Combinatorics, Probability and Computing (2026), 1–21
doi:10.1017/S0963548326100571

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Low-complexity approximations for sets defined by generalizations of affine conditions

W. T. Gowers[1,2] and Thomas Karam[3]

[1]Collège de France, Paris, France, [2]University of Cambridge, Cambridge, UK, and [3]School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai, China
Corresponding author: Thomas Karam; Email: thomaskaram@sjtu.edu.cn

(Received 6 October 2024 UTC; revised 18 May 2026 UTC; accepted 23 August 2026 UTC)

Abstract

Let p be a prime, let S be a non-empty subset of $\mathbb{F}_p$ and let $0 < \epsilon \leq 1$. We show that there exists a constant $C = C(p, \epsilon)$ such that for every positive integer k, whenever $\phi_1, \dots, \phi_k : \mathbb{F}_p^n \to \mathbb{F}_p$ are linear forms and $E_1, \dots, E_k$ are subsets of $\mathbb{F}_p$, there exist linear forms $\psi_1, \dots, \psi_C : \mathbb{F}_p^n \to \mathbb{F}_p$ and subsets $F_1, \dots, F_C$ of $\mathbb{F}_p$ such that the set $U = \{x \in S^n : \psi_1(x) \in F_1, \dots, \psi_C(x) \in F_C\}$ is contained inside the set $V = \{x \in S^n : \phi_1(x) \in E_1, \dots, \phi_k(x) \in E_k\}$, and the difference $V \setminus U$ has density at most ϵ inside S^n. We then generalize this result to one where $\phi_1, \dots, \phi_k$ are replaced by homomorphisms $G^n \to H$ for some pair of finite Abelian groups G and H, and to another where they are replaced by polynomial maps $\mathbb{F}_p^n \to \mathbb{F}_p$ of small degree.

Keywords: Finite fields; Polynomials; Linear forms; Boolean cube

2020 MSC Codes: Primary: 11T06

1. Introduction

Let p be a prime. If $\phi_i(x) = b_i$, $i = 1, \dots, k$ is a system of linear equations over some finite-dimensional vector space over $\mathbb{F}_p$, then the solution set is either empty or an affine subspace with codimension equal to $\dim\langle\phi_1, \dots, \phi_k\rangle$, and therefore with density equal to

$$p^{-\dim\langle\phi_1,\dots,\phi_k\rangle}.$$

It follows that if the density of the set of solutions is at least δ, then there is a set of at most $\log_p \delta^{-1}$ equations with the same solution set.

In this paper we shall generalize this simple observation in a number of ways. Our initial motivation for doing this was that in the context of results such as the density Hales-Jewett theorem [1, 6] certain generalizations of affine subspaces occur naturally. Consider, for instance, the problem of understanding triples of dense subsets A, B and C of $\{0, 1, 2\}^n$ such that the set $A \times B \times C$ contains no combinatorial line. (One might wish to do this with a view to obtaining a density increment argument.) One large source of examples is constructed as follows. Pick a prime p and three subsets U, V and W of $\mathbb{F}_p$ such that $U \times V \times W$ contains no arithmetic progression, let $\phi : \{0, 1, 2\}^n \to \mathbb{F}_p$ be the function where $\phi(x)$ is the mod-p sum of the coordinates of x, and let $A = \phi^{-1}(U)$, $B = \phi^{-1}(V)$ and $C = \phi^{-1}(W)$. Since ϕ takes combinatorial lines to arithmetic progressions in $\mathbb{F}_p$, it follows that $A \times B \times C$ has the desired property.

© The Author(s), 2026. Published by Cambridge University Press.

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

The above construction can be generalized further, and led us to consider sets of the following form. Let G be a finite Abelian group, let S be a subset of G, let $\phi_1, \dots, \phi_k : G^n \to G$ be homomorphisms, and let $E_1, \dots, E_k$ be subsets of G. Then we take the subset

$$\{x \in G^n : \phi_i(x) \in E_i, i = 1, 2, \dots, k\},$$

which is subspace-like, in that it is the intersection of sets defined by linear-like conditions. One of our main results will be that, as in the simple case where $G = \mathbb{Z}_p$ and $E_1, \dots, E_k$ are singletons, if the above subset is dense, then we may replace it by a subset defined by a number of conditions that is bounded above in terms of the density. However, the new subset is not identical to the old one: in this set-up, the best one can hope for is a good approximation, and that is what we shall obtain. The precise statement will be given later.

We also prove a generalization in a different direction, considering subsets of S^n that are defined by low-degree polynomial conditions on $\mathbb{F}_p^n$. Again we shall show that if such a subset is dense, then it can be approximated by a subset of bounded complexity, meaning that it is defined using a bounded number of low-degree polynomial conditions. Again, we defer the precise statement till later.

In a separate paper [4] we shall use these results to rule out a class of potential counterexamples to the polynomial density Hales-Jewett conjecture that are based on modular obstructions.

The following examples indicate some of the complexities that arise in these generalized settings: when our conditions are either restricted to subsets of the form S^n, defined in Abelian groups, or defined by polynomials. For brevity, if $\phi : G^n \to G$ and $E \subset G$, we write (ϕ, E) for the condition $\phi(x) \in E$.

Example 1. *Assume $p \neq 2$. For $i = 2, 3, \dots, k$, let $\phi_i : \{0,1\}^n \to \mathbb{F}_p$ be the linear form defined by $\phi_i(x) = x_1 + x_i$ and let $E_i \subset \mathbb{F}_p$ be the set $\{0, 1\}$. Then the set of points of $\{0,1\}^n$ satisfying the conditions (ϕ_i, E_i) consists of all $x \in \{0,1\}^n$ such that either $x_1 = 0$, or $x_1 = 1$ and $x_2 = \cdots = x_k = 0$. The conditions are linearly independent, and the density of this set is $1/2 + 1/2^k$. Given $\epsilon > 0$, either $k \leq \log_2(\epsilon^{-1})$ or we can approximate the set to within density ϵ by the set of points satisfying the single condition $x_1 = 0$.*

Example 2. *For $i = 2, 3, \dots, k$, let $\phi_i : \mathbb{Z}_p^n \to \mathbb{Z}_p^2$ be the group homomorphism defined by $\phi_i(x) = (x_1, x_i)$ and let $E_i \subset \mathbb{Z}_p^2$ be the set $(\{0\} \times \mathbb{Z}_p) \cup \{(1,1)\}$. Then the set of points of $\mathbb{Z}_p^n$ satisfying the conditions (ϕ_i, E_i) consists of all $x \in \mathbb{Z}_p^n$ such that either $x_1 = 0$, or $x_1 = x_2 = \cdots = x_k = 1$. The density of this set is $1/p + 1/p^k$, so it can be approximated in a similar way.*

Example 3. *For $i = 2, 3, \dots, k$, let $P : \mathbb{F}_p^n \to \mathbb{F}_p$ be the degree-2 polynomial defined by $P(x) = x_1 x_i$ and let $E_i \subset \mathbb{F}_p$ be the set $\{0\}$. Then the set of points of $\mathbb{F}_p^n$ satisfying the conditions (ϕ_i, E_i) consists of all $x \in \mathbb{F}_p^n$ such that either $x_1 = 0$, or $x_1 \neq 0$ and $x_2 = \cdots = x_k = 0$. The density of this set is $1/p + (p-1)/p^k$, so again it can be approximated in a similar way.*

Note that in each of the three examples, the low-complexity set that approximates the solution set – that is, the set of points that satisfy the conditions – is a subset of the solution set. Our main results will show that this is a general phenomenon. To formulate them it will be useful to have a few definitions.

Definition 1.1. *Let X be a finite set and let U, V be subsets of X. Then U* internally ϵ-approximates *V if $U \subset V$ and $|V \setminus U|$ has density at most ϵ in X. Given a set Z, a function $f : X \to Z$, and a subset $E \subset Z$, we call the pair (f, E) a* condition *on X, and we say that x* satisfies the condition *(f, E) if $f(x) \in E$. If $\mathcal{F} = \{(f_1, E_1), \dots, (f_k, E_k)\}$ is a set of conditions on X, the* satisfying set *of the conditions of $\mathcal{F}$ is the set of $x \in X$ that satisfy all the conditions in $\mathcal{F}$. If $\mathcal{F}$ and $\mathcal{G}$ are sets of conditions on X, we say that $\mathcal{G}$* internally ϵ-approximates *$\mathcal{F}$ if the satisfying set of $\mathcal{G}$ internally ϵ-approximates the*

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

satisfying set of $\mathcal{F}$. *If* $S \subset \mathbb{F}_p$, *then a* mod-p condition *on* S^n *is a condition* (ϕ, E) *on* S^n, *where* ϕ *is the restriction to* S^n *of a linear form defined on* $\mathbb{F}_p^n$ *and* E *is a subset of* $\mathbb{F}_p$.

Our first main result is the following theorem.

Theorem 1.2. *Let* $\epsilon > 0$, *let* p *be a prime and let* S *be a non-empty subset of* $\mathbb{F}_p$. *There exists* K *such that for every* n *the following property is satisfied. Every set of mod-*p *conditions on* S^n *can be internally* ϵ*-approximated by a set of at most* K *mod-*p *conditions.*

In order to give some idea of why this might be an interesting statement, we make a few remarks. Note first that if we set $p = 2$, then the result is easy. We can identify $\{0,1\}^n$ with $\mathbb{F}_2^n$, and the satisfying set will take the form $\bigcap_{i=1}^r \{x : \phi_i(x) \in E_i\}$, where each ϕ_i is a linear form on $\mathbb{F}_2^n$ and each E_i is a subset of $\mathbb{F}_2$. If $E_i = \mathbb{F}_2$, then we can discard the condition (ϕ_i, E_i) without affecting the satisfying set, and if any E_i is empty then the satisfying set is empty, so we can obtain it using just the condition (ϕ_i, E_i). So without loss of generality all the sets E_i are singletons and the satisfying set is an affine subspace. If the codimension of the subspace is greater than $\log_2(\epsilon^{-1})$ then we can internally ϵ-approximate it by the empty set, which as above just needs one condition. Otherwise, we can pick k linearly independent conditions that specify the subspace and we have a set of at most $\log_2(\epsilon^{-1})$ conditions with the same satisfying set.

It is only slightly harder to prove the following closely related statement about linear conditions on $\mathbb{F}_p^n$ with $p > 2$, by which we mean conditions (ϕ, E), where ϕ is a linear form on $\mathbb{F}_p^n$ and $E \subset \mathbb{F}_p$. However, the bound is slightly weaker, owing to a crude argument at the end of the proof. We have not tried to improve the bound, and indeed it is not obviously possible, because in this case satisfying sets do not have to be affine subspaces.

Proposition 1.3. *For every* $\epsilon > 0$ *and every* n *the following property is satisfied. Every set of linear conditions on* $\mathbb{F}_p^n$ *can be internally* ϵ*-approximated by a set of at most* $\epsilon^{-p \log p}$ *linear conditions.*

Proof. Let the conditions be (ϕ_i, E_i) for $i = 1, \dots, r$. As in the $p = 2$ case we may assume that each E_i is a non-empty proper subset of $\mathbb{F}_p$. We may also assume that the ϕ_i are distinct, since if $\phi_i = \phi_j$ then we may replace the two conditions (ϕ_i, E_i) and (ϕ_j, E_j) by the single condition $(\phi_i, E_i \cap E_j)$.

Now choose a maximal set I such that the forms ϕ_i with $i \in I$ are linearly independent. Then if we choose $x \in \mathbb{F}_p^n$ uniformly at random, the events $\phi_i(x) \in E_i$ with $i \in I$ are independent, so the probability that x belongs to the satisfying set is at most $\prod_{i \in I} |E_i|/p$, which is at most $(1 - 1/p)^{|I|}$. If $|I| > p \log(\epsilon^{-1})$, then this is at most ϵ, so we can internally ϵ-approximate the satisfying set by the empty set, which requires just one condition. Otherwise, the forms $\phi_1, \dots, \phi_r$ are contained in a linear subspace of dimension at most $p \log(\epsilon^{-1})$, and there is at most one condition for each i, so $r \leq p^{p \log(\epsilon^{-1})}$, which proves the result. □

The key to the proofs above was that if a set satisfies k non-trivial linearly independent conditions, then its density must be at most α^k for some $\alpha < 1$ that is independent of n. What makes Theorem 1.2 harder is that this is no longer true for mod-p conditions on $\{0,1\}^n$, as Example 1 shows. Example 1 is a simple representative of a large class of examples, but it illustrates the basic challenge that we have to deal with.

