A Weak Regularity Lemma for Polynomials

Guy Moshkovitz
City University of New York, NY, USA

Dora Woodruff
Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

A regularity lemma for polynomials provides a decomposition in terms of a bounded number of approximately independent polynomials. Such regularity lemmas play an important role in numerous results, yet suffer from the familiar shortcoming of having tower-type bounds or worse. In this paper we design a new, weaker regularity lemma with strong bounds. The new regularity lemma in particular provides means for quantitatively studying the curves contained in the image of a polynomial map, which is beyond the reach of standard methods.

The weak regularity lemma turns out to be powerful enough to yield results on arithmetic circuits and polynomial ranks that may be of independent interest:

- A general upper bound on the arithmetic circuit size of low-degree polynomials based solely on their image.
- An upper bound on the top fan-in of depth-4 arithmetic formulas under similar conditions.
- A quantitative bound for the Green-Tao notion of rank for polynomials, significantly improving on a result of Karam.

2012 ACM Subject Classification Theory of computation → Algebraic complexity theory

Keywords and phrases weak regularity lemma, finite-field polynomials, polynomial maps, structure-versus-randomness, arithmetic circuits

Digital Object Identifier 10.4230/LIPIcs.CCC.2026.14

Related Version *Full Version*: `https://arxiv.org/abs/2509.21536`

Funding *Guy Moshkovitz*: NSF Award DMS-2302988 and a PSC-CUNY award. The project began while D.W. was a summer graduate research assistant supported by this NSF award.

Acknowledgements We are grateful to the anonymous referees for multiple useful comments.

1 Introduction

Regularity lemmas for polynomials have served as a key component in a variety of important results, as in the proof of the finite-field Gowers inverse conjecture in additive combinatorics and higher-order Fourier analysis (e.g. [15, 38]), Stillman's conjecture in commutative algebra [1], and the Reed-Muller weight distribution problem in coding theory [3, 23], to name but a few. They serve as a key component in higher-order arithmetic regularity lemmas (e.g., [13, 14, 15, 17, 22]).[1] As is usually the case with regularity lemmas, they unfortunately only have weak bounds, which are typically inherited by applications. In this paper we introduce a weaker regularity lemma for polynomials, with far better bounds, that is nevertheless powerful enough for some applications. For example, we obtain strong bounds for a result of Karam [21] on the relation between the generalized rank of a polynomial and its range. Our results also yield a method to prove *upper* bounds for arithmetic circuits.

[1] Roughly speaking, arithmetic regularity lemmas approximate general functions in terms of low-degree polynomials that are approximately independent. (This approximate independence is indispensable; to quote Green [14]: second-order arithmetic regularity lemmas "reduce the study of general functions... to the study of projections... onto quadratic factors. This, however, is of little consequence unless we can study those supposedly simple objects.")

© Guy Moshkovitz and Dora Woodruff;
licensed under Creative Commons License CC-BY 4.0

Roughly speaking, the usual regularity lemmas for polynomials (see, e.g., Proposition 8.1 in [9], Lemma 2.4 in [15], or Lemma 3.11 in [14]) can be described as results about polynomial decompositions. Namely, they express any n-variate polynomial $P\colon \mathbb{F}^n \to \mathbb{F}$ – and more generally, any polynomial map $\mathbf{P}\colon \mathbb{F}^n \to \mathbb{F}^m$ – in terms of a bounded number (independent of n) of polynomials $\mathbf{X} = (X_1, \dots, X_k)$, that is, $P = F(\mathbf{X})$, with the property that the polynomials in $\mathbf{X}$ "behave" approximately like k (jointly independent) variables. Although the precise definition of "behave" can differ depending on the context, the intuition is as follows: for many purposes, any polynomial can be analyzed as though it had only k variables. The number k of *parts* $X_1, \dots, X_k$ in the decomposition depends only on the degree d and the number m of the given polynomials $\mathbf{P} = (P_1, \dots, P_m)$, and as mentioned above, this dependency – befitting a regularity lemma – can be quite astronomical, often having worse than a tower-type growth (see, e.g., the Remark following Lemma 2.4 in [15]).[2] A main reason for these weak bounds is that in order to guarantee approximate independence, the polynomial parts $X_1, \dots, X_k$, as well as all their linear combinations, are required to have *rank* (see Section 2) that is large as a function of k, their number. Thus, the larger the number of parts, the more difficult it is to satisfy the rank requirement, which translates to an extremely large number of steps in the process to construct them.[3]

In the setting of graphs, Frieze and Kannan [11] famously obtained a weak regularity lemma that has both strong bounds and multiple interesting applications. For comparison, the usual graph regularity lemma of Szemerédi decomposes any graph – by partitioning its vertices – into a bounded number of parts (independent of the number of vertices) with the guarantee that almost *all* graph parts are "approximately" random. The weak graph regularity lemma decomposes any graph in the same sense, but comes with only one, global guarantee. Our weak regularity lemma for polynomials is partly inspired by this approach.

1.1 Weak regularity for polynomials

The basic idea of our weak regularity lemma is to relax the requirement on the k polynomials $X_1, \dots, X_k$, and only require that at least one of them, of maximal degree, behaves like a free variable with respect to all the others. That is, conditioned on $X_2, \dots, X_k$ taking *any* given value $y \in \mathbb{F}^{k-1}$, we require that X_1 takes any value in $\mathbb{F}$ with approximately the same frequency. This imitates the case of a genuinely free variable; or more concretely, the case where $X_1 = x_1$, and $X_i = X_i(x_2, \dots, x_n)$ for $i > 1$, where $x_1, \dots, x_n$ are variables.

