy-Schwarz inequality gives

$$\left(\sum_{\boldsymbol{x}\in S\setminus T} c_{\boldsymbol{x}}\right)^2 \le |S \setminus T| \cdot \sum_{\boldsymbol{x}\in S\setminus T} c_{\boldsymbol{x}}^2 \le (1 - 4\beta/5)\frac{|V|^{2\ell}}{|\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|^2}.$$

Taking the square root and combining the above two inequalities gives

$$p_5 \ge \frac{\beta}{10}\frac{1}{|\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|}.$$

Combining all the above inequalities we see that p_1, the H-density in f, is bounded by

$$p_1 \geq \left(\left(\frac{\epsilon}{4|\mathcal{S}|}\right)^{|E_\ell|} \frac{\beta(\deg \mathfrak{B}, \|\mathfrak{B}\|)}{10} - 2^{|E_\ell|+1}\theta(\deg \mathfrak{B}, \|\mathfrak{B}\|)\right) \frac{1}{|\Phi_{I'}(\overline{\boldsymbol{L}}^\ell)|} - 3^{|E_\ell|}\eta(\deg \mathfrak{B}', \|\mathfrak{B}'\|) > \delta(\epsilon, \mathcal{H}).$$

This provides the desired contradiction. Therefore we conclude that the recoloring $g\colon V \to \mathcal{S}$ has no generic H-instances for every $H \in \mathcal{H}$. □

Proof of Theorem 2.4. Define $\overline{\mathcal{S}} := \mathcal{S}^{\mathbb{F}_p^\times}$ with $\mathbb{F}_p^\times$-action defined by

$$b' \cdot (c_b)_{b\in\mathbb{F}_p^\times} := (c_{b'b})_{b\in\mathbb{F}_p^\times}.$$

First we partition

$$\mathcal{H} = \bigsqcup_{c\in\mathcal{S}\sqcup\{0\}} \mathcal{H}_c$$

as follows. If $L_i \equiv 0$ for some $i \in [m]$, place H in the set $\mathcal{H}_{\psi(i)}$. Otherwise, place H in $\mathcal{H}_0$. (Without loss of generality, we can assume that no pattern in $\mathcal{H}$ has multiple linear forms that are identically equal to 0.)

Now we define sets $\overline{\mathcal{H}}_c$ of $\overline{\mathcal{S}}$-colored patterns for each $c \in \mathcal{S}$. Let $H = (\boldsymbol{L}, \psi) \in \mathcal{H}_0 \cup \mathcal{H}_c$ be an $\mathcal{S}$-colored pattern over $\mathbb{F}_p$ consisting of m linear forms in ℓ variables. We can write $\boldsymbol{L} = (L_i^\ell)_{i\in J}$ for some set $J \subseteq \mathbb{F}_p^\ell$ of size m. We convert H to a set of patterns defined by the system $\overline{\boldsymbol{L}}^\ell = (L_i^\ell)_{i\in E_\ell}$ as follows. (See Definition 2.7 for the definitions of L_i^ℓ and E_ℓ.) For each function $\overline{\psi}\colon E_\ell \to \overline{\mathcal{S}}$ that satisfies $\overline{\psi}(i)_b = \psi(bi)$ whenever $i \in E_\ell \subseteq \mathbb{F}_p^\ell$ and $b \in \mathbb{F}_p^\times$ are such that $bi \in J \subseteq \mathbb{F}_p^\ell$, we add $(\overline{\boldsymbol{L}}^\ell, \overline{\psi})$ to $\overline{\mathcal{H}}_c$.

For each $c \in \mathcal{S}$, we apply Theorem 2.8 to $\overline{\mathcal{H}}_c$ with parameter ϵ. This produces a finite subset $\overline{\mathcal{H}}_{c,\epsilon} \subseteq \overline{\mathcal{H}}_c$ and $\delta_c = \delta(\epsilon, \overline{\mathcal{H}}_c) > 0$ with several desirable properties.

Let $\mathcal{H}_\epsilon \subseteq \mathcal{H}$ be the finite subset consisting of all patterns H such that some pattern $\overline{H}$ corresponding to H lies in $\overline{\mathcal{H}}_{c,\epsilon}$ for some $c \in \mathcal{S}$. Let $\delta = \min_{c\in\mathcal{S}} \delta_c > 0$. We claim that $\mathcal{H}_\epsilon, \delta$ satisfy the desired conclusion.