Once we have proved Theorem 1.2, we shall generalize it to a similar statement with $\mathbb{F}_p$ replaced by an arbitrary finite Abelian group. If G and H are finite Abelian groups, $S \subset G$ and n is a positive integer, then an *H-condition* on S^n is a pair (ϕ, E), where ϕ is the restriction to S^n of a homomorphism from G^n to H, and $E \subset H$. (To avoid excessively heavy notation we will also usually denote the homomorphism itself by ϕ, as it will be clear from the context whether the homomorphism or its restriction is being considered.) The definition of an internal ϵ-approximation carries over in the obvious way: we shall say that a set $\mathcal{F}$ of H-conditions on S^n *internally ϵ-approximates* a set $\mathcal{G}$ of H-conditions on S^n if every point that satisfies $\mathcal{F}$ satisfies $\mathcal{G}$ and the difference of the satisfying sets has density at most ϵ inside S^n.

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

Theorem 1.4. *Let G and H be finite Abelian groups, let S be a non-empty subset of G, and let $\epsilon > 0$. Then there exists $A = A(\epsilon, G, S, H)$ such that every set of H-conditions can be ϵ-approximated on S^n by a set of at most A H-conditions.*

This theorem turns out to need additional ideas even when $G = \mathbb{Z}_p$ and $H = \mathbb{Z}_p^2$.

The other generalization we shall prove is to polynomial conditions. If $S \subset \mathbb{F}_p$, then a *k-dimensional condition of degree at most d* on S^n is a pair (ϕ, E), where $\phi : \mathbb{F}_p^n \to \mathbb{F}_p^k$ is a polynomial map of degree at most d (that is, each coordinate of $\phi(x)$ is obtained by applying a polynomial with degree at most d to x) and $E \subset \mathbb{F}_p^k$. We prove the following theorem, which has a weaker conclusion than in the linear case, but still shows that a set defined by polynomial conditions can be approximated by a set of bounded complexity defined by polynomial conditions.

Theorem 1.5. *Let $\epsilon > 0$, let p be a prime, and let k, d be positive integers with $d < p$. Then there exists M such that for every non-empty $S \subset \mathbb{F}_p$ the following property holds. Every set of k-dimensional conditions with degree at most d on S^n can be internally ϵ-approximated by a single M-dimensional condition with degree at most d on S^n.*

2. Preliminary lemmas

The next section will be devoted to a proof of Theorem 1.2. To prepare for this, we collect together some more technical statements that will be useful in the proof. They will mainly concern the joint distribution of the values $\phi_1(x), \dots, \phi_k(x)$, where each ϕ_i is a linear form from $\mathbb{F}_p^n$ to $\mathbb{F}_p$ and x is chosen uniformly from S^n for some subset S of $\mathbb{F}_p$ of size at least 2. We begin with a result that will help us with the case $k = 1$. We write ω_p for $\exp(2\pi i/p)$.

Lemma 2.1. *Let p be a prime and let S be a subset of $\mathbb{F}_p$ that contains at least two elements. Then*

$$\Big| \mathop{\mathbb{E}}_{x \in S} \omega_p^{tx} \Big| \le 1 - p^{-2}$$

for each $t \in \mathbb{F}_p^*$.

Proof. By the triangle inequality, it is sufficient to prove the result in the case $|S| = 2$, since

$$\mathop{\mathbb{E}}_{x \in S} \omega_p^{tx} = \mathop{\mathbb{E}}_{\substack{T \subset S \\ |T| = 2}} \mathop{\mathbb{E}}_{x \in T} \omega_p^{tx}.$$

But if $y \neq 0$, then

$$\big|1 + \omega_p^y\big|^2 = 2 + 2\cos(2\pi y/p) \le 2 + 2\cos(2\pi/p).$$

Using the bound $\cos\theta \le 1 - \theta^2/2 + \theta^4/24$, which is at most $1 - \theta^2/4$ when $\theta \le 2\pi/3$, we find that this is at most $4 - 4\pi^2/4p^2 = 4 - \pi^2/p^2 = 4(1 - \pi^2/4p^2)$. It follows that $|1 + \omega_p^y|/2 \le 1 - \pi^2/8p^2 \le 1 - 1/p^2$, which implies the result. □

From this lemma we deduce an upper bound on the Fourier coefficients of a linear phase function restricted to S^n. Given a linear form ϕ defined by a formula $x \mapsto \sum_{i=1}^n \lambda_i x_i$, we define $Z(\phi)$ to be $\{i : \lambda_i \neq 0\}$. We shall call this set the *support* of ϕ, even though technically it is the support of the coefficients of ϕ with respect to the standard basis. This should cause no confusion, since the complement of the kernel of ϕ will play no role in the paper.

Our bound shows that the Fourier coefficients are exponentially small in the size of the support.

Proposition 2.2. *Let p be a prime and let S be a subset of $\mathbb{F}_p$ that contains at least two elements. Let $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ be a linear form. Then*

$$\Big| \mathop{\mathbb{E}}_{x \in S^n} \omega_p^{\phi(x)} \Big| \le (1 - p^{-2})^{|Z(\phi)|}.$$

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

Proof. If $\phi(x) = \sum_i \lambda_i x_i$, then the left-hand side splits as a product

$$\mathop{\mathbb{E}}_{x\in S^n} \omega_p^{\phi(x)} = \prod_{i=1}^n \mathop{\mathbb{E}}_{x_i\in S} \omega_p^{\lambda_i x_i}.$$

The inner expectation of the right-hand side is equal to 1 for every $i \notin Z(\phi)$, and by Lemma 2.1 it has modulus at most $1 - p^{-2}$ for every $z \in Z(\phi)$. The result follows. □

Remark 2.3. *Note that the same conclusion holds for $t\phi$ for every non-zero t, since in that case $Z(t\phi) = Z(\phi)$.*

A standard calculation now gives us the equidistribution result we shall need. Given a linear map $\phi : \mathbb{F}_p^n \to \mathbb{F}_p^k$ and an element $a \in \mathbb{F}_p^k$, we write $a.\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ for the map that takes x to $a.\phi(x)$, where $a.b$ is notation for $\sum_{i=1}^k a_i b_i$.

Proposition 2.4. *Let p be a prime, let k be a positive integer, and let S be a subset of $\mathbb{F}_p$ that contains at least two elements. Let $\phi : \mathbb{F}_p^n \to \mathbb{F}_p^k$ be a linear map and let x be chosen uniformly at random from S^n. Suppose that $|Z(a.\phi)| \geq r$ for every non-zero $a \in \mathbb{F}_p^k$. Then for each $z \in \mathbb{F}_p^k$,*

$$|\mathbb{P}[\phi(x) = z] - p^{-k}| \leq (1 - p^{-2})^r.$$

Proof. Let δ_z be the function that takes the value 1 at z and 0 everywhere else. Then

$$\mathbb{P}[\phi(x) = z] = \mathop{\mathbb{E}}_{x\in S^n} \delta_z(\phi(x)),$$

which by the Fourier inversion formula is equal to

$$\mathop{\mathbb{E}}_{x\in S^n} \sum_{a\in\mathbb{F}_p^k} \hat{\delta_z}(a)\omega_p^{a.\phi(x)} = \sum_{a\in\mathbb{F}_p^k} \hat{\delta_z}(a) \mathop{\mathbb{E}}_{x\in S^n} \omega_p^{a.\phi(x)}.$$

But $\hat{\delta_z}(a) = p^{-k}\omega_p^{-a.z}$, so when $t = 0$ it is equal to p^{-k}, and otherwise it has modulus at most p^{-k}. The result now follows from Proposition 2.2 and the triangle inequality. □

If a linear form $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ takes a value once, then it must take that value with probability at least p^{-1}. We now show that something like this, but weaker, holds for the restriction of ϕ to a set of the form S^n.

Proposition 2.5. *Let p be a prime and let S be a non-empty subset of $\mathbb{F}_p$ that contains at least two elements. For every linear form $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ and every $y \in \mathbb{F}_p$ the probability $\mathbb{P}_{x\in S^n}(\phi(x) = y)$ is either 0 or at least $|S|^{-t}$, where $t = \lceil (p-1)/(|S|-1)\rceil$.*

Proof. Let $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ be a linear form given by the formula $\phi(x) = \sum_{i=1}^n \lambda_i x_i$. Without loss of generality there exists m such that $\lambda_1, \dots, \lambda_m$ are non-zero and $\lambda_{m+1} = \cdots = \lambda_n = 0$.

Since there are $|S|^m$ distinct choices for $(x_1, \dots, x_m)$ if $x \in S^n$, each value of $\phi(x)$ that is taken at all is taken with probability at least $|S|^{-m}$.

We now prove a second upper bound that is less trivial. Write $\lambda.S$ for the dilate $\{\lambda u : u \in S\}$ of S by λ. By the Cauchy-Davenport theorem, for every r we have the inequality

$$|\lambda_1.S + \lambda_2.S + \cdots + \lambda_r.S| \geq \min\{r(|S|-1)+1, p\}.$$

In particular, if $r(|S|-1) \geq p-1$, then $\lambda_1.S + \cdots + \lambda_r.S = \mathbb{F}_p$.

For such r, it follows that if x is chosen uniformly at random from S^n, then for every $u \in \mathbb{F}_p$, the probability that $\lambda_1 x_1 + \cdots + \lambda_r x_r = u$ is at least $|S|^{-r}$. By conditioning on the value taken by $\lambda_{r+1}x_{r+1} + \cdots + \lambda_n x_n$ and applying the law of total probability, we deduce that each element of $\mathbb{F}_p$ is equal to $\phi(x)$ with probability at least $|S|^{-r}$.

Combining this with the first bound gives the result claimed. □

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

If $S = \{0, 1, \dots, h\}$ and $h|(p-1)$, then the bound in Proposition 2.5 is optimal, since the linear form ϕ defined by $\phi(x) = x_1 + \cdots + x_{(p-1)/h}$ takes the value $p-1$ with probability $|S|^{-(p-1)/h}$. Note that when $|S| = 2$, the lower bound we obtain for each non-zero probability is $2^{-(p-1)}$.

3. Approximating sets of mod-*p* conditions

In this section we shall prove Theorem 1.2. From now on, we shall write 'ϵ-approximated' to mean 'internally ϵ-approximated', since all the approximations we consider will be internal ones. (Of course, it is then important to keep in mind that the relation 'can be ϵ-approximated by' is not symmetric.)

Our basic strategy will be to start with a set of mod-p conditions and approximate it by sets with more and more structure until eventually we arrive at a set of bounded size. In order to explain this strategy in more detail, we need some definitions.

Definition 3.1. *Let p be a prime. If $\phi_0 : \mathbb{F}_p^n \to \mathbb{F}_p$ is a linear form, then a* sunflower with centre ϕ_0 *is a set of forms $\{\phi_0 + \phi_i : i \in I\}$ such that the forms ϕ_i are disjointly supported. A* symmetric family *of forms is a set $\{\phi_i : i \in I\}$ that are equal up to permutations of the coordinates – that is, for each y the number of coefficients of ϕ_i equal to y is the same for each i. A* symmetric sunflower *is a sunflower $\{\phi_0 + \phi_i : i \in I\}$ such that the set $\{\phi_i : i \in I\}$ is a symmetric family. A set of forms $\{\phi_i : i \in I\}$ is* r-separated *if $|Z(\sum_{i\in I} u_i\phi_i)| \geq r$ for every non-zero $u \in \mathbb{F}_p^I$ – that is, if the support size of every non-trivial linear combination of the ϕ_i is at least r. The* support distance *on the set of all linear forms $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ is the metric given by the formula $d(\phi, \psi) = |Z(\phi - \psi)|$. The* ball of radius r about ϕ_0 *is the set of all ϕ such that $d(\phi_0, \phi) \leq r$. We say that a sunflower $\{\phi_0 + \phi_i : i \in I\}$* has radius at most r *if it is contained in the ball of radius r about ϕ_0.*

Note that, contrary to what our terminology might at first suggest, the condition that the forms are r-separated is stronger than simply saying that $d(\phi_i, \phi_j) \geq r$ for every $i \neq j$: we need all non-trivial linear combinations to have large support and not just differences between two distinct forms.