Intuitively, the benefit of having a weak regular decomposition for a polynomial map $\mathbf{P}$ is that it reveals some information about the univariate polynomials "induced" by $\mathbf{P}$. In particular, it enables us to argue about the image of $\mathbf{P}$ (for which standard rank-based tools are unsuitable, as they typically give information only about equidistribution); indeed, a corollary of the weak regularity lemma will be that lines that $\mathbf{X}$ intersects somewhat often must in fact be fully contained in the image of $\mathbf{X}$. This translates to a polynomial curve (univariate polynomial map) contained in the image of $\mathbf{P}$. See below for more about applications.

[2] Interestingly, recent regularization results of Lampert and Ziegler [29, 27] give strong bounds; however, unlike regularity lemmas, these do not provide a decomposition of $\mathbf{P}$ as a polynomial in $X_1, \dots, X_k$, but rather place $\mathbf{P}$ in the ideal generated by them.

[3] See, e.g., Proposition 8.1 in [9] for a clean example of this phenomenon: the reason that the process terminates at all is a decrease in the lexicographic order of the tuple of degrees of $\mathbf{X}$ in each step.

We note that throughout the paper we work with finite fields $\mathbb{F}$, for which the notion of weak regularity can be most elementarily defined. Moreover, rank-based methods, which we rely on extensively in the proof (as do proofs of other regularity lemmas for polynomials), tend to be most challenging in the finite-field setting.[4]

Formally, we define weak regularity as follows. We refer the reader to Section 2 for a more formal discussion, and the proofs. For tuples of polynomials $\mathbf{P} = (P_1, \ldots, P_m)$ and $\mathbf{X} = (X_1, \ldots, X_k)$ over $\mathbb{F}$, we write $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$ if every component P_i of $\mathbf{P}$ can be written as $F_i(X_1, \ldots, X_k)$ for some polynomial F_i. Henceforth, for any k-tuple $v = (v_1, \ldots, v_k)$, we use the notation v' for the $(k-1)$-tuple $v' = (v_2, \ldots, v_k)$.

▶ **Definition** (weak regularity). A tuple of polynomials $\mathbf{P}$ over $\mathbb{F} = \mathbb{F}_q$ has a *weak ϵ-regular decomposition* of size k if $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$, for some k homogeneous polynomials $\mathbf{X} = (X_1, \ldots, X_k)$, such that:

- for every $y \in \mathbb{F}^k$, $|\Pr[\mathbf{X} = y] - q^{-1}\Pr[\mathbf{X}' = y']| \le \epsilon q^{-k}$,
- $\mathbf{P} \not\subseteq \mathbb{F}[\mathbf{X}']$, and
- $\deg(X_1) = \max_i \deg(X_i)$.

The probabilities here are with respect to a uniformly random input; that is, $\Pr[\mathbf{X} = y]$ is a shorthand for $\Pr_{x \in \mathbb{F}^n}[\mathbf{X}(x) = y]$, with x chosen uniformly at random from $\mathbb{F}^n$. Intuitively, the probabilistic condition can be motivated as follows. If X_1 were truly an independent variable x_1, with $X_i = X_i(x_2, \ldots, x_n)$ for $i > 1$, we would certainly have both $\Pr[X_1 = y_1] = \frac{1}{q}$ and $\Pr[\mathbf{X} = y] = \Pr[X_1 = y_1]\Pr[\mathbf{X}' = y']$. Hence, the probabilistic condition mimics the behavior we would expect for an independent variable x_1.

Note that the two additional conditions at the end of the definition are necessary to avoid trivialities. This is clearly the case for the first, which guarantees that the decomposition of $\mathbf{P}$ depends on X_1. As for the other additional condition, that X_1 is of maximal degree – note that we can decompose any polynomial $P = P(x_1, \ldots, x_n)$, of degree d in x_1 say, as $\sum_{i=0}^{d} x_1^i C_i(x_2, \ldots, x_n)$; that is, we always have a small decomposition $P = F(x_1, C_0, \ldots, C_d)$ (where $F(z, y_0, y_1, \ldots, y_d) = \sum_{i=0}^{d} z^i y_i$). Although x_1 is genuinely independent of the other parts of the decomposition $C_0, \ldots, C_d$, its degree $\deg(x_1) = 1$ is generally far from the largest degree among them.

Let us also mention two extreme cases in the definition of weak regularity. First, if $\mathbf{P}$ is a tuple of forms that, viewed as a map $\mathbf{P} \colon \mathbb{F}^n \to \mathbb{F}^m$, is already equidistributed, that is, $|\Pr[\mathbf{P} = y] - q^{-m}| \le \epsilon q^{-m}$ for every $y \in \mathbb{F}^m$, then $\mathbf{P}$ trivially has a small weak ϵ-regular decomposition. Second, $\mathbf{P}$ always has the trivial decomposition $\mathbf{P} = \mathbf{P}(\mathbf{X})$ with $\mathbf{X} = (x_1, \ldots, x_n)$ its variables, which is a weak 0-regular (i.e., a perfectly weak regular decomposition; assuming without loss of generality that $\mathbf{P}$ depends on x_1), though of a trivially large size n.