Let V be a finite-dimensional $\mathbb{F}_p$-vector space. Let $f\colon V \to \mathcal{S}$ be a function. Suppose that the H-density in f is at most δ for every $H \in \mathcal{H}_\epsilon$. Define $\overline{f}\colon V \to \overline{\mathcal{S}}$ by

$$\overline{f}(x) := (f(bx))_{b\in\mathbb{F}_p^\times}.$$

Note that $\overline{f}$ is a projective function. Furthermore, we claim that the $\overline{H}$-density in $\overline{f}$ is at most δ for all $\overline{H} \in \overline{\mathcal{H}}_{f(0)}$. This is true simply because if $\boldsymbol{x} = (x_1, \ldots, x_\ell) \in V^\ell$ is an $\overline{H}$-instance in $\overline{f}$, then $\boldsymbol{x}$ is also an H-instance in f where $\overline{H} \in \overline{\mathcal{H}}_{f(0)}$ is any patterns corresponding to $H \in \mathcal{H}_0 \sqcup \mathcal{H}_{f(0)}$.

Thus by assumption, there exists a projective recoloring $\overline{g}\colon V \to \overline{\mathcal{S}}$ such that $\overline{g}$ agrees with $\overline{f}$ on all but an at most ϵ-fraction of V and $\overline{g}$ has no generic $\overline{H}$-instances for every $\overline{H} \in \overline{\mathcal{H}}_{f(0)}$.

Define $g\colon V \to \mathcal{S}$ by $g(x) := \overline{g}(x)_1$ for $x \neq 0$ and $g(0) = f(0)$. Note that g agrees with f on all but an at most ϵ-fraction of V. Furthermore, note that g has no H-instances for $H \in \mathcal{H}_c$ with $c \neq f(0)$ since $f(0) = g(0)$. Finally, g has no generic H-instances for $H \in \mathcal{H}_0 \sqcup \mathcal{H}_{f(0)}$ since any such generic H-instance in g is a generic $\overline{H}$-instance in $\overline{g}$ for some $\overline{H}$ corresponding H in $\overline{\mathcal{H}}_{f(0)}$, which we assumed was not the case. □

7. Proof of property testing results

Proof of Theorem 1.3. It follows from [9, Theorem 10] that a linear-invariant property is testable only if it is semi subspace-hereditary.

Now suppose that $\mathcal{P}$ is a linear-invariant semi subspace-hereditary property. By definition, there exists a subspace-hereditary property $\mathcal{Q}$ such that

(i) every function satisfying $\mathcal{P}$ also satisfies $\mathcal{Q}$;

(ii) for all $\epsilon > 0$, there exists $N(\epsilon)$ such that if $f\colon V \to \mathcal{S}$ satisfies $\mathcal{Q}$ and is ϵ-far from satisfying $\mathcal{P}$, then $\dim V < N(\epsilon)$.

We define $\mathcal{H}$ a (possibly infinite) set of $\mathcal{S}$-colored patterns. For each $f\colon \mathbb{F}_p^\ell \to \mathcal{S}$ that does not satisfy $\mathcal{Q}$, include $H = (\boldsymbol{L}^\ell, f)$ in $\mathcal{H}$ ($\boldsymbol{L}^\ell$ is the system of linear forms that defines an ℓ-dimensional subspace, defined in Definition 2.7). Since $\mathcal{Q}$ is subspace-hereditary, it immediately follows that $\mathcal{Q}$ consists exactly of the functions with no generic H-instances for any $H \in \mathcal{H}$.

By Theorem 2.4, there exist a finite subset $\mathcal{H}_\epsilon \subseteq \mathcal{H}$ and some $\delta(\epsilon, \mathcal{H}) > 0$ such that the following holds. If $f\colon V \to \mathcal{S}$ has H-density at most $\delta(\epsilon, \mathcal{H})$ for every $H \in \mathcal{H}_\epsilon$, then f is ϵ-close to $\mathcal{Q}$. Define $\ell(\epsilon)$ to be the largest ℓ such that some pattern defined by the system $\boldsymbol{L}^\ell$ is present in $\mathcal{H}_\epsilon$.