We begin with the following simple observation, which we shall use repeatedly.

Remark 3.2. *Let X be a finite set, let $q \geq 1$ be a positive integer and let $\epsilon > 0$. Let $\mathcal{F}_1, \dots, \mathcal{F}_q, \mathcal{G}_1, \dots, \mathcal{G}_q$ be sets of conditions on X. If $\mathcal{F}_i$ internally ϵ-approximates $\mathcal{G}_i$ for each $i \in [q]$ then $\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_q$ internally $q\epsilon$-approximates $\mathcal{G}_1 \cup \cdots \cup \mathcal{G}_q$.*

Theorem 1.2 concerns arbitrary sets of mod-p conditions. We shall begin by proving it under a strong additional hypothesis concerning those conditions, and then we shall progressively weaken this hypothesis until we end up proving the result in full generality. At each stage, if we have a proof for families $\mathcal{F}$ of conditions that satisfy an additional property P and we would like a proof for families $\mathcal{E}$ of conditions that satisfy a weaker property Q, it is sufficient to prove that every family $\mathcal{E}$ that satisfies Q can be approximated by a family $\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_q$ such that q is bounded and each $\mathcal{F}_i$ satisfies P. Then using the remark above and the result for families that satisfy P, we end up with a bounded number of conditions that approximate $\mathcal{E}$.

Our first additional hypothesis will be that the set of forms is a symmetric sunflower and that all the subsets of $\mathbb{F}_p$ are the same. Having proved the result under that hypothesis, we shall then weaken the hypothesis by no longer insisting that the set is a sunflower, but keeping the symmetry and the assumption that the subsets of $\mathbb{F}_p$ are the same and adding the condition that the set of forms has bounded radius. Finally, we shall prove the result in full generality. These stages will be the contents of Propositions 3.3, 3.4, and Theorem 1.2 respectively.

Given a linear form $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ with coefficients $(\lambda_1, \dots, \lambda_n)$, let $D_\phi : \mathbb{F}_p \to \mathbb{N}$ be defined by $D_\phi(y) = |\{i \in [n] : \lambda_i = y\}|$. We shall call D_ϕ the *coefficient distribution* of ϕ. Note that a symmetric family of forms is one where all the forms have the same coefficient distribution.

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

We begin, then, by proving the result for symmetric sunflowers and identical subsets of $\mathbb{F}_p$. This is a generalization of the example we gave earlier where the forms were $x \mapsto x_1 + x_i$, and as with that example it will turn out that as long as there are enough forms, we can ϵ-approximate the set of conditions by a set of the form (ϕ_0, H), where ϕ_0 is the centre of the sunflower.

Proposition 3.3. *Let p be a prime, let S be a subset of $\mathbb{F}_p$ that contains at least two elements, let $\epsilon > 0$, let $\{\phi_0 + \phi_i : i \in I\}$ be a symmetric sunflower, and let $E \subset \mathbb{F}_p$. Then there exists $H \subset \mathbb{F}_p$ such that on S^n the condition (ϕ_0, H) $p(1 - 2^{-(p-1)})^{|I|}$-approximates the set of conditions $(\phi_0 + \phi_i, E)$.*

Proof. Let $V \subset \mathbb{F}_p$ be the set of all possible values taken by $\phi_i(x)$ when $x \in S^n$, and note that by the symmetry condition V is the same for all $i \in I$. We let H be the set $\{t \in \mathbb{F}_p : V + t \subset E\}$ and we let T be the complement of H. If $x \in S^n$ and $\phi_0(x) \in H$, then by the definition of H, $\phi_0(x) + \phi_i(x) \in E$. This establishes that the satisfying set for the single condition (ϕ_0, H) is contained in the satisfying set for the conditions $(\phi_0 + \phi_i, E)$, so it remains to obtain an upper bound for the probability that $\phi_0(x) \in T$ and $\phi_0(x) + \phi_i(x) \in E$ for every $i \in I$.

Let $t \in T$. Then

$$\mathbb{P}[\phi_0(x) = t \ \wedge \ \forall i \in I \ \phi_0(x) + \phi_i(x) \in E] \leq \mathbb{P}[\forall i \in I \ \phi_i(x) \in E - t] = \prod_{i \in I} \mathbb{P}[\phi_i(x) \in E - t],$$

where the last equality follows from the fact that the ϕ_i are disjointly supported and hence the events $\phi_i(x) \in E - t$ are independent (when x is chosen uniformly at random from S^n).

Since $V + t$ is not contained in E, the probability that $\phi_i(x) \notin E - t$ is non-zero, and therefore by Lemma 2.5 it is at least $2^{-(p-1)}$. It follows that the product above is at most $(1 - 2^{-(p-1)})^{|I|}$. There are at most p possibilities for t, so the result follows. □

The proposition above implies that if the set of forms is a symmetric sunflower, then either it is already a family of bounded size, or it can be approximated by a family of size 1.

Proposition 3.4. *For every prime p, every non-negative integer r and every $\epsilon > 0$ there is a constant $A = A(p, r, \epsilon)$ such that for every pair of subsets S and E of $\mathbb{F}_p$ with $|S| \geq 2$ the following property holds. For every symmetric family $\{\phi_0 + \phi_i : i \in I\}$ contained in the ball about ϕ_0 of radius r, the conditions $(\phi_0 + \phi_i, E)$ on S^n can be ϵ-approximated by a set of at most A mod-p conditions.*

Proof. We prove the result by induction on r. The proof will be similar to the proof by Erdős and Rado of the sunflower lemma (that is, their argument that gives a factorial-type bound for the sunflower conjecture). When $r = 0$, the ball of radius r about ϕ_0 consists of ϕ_0 only, so the result holds with $A(p, 0, \epsilon) = 1$. Now let $r \geq 1$ and let us assume that the result holds up to $r - 1$. Let M be a maximal subset of I such that the forms ϕ_i with $i \in M$ have disjoint supports, and let $Z = \bigcup_{i \in M} Z(\phi_i)$ be the union of those supports.

If $|M| \geq 2^{p-1} \log(p\epsilon^{-1})$ then by Proposition 3.3 there exists a subset $H \subset \mathbb{F}_p$ such that the condition (ϕ_0, H) ϵ-approximates the conditions $(\phi_0 + \phi_i, E)$ with $i \in M$. From the proof of Proposition 3.3 and the symmetry condition, (ϕ_0, H) implies every condition $(\phi_0 + \phi_i, E)$ with $i \in I$, so it follows that the condition (ϕ_0, H) ϵ-approximates the conditions $(\phi_0 + \phi_i, E)$ with $i \in I$.

For each i, let the coefficients of ϕ_i be $(\lambda_{i1}, \dots, \lambda_{in})$. If instead $|M| \leq 2^{p-1} \log(p\epsilon^{-1})$ then for each $i \in I$ the support of ϕ_i intersects Z, so there is some $j \in Z$ and some non-zero $y \in \mathbb{F}_p$ such that $\lambda_{ij} = y$. That is, if we set $I_{j,y}$ to be the set of $i \in I$ such that $\lambda_{ij} = y$, then $I = \bigcup_{y \in \mathbb{F}_p^*} \bigcup_{j \in Z} I_{j,y}$.

Now let us fix y and j and consider the system of forms $\{\phi_0 + \phi_i : i \in I_{j,y}\}$. Since each λ_{ij} with $i \in I_{j,y}$ is equal to y, we can obtain the same system of forms by replacing each λ_{ij} with $i \in I_{j,y}$ by $\lambda_{ij} - y$ and replacing λ_{0j} by $\lambda_{0j} + y$. Since for each ϕ_i with $i \in I_{j,y}$ the support size has been reduced by 1 and the coefficient we have removed is y, the result is a symmetric family contained in a ball of radius $r - 1$ about the modification of ϕ_0.

By the inductive hypothesis, we can therefore $\epsilon/(2^{p-1}pr \log(p\epsilon^{-1}))$-approximate each family $\{\phi_0 + \phi_i : i \in I_{j,y}\}$ by a family of at most $A(p, r-1, \epsilon/(2^{p-1}pr \log(p\epsilon^{-1})))$ mod-p conditions. By

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

Remark 3.2 and the fact that $|\mathbb{F}_p^* \times Z| \leq 2^{p-1}pr\log(p\epsilon^{-1})$, it follows that we can ϵ-approximate the family $\{\phi_0 + \phi_i : i \in I\}$ by a family of at most

$$A(p, r, \epsilon) = 2^{p-1}pr\log(p\epsilon^{-1})A(p, r-1, \epsilon/(2^{p-1}pr\log(p\epsilon^{-1}))$$

mod-p conditions. □

A back-of-envelope calculation shows that $A(p, r, \epsilon)$ is at most $(2p)^{2r^2p}(\log(\epsilon^{-1}))^r$. (We have not attempted to optimize this bound.)

We have now shown that symmetric families of bounded radius can be approximated. It remains to prove that an arbitrary family can be approximated by a union of boundedly many symmetric families of bounded radius. That will be the content of the rest of the argument in this section.

Proof of Theorem 1.2. Let the set of conditions to be approximated be $\{(\phi_i, E_i) : i \in I\}$. We may assume that each E_i is a proper subset of $\mathbb{F}_p$. Let $r = 2p^3(\log p)\log(2/\epsilon)$. □

Case 1. Suppose that there exists a subset I' of I of size at least $p\log(2/\epsilon)$ such that the linear forms ϕ_i with $i \in I'$ are r-separated. Pick such a set that also satisfies the upper bound $|I'| \leq 2p\log(2/\epsilon)$. By Proposition 2.4 and our choice of r,

$$\mathbb{P}[\forall i \in I', \phi_i = y_i] \leq p^{-|I'|} + (1 - p^{-2})^r \leq 2p^{-|I'|}$$

for each $y \in \mathbb{F}_p^{I'}$.

It follows that

$$\mathbb{P}[\forall i \in I'\ \phi_i \in E_i] \leq \left(\prod_{i \in I'} |E_i|\right)(2p^{-|I'|}) \leq 2(1 - p^{-1})^{|I'|} \leq \epsilon$$

so the set of conditions $\{(\phi_i, E_i) : i \in I\}$ is ϵ-approximated by the condition (x_1, S^c) (or the condition $(\phi, \emptyset)$ for any form ϕ).

Case 2. Now suppose that there does not exist such a subset I'. Let M be a maximal r-separated subset of I. Then by hypothesis M has size at most $p\log(2/\epsilon)$, and for every $i \in I$ there exists $a \in \mathbb{F}^M$ such that $\phi_i - \sum_{m \in M} a_m\phi_m$ has support size at most r.

As in the proof of Proposition 3.4, we now divide the set of conditions into families. This time, given a subset $E \subset \mathbb{F}_p$, an element $a \in \mathbb{F}^M$, and a function $D : \mathbb{F}_p \to \mathbb{N}$ such that $\sum_y D(y) = n$ and $\sum_{y \neq 0} D(y) \leq r$, we let $I_{E,a,D}$ be the set of i such that $E_i = E$, and $\phi_i - \sum_{m \in M} a_m\phi_m$ has coefficient distribution D. There are at most 2^p choices of E, at most $p^{|M|}$ choices of a, and at most r^p choices of $D : \mathbb{F}_p \to \mathbb{N}$ such that $\sum_y D(y) = n$ and $\sum_{y \neq 0} D(y) \leq r$, so the number of families $I_{E,a,D}$ we need to approximate is bounded independently of n. Also, the forms in each family are a symmetric family of radius at most r and within each family the subsets of $\mathbb{F}_p$ are the same.

Therefore, by Proposition 3.4 we can $(2^p p^{|M|} r^p)^{-1}\epsilon$-approximate each set of conditions $\{(\phi_i, E_i) : i \in I_{E,a,D}\}$ by a set of at most $A(p, r, (2^p p^{|M|} r^p)^{-1}\epsilon)$ conditions. The number of triples (E, a, D) is at most

$$2^p p^{|M|} r^p \leq 2^p p^{p\log(2/\epsilon)}(2p^3(\log p)\log(2/\epsilon))^p \leq p^{4p\log(2/\epsilon)},$$

so the result now follows by Remark 3.2, and the number of conditions needed for the approximation is at most

$$p^{4p\log(2/\epsilon)}A(p, 2p^3\log p\log(2/\epsilon), p^{-4p\log(2/\epsilon)}\epsilon).$$

After a back-of-envelope calculation we can bound this above by $2^{16p^8\log(2/\epsilon)^2}$. Thus, for fixed p the number of conditions needed for the approximation depends quasipolynomially on $1/\epsilon$.