Our main result is that a rather small weak regular decomposition always exists (see Theorem 8).

▶ **Theorem 1** (weak regularity lemma). *Let $\epsilon = q^{-c}$ with $c > 1$. Every polynomial m-tuple $\mathbf{P}$ over $\mathbb{F} = \mathbb{F}_q$, of degrees at most d with $d < \mathrm{char}(\mathbb{F})$, has a weak ϵ-regular decomposition of size at most $(mc)^{2^{d(1+o(1))}}$.*[5]

[4] To get an analogue of our weak regularity lemma for infinite fields would require replacing the probabilistic notion of approximate independence in the definition of weak regularity below, for example by a geometric or by a commutative algebra analogue (see, e.g., [28, 9]), and adapting the proof accordingly. Technically, this would require replacing the notion of bias of a polynomial (see Section 2.2) by standard analogues for infinite fields. For the structure-vs-randomness theorem we rely on (Theorem 17), a geometric analogue over infinite fields has long been known [25, 36] with, in fact, a slightly better, linear bound.

[5] Throughout, the $o(1)$ term in $d(1 + o(1))$ goes to 0 as d goes to infinity.

For comparison, the usual regularity lemmas for polynomials only give tower-type bounds or worse; namely, the bounds are at the very least a tower of exponents, of height depending on d, with m at the top. Our weak regularity lemma, in particular, is polynomial in m (which is especially relevant in the important case of low-degree polynomials $d = O(1)$).

The proof of our weak regularity lemma relies on several components, including a combinatorial regularization process over "punctured" spaces of polynomials, and a Fourier-analytic quasirandomness argument over such spaces. See Section 1.4 for more information.

1.2 Applications for arithmetic circuits

Using the weak regularity lemma, one can prove results that go beyond standard notions of structure for polynomials such as rank or partition rank. These standard notions only give "shallow" decompositions – their building blocks are generally just one degree lower – whereas we will be interested in decompositions that go "deeper". Concretely, in the next two sections we present applications that give novel structural results for polynomials based solely on their image.

Inspired by [21], we make the following definition for polynomial maps (that is, tuples of polynomials). Define the *univariate degree* of a polynomial map $\mathbf{P}\colon \mathbb{F}^n \to \mathbb{F}^m$ to be the smallest degree of a non-constant curve contained in its image:

$$\mathrm{udeg}(\mathbf{P}) = \min\{\deg(\mathbf{U}) \mid \mathbf{U}\colon \mathbb{F} \to \mathbb{F}^m \text{ non-constant: } \operatorname{Im}\mathbf{U} \subseteq \operatorname{Im}\mathbf{P}\}.$$

By convention, if $\mathbf{P}$ is constant then $\mathrm{udeg}(\mathbf{P}) = \infty$. Note that $1 \le \mathrm{udeg}(\mathbf{P}) \le \deg(\mathbf{P})$ $(:= \max_i \deg(P_i))$ for any non-constant $\mathbf{P}$;[6] see [21] and Section 3.1 for more about this.

▶ Remark 2. Univariate degree is related to the concepts of subspace-evasive sets and variety-evasive sets over finite fields, which are sets that have small intersections with, in particular, any line and any curve of low degree, respectively. (Explicit constructions were given by Dvir, Kollár, and Lovett [7, 6].) In particular, if $\mathrm{Im}(\mathbf{P})$ evades curves in the sense that $|\mathrm{Im}(\mathbf{P}) \cap C| < |\mathbb{F}|/u$ for every curve $C \subseteq \mathbb{F}^m$ of degree u, then $\mathrm{udeg}(\mathbf{P}) > u$. (Indeed, otherwise $C \subseteq \mathrm{Im}(\mathbf{P})$ with $\deg(C) = u$, and $|C| \ge |\mathbb{F}|/u$.) However, the converse does not hold – large $\mathrm{udeg}(\mathbf{P})$ is about $\mathrm{Im}(\mathbf{P})$ avoiding containment of entire curves, and so is a weaker property than $\mathrm{Im}(\mathbf{P})$ avoiding any significant intersections with curves. Thus, applications requiring large $\mathrm{udeg}(\mathbf{P})$, as we give in this paper, require an even weaker condition on $\mathrm{Im}(\mathbf{P})$ than being variety evasive.

Let $L(P)$ denote the arithmetic formula size of a polynomial P. It is known that for most degree-d polynomials P in n variables, over any field, $L(P) \ge \binom{n+d}{d}^{1-o(1)}$. Indeed, a nice argument of Hrubeš and Yehudayoff (Theorem 3.7 in [19]) shows that this holds even when restricted to polynomials with coefficients in $\{0, 1\}$. In fact, a similar conclusion holds more generally for arithmetic circuits.[7]

We use the weak regularity lemma to show that for low-degree polynomials, high univariate degree implies the existence of formulas that are considerably smaller than the typical $n^{d-o(1)}$, with a power-saving bound.

▶ **Theorem 3.** *Let P be an n-variate polynomial of degree at most d and univariate degree at least u, over any finite field $\mathbb{F}$ with* $\mathrm{char}(\mathbb{F}) > d$. *If $d \le \frac{1}{2}\log_2\log_2(n)$ then*

$$L(P) \le n^{\lfloor d/u \rfloor + o(1)}.$$

[6] Indeed, if $\mathbf{P}(u) \ne \mathbf{P}(v)$ then $\mathbf{U}(t) = \mathbf{P}(u + t(v - u))$ is non-constant and $\deg(\mathbf{U}) \le \deg(\mathbf{P})$.