Now we define the oblivious tester for $\mathcal{P}$. Given $\epsilon > 0$, the tester produces

$$d(\epsilon) := \max\left\{N(\epsilon/2), \lceil \log_p(2/\delta(\epsilon/2, \mathcal{H}))\rceil + \ell(\epsilon/2)\right\}.$$

Given a function $f\colon V \to \mathcal{S}$ our tester receives oracle access to $f|_U$ where

(i) if $\dim V \geq d(\epsilon)$, then U is a uniform random affine subspace of dimension $d(\epsilon)$;
(ii) else, $U = V$.

Our tester works as follows. If $\dim U < d(\epsilon)$ the tester accepts if $f|_U \in \mathcal{P}$. If $\dim U \geq d(\epsilon)$ the tester accepts in $f|_U \in \mathcal{Q}$.

Suppose $f \in \mathcal{P}$. If $\dim U < d(\epsilon)$, then $U = V$, so $f|_U = f \in \mathcal{P}$. Thus the tester always accepts in this case. In the other case, note that since $f \in \mathcal{P}$, it follows that $f \in \mathcal{Q}$, and since $\mathcal{Q}$ is subspace-hereditary, $f|_U \in \mathcal{Q}$. Thus the tester also always accepts in this case.

Now suppose that f is ϵ-far from $\mathcal{P}$. By the definition of $\mathcal{Q}$ we know that either $\dim V < N(\epsilon/2) \leq d(\epsilon)$ or f is $\epsilon/2$-far from $\mathcal{Q}$. Consider the action of the tester. If $\dim U < d(\epsilon)$, then $U = V$ so $f|_U = f \notin \mathcal{P}$. Thus the tester always rejects in this case. In the other case, note that since f is $\epsilon/2$-far from $\mathcal{Q}$, by assumption there is some $H \in \mathcal{H}_{\epsilon/2}$ such that f has H-density more than $\delta(\epsilon/2, \mathcal{H})$. Let $H = (\boldsymbol{L}^\ell, \psi)$ for some $\ell \leq \ell(\epsilon/2)$. We claim the fact that the H-density in f is large implies that there is at least a $\delta(\epsilon/2, \mathcal{H})/2$-fraction of ℓ-dimensional subspaces that f colors by ψ. (Note that this does not immediately follow since the H-density includes the contribution of H-instances that are not generic.) We can compute that the probability a uniform random $\boldsymbol{L}^\ell$-instance in V is not generic is at most $p^{\ell - \dim V}$. It follows that the fraction of ℓ-dimensional subspaces that f colors by ψ is at least

$$\delta(\epsilon/2, \mathcal{H}) - p^{\ell - \dim V} \geq \delta(\epsilon/2, \mathcal{H}) - p^{-\lceil \log_p(2/\delta(\epsilon/2, \mathcal{H}))\rceil} \geq \delta(\epsilon/2, \mathcal{H})/2.$$

We conclude that in this case the tester rejects with probability at least $\delta(\epsilon/2, \mathcal{H})/2$, as desired. □

Proof of Theorem 1.4. Suppose $\mathcal{P}$ is a linear-invariant property that is PO-testable. By definition, there exists some d, independent of ϵ and $\dim V$, such that to test $f\colon V \to \mathcal{S}$, the tester receives $f|_U$ where U is

(i) if $\dim V \geq d$, then U is a uniform random linear subspace of dimension d;
(ii) else, $U = V$.

We define $\mathcal{H}$ to be the set of patterns of the form $(\boldsymbol{L}^d, \psi)$ where $\psi\colon \mathbb{F}_p^d \to \mathcal{S}$ is a restriction that the tester rejects on ($\boldsymbol{L}^d$ is the pattern that defines a d-dimensional subspace, defined in Definition 2.7). We claim that for every $f\colon V \to \mathcal{S}$ with $\dim V \geq d$, it holds that $f \in \mathcal{P}$ if and only if f has no generic $\mathcal{H}$-instances. This claim suffices to prove that $\mathcal{H}$ is subspace-hereditary and locally characterized.