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

4. More general finite Abelian groups

In this and the next section we shall prove Theorem 1.4. This needs extra ideas, but it turns out that those ideas are already needed for the simpler case where $S = G$, so in this section we shall prove that special case, and then in the next section we shall combine the proof method of this section with the ideas about sunflowers from the previous section in order to obtain a proof for general S.

The following definition will be central to the proof.

Definition 4.1. *Given homomorphisms $\phi_1, \dots, \phi_k : G^b \to H$ and characters $\chi_1, \dots, \chi_k \in \hat{H}$, the map $x \mapsto \prod_{i=1}^k \chi_i(\phi_i(x))$ defines a character on G. Let us call $\phi_1, \dots, \phi_k$* independent *if the only way for $\prod_i \chi_i \circ \phi_i$ to be the trivial character is if all the χ_i are the trivial character.*

Proposition 4.2. *Let $\phi_1, \dots, \phi_k$ be independent homomorphisms from G^n to H. Then for every $y \in H^k$,*

$$\mathbb{P}_{x \in G^n}[\forall i\, \phi_i(x) = y_i] = |H|^{-k}.$$

Proof. The probability in question is equal to

$$\mathop{\mathbb{E}}_{x} \mathop{\mathbb{E}}_{\chi_1,\dots,\chi_k \in \hat{H}} \prod_{i=1}^k \chi_i(\phi_i(x) - y_i)$$

since if $\phi_i(x) = y_i$ for each i, then the inner expectation is equal to 1, and if for some i we have $\phi_i(x) \neq y_i$, then $\mathbb{E}_{\chi_i} \chi_i(\phi_i(x) - y_i) = 0$, so the whole expectation is 0.

We can rewrite this expression as

$$\mathop{\mathbb{E}}_{\chi_1,\dots,\chi_k \in \hat{H}} \prod_{i=1}^k \chi_i(-y_i) \mathop{\mathbb{E}}_{x} \prod_{i=1}^k \chi_i(\phi_i(x)).$$

If $\chi_1, \dots, \chi_k$ are all equal to the trivial character, then the expression being averaged over is equal to 1. Otherwise, by hypothesis, $\prod_i \chi_i \circ \phi_i$ is not the trivial character on G^n, which implies that $\mathbb{E}_x \prod_i \chi_i \circ \phi_i = 0$. Therefore, the probability is $|\hat{H}|^{-k} = |H|^{-k}$. □

The next definition and lemma are standard, but we recall them here for convenience.

Definition 4.3. *Given an Abelian group G, a subgroup $H \subset G$, and a character $\chi \in \hat{G}$, say that χ* annihilates*R H if $\chi(x) = 1$ for every $x \in H$. The* annihilator *of H, denoted $H^\perp$, is the subgroup of $\hat{G}$ that consists of all χ that annihilate H. Similarly, if $\hat{H}$ is a subgroup of $\hat{G}$, we say that $x \in G$* annihilates *$\hat{H}$ if $\chi(x) = 1$ for every $\chi \in \hat{H}$, and the* annihilator *of $\hat{H}$ is the subgroup of all $x \in G$ that annihilate $\hat{H}$.*

Lemma 4.4. *Let G be a finite Abelian group and let H be a subgroup of G. Then $H^\perp \cong \widehat{G/H}$.*

Proof. Define a map $\iota : H^\perp \to \widehat{G/H}$ by $\iota(\chi)(xH) = \chi(x)$. Note that this is well defined, since if $xy^{-1} \in H$, then $\chi(y) = \chi((xy^{-1})y) = \chi(x)$. It is also a homomorphism, since

$$\iota(\chi_1\chi_2)(xH) = \chi_1\chi_2(x) = \chi_1(x)\chi_2(x) = \iota(\chi_1)(xH)\iota(\chi_2)(xH)$$

for every $x \in G$.

Now let $\psi : G/H \to \mathbb{C}$ be a character. Define $\theta(\psi) : G \to \mathbb{C}$ by $\theta(\psi)(x) = \psi(xH)$. Then $\theta(\psi)$ is a character on G, since

$$\theta(\psi)(xy) = \psi(xyH) = \psi(xHyH) = \psi(xH)\psi(yH) = \theta(\psi)(x)\theta(\psi)(y).$$

It also annihilates H, since if $x \in H$, then $\theta(\psi)(x) = \psi(H) = 1$. Finally, for any character χ that annihilates H, we have that $\theta\iota\chi(x) = \iota\chi(xH) = \chi(x)$ for every x, so θ is a left inverse for ι. In

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

the other direction, if $\psi : G/H \to \mathbb{C}$ is a character, then $\iota\theta\psi(xH) = \theta\psi(x) = \psi(xH)$, so θ is also a right inverse for ι. □

Proposition 4.5. *Let $\phi_1, \dots, \phi_k$ be homomorphisms from G^n to H. Let $K \subset \hat{H}^k$ be the subgroup that consists of all k-tuples $(\chi_1, \dots, \chi_k)$ such that $\prod_{i=1}^k \chi_i \circ \phi_i$ is the trivial character on G^n. Let $Y \subset H^k$ be the annihilator of K. Then*

$$\mathbb{P}[\forall i\, \phi_i(x) = y_i] = \begin{cases} |Y|^{-1} & y \in Y \\ 0 & y \notin Y \end{cases}$$

Proof. Again, the probability in question is equal to

$$\mathop{\mathbb{E}}_{\chi_1,\dots,\chi_k \in \hat{H}} \prod_{i=1}^k \chi_i(-y_i) \mathop{\mathbb{E}}_{x} \prod_{i=1}^k \chi_i(\phi_i(x)).$$

If $(\chi_1, \dots, \chi_k) \notin K$, then $\prod_{i=1}^k \chi_i \circ \phi_i$ is not the trivial character, so the expectation over x is 0. Therefore, the probability is equal to

$$(|K|/|\hat{H}|^k) \mathop{\mathbb{E}}_{(\chi_1,\dots,\chi_k)\in K} \prod_{i=1}^k \chi_i(-y_i).$$

If $(y_1, \dots, y_k) \in Y$, then $\prod_{i=1}^k \chi_i(y_i) = 1$ for every $(\chi_1, \dots, \chi_k) \in K$, and therefore we end up with $(|K|/|\hat{H}|^k)$. But by Lemma 4.4 (and Pontryagin duality), $|\hat{H}|^k/|K| = |Y|$, so this equals $|Y|^{-1}$.

If $(y_1, \dots, y_k) \notin Y$, then the evaluation map $(\chi_1, \dots, \chi_k) \mapsto \prod_{i=1}^k \chi_i(y_i)$ is a non-trivial character on K, and therefore its average is 0. □

Proposition 4.6. *Let $\phi + \phi_1, \dots, \phi + \phi_k$ be homomorphisms from G^n to H and let $E \subset H$. Suppose that each ϕ_i takes values in a subgroup H_1 of H and that the ϕ_i are independent when considered as homomorphisms to H_1. Then there is a subset $F \subset H$ such that the condition (ϕ, F) ϵ-approximates the set of conditions $(\phi + \phi_i, E)$, where $\epsilon = (1 - 1/|H_1|)^k$.*

Proof. Let $F = \{z \in H : z + H_1 \subset E\}$. Then if $\phi(x) \in F$, we have for all i that $\phi(x) + \phi_1(x) \in F + H_1 \subset E$. Thus, the single condition (ϕ, F) implies all the conditions $(\phi + \phi_i, E)$.

Now let us fix some $z \notin F$ and condition on the event that $\phi(x) = z$. Let $w \in H_1$ be such that $z + w \notin E$. Then the probability that $\phi(x) + \phi_1(x), \dots, \phi(x) + \phi_k(x) \in E$ given that $\phi(x) = z$ is at most the probability that none of the $\phi_i(x)$ is equal to w.

But each ψ_i takes values in H_1, and by hypothesis the ϕ_i are independent when considered as homomorphisms to H_1, so by Proposition 4.2 the probability that none of them take the value w is at most $(1 - 1/|H_1|)^k$. □

Theorem 4.7. *Let G and H be finite Abelian groups and let $\epsilon > 0$. Then there exists a constant C such that for every n, every set of H-conditions on G^n can be ϵ-approximated by at most C H-conditions.*

Proof. We note first that the number of subsets of H is bounded, so by the usual partitioning argument it is sufficient to prove the result in the case that they are all equal to a single set E.

We prove the result by induction on the size of a subgroup H_1 of H. The inductive hypothesis is that if the homomorphisms are of the form $\phi + \phi_i$, where each ϕ_i takes values in H_1, then the result holds. Let us write $C(G, H_1, \epsilon)$ for the number of conditions needed for an ϵ-approximation in this case.

If H_1 is the trivial subgroup $\{0\}$, then all the homomorphisms are the same, so all the conditions are the same, and the result is trivial.

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

Suppose now that the homomorphisms are of the given form and that we know the result for all proper subgroups of H_1. If it is possible to find k of the homomorphisms ϕ_i that are independent, where $(1-|H_1|^{-1})^k \le \epsilon$, then by the first paragraph of the proof of Lemma 4.6 there is a single condition that approximates the entire set of conditions.

If it is not possible to find more than k of the ϕ_i that are independent, then pick a maximal independent set $\{\psi_1, \dots, \psi_{k'}\}$, with $k' \le k$. Then for every ϕ_i there is a combination $(\chi \circ \phi_i) \prod_{j=1}^{k'} (\chi_j \circ \psi_j)$ that equals the trivial character, where all of $\chi_1, \dots, \chi_{k'}$ and χ are characters on H_1 and χ is not the trivial character. By the usual partitioning argument we can assume that the $(k'+1)$-tuple $(\chi, \chi_1, \dots, \chi_{k'})$ is the same for every ϕ_i. So that gives us a character $\gamma: G^n \to \mathbb{C}$ and a character $\chi \in \hat{H}_1$ such that $\chi \circ \phi_i = \gamma$ for every i.

Let us write $\xi_i = \phi_i - \phi_1$ for each i. Then $\phi_i = \phi_1 + \xi_i$, and $\chi \circ \xi_i = 0$ for every i. It follows that each ξ_i takes values in the subgroup $H_2 = \{y \in H_1 : \chi(y) = 0\}$, which is a proper subgroup of H_1, as χ is a non-trivial character on H_1. Therefore, by the inductive hypothesis, the conditions (ϕ_i, E) can be ϵ-approximated (by at most $C(G, H_2, \epsilon)$ conditions), and we are done. □

5. General finite Abelian groups with restricted alphabets

We now show how to combine the methods of proof of the last two sections to give a proof of Theorem 1.4, which generalizes Theorem 4.7 to the corresponding statement for restricted alphabets.

Before we start on this, we briefly mention a set-up that appears to be more general than that of Theorem 1.4 but is in fact equivalent to it. Suppose that H is a finite Abelian group, S is a non-empty finite set (so not necessarily a subset of a finite Abelian group), and $\phi_1, \dots, \phi_n$ are functions from S to H. We can define a function $\phi : S^n \to H$ by $\phi(x) = \sum_{i=1}^n \phi_i(x_i)$, and if $E \subset H$, then we can consider the condition (ϕ, E) on S^n, which is the condition $\phi(x) \in E$. This more general looking situation is covered by Theorem 1.4, because for every k we can find a finite Abelian group G and a subset $S \subset G$ of size k such that every function from S to H can be extended to a homomorphism from G to H. Then if we extend functions $\phi_1, \dots, \phi_n$ to homomorphisms $\psi_1, \dots, \psi_n$, we find that the map ϕ defined above is the restriction to S^n of the homomorphism $\psi : G^n \to H$ defined by $\psi(x) = \sum_{i=1}^n \psi_i(x)$. Thus, in a certain sense Theorem 1.4 is not really about subsets of G but more about particular kinds of functions from a product S^n to a finite Abelian group H. However, the embedding of S into a group G is convenient.