[7] This traces back to Stoß [37], at least in the univariate case. See also Lovett [30].

We note that any improvement to the upper bound in the weak regularity lemma would translate to a less stringent condition on d in Theorem 3 (e.g., a polynomial upper bound d^C in Theorem 8 would translate to the condition $d \le n^{o(1)}$).

Our machinery also has a direct implication for the study of depth-4 arithmetic formulas. A $\sum^{[r]}\prod^{[a]}\sum\prod^{[b]}$-formula (e.g., [24, 26]) for a polynomial P is a decomposition of the form $P = \sum_{i=1}^{r}\prod_{j=1}^{a_i} Q_{i,j}$ with $a_i \le a$ and $\deg(Q_{i,j}) \le b$. It is an important challenge to understand the interplay between any of the different "fan-in" parameters (r, a, b). Trivially, if the bottom fan-in b is allowed to be as large as $\deg(P)$, then P has a $\sum^{[r]}\prod^{[a]}\sum\prod^{[b]}$-formula with $r = a = 1$. The question becomes interesting when the bottom fan-in b is bounded away from $\deg(P)$. Using the weak regularity lemma, we deduce an *upper bound* on the top fan-in r from a lower bound on the univariate degree. (It is worth noting that the model of depth-4 homogeneous arithmetic formulas with top fan-in r was shown by Kumar and Saraf [26] to be already powerful for $r = \Theta(\log n)$.)

▶ **Theorem 4.** *Let $P\colon \mathbb{F}^n \to \mathbb{F}$ with $\deg(P) \le d$, over any finite field $\mathbb{F}$ with $d < \mathrm{char}(\mathbb{F})$. Suppose $P \in \mathbb{F}[\mathbf{P}]$ with $\mathbf{P} = (P_1, \ldots, P_m)$, $\deg(\mathbf{P}) \le d$, such that $\mathrm{udeg}(\mathbf{P}) = u$. Then there is a $\sum^{[r]}\prod^{[2u]}\sum\prod^{[d/u]}$-formula for P with top fan-in $r \le (2m)^{2^{d(1+o(1))}}$. Moreover, the formula is homogeneous*[8] *if P is.*

Contrapositively, the above results can also be interpreted as implying a barrier for lower bound results: roughly, any strong enough lower bound method must somehow avoid applying to polynomial maps of high univariate degree. It therefore seems that the notion of univariate degree, and more generally a qualitative understanding of the curves contained in the image of a polynomial map, could be an important data point to consider when studying questions about arithmetic circuits.[9]

There are also connections with polynomial identity testing. For example, a line of work by Oliveira, Sengupta, and others (e.g., [34, 12]) uses a regularity lemma for polynomials to obtain efficient identity testing algorithms for depth-4 circuits with bounded top and bottom fan-ins. We leave it for future work to determine whether our weak regularity lemma could suffice here.

1.3 Applications for generalized ranks of polynomials

Is there an algebraic explanation for non-surjectivity of (multivariate) polynomials? An important class of non-surjective polynomials are those of the form $P = F(Q)$ with $1 < \deg(Q) < \deg(P)$ – sometimes called *uni-multivariate decomposable* [16] – for some non-surjective univariate F (for example, $P = (Q(x_1, \ldots, x_n))^t$ with $\gcd(t, |\mathbb{F}| - 1) > 1$). For such a polynomial of degree d, it must be that $\deg(Q) \le d/2$ (as $\deg(F) \ge 2$). We show that this is always the case, in the sense that every non-surjective polynomial can be written in terms of a rather small number of polynomials all of degree at most $d/2$ (despite most polynomials not being uni-multivariate decomposable [39]). In fact, we also show a significant generalization of this to polynomial maps (note that the image of a polynomial map holds information about the dependencies between its various polynomials).

[8] That is, each $Q_{i,j}$ is homogeneous, and all products $\prod_j Q_{i,j}$ are of the same degree.

[9] Univariate degree should not be confused with the concept of an elusive function of Raz [35], which is a polynomial map $\mathbf{P}$ satisfying $\mathrm{Im}\,\mathbf{P} \not\subseteq \mathrm{Im}\,\mathbf{\Gamma}$, for every polynomial map $\mathbf{\Gamma}$ with given parameters. Indeed, univariate degree is about containment in the other direction: a lower bound on $\mathrm{udeg}(\mathbf{P})$ means that $\mathrm{Im}\,\mathbf{U} \not\subseteq \mathrm{Im}\,\mathbf{P}$, for every polynomial curve $\mathbf{U}$ of a given degree.

Following Green and Tao ([15], Definition 1.6), let $\operatorname{rank}_t(P)$, for a polynomial P and $t \geq 0$, denote the least number of polynomials of degree at most t such that P can be expressed as a function of only them. Note that this extends naturally to polynomial maps $\mathbf{P}$. One easy-to-state application of the weak regularity lemma is that any polynomial that is not surjective, and more generally any polynomial map $\mathbf{P}\colon \mathbb{F}^n \to \mathbb{F}^m$ whose image $\operatorname{Im}\mathbf{P}$ ($\subseteq \mathbb{F}^m$) does not contain a line, must have low $\operatorname{rank}_t(\mathbf{P})$, for some t bounded away from $\deg(\mathbf{P})$.