Suppose $f\colon V \to \mathcal{S}$ satisfies $\mathcal{P}$ and $\dim V \geq d$. By the definition of PO-testable, the tester must accept f with probability 1. Thus the tester must accept $f|_U$ for every $U \leq V$ of dimension d. This implies that f has no generic $\mathcal{H}$-instances. Now suppose that $f\colon V \to \mathcal{S}$ does not satisfy $\mathcal{P}$ and $\dim V \geq d$ holds. By definition, the tester must accept f with positive probability. Thus there

must be some $U \leq V$ of dimension d such that $f|_U$ rejects. This is equivalent to the fact that f contains a generic H-instance for some $H \in \mathcal{H}$, proving the desired result.

Now we show that every linear-invariant subspace-hereditary locally characterized property is testable. Suppose $\mathcal{P}$ is such a property. It follows that there is some d and (finite) $\mathcal{H}$ consisting of patterns of the form $(\boldsymbol{L}^d, \psi)$ such that for $f\colon V \to \mathcal{S}$ with $\dim V \geq d$, we have f satisfies $\mathcal{P}$ if and only if f has no generic $\mathcal{H}$-instances.

The PO-tester for $\mathcal{P}$ proceeds in the obvious way. The tester is given oracle access to $f|_U$ where U is

(i) if $\dim V \geq d$, then U is a uniform random linear subspace of dimension d;
(ii) else, $U = V$.

The tester accepts if and only if $f|_U \in \mathcal{P}$.

Suppose $f\colon V \to \mathcal{S}$ satisfies $\mathcal{P}$. If $\dim V < d$, then $f|_U = f \in \mathcal{P}$, so the tester accepts f. If $\dim V \geq d$, then by the fact that $\mathcal{P}$ is subspace-hereditary and locally characterized, it follows that $f|_U \in \mathcal{P}$ for all d-dimensional $U \leq V$. Thus the tester accepts f in this case as well.

Now suppose that $f\colon V \to \mathcal{S}$ is ϵ-far from $\mathcal{P}$. If $\dim V < d$, then $f|_U = f \notin \mathcal{P}$, so the tester rejects f. If $\dim V \geq d$, by Theorem 2.4, there must be some $H = (\boldsymbol{L}^d, \psi) \in \mathcal{H}$ such that the H-density in f is more than $\delta(\epsilon, \mathcal{H})$. We claim that this implies that there is a large fraction of d-dimensional subspaces that f colors by some $H \in \mathcal{H}$. (Note that this does not immediately follow since the H-density includes the contribution of H-instances that are not generic.) We can compute that at most a $p^{d-\dim V}$-fraction of H-instances are non-generic. Thus at least the fraction of d-dimensional subspaces that f colors by ψ is at least

$$\delta(\epsilon, \mathcal{H}) - p^{d-\dim V}.$$

This parameter is negative for small values of $\dim V$, so we can also use the fact that since f does not satisfy $\mathcal{P}$, there is at least one d-dimensional subspaces that is colored by $\mathcal{H}$. Thus the rejection probability of this tester is at least

$$\max\left\{\delta(\epsilon) - p^{d-\dim V}, p^{-d\cdot\dim V}\right\}.$$

Note that for $\dim V \geq d$, this parameter is uniformly bounded away from 0, as desired. □

References

[1] Noga Alon, Richard A. Duke, Hanno Lefmann, Vojtěch Rödl, and Raphael Yuster, *The algorithmic aspects of the regularity lemma*, J. Algorithms **16** (1994), 80–109.

[2] Noga Alon, Eldar Fischer, Michael Krivelevich, and Mario Szegedy, *Efficient testing of large graphs*, Combinatorica **20** (2000), 451–476.

[3] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron, *Testing Reed-Muller codes*, IEEE Trans. Inform. Theory **51** (2005), 4032–4039.

[4] Noga Alon and Asaf Shapira, *A characterization of the (natural) graph properties testable with one-sided error*, SIAM J. Comput. **37** (2008), 1703–1727.

[5] Vitaly Bergelson, Terence Tao, and Tamar Ziegler, *An inverse theorem for the uniformity seminorms associated with the action of* $\mathbb{F}_p^\infty$, Geom. Funct. Anal. **19** (2010), 1539–1596.

[6] Arnab Bhattacharyya, Victor Chen, Madhu Sudan, and Ning Xie, *Testing linear-invariant non-linear properties*, Theory Comput. **7** (2011), 75–99.