Now let us turn to the proof. We begin with some definitions that will generalize those that we used in the case $G = H = \mathbb{F}_p$. Given a homomorphism $\phi : G^n \to H$, we can express it in the form $\phi(x) = \sum_i \phi_i(x_i)$, where each ϕ_i is a homomorphism from G to H. If $S \subset G$ is non-empty, then we say that the *support of ϕ with respect to S* (or just the support of ϕ if S is clear from the context) is the set of all i such that ϕ_i is not constant on S. We call the maps $\phi_i|_S$ the *coefficients* of ϕ (with respect to S), and the *coefficient distribution* of ϕ is the function that takes each function $\psi : S \to H$ to the number of i such that $\phi_i|_S = \psi$. The definitions of 'sunflower', 'ball', and 'symmetric sunflower' carry over word for word from our earlier definitions (see Definition 3.1) once the words and phrases 'support' and 'coefficient distribution' are interpreted in these generalized senses. We say that homomorphisms $\phi_i : G^n \to H$ defined by $\phi_i(x) = \sum_{j=1}^n \phi_{ij}(x_j)$ for every $i \in I$ are r-separated if whenever $\chi_i \in \hat{H}$ are characters for each $i \in I$ that are not all trivial, the homomorphism

$$\prod_{i \in I} \chi_i \circ \phi_i : x \to \prod_{j=1}^n \prod_{i \in I} \chi_i(\phi_{ij}(x_j))$$

has support size at least r (with the support being the number of j such that $\prod_{i \in I} \chi_i(\phi_{ij}(x_j))$ is not constant on S).

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

We begin by proving an analogue of Proposition 2.5, but with a less satisfactory bound.

Proposition 5.1. *Let G and H be finite Abelian groups, let S be a non-empty subset of G, let n be a positive integer, and let $\phi : G^n \to H$ be a homomorphism. Let x be an element of S^n, chosen uniformly at random. Then for every $y \in H$, the probability that $\phi(x) = y$ is either 0 or at least c, where $c = c(H)$ is a positive constant that depends only on H (and on S).*

Proof. Let Z be the support of ϕ. Then $\phi(x)$ depends only on $x|_Z$, so each value y that is taken with non-zero probability is taken with probability at least $|S|^{-|Z|}$.

We also have that

$$\mathbb{P}[\phi(x) = y] = \mathop{\mathbb{E}}_{x \in S^n} \mathop{\mathbb{E}}_{\chi \in \hat{H}} \chi(\phi(x) - y) = \mathop{\mathbb{E}}_{\chi \in \hat{H}} \chi(-y) \mathop{\mathbb{E}}_{x \in S^n} \chi(\phi(x)) = \mathop{\mathbb{E}}_{\chi \in \hat{H}} \chi(-y) \prod_i \mathop{\mathbb{E}}_{x_i \in S} \chi(\phi_i(x_i)).$$

Note that $\chi \circ \phi$ is a homomorphism and its support is the number of i such that $\chi \circ \phi_i$ is not constant on S. For each such i, $|\mathbb{E}_{x_i} \chi(\phi_i(x_i))|$ is at most $||S| - 1 + \exp(2\pi i/k)|/|S|$, where k is the order of χ in $\hat{H}$. It is therefore at most $1 - e(H)^{-2}/|S|$, where $e(H)$ is the exponent of H.

It follows that if for each non-trivial character χ the support size of $\chi \circ \phi$ is at least r, then for every y the difference between $\mathbb{P}[\phi(x) = y]$ and $|H|^{-1}$ is at most $\exp(-re(H)^{-2}/|S|)$. So if r is at least $|S|e(H)^2 \log(2|H|)$, then the probability that $\phi(x) = y$ is at least $|H|^{-1}/2$ for all y.

With these observations in mind, let us prove Proposition 5.1 by induction on $|H|$. Suppose that there is some non-trivial character χ such that $\chi \circ \phi$ has support size less than r. Without loss of generality the support of χ is $\{1, 2, \dots, m\}$ for some $m \leq r$. Let us write $x = x^1 + x^2$ where $x^1 = (x_1, \dots, x_m, 0, \dots, 0)$ and $x^2 = (0, \dots, 0, x_{m+1}, \dots, x_n)$. Then $\chi(\phi(x))$ depends only on x^1, or to put it another way, $\phi(x^2)$ belongs to a certain coset of $H_1 = \ker \chi$. Thus, for each coset $z + H_1$, either there is no x with $\phi(x) \in z + H_1$ or the probability that $\phi(x) \in z + H_1$ is at least $|S|^{-r}$. Suppose now that there exists x with $\phi(x) \in z + H_1$ and pick x^1 such that $\phi(x^1 + x^2) \in z + H_1$ for every x^2. Then for each $y \in z + H_1$ such that there exists x^2 with $\phi(x^1 + x^2) = y$, the probability (over x^2 chosen at random in S^{n-m}) that $\phi(x^1 + x^2) = y$ is, by the inductive hypothesis, at least $c(H_1)$. It follows that we may take $c(H)$ to be $|S|^{-|S|e(H)^2 \log(2|H|)} c(H_1)$. Since $|H_1| \leq |H|/2$ and $e(H_1) \leq e(H)$, we may therefore take $c(H)$ to be at least $|S|^{-|S|e(H)^2(\log(2|H|))^2}$. □

The next lemma and its proof combine Proposition 3.3 with Proposition 4.6. We stress that, consistent with the definition that we have given, maps θ_i consisting merely of the same map repeated several times form a sunflower.

Lemma 5.2. *Let G and H be finite Abelian groups, let S be a non-empty subset of G, and let H_1 be a subgroup of H. Let $\Phi = \{\phi_1, \dots, \phi_m\}$ be a set of m homomorphisms from G^n to H and suppose that we can write $\phi_i = \theta_i + \psi_i$ in such a way that the maps θ_i form a symmetric sunflower contained in a ball of radius r, that the maps ψ_i take values in H_1 and are $(rm + r')$-separated as homomorphisms $G^n \to H_1$. Then for every proper subset $E \subset H$ the conditions (ϕ_i, E), regarded as conditions on S^n, can be ϵ-approximated by a single condition, where*

$$\epsilon = |H|\big(\exp\big(-m|S|^{-|S|e(H)^2(\log(2|H|))^2}|H_1|\big) + |H_1|^m \exp\big(-r'e(H_1)^{-2}/|S|\big)\big).$$

Proof. Let $\theta_i = \gamma + \zeta_i$, where the ζ_i have disjoint supports of size at most r and the same coefficient distribution. For each $x \in H$, let $E_x = \{y \in H_1 : x + y \in E\}$. We now let $F \subset H$ be the set of $t \in H$ such that $E_{t+\zeta_i(x)} = H_1$ for every $x \in S^n$. Note that since the ζ_i have the same coefficient distribution, their images are the same, so F does not depend on i.

We shall now show that the single condition (γ, F) ϵ-approximates the set of conditions (ϕ_i, E). First, observe that if $\gamma(x) \in F$, then for every i,

$$E_{\theta_i(x)} = E_{\gamma(x)+\zeta_i(x)} = H_1,$$

and therefore, since ψ_i takes values in H_1, $\phi_i(x) = \theta_i(x) + \psi_i(x) \in E$. Thus the condition (γ, F) implies all the other conditions.

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

It remains to prove that the set of $x \in S^n$ such that $\gamma(x) \notin F$ and $\phi_i(x) \in E$ for every i has density at most ϵ. For this purpose let us fix $y \notin F$ and obtain an upper bound for the probability that $\gamma(x) = y$ and $\phi_i(x) \in E$ for every i. Note that if $\gamma(x) = y$, then $\phi_i(x) \in E$ if and only if $\psi_i(x) \in E_{y+\zeta_i(x)}$ for every i. Since $y \notin F$, there exists $x \in S^n$ such that $E_{y+\zeta_i(x)} \neq H_1$. That is, there exist $x \in S^n$ and $w \in H$ such that $\zeta_i(x) = w$ and $E_{y+w} \neq H_1$. By Proposition 5.1 the probability that $\zeta_i(x) = w$, when x is chosen uniformly from S^n, is at least $|S|^{-|S|e(H)^2(\log(2|H|))^2}$, from which it follows that the probability that $E_{y+\zeta_i(x)} \neq H_1$ is at least $|S|^{-|S|e(H)^2(\log(2|H|))^2}$. Let $N(x)$ be the number of i such that $E_{y+\zeta_i(x)} \neq H_1$.

Now let Z be the union of the supports of the maps ζ_i. Then Z has cardinality at most rm. Let us write W for Z^c and π_Z and π_W for the coordinate projections onto Z and W. Since the maps ψ_i are $(rm + r')$-separated, it follows that the maps $\psi_i \circ \pi_W$ are r'-separated. Applying the beginning of the proof of Proposition 5.1 to the map $\psi = (\psi_1 \circ \pi_W, \dots, \psi_m \circ \pi_W)$, we find that for each $(w_1, \dots, w_m) \in H_1^m$ the probability that $\psi_i(\pi_W x) = w_i$ for $i = 1, 2, \dots, m$ differs from $|H_1|^{-m}$ by at most $\exp(-r'e(H_1)^{-2}/|S|)$, since $e(H^m) = e(H)$. Let us condition on the restriction $\pi_Z(x)$ of x to Z. In the following calculation, we emphasize that the values $N(x)$ and $\zeta_i(x)$ only depend on $\pi_Z(x)$. Applying the law of total probability with respect to the possibilities for $\pi_W(x)$ we find that the probability that $\psi_i(x) \in E_{y+\zeta_i(x)}$ for every i is at most

$$\sum_{w_1 \in E_{y+\zeta_1(x)}, \dots, w_m \in E_{y+\zeta_m(x)}} (|H_1|^{-m} + \exp(-r'e(H_1)^{-2}/|S|))$$

which in turn is bounded above by

$$(1 - |H_1|^{-1})^{N(x)} + |H_1|^m \exp(-r'e(H_1)^{-2}/|S|)$$

since the number $|E_{y+\zeta_1(x)}| \dots |E_{y+\zeta_m(x)}|$ of summands is at most $|H_1|^m(1 - |H_1|^{-1})^{N(x)}$.

We now drop the conditioning on $\pi_Z(x)$. Let $q = |S|^{-|S|e(H)^2(\log(2|H|))^2}$. The probability that $\psi_i(x) \in E_{y+\zeta_i(x)}$ for every i is at most

$$\sum_{j=0}^{m} \mathbb{P}_{z \in S^Z}(N(z) = j)(1 - |H_1|^{-1})^j + |H_1|^m \exp(-r'e(H_1)^{-2}/|S|).$$

Because $(1 - |H_1|^{-1})^j$ decreases with j, the first sum is at most $\sum_{j=0}^{m} A(j)(1 - |H_1|^{-1})^j$ whenever a probability distribution A satisfies

$$\sum_{u=0}^{j} \mathbb{P}_{z \in S^Z}(N(z) = u) \leq \sum_{u=0}^{j} A(u)$$

for every $j = 0, \dots, m$. By definition of q (and because $q \leq 1/2$) we may take $A(j) = \binom{m}{j}(1 - q)^{m-j}q^j$. The probability that $\psi_i(x) \in E_{y+\zeta_i(x)}$ for every i is therefore at most

$$\begin{aligned}
&\sum_{j=0}^{m} \binom{m}{j}(1 - q)^{m-j}q^j(1 - |H_1|^{-1})^j + |H_1|^m \exp(-r'e(H_1)^{-2}/|S|) \\
&\quad = (1 - q|H_1|^{-1})^m + |H_1|^m \exp(-r'e(H_1)^{-2}/|S|) \\
&\quad \leq \exp(-mq|H_1|^{-1}) + |H_1|^m \exp(-r'e(H_1)^{-2}/|S|),
\end{aligned}$$

where the equality follows from the binomial theorem.