▶ **Theorem** (see Theorem 21)**.** Let $\mathbf{P}\colon \mathbb{F}^n \to \mathbb{F}^m$ be a polynomial map with $\deg(\mathbf{P}) \leq d$, over any finite field $\mathbb{F}$ with $d < \operatorname{char}(\mathbb{F})$. If $\operatorname{Im}\mathbf{P}$ does not contain a line then

$$\operatorname{rank}_{d/2}(\mathbf{P}) \leq (2m)^{2^{d(1+o(1))}}.$$

▶ Remark 5. Just for the special case of $\mathbb{F}_3$, there is a vast literature on subsets of $\mathbb{F}_3^m$ avoiding a line, also known as cap sets. The most basic example of a cap set is $\{0,1\}^m$; this cap set leads to the following easy observation: for any (low-degree) polynomial map $\mathbf{P}\colon \mathbb{F}_3^n \to \mathbb{F}_3^m$, in any number of variables n, its "square" $\mathbf{P}^2 := (P_1^2, \ldots, P_m^2)\colon \mathbb{F}_3^n \to \mathbb{F}_3^m$ is a (low-degree) polynomial map with a line-free image. More generally for $\mathbb{F}_p$, with any prime $p \geq 3$, consider for example the analogue of Behrend's construction (see, e.g., Lemma 17 in [10]); it corresponds to the following observation: if $\operatorname{Im}(P_i) \subseteq \{0, 1, \ldots, (p-1)/2\}$ for every i, and the real number $\sum_{i=1}^m P_i^2$ is constant, then the image of $\mathbf{P} = (P_1, \ldots, P_m)$ contains no three points on a line, and in particular is line-free. See, e.g., [8] and the references within for more constructions of cap sets.

One can significantly extend this result, by considering the univariate degree udeg, which we recall is the smallest degree of a curve contained in the image of $\mathbf{P}$. We note that $\operatorname{udeg}(\mathbf{P}) = 1$ if and only if $\operatorname{Im}\mathbf{P}$ contains a line (and for $m = 1$ this means surjectivity).

It is not hard to show that, as n tends to infinity, we cannot hope to bound $\operatorname{rank}_t(P)$ for $t < \deg(P)/\operatorname{udeg}(P)$ (see, e.g., Section 3.1). Karam proved (see Theorem 1.3[10] in [21]) that we indeed can for $t = \deg(P)/\operatorname{udeg}(P)$.

▶ **Theorem 6** ([21])**.** *Let $P\colon \mathbb{F}^n \to \mathbb{F}$ be a polynomial with $\deg(P) \leq d$, over any prime finite field $\mathbb{F} = \mathbb{F}_p$ with $d < p$. For $t = \deg(P)/\operatorname{udeg}(P)$, we have $\operatorname{rank}_t(P) \leq \gamma(p,d)$, for some function γ.*

The proof uses a variant of the usual regularity lemma for polynomials, and thus yields an extremely weak bound $\gamma(p,d)$ on $\operatorname{rank}_t(P)$.

We show that the weak regularity lemma is sufficiently powerful to bound rank_t for any polynomial, and in fact also for any polynomial map and over non-prime fields too, yielding a correspondingly strong bound.

▶ **Theorem** (see Theorem 24)**.** Let $\mathbf{P}\colon \mathbb{F}^n \to \mathbb{F}^m$ be a polynomial map with $\deg(\mathbf{P}) \leq d$, over any finite field $\mathbb{F}$ with $d < \operatorname{char}(\mathbb{F})$. For $t = \deg(\mathbf{P})/\operatorname{udeg}(\mathbf{P})$,

$$\operatorname{rank}_t(\mathbf{P}) \leq (2m)^{2^{d(1+o(1))}}.$$

Roughly speaking, our proof follows by showing that for any weak ϵ-regular decomposition $\mathbf{P} = \mathbf{F}(\mathbf{X})$, for ϵ bounded away from 1, we have $\operatorname{udeg}(\mathbf{P}) \leq \deg_1(\mathbf{F})$, the degree of $\mathbf{F}$ in its first variable (so $\operatorname{Im}\mathbf{U} \subseteq \operatorname{Im}\mathbf{P}$ for some non-constant curve $\mathbf{U}$ with $\deg(\mathbf{U}) = \deg_1(\mathbf{F})$). Using the guarantees of the homogeneity of $\mathbf{X}$ and of the maximality of $\deg(X_1)$ ($= \deg(\mathbf{X})$), and removing "redundant" monomials from $\mathbf{F}$, we will deduce $\deg_1(\mathbf{F}) \leq \deg(\mathbf{P})/\deg(\mathbf{X})$, that is, $\deg(\mathbf{X}) \leq \deg(\mathbf{P})/\operatorname{udeg}(\mathbf{P}) = t$, which would allow us to bound rank_t.

---

[10] More precisely, Karam's result is stated for the closely related notion of rk_t; see Section 3.