[7] Arnab Bhattacharyya, Eldar Fischer, Hamed Hatami, Pooya Hatami, and Shachar Lovett, *Every locally characterized affine-invariant property is testable*, STOC'13—Proceedings of the 2013 ACM Symposium on Theory of Computing, ACM, New York, 2013, pp. 429–435.

[8] Arnab Bhattacharyya, Eldar Fischer, and Shachar Lovett, *Testing low complexity affine-invariant properties*, Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA, 2012, pp. 1337–1355.

[9] Arnab Bhattacharyya, Elena Grigorescu, and Asaf Shapira, *A unified framework for testing linear-invariant properties*, Random Structures Algorithms **46** (2015), 232–260.

[10] Manuel Blum, Michael Luby, and Ronitt Rubinfeld, *Self-testing/correcting with applications to numerical problems*, Proceedings of the 22nd Annual ACM Symposium on Theory of Computing (Baltimore, MD, 1990), vol. 47, 1993, pp. 549–595.
[11] David Conlon and Jacob Fox, *Graph removal lemmas*, Surveys in combinatorics 2013, London Math. Soc. Lecture Note Ser., vol. 409, Cambridge Univ. Press, Cambridge, 2013, pp. 1–49.
[12] Jacob Fox, Jonathan Tidor, and Yufei Zhao, *Induced arithmetic removal: complexity 1 patterns over finite fields*, arXiv:1911.03427.
[13] Zoltán Füredi, *Extremal hypergraphs and combinatorial geometry*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, pp. 1343–1352.
[14] Oded Goldreich, Shafi Goldwasser, and Dana Ron, *Property testing and its connection to learning and approximation*, J. ACM **45** (1998), 653–750.
[15] Oded Goldreich and Dana Ron, *On proximity-oblivious testing*, SIAM J. Comput. **40** (2011), 534–566.
[16] W. Timothy Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465–588.
[17] W. Timothy Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. (2) **166** (2007), 897–946.
[18] W. Timothy Gowers and Julia Wolf, *Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$*, Geom. Funct. Anal. **21** (2011), 36–69.
[19] Ben Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), 340–376.
[20] Ben Green and Tom Sanders, *Fourier uniformity on subspaces*, arXiv:1510.08739.
[21] Ben Green and Terence Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), 481–547.
[22] Ben Green and Terence Tao, *Linear equations in primes*, Ann. of Math. (2) **171** (2010), 1753–1850.
[23] Ben Green, Terence Tao, and Tamar Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$-norm*, Ann. of Math. (2) **176** (2012), 1231–1372.
[24] Hamed Hatami, Pooya Hatami, and Shachar Lovett, *General systems of linear forms: equidistribution and true complexity*, Adv. Math. **292** (2016), 446–477.
[25] Hamed Hatami, Pooya Hatami, and Shachar Lovett, *Higher-order fourier analysis and applications*, Foundations and Trends® in Theoretical Computer Science **13** (2019), 247–448.
[26] Tali Kaufman and Madhu Sudan, *Algebraic property testing: the role of invariance*, STOC'08, ACM, New York, 2008, pp. 403–412.
[27] Daniel KráĬ, Oriol Serra, and Lluís Vena, *A removal lemma for systems of linear equations over finite fields*, Israel J. Math. **187** (2012), 193–207.
[28] Vojtěch Rödl, Brendan Nagle, Jozef Skokan, Mathias Schacht, and Yoshiharu Kohayakawa, *The hypergraph regularity method and its applications*, Proc. Natl. Acad. Sci. USA **102** (2005), 8109–8113.
[29] Vojtěch Rödl and Mathias Schacht, *Property testing in hypergraphs and the removal lemma [extended abstract]*, STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing, ACM, New York, 2007, pp. 488–495.
[30] Imre Z. Ruzsa and Endre 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.
[31] 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.
[32] Terence Tao and Tamar Ziegler, *The inverse conjecture for the Gowers norm over finite fields via the correspondence principle*, Anal. PDE **3** (2010), 1–20.
[33] Terence Tao and Tamar Ziegler, *The inverse conjecture for the Gowers norm over finite fields in low characteristic*, Ann. Comb. **16** (2012), 121–188.