Since there are at most $|H|$ possibilities for y, we obtain the bound stated. □

Remark 5.3. *Note that if we wish to obtain an ϵ-approximation, then it is sufficient if $m \geq \log(2|H|/\epsilon)|S|^{|S|e(H)^2(\log(2|H|))^2}|H_1|$ and $r' \geq \log(2|H||H_1|^m/\epsilon)|S|e(H_1)^2$. Note also that the proof*

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

that the condition (γ, F) implies all the other conditions used only the fact that the θ_i have the same coefficient distribution, so if $t = \lceil \log(2|H|/\epsilon)|S|^{|S|e(H)^2(\log(2|H|))^2}|H_1|\rceil$, then to obtain an ϵ-approximation it is sufficient if the θ_i belong to a symmetric ball of radius r and we can find t of them that form a sunflower. Taking the contrapositive, if the θ_i belong to a symmetric ball of radius r and (γ, F) does not *ϵ-approximate the set of conditions (ϕ_i, E), then it is not possible to find a set I of size t such that the conditions θ_i with $i \in I$ form a sunflower and the conditions ψ_i with $i \in I$ (viewed as homomorphisms $G^n \to H_1$) are $rm + r'$-separated with r', m both obeying the above lower bounds.*

We are now ready to prove the main result of the section.

Proof of Theorem 1.4. We prove by induction on the size of $|H_1|$ and on the radius r_{H_1} that for every $\epsilon > 0$ the theorem holds (with bounds depending only on ϵ, G, S, H, H_1, r_{H_1}) for every set of conditions $(\theta_i + \psi_i, E)$ such that each θ_i is a map from G^n to H, each ψ_i is a map from G^n to H_1, and the θ_i are contained in a symmetric ball of radius at most r_{H_1}.

We again write $\theta_i = \gamma + \zeta_i$, where the ζ_i have disjoint supports of size at most r_{H_1} and the same coefficient distribution.

We first describe the inductive step in the case where $r_{H_1} > 0$ and H_1 is not the trivial subgroup. The cases where $H_1 = \{e\}$ or $r_{H_1} = 0$ can be dealt with by a degenerate version of the same argument, as we will explain at the end of the proof.

Suppose that the result holds for all $\epsilon > 0$ and for all pairs (H_2, r) such that H_2 is a strict subgroup of H_1, or such that $H_2 = H_1$ and $r < r_{H_1}$. Let $\epsilon > 0$. Then by Lemma 5.2 (and Remark 5.3) we are done if we can find $t_{H_1} = \lceil \log(2|H|/\epsilon)|S|^{|S|e(H)^2(\log(2|H|))^2}|H_1|\rceil$ of the maps such that the maps θ_i form a sunflower with centre γ and the maps ψ_i are $(r_{H_1}t_{H_1} + r'_{H_1})$-separated for $r'_{H_1} = \lceil \log(2|H||H_1|^m/\epsilon)|S|e(H_1)^2\rceil$. If we cannot find such a set of maps which is that large, then let M be a maximal set of such maps.

Then for each i, the support of the map ζ_i contains a support element inside the union Z of the supports of the maps ζ_j with $j \in M$, or there exists a non-trivial character $\chi \in \hat{H}_1$ such that $\chi \circ \psi_i$ has support distance at most $r_{H_1}t_{H_1} + r'_{H_1}$ from a product of the form $\prod_{m\in M} \chi_m \circ \psi_m$, meaning that the set of j such that

$$x_j \to \chi(\psi_{ij}(x_j))\Big(\prod_{m\in M} \chi_m(\psi_{mj}(x_j))\Big)^{-1}$$

is non-constant on S has size at most $r_{H_1}t_{H_1} + r'_{H_1}$. By Remark 3.2 we may assume that the first situation occurs for all i, or that the second situation occurs for all i.

Assume that we are in the first situation. Let d be the minimal size of a generating set of G. The number of homomorphisms from G to H is at most $|H|^d$, so the number of pairs (ξ, z) where ξ is a non-zero homomorphism from G to H and z is an element of Z is at most $r_{H_1}|H|^d|M|$. We now partition the set of indices i into sets $I_{(\xi,z)}$ in such a way that for every pair (ξ, z), the coefficient of the map ζ_i at z is equal to ξ, in our generalized sense of the word 'coefficient'.

We claim now that each set $\{\theta_i : i \in I_{(\xi,z)}\}$ is contained in a ball of radius at most $r_{H_1} - 1$. To see this, pick ξ and z as above, and let $\gamma' : G^n \to H$ be the homomorphism that takes x to $\gamma + \xi(x_z)$. Then each map $\theta_i - \gamma'$ with $i \in I_{(\xi,z)}$ is a homomorphism of support size at most $r_{H_1} - 1$, so the homomorphisms $\theta_i = \gamma' + (\theta_i - \gamma')$ live in a ball of radius at most $r_{H_1} - 1$ with centre γ'. By Remark 5.3 it suffices to approximate each set $\{(\theta_i + \psi_i, E) : i \in I_{(\xi,z)}\}$, which in turn we can do by the inductive hypothesis.

Assume that we are instead in the second situation. There are at most $|H_1|^{t_{H_1}+1}$ choices for χ and the χ_m with $m \in M$, and therefore by Remark 3.2 we may assume that these choices are the same for each j. We write Φ for the corresponding set of maps $\phi_i = \theta_i + \psi_i$. We now have that there exist a non-trivial character $\chi \in \hat{H}_1$ and a character χ_m for each $m \in M$ such that for every $\phi = \theta + \psi \in \Phi$ the composition $\chi \circ \psi$ is within support distance $r_{H_1}t_{H_1} + r'_{H_1}$ of $\prod_{m\in M} \chi_m \circ \psi_m$.

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

Let us now choose an arbitrary $\phi_0 = \theta_0 + \psi_0 \in \Phi$. Then for each $\phi = \theta + \psi \in \Phi$, the homomorphism $\chi \circ (\psi - \psi_0)$ has support size at most $2(r_{H_1} t_{H_1} + r'_{H_1})$, and therefore we can write $\psi - \psi_0$ as a sum of a map θ' of support size at most $2(r_{H_1} t_{H_1} + r'_{H_1})$ and a map ψ' that takes values in the subgroup $H_2 = \ker \chi$ of H_1. Since χ is not the trivial character, H_2 is not the whole of H_1.

To summarize, we can write each $\phi_i = \theta_i + \psi_i$ as $\theta_i + \theta'_i + \psi_0 + \psi'_i$, where the θ_i form a symmetric sunflower of radius at most r_{H_1}, the θ'_i have support size at most $2(r_{H_1} t_{H_1} + r'_{H_1})$, ψ_0 is fixed, and the ψ'_i take values in H_2. Note that the maps $\theta_i + \theta'_i + \psi_0$ are contained in a ball of radius at most $r_{H_2} = r_{H_1}(2t_{H_1} + 1) + 2r'_{H_1}$.

Let us write $\theta_i + \theta'_i + \psi_0 = \gamma + \eta_i$, where γ is the centre of the ball and η_i is a map of support size at most r_{H_2}. There are at most $|H|^d$ homomorphisms from G to H (where again d is the minimal size of a generating set of G), and therefore at most $|H|^{dr_{H_2}}$ possible coefficient distributions for the η_i. So by Remark 3.2 we may assume that the η_i have the same coefficient distribution, and then conclude by the inductive hypothesis. The inductive step is now complete.

In the case where $H_1 = \{e\}$ and $r_{H_1} = 0$, which is the base case of the induction, the result is immediate. If only the first or only the second of these equalities hold, then the inductive step is as described above, but with only the second situation or only the first situation, respectively, rather than both of them.

It remains to remark that we assumed that there was a set E such that $E_i = E$ for all the conditions (ϕ_i, E_i) in the collection. As usual, this is permissible by Remark 3.2. □

6. Polynomial conditions with restricted alphabets

The approximation results we have proved so far have concerned linear conditions. The aim of this section is to prove an approximation result for polynomial conditions. Let V be a k-dimensional vector space over $\mathbb{F}_p$. We shall call a map $\phi : \mathbb{F}_p^n \to V$ a *polynomial map of degree at most d* if the map $a^*\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ defined by $x \mapsto a^*\phi(x)$ is a polynomial of degree at most d for each $a^* \in V^*$. If $V = \mathbb{F}_p^k$, then we can regard a polynomial map $\phi : \mathbb{F}_p^n \to \mathbb{F}_p^k$ as a k-tuple $(\phi_1, \dots, \phi_k)$ of polynomials on $\mathbb{F}_p^n$, as we did in the introduction, in which case the degree is simply the maximum degree of the ϕ_i, but for most of the argument it will be convenient to think of polynomials in this more basis-free way. (However, we do not replace $\mathbb{F}_p^n$ by an arbitrary n-dimensional vector space, since we are considering sets of the form S^n, which are tied to the standard basis of $\mathbb{F}_p^n$.)

Thus, we now define a *k-dimensional condition with degree at most d* to be a pair (ϕ, E), where $\phi : \mathbb{F}_p^n \to V$ for some k-dimensional vector space V over $\mathbb{F}_p$, ϕ is a polynomial map of degree at most d, and $E \subset V$. We shall prove Theorem 1.5, which stated that for every $\epsilon > 0$, every prime p, every subset $S \subset \mathbb{F}_p$, every positive integer k and every positive integer d there exists K such that every set of k-dimensional conditions with degree at most d on S^n can be ϵ-approximated by a polynomial condition of degree at most d and dimension at most K.

As commented in the introduction, this result is a little different from our earlier approximation results, since we approximate by a single condition, but to pay for that we allow the dimension of that condition to be larger than the dimensions of the conditions we started with. It is not clear whether the natural direct generalization of our linear results is true – there does not seem to be an obvious analogue of the sunflower structure that played an important role in the proofs.

Given a k-dimensional space V and a polynomial map $\phi : \mathbb{F}_p^n \to V$ of degree at most d, for each $0 \leq i \leq d$ let $W_i^* \subset V^*$ be the space of all $a^* \in V^*$ such that $a^*\phi$ has degree at most i. We can take a basis of W_0^*, extend it to a basis of W_1^*, and so on. Let V_i^* be the subspace of V^* spanned by the basis vectors that belong to $W_i^* \setminus W_{i-1}^*$. (Since the basis is not uniquely determined, the subspaces V_i^* are not uniquely determined either, but this will not matter to us.) Let $k_i = \dim V_i^*$ and note that $V^* = V_0^* \oplus \cdots \oplus V_d^*$, so $k_0 + k_1 + \cdots + k_d = k$. We shall call the sequence $(k_0, k_1, \dots, k_d)$

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

the *degree profile* of the polynomial ϕ. Let $V = V_0 \oplus \cdots \oplus V_d$ be the corresponding decomposition of V – that is, we take the dual basis and the subspaces spanned by the corresponding subsets. Let P_i be the projection from V to V_i (that is, the coordinate projection with respect to this basis) and note that if a^* is a non-zero element of V_i^*, then $a^* P_i \phi$ has degree exactly i.

We shall prove Theorem 1.5 by induction on the degree profile in the colex ordering. That is, we shall prove the result for polynomial conditions where the polynomials have a given degree profile $(k_0, k_1, \dots, k_d)$ assuming the result for polynomial conditions with degree profile $(k'_0, k'_1, \dots, k'_d)$ such that there exists i for which $k'_i < k_i$ and $k'_j = k_j$ for every $j > i$.

The proof will appear somewhat simpler than several of the proofs in earlier sections. This is partly because allowing the dimension to increase gives us considerable extra flexibility, but the simplicity is also partly illusory, as the proof relies on the main result of [2], which in turn relies on one of the main results of [5].

The result from [2] concerns the equidistribution of high rank polynomials on sets of the form S^n, where $S \subset \mathbb{F}_p$ is non-empty. When $S = \mathbb{F}_p$ it is a result of Green and Tao [3, Theorem 1.7]. Before stating it, let us define a notion of rank for polynomials on restricted alphabets. The definition we give here is less granular than the one we gave in the paper [2], but it suffices for our purposes and will be slightly more convenient to use.