1.4 The proof

Towards the proof of the weak regularity lemma, we obtain results on linear spaces of polynomials and forms (homogeneous polynomials) that may be of independent interest. In particular, we analyze subspaces of high-rank polynomials (where the rank is allowed to grow with the number of spanning polynomials, i.e., the dimension of the space). More particularly, we analyze *punctured* spaces of high-rank polynomials. By a punctured space we mean the complement of a linear subspace: $V \setminus U$ where V is a linear space (of polynomials) and U is a strict subspace.[11] For comparison, the classical regularity lemmas for polynomials work only with the "trivial" punctured space $V \setminus \{0\}$; or more concretely, *all* nonzero linear combinations of the parts $\mathbf{X}$ are required to have high rank. Obtaining a high-rank punctured space $V \setminus U$, rather than a high-rank *trivial* punctured space $V \setminus \{0\}$, turns out to be achievable with far better bounds. We call this result a "rank-regularity lemma", and note that its strong bounds eventually translate to strong bounds for our weak regularity lemma. Using it, the rest of our proof has, very roughly, the following logical structure (see Section 2 for all definitions):

$$\begin{aligned}\mathbf{X}\ t\text{-rank-regular} &\Leftrightarrow \operatorname{rk}(X_1 + \operatorname{Sp}\mathbf{X}') > t|\mathbf{X}| \\ &\Rightarrow |\operatorname{bias}(X_1 + \operatorname{Sp}\mathbf{X}')| < \epsilon q^{-|\mathbf{X}|} \\ &\Rightarrow \mathbf{X} \text{ weak } \epsilon\text{-regular.}\end{aligned}$$

The proof moreover guarantees the additional required conditions, e.g., that X_1 has maximal degree. To translate from high rank to low bias, our proof relies on recent results on the relationship between structure and randomness for polynomials (see e.g. [20, 31, 5, 32]). These results are proved by going through the corresponding multilinear forms – which requires the characteristic of the field to be larger than the degree of the parts – by relating analytic rank and partition rank (defined by Gowers and Wolf [13] and by Naslund [33], respectively).

1.5 Organization

In Section 2 we state and prove our weak regularity lemma. The proof has several components. In Section 2.1 we prove a rank-regularity lemma, which finds a small decomposition where every linear combination outside of a strict subspace has high rank. In Section 2.2 we show that a similar condition in terms of bias (rather than rank) suffices to get weak regularity, using Fourier-analytic techniques and known structure-vs-randomness results. Finally, in Section 2.3, we combine the aforementioned pieces to prove the weak regularity lemma. In Section 3 we obtain the applications of the weak regularity lemma outlined above.

1.6 Preliminaries

An important aspect of regularity lemmas for polynomials is that they are independent of the number of variables n (similarly to how, for example, the graph regularity lemma is independent of the number of vertices of the graph). As such, it will be convenient to use the notation $\operatorname{Poly}(\mathbb{F})$ for the set of all polynomials over $\mathbb{F}$ (meaning with coefficients in $\mathbb{F}$) in any (finite) number of variables $x_1, \ldots, x_n$ ($n \in \mathbb{N}$). While we work with (formal) polynomials in this paper, note that any n-variate polynomial over $\mathbb{F}$ determines a function $\mathbb{F}^n \to \mathbb{F}$

[11] For example, a punctured space of codimension 1 ($V \setminus U$ with $\operatorname{codim}_V(U) = 1$) is of the form $\{cw + u \mid c \neq 0, u \in U\}$ for some $w \in V$.

(uniquely so if every variable has degree smaller than $|\mathbb{F}|$). We write $\mathrm{Poly}_d(\mathbb{F})$ for the set of polynomials of (total) degree at most d. We henceforth omit $\mathbb{F}$ from our notation when it is clear from context. Naturally, $\mathrm{Poly}^m = \{(P_1,\ldots,P_m) \mid \forall i\colon P_i \in \mathrm{Poly}\}$ stands for the set of m-tuples of polynomials, or equivalently, polynomial maps into $\mathbb{F}^m$; and analogously for Poly_d^m. Similarly, we denote by Form the set of all homogeneous polynomials (or forms for short) of *positive* degree, in any (finite) number of variables; we denote by Form_d the set of forms of degree d; and by Form_d^m the set of m-tuples of forms of degree d. For polynomials $X_1,\ldots,X_r \in \mathrm{Poly}(\mathbb{F})$, we denote by $\mathbb{F}[X_1,\ldots,X_r]$ the algebra they generate, that is, $\{F(X_1,\ldots,X_r) \mid r\text{-variate } F \in \mathrm{Poly}\}$. Of course, if $x_1,\ldots,x_n$ are independent variables then $\mathbb{F}[x_1,\ldots,x_n]$ is the usual ring of n-variate polynomials in $x_1,\ldots,x_n$. We denote by $\deg_i(P)$ the degree of P in the variable x_i (that is, its degree when viewed as a univariate polynomial in x_i).

We will frequently use boldface letters to emphasize that an object is a tuple; so for example, we will use $\mathbf{P}$, instead of P, when dealing with a tuple of polynomials (or a polynomial map). We will use the convention that for a tuple $\mathbf{X} = (X_1,\ldots,X_k)$ and a set A, the notation $\mathbf{X} \subseteq A$ is a shorthand for $X_i \in A$ for all i. For $\mathbf{P} \subseteq \mathrm{Poly}$ we denote $\deg(\mathbf{P}) = \max_i \deg(P_i)$. More generally, for a set $V \subseteq \mathrm{Poly}$ we denote $\deg(V) = \max_{P\in V} \deg(P)$.

Probabilities, unless otherwise specified, are to be understood with respect to a uniformly random input; that is, for $\mathbf{X}\colon \mathbb{F}^n \to \mathbb{F}^k$, $\Pr[\mathbf{X} = y]$ is a shorthand for $\Pr_{x\in\mathbb{F}^n}[\mathbf{X}(x) = y]$ with x chosen uniformly at random from $\mathbb{F}^n$. The image (or range) of a polynomial map $\mathbf{P}\colon \mathbb{F}^n \to \mathbb{F}^m$ is denoted $\mathrm{Im}\,\mathbf{P} = \{\mathbf{P}(x) | x \in \mathbb{F}^n\}$ ($\subseteq \mathbb{F}^m$). For a subset $S \subseteq \mathbb{F}^n$, we write $\mathbf{P}(S) = \{\mathbf{P}(x) | x \in S\}$.

Note that $\mathrm{Poly}_d(\mathbb{F})$ is a linear space over $\mathbb{F}$. We write $U \le V$ when U is a subspace of V. We denote the linear span of $\mathbf{X} \in \mathrm{Poly}_d^k(\mathbb{F})$ by $\mathrm{Sp}\,\mathbf{X} = \mathrm{Sp}\{X_1,\ldots,X_k\} = \{a\cdot\mathbf{X} | a \in \mathbb{F}^k\}$, which is of course a subspace of $\mathrm{Poly}_d(\mathbb{F})$.

For $\mathbf{P} \in \mathrm{Poly}^m(\mathbb{F})$, we refer to an expression of the form $\mathbf{P} = \mathbf{F}(\mathbf{X})$ (or $\mathbf{P} = \mathbf{F}\circ\mathbf{X}$), with $\mathbf{F} \in \mathrm{Poly}^m(\mathbb{F})$ and $\mathbf{X} \in \mathrm{Poly}^k(\mathbb{F})$, as a *decomposition* of $\mathbf{P}$ (so that for every component P_i of $\mathbf{P}$, we have the decomposition $P_i = F_i(\mathbf{X})$). Note that this is a composition of (formal) polynomials. When the map $\mathbf{F}$ is immaterial, we denote the decomposition by $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$.

As mentioned above, all fields considered in this paper are finite; we denote by $\mathbb{F}_q$ the finite field of size q (so q is a power of the characteristic $\mathrm{char}(\mathbb{F})$).

2 Weak Regularity

For convenience, we repeat here the definitions of weak regularity and the weak regularity lemma. Recall that for a k-tuple $\mathbf{y} = (y_1,\ldots,y_k)$, we denote by $\mathbf{y}'$ the $(k-1)$-tuple $\mathbf{y}' = (y_2,\ldots,y_k)$.

▶ **Definition 7** (weak ϵ-regularity). *Let $\mathbb{F} = \mathbb{F}_q$. We say that $\mathbf{X} \in \mathrm{Form}^k(\mathbb{F})$ is* weak ϵ-regular *if $\deg(X_1) = \deg(\mathbf{X})$, and for every $y \in \mathbb{F}^k$,*

$$\left|\Pr[\mathbf{X} = \mathbf{y}] - q^{-1}\Pr[\mathbf{X}' = \mathbf{y}']\right| \le \epsilon q^{-k}.$$

For $\mathbf{P} \subseteq \mathrm{Poly}(\mathbb{F})$, we say that $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$ is a weak ϵ-regular decomposition *if $\mathbf{X} \subseteq \mathrm{Form}$ is weak ϵ-regular and $\mathbf{P} \not\subseteq \mathbb{F}[\mathbf{X}']$.*[12]

[12] For completeness, the empty decomposition ($k = 0$) is trivially weak ϵ-regular for any ϵ. Thus, constant polynomial maps $\mathbf{P}$ trivially have a weak ϵ-regular decomposition.

Roughly speaking, a set of forms $\mathbf{X} \subseteq \text{Form}$ is weak ϵ-regular if at least one of its members, of largest degree, is approximately "unbiased" and "independent" of the rest; moreover, a polynomial map $\mathbf{P}$ has a weak ϵ-regular decomposition if it can be expressed in terms of a weak ϵ-regular $\mathbf{X}$, and depends on X_1. Intuitively, P can be thought of as a polynomial in only k variables, and although these may not be independent, free variables, at least one of them approximately is.

Let us also note that the probabilistic condition in Definition 7 means that $\Pr[\mathbf{X} = \mathbf{y}]/q^{-k} = \alpha \pm \epsilon$ for every $\mathbf{y} \in \mathbb{F}^k$ with $\Pr[\mathbf{X}' = \mathbf{y}']/q^{-(k-1)} = \alpha$ (recall $\mathbf{X}'$ is a $(k-1)$-tuple). Intuitively, this means $\mathbf{X}$ is distributed like $\mathbf{X}'$ up to an additive error of ϵ. We also mention that Definition 7 implies $\deg(\mathbf{X}) \le \deg(\mathbf{P})$ (see Corollary 29).

▶ **Theorem 8** (weak regularity lemma). *Let $c \ge 1$. Every $\mathbf{P} \in \text{Poly}_d^m(\mathbb{F}_q)$, with $d < \text{char}(\mathbb{F}_q)$, has a weak ϵ-regular decomposition of size*

$$k \le (2mc)^{2^{d(1+o(1))}},$$

where we write $\epsilon q^{-k} = q^{-ck}$ for the error term.

Note that the decomposition size is polynomial in m. We remark that Theorem 8 is even stronger than the statement in the Introduction (Theorem 1), since here ϵ is a function of the decomposition size k (namely, $\epsilon = q^{-(c-1)k}$).

2.1 Rank-regularity

The proof of the weak regularity lemma has several components. At the core of the proof is what we call a *rank-regularity lemma*[13] (see Definition 10). Roughly, the requirement for the parts $\mathbf{X}$ in a rank-regular decomposition is that they are homogeneous (not necessarily all of the same degree), and that every linear combination of $\mathbf{X}$ outside of a strict subspace has high rank (see Definition 9). We will later convert a rank-regular decomposition into a weak regular decomposition. We begin by setting up the basic definitions we need, and state the results that form the steps of our overall strategy.

An important definition in the study of polynomials is that of *rank*, sometimes also called strength [1] or Schmidt rank [36]. Recall that a polynomial is *reducible* if it can be factorized as the product of two non-constant polynomials (all polynomials over the same field).

▶ **Definition 9** (rank). *The* rank *of a form $P \in \text{Form}_d(\mathbb{F})$ is:*

- *For $d \ge 2$: the least number of reducible forms that sum up to P:*

$$\text{rk}(P) = \min\left\{r \,\middle|\, \exists R_1, \dots, R_r \in \text{Form}_d(\mathbb{F}) \text{ } \textit{reducible}\text{: } P = R_1 + \dots + R_r\right\}.$$

- *For $d = 1$, $\text{rk}(P) = \infty$.*

More generally, the rank of a polynomial $P \in \text{Poly}(\mathbb{F})$ is the rank of its highest-degree homogeneous part,[14] *or 0 if P is a constant.*

Note that for a homogeneous $P \in \text{Form}_d$ with $\text{rk}(P) \le r$ $(< \infty)$, we can write $P = Q(Y_1, \dots, Y_k)$ with $k \le 2r$, for some homogeneous $Q \in \text{Form}_2$ and $Y_i \in \text{Form}$ with $\deg(Y_i) < d$. This uses the fact that a reducible form R factorizes as the product of two forms (of lower degree); indeed, the product of the lowest-degree homogeneous parts of the factors is a homogeneous part of R.

[13] Or "rankularity" lemma for short.

[14] If $\deg(P) = d$ and we write $P = \sum_{i=0}^{d} P_i$ with P_i a form of degree i, then $\text{rk}(P) = \text{rk}(P_d)$.

The core of the proof of our weak regularity lemma is a weak regularity lemma for rk. For $\mathbf{P} \subseteq \text{Poly}$, call a decomposition $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$ *minimal* if $\mathbf{P} \nsubseteq \mathbb{F}[\mathbf{X}^*]$ for every $\mathbf{X}^* \subseteq \text{Form}$ with $\text{Sp}\,\mathbf{X}^* \subsetneq \text{Sp}\,\mathbf{X}$. Put differently, no strict subspace $V' \lneq \text{Sp}\,\mathbf{X}$ generates $\mathbf{P}$ (where V' has a basis of forms $\mathbf{X}^*$).

▶ **Definition 10** (rank-regularity)**.** *For any $t \geq 0$:*

- $\mathbf{X} \in \text{Form}^r$ *is* t-rank-regular *if* $\text{rk}(X) \geq tr$ *for all* $X \in \text{Sp}\,\mathbf{X}$ *outside of a strict subspace.*
- $\mathbf{P} \in \text{Poly}^m(\mathbb{F})$ *has a t-rank-regular decomposition of size r if $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$ is a minimal decomposition for some t-rank-regular $\mathbf{X} \in \text{Form}^r(\mathbb{F})$.*

Note that "all $X \in \text{Sp}\,\mathbf{X}$ outside of a strict subspace" refers, explicitly, to all $X \in (\text{Sp}\,\mathbf{X}) \setminus U$ for some strict subspace $U \lneq \text{Sp}\,\mathbf{X}$.[15] In other words, $\mathbf{X}$ is t-rank-regular if $\text{Sp}\{X \in \text{Sp}\,\mathbf{X} | \,\text{rk}(X) < tr\} \neq \text{Sp}\,\mathbf{X}$. This aspect[16] of the definition is one reason that we are able to obtain strong bounds for this regularity lemma. We also mention that Definition 10 implies $\deg(\mathbf{X}) \leq \deg(\mathbf{P})$ (see Corollary 29).

▶ **Theorem 11** (rank-regularity lemma)**.** *Every $\mathbf{P} \in \text{Poly}_d^m$ has a t-rank-regular decomposition of size at most $((2t+1)dm)^{2^d}$.*

2.1.1 Linear spaces of polynomials

Here we prove several useful properties of linear spaces of polynomials, which will be used in the proof of the rank-regularity lemma and the weak regularity lemma. We begin with some useful notation.

Henceforth, for a linear space $V \leq \text{Poly}$ and $R > 0$, we denote the span of its low-rank polynomials by

$$U_R(V) = \text{Sp}\{X \in V \mid \text{rk}(X) < R\}.$$

Note that $\mathbf{X} \in \text{Form}^k$ is t-rank-regular if and only if $U_R(V) \neq V$ for $V = \text{Sp}\,\mathbf{X}$ and $R = tk$.

We say that a finite-dimensional linear space $V \leq \text{Poly}$ is *homogeneous* if V is spanned by forms.[17] Equivalently, V is the sum – and thus the direct sum – of its homogeneous parts: $V = \sum_{i \geq 0} V_i$ where $V_i := V \cap \