Definition 6.1. *Let p be a prime, let d be a nonnegative integer, and let $P : \mathbb{F}_p^n \to \mathbb{F}_p$ be a polynomial with degree at most d. We define the* degree-d rank $\mathrm{rk}_d P$ *of P as follows.*

1. *If $d = 0$ then $\mathrm{rk}_d P = 0$.*
2. *If $d = 1$ then $\mathrm{rk}_d P$ is the number of indices $i \in [n]$ such that $a_i \neq 0$ when we represent P in the form $P(x) = c + \sum_{i=1}^n a_i x_i$, where $c, a_1, \dots, a_n \in \mathbb{F}_p$ – that is, the size of the support of the linear form that differs from P by an additive constant.*
3. *If $d \geq 2$ then $\mathrm{rk}_d P$ is the smallest nonnegative integer r such that there exists a function $F : \mathbb{F}_p^r \to \mathbb{F}_p$ and polynomials $P_1, \dots, P_r : \mathbb{F}_p^n \to \mathbb{F}_p$ each with degree strictly smaller than $\deg P$ such that $P = F(P_1, \dots, P_r)$.*

If S is a non-empty subset of $\mathbb{F}_p$, then the degree-d rank of P with respect to S, *denoted by* $\mathrm{rk}_{d,S} P$, *is* $\min_{P_0} \mathrm{rk}_d (P - P_0)$, *where the minimum is taken over all polynomials $P_0 : \mathbb{F}_p^n \to \mathbb{F}_p$ with degree at most d such that $P_0(S^n) = \{0\}$. Equivalently, it is the minimum degree-d rank of any polynomial Q with degree at most d that agrees with P on S^n.*

The next result is very similar to Theorem 1.4 of [2] and is an immediate consequence of it. It shows that the values of a high-rank polynomial are approximately equidistributed on sets of the form S^n.

Theorem 6.2. *For every prime p, every non-empty subset $S \subset \mathbb{F}_p$, every positive integer $d < p$, and every $\epsilon > 0$ there exists τ such that the following holds. If $P : \mathbb{F}_p^n \to \mathbb{F}_p$ is any polynomial of degree d with $\mathrm{rk}_{d,S}(P) \geq \tau$, then $|\mathbb{E}_{x \in S^n} \omega_p^{P(x)}| \leq \epsilon$.*

We shall also simplify the argument in a superficial way by being less explicit about how various bounds depend on each other. Given a property κ of polynomial conditions such that every condition that satisfies κ has degree at most d, let us say that conditions that satisfy κ are *boundedly approximable on S^n*, or simply that they are *approximable*, if for every $\epsilon > 0$ there exists $K = K(\epsilon, \kappa, S)$ such that every set of polynomial conditions on S^n that satisfy κ can be ϵ-internally approximated by a polynomial condition of degree at most d and dimension at most K. We shall frequently make implicit use of the following lemma, which is very similar to Remark 3.2.

Lemma 6.3. *Let κ be a property of polynomial conditions, and let $\mathcal{P}$ be a family of properties of polynomial conditions. Suppose that there exists a constant $M = M(\kappa, S)$ such that every set of conditions $\mathcal{C}$ that satisfy κ is a union of at most M sets of conditions each satisfying a property of $\mathcal{P}$.*

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

If for every $\lambda \in \mathcal{P}$, conditions that satisfy λ are approximable on S^n, then conditions that satisfy κ are approximable on S^n.

Proof. Let $\epsilon > 0$ and let $\mathcal{C}$ be a set of conditions that satisfy κ and write $\mathcal{C}$ as $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_m$, where $m \leq M$ and the conditions in each $\mathcal{C}_i$ satisfy λ_i. For each set $\mathcal{C}_i$, let (ϕ_i, E_i) be a condition that (ϵ/m)-approximates $\mathcal{C}_i$ (or more precisely, the set $\{(\phi_i, E_i)\}$ approximates $\mathcal{C}_i$), where $E_i \subset \mathbb{F}_p^{k_i}$ for some $k_i \leq K$. Then let $k = k_1 + \cdots + k_m$, let $\phi : S^n \to \mathbb{F}_p^k$ be defined by $\phi(x) = (\phi_1(x), \ldots, \phi_m(x))$, and let $E = E_1 \times \cdots \times E_m$. Then x satisfies (ϕ, E) if and only if it satisfies (ϕ_i, E_i) for each i. Also, if x satisfies $\mathcal{C}$ but not (ϕ, E), then there is some i such that it fails to satisfy (ϕ_i, E_i) but satisfies $\mathcal{C}_i$. For each i, the density of such x is at most ϵ/m, so we are done by a union bound. □

If we have a set of conditions that we wish to approximate, this lemma allows us to partition it into a bounded number of pieces and approximate each piece. We have done this several times already in the paper, but from now on we shall not bother to give a careful justification each time we do so.

We shall also use the following more elaborate version of the principle, which relies heavily on the fact that we allow ourselves to increase the dimensions of conditions as we approximate. If κ is a property of polynomials, we shall say that a set of conditions satisfies κ if the corresponding polynomials satisfy κ.

Lemma 6.4. *Let p be a prime, let S be a non-empty subset of $\mathbb{F}_p$, let k and d be positive integers, let V be a k-dimensional vector space over $\mathbb{F}_p$, let $\rho : \mathbb{F}_p^n \to V$ be a polynomial map of degree at most d, and let κ and λ be properties of polynomials that imply that they have degree at most d. Suppose that k is bounded above by a function of p, S and d. Suppose also that for every $y \in V$ and every polynomial ϕ that satisfies κ there is a polynomial ψ that satisfies λ such that ϕ and ψ agree on the set $\{x \in \mathbb{F}_p^n : \rho(x) = y\}$. Finally, suppose that polynomial conditions that satisfy λ are approximable on S^n. Then polynomial conditions that satisfy κ are approximable on S^n.*

Proof. Let $\mathcal{C}$ be a set of conditions that satisfy κ. For each $y \in V$ and each condition $(\phi, E) \in \mathcal{C}$, where E is a subset of a vector space W, let

$$E_y = \{(v, w) \in V \times W : v = y \implies w \in E\},$$

so that $((\rho, \phi), E_y)$ is the condition $\rho(x) = y \implies \phi(x) \in E$.

For each $y \in V$, let $\mathcal{C}_y$ be the set of conditions $\{((\rho, \phi), E_y) : (\phi, E) \in \mathcal{C}\}$. If x satisfies (ϕ, E) then it clearly satisfies $((\rho, \phi), E_y)$ for every $y \in V$, and the converse is true as well, since if we let $y = \rho(x)$, then the fact that x satisfies $((\rho, \phi), E_y)$ guarantees that $\phi(x) \in E$, as noted above. Therefore, the satisfying sets of $\mathcal{C}$ and $\bigcup_{y \in V} \mathcal{C}_y$ are the same.

By Lemma 6.3 it is therefore sufficient to prove that each set $\mathcal{C}_y$ of conditions is approximable. Let us fix $y \in V$, and for each $(\phi, E) \in \mathcal{C}$ let $\tilde{\phi}$ be a polynomial that satisfies λ and that agrees with ϕ on $\{x \in \mathbb{F}_p^n : \rho(x) = y\}$. Let $\mathcal{D}_y$ be the set of conditions $\{(\tilde{\phi}, E) : (\phi, E) \in \mathcal{C}\}$ and let $\tilde{\mathcal{C}}_y$ be the set of conditions $\{((\rho, \tilde{\phi}), E_y) : (\phi, E) \in \mathcal{C}_y\}$. Note that the satisfying set of $\tilde{\mathcal{C}}_y$ is the same as that of $\mathcal{C}_y$.

Since the polynomials $\tilde{\phi}$ satisfy λ, and polynomial conditions that satisfy λ are approximable, $\mathcal{D}_y$ is approximable by a condition (θ, F). It follows that $\tilde{\mathcal{C}}_y$ is approximable by the condition $((\rho, \theta), F_y)$. □

We are now ready to begin the proof in earnest. The basic strategy is the same as it has been with our earlier proofs: either the conditions we wish to approximate contain a large set of 'sufficiently independent' conditions, in which case the satisfying set has small density, or there is a small set of conditions that in some suitable sense 'approximately generates' all the other conditions, in which case we can use Lemma 6.4 to reduce to a simpler case that is covered by an inductive hypothesis.

Let us begin with a standard condition for determining whether a function taking values in a k-dimensional vector space over $\mathbb{F}_p$ is approximately equidistributed.

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

Lemma 6.5. *Let X be a finite set, let p be a prime, let k be a positive integer, let V be a k-dimensional vector space over $\mathbb{F}_p$, and let $\phi : X \to V$. Then for every $y \in V$,*

$$|\mathbb{P}[\phi(x) = y] - p^{-k}| \leq p^{-k} \sum_{\substack{a^* \in V^* \\ a^* \neq 0}} \big| \mathop{\mathbb{E}}_x \omega_p^{a^*(\phi(x))} \big|.$$

Proof. Note first that $\mathbb{P}[\phi(x) = y] = \mathbb{E}_x \mathbb{1}_y(\phi(x))$, where we write $\mathbb{1}_y$ for the characteristic function of the singleton set $\{y\}$. By the Fourier inversion formula,

$$\mathop{\mathbb{E}}_x \mathbb{1}_y(\phi(x)) = \mathop{\mathbb{E}}_x \sum_{a^* \in V^*} \widehat{\mathbb{1}_y}(a^*) \omega_p^{a^*(\phi(x))} = p^{-k} \sum_{a^* \in V^*} \mathop{\mathbb{E}}_x \omega_p^{a^*(\phi(x)-y)}.$$

The result follows. □

The above lemma immediately implies a version for several functions, but it is convenient to state it separately.

Corollary 6.6. *Let X be a finite set, let p be a prime, let r and $k_1, \dots, k_r$ be positive integers, let $k = \sum_i k_i$, for each $i \leq r$ let V_i be a vector space over $\mathbb{F}_p$ of dimension k_i, and let $\phi_i : X \to V_i$. Then for every $(y_1, \dots, y_r) \in V_1 \times \cdots \times V_r$,*

$$|\mathbb{P}[\forall i\ \phi_i(x) = y_i] - p^{-k}| \leq p^{-k} \sum_{\substack{a^* \in V_1^* \times \cdots \times V_r^* \\ a^* \neq 0}} \big| \mathop{\mathbb{E}}_x \omega_p^{\sum_i a_i^*(\phi_i(x))} \big|.$$

Proof. We can apply the lemma to the map $\phi : X \to V_1 \times \cdots \times V_r$ defined by $x \mapsto (\phi_1(x), \dots, \phi_r(x))$. □

A simple further consequence of this corollary and Theorem 6.2 is the following result.

Corollary 6.7. *Let $\epsilon > 0$, let p be a prime, let $d < p$ and k be positive integers, and let S be a non-empty subset of $\mathbb{F}_p$. There exist R, τ such that for every $r \geq R$, every sequence $V_1, \dots, V_r$ of vector spaces over $\mathbb{F}_p$ with $\dim V_i \leq k$ for each i, every sequence of polynomial maps $\phi_1, \dots, \phi_r$ of degree at most d with $\phi_i : \mathbb{F}_p^n \to V_i$, and every sequence $E_1, \dots, E_r$ with E_i a proper subset of V_i the following holds. If the satisfying set of the conditions (ϕ_i, E_i) on S^n has density greater than ϵ, then there exist $a_1^*, \dots, a_r^*$ with $a_i^* \in V_i^*$, not all zero, such that $\sum_i a_i^* \phi_i$ is a polynomial of degree $d' \leq d$ with $\mathrm{rk}_{d',S}(\sum_i a_i^* \phi_i) \leq \tau$.*

Proof. Let R be the smallest integer such that $(1 - p^{-k})^R \leq \epsilon/2$. Then the density of the set $E_1 \times \cdots \times E_R$ in $V_1 \times \cdots \times V_R$ is at most $\epsilon/2$, so by Corollary 6.6 with $X = S^n$ we find that

$$\mathbb{P}_{x \in S^n}[(\phi_1(x), \dots, \phi_R(x)) \in E_1 \times \cdots \times E_R] \leq \epsilon/2 + (\epsilon/2) \sum_{\substack{a^* \in V_1^* \times \cdots \times V_R^* \\ a^* \neq 0}} \big| \mathop{\mathbb{E}}_{x \in S^n} \omega_p^{\sum_i a_i^*(\phi_i(x))} \big|.$$

If the satisfying set of the conditions (ϕ_i, E_i) with $i \leq r$ has density at least ϵ in S^n, then this is also true with $i \leq R$ instead, so by the above inequality there exists a non-zero $a^* \in V_1^* \times \cdots \times V_R^*$ such that

$$\big| \mathop{\mathbb{E}}_{x \in S^n} \omega_p^{\sum_i a_i^*(\phi_i(x))} \big| \geq 1/p^{Rk}.$$

Let $\psi = \sum_i a_i^* \phi_i$ and let d' be the degree of ψ. By Theorem 6.2 there is a constant τ depending only on p, S, ϵ, k such that $\mathrm{rk}_{d',S}(\psi) \leq \tau$. This proves the result. □

We now prove an easy result that will turn out to be the base case of an inductive argument. It also has a very similar structure to that of the main proof. Let us define the *support* $Z(\phi)$ of a linear map $\phi : \mathbb{F}_p^n \to V$ to be the set of coordinates on which ϕ depends. That is, if $e_1, \dots, e_n$ is

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

the standard basis, then $Z(\phi) = \{i : \phi(e_i) \neq 0\}$. If ϕ is an affine map, then we define $Z(\phi)$ to be the support of its linearization, and we also call it the support.

Lemma 6.8. *Let p be a prime, let s and k be positive integers and let S be a non-empty subset of $\mathbb{F}_p$. Then conditions of the form (ϕ, E) where ϕ is an affine map from $\mathbb{F}_p^n$ to some vector space V of dimension at most k, $E \subset V$, and ϕ has support size at most s, are approximable on S^n.*

Proof. We prove the result by induction on s, observing that it is trivial when $s = 0$. We now assume that the result holds for $s - 1$ and prove it for s.

Let $\epsilon > 0$ and let $\mathcal{C}$ be a set of k-dimensional affine conditions of support size at most s. We may assume that each (ϕ, E) is such that there exists $x \in S^n$ with $\phi(x) \notin E$, since we can throw away all conditions that do not satisfy this property.

Now let $\{(\phi_1, E_1), \dots, (\phi_r, E_r)\}$ be a maximal subset of $\mathcal{C}$ such that the maps $\phi_1, \dots, \phi_r$ are disjointly supported. Then the events $\phi_i(x) \in E_i$ are independent, and each holds with probability at most $1 - |S|^{-s}$. Therefore, if $(1 - |S|^{-s})^r < \epsilon$, we can approximate $\mathcal{C}$ by a single condition $(\phi, \emptyset)$.

If $(1 - |S|^{-s})^r \geq \epsilon$, then for every $(\phi, E) \in \mathcal{C}$ the support of ϕ intersects the union Z of the supports of $\phi_1, \dots, \phi_r$. It follows that ϕ agrees with a map of support size at most $s - 1$ on each set of the form $\{x : \forall j \in Z\ x_j = y_j\}$. By the inductive hypothesis and Lemma 6.4 we are done. □

Proof of Theorem 1.5. There are at most $(d+1)^k$ degree profiles that the polynomials can have, so by Lemma 6.3 we may assume that they all have the same degree profile. We may also throw away all conditions (ϕ, E) where $E = \mathbb{F}_p^k$. As promised in the introduction to this section, we now prove the result by induction on the degree profile, with respect to the colex order.

Let $\epsilon > 0$ and let $\mathcal{C}$ be a set of k-dimensional polynomial conditions (ϕ, E) such that each polynomial ϕ has degree profile $(k_0, k_1, \dots, k_d)$. If the satisfying set of $\mathcal{C}$ in S^n has density less than ϵ, then we can approximate it with a single condition of the form $(\phi, \emptyset)$. Otherwise, Corollary 6.7 gives us R and τ such that for any r polynomials $\phi_1, \dots, \phi_r$ involved in the conditions in $\mathcal{C}$, if $r > R$ then there exist $a_1^*, \dots, a_r^*$ not all zero such that $\sum_{i=1}^r a_i^* \phi_i$ has rank at most τ. (By this we mean it has degree d' for some $d' \leq d$ and (d', S)-rank at most τ.)

Let $\phi_1, \dots, \phi_r$ be a maximal set of polynomials coming from the conditions in $\mathcal{C}$ for which $\sum_{i=1}^r a_i^* \phi_i$ has rank above τ for all $(a_1^*, \dots, a_r^*) \neq 0$, and suppose that $\phi_i : \mathbb{F}_p^n \to V_i$. Then for every $(\phi, E) \in \mathcal{C}$ with $\phi : \mathbb{F}_p^n \to V$ there exist $a_1^*, \dots, a_r^*$ and b^* with $a_i^* \in V_i^*$ and $b^* \in V^* \setminus \{0\}$ such that $\sum_{i=1}^r a_i^* \phi_i + b^* \phi$ has rank at most τ.

After a further application of Lemma 6.3, we may assume that $(a_1^*, \dots, a_r^*, b^*)$ is the same for each $(\phi, E) \in \mathcal{C}$. Let ψ be the polynomial $\sum_{i=1}^r a_i^* \phi_i$, so $\psi + b^* \phi$ has rank at most τ for each $(\phi, E) \in \mathcal{C}$.

Let us now fix $(\phi, E) \in \mathcal{C}$. On each set $\{x \in S^n : \psi(x) = y\}$, the polynomials $b^*\phi$ and $\psi + b^*\phi - y$ agree, so $b^*\phi$ agrees with a polynomial of rank at most τ. Let $V^* = V_0^* \oplus \cdots \oplus V_d^*$ be a decomposition as described at the beginning of the section, so $\dim V_i^* = k_i$ and $w^*\phi$ has degree exactly i for every $w^* \in V_i^*$. Let $b^* = b_0^* + b_1^* + \cdots + b_d^*$ with $b_i^* \in V_i^*$.

Then if $b^*\phi$ has degree d', it follows that $b_{d'}^* \neq 0$ and $b_i^* = 0$ for every $i > d'$. Now pick $w \in V_{d'}$ such that $b_{d'}^*(w) = 1$. Then we can write

$$\phi(x) = b_{d'}^*\phi(x)w + (\phi(x) - b_{d'}^*\phi(x)w)$$

and the image of the polynomial map $\phi(x) - b_{d'}^*\phi(x)w$ lies in the kernel of $b_{d'}^*$.

Since $b^*\phi$ agrees with a polynomial of degree d' and rank at most τ, and since $(b_0^* + \cdots + b_{d'-1}^*)\phi$ has degree less than d', $b_{d'}^*\phi$ has rank at most $\tau + 1$. We can therefore write it in the form $F(\gamma_1, \dots, \gamma_t)$ with $t \leq \tau + 1$ and with $\gamma_1, \dots, \gamma_t$ all of degree less than d'. This allows us to replace the condition (ϕ, E) by a condition with a degree profile that comes earlier than $(k_0, k_1, \dots, k_d)$ and that agrees with ϕ on $\{x \in S^n : \psi(x) = y\}$. Let $\tilde{\phi}$ be the polynomial $\phi - (b_{d'}^*\phi)w$, regarded as taking values in $\ker(b_{d'}^*)$, and note that $\tilde{\phi}$ has degree profile $(k_0, k_1, \dots, k_{d'-1}, k_{d'} -$

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

$1, k_{d'+1}, \dots, k_d$). Let $\iota : \ker(b^*_{d'}) \to V$ be the inclusion map, and let $\theta : \ker(b^*_{d'}) \times \mathbb{F}_p^t \to V$ be the map $(u, a) \mapsto \iota u + aw$. Now let us take the polynomial $(\tilde{\phi}, \gamma_1, \dots, \gamma_t)$ and let $H \subset \ker(b^*_{d'}) \times \mathbb{F}_p^t$ be the set $\{(v, u) : \theta(v, F(u)) \in E\}$.

Then $(\tilde{\phi}, \gamma_1, \dots, \gamma_t)(x) \in H$ if and only if $\iota\tilde{\phi}(x) + F(\gamma_1(x), \dots, \gamma_t(x)) \in E$, which is true if and only if $\phi(x) - b^*_{d'}\phi(x)w + F(\gamma_1(x), \dots, \gamma_t(x))w \in E$, which is true if and only if $\phi(x) \in E$. Thus, the condition (ϕ, E) is equivalent to the condition $((\tilde{\phi}, \gamma_1, \dots, \gamma_t), H)$. But $\tilde{\phi}$ has degree profile $(k_0, k_1, \dots, k_{d'-1}, k_{d'} - 1, k_{d'+1}, \dots, k_d)$, as already noted, and since each γ_i has degree less than d', the degree profile of $(\tilde{\phi}, \gamma_1, \dots, \gamma_t)$ precedes that of ϕ in colex order. Furthermore, since t is bounded independently of n, all the degrees in the degree profile of $(\tilde{\phi}, \gamma_1, \dots, \gamma_t)$ are bounded independently of n. Hence we are done by the inductive hypothesis and Lemma 6.4, once we have established the base case.

Recall that for polynomials of degree 1, we define the rank to be the size of the support of the linearization. So the inductive argument just given allows us to reduce to the case where all the polynomials are linear and have bounded support size. This case is dealt with by Lemma 6.8. □

7. Conclusion and open questions

We finish by mentioning two questions to which we do not know the answers. One is the obvious one of improving the bounds in our approximation results. Even in the most basic approximation result, Proposition 1.3, the question of the optimal bounds is not yet settled.

Question 7.1. *Can we replace the bound* $(p \log \epsilon^{-1})^p$ *by* $O_p(\log \epsilon^{-1})$ *in Proposition 1.3 question-mark*

If this is indeed possible, then we may ask whether we can also take bounds $O(\log \epsilon^{-1})$ in Theorems 1.2, 1.4, and 1.5. In the case of Theorem 1.5 it would probably help to first have linear bounds $H_{p,d,S}(\epsilon) = O_{p,d,S}(\log \epsilon^{-1})$ in Theorem 6.2 on the equidistribution of polynomials.

In a second direction, our first approximation result for linear forms $\mathbb{F}_p^n \to \mathbb{F}_p$, Proposition 1.2, involves internal approximation. We can also ask for an external approximation, where it also does not seem obviously unreasonable to require the conditions to belong to the original family of conditions. (In the case of internal approximations, it is not possible to add this additional requirement, as it would imply that the set of $x \in S^n$ satisfying the approximating set of conditions is exactly the same as that satisfying the original set of conditions, which is not always possible, as can be seen with an example we gave earlier, where we take S to be $\{0, 1\}$ and the family of linear forms to be $\{(x_1 + x_i, \{0, 1\}) : i \in [2, n]\}$.)

Conjecture 7.2. *Let p be a prime, let S be a non-empty subset of $\mathbb{F}_p$, and let $\epsilon > 0$. There exists a constant $A = A(p, S, \epsilon)$ such that for every set $\mathcal{F}$ of mod-p conditions on S^n there is a subset $\mathcal{G}$ of $\mathcal{F}$ such that the satisfying set of $\mathcal{G}$ ϵ-externally approximates that of $\mathcal{F}$ (meaning that the density of points that satisfy $\mathcal{G}$ but not $\mathcal{F}$ is at most ϵ).*

Data availability statement

This paper is available on arXiv at https://arxiv.org/abs/2306.00747.

Acknowledgements

The preparation of this paper until an earlier version was released on the arXiv (in June 2023) and subsequently submitted to the journal began while the second author was affiliated with the University of Cambridge, supported by an EPSRC International Doctoral Scholarship (project

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press

reference 2114524), and ended while he was affiliated with the University of Oxford, supported by ERC Advanced Grant 883810.

Competing interests

The authors declare none.

References

[1] Furstenberg, H. and Katznelson, Y. (1991) A density version of the Hales-Jewett theorem. *J. Anal. Math.* **57** 64–119.
[2] Gowers, W. T. and Karam, T. (2022) Equidistribution of high-rank polynomials with variables restricted to subsets of $\mathbb{F}_p$. arXiv:2209.04932.
[3] Green, B. and Tao, T. (2009) The distribution of polynomials over finite fields, with applications to the Gowers norms. *Contrib. Discret. Math.* **4**(2) 1–36.
[4] Gowers, W. T. and Karam, T. Modular obstructions to uniformity in the polynomial density Hales-Jewett conjecture, in preparation.
[5] Karam, T. (2026) High-rank subtensors of high-rank tensors. *Adv. Comb.* **10**, 70 pp.
[6] Polymath, D. H. J. (2012) A new proof of the density Hales-Jewett theorem. *Ann. Math.* **175** 1283–1327.

Cite this article: Gowers WT and Karam T (2026). Low-complexity approximations for sets defined by generalizations of affine conditions. *Combinatorics, Probability and Computing*. https://doi.org/10.1017/S0963548326100571

https://doi.org/10.1017/S0963548326100571 Published online by Cambridge University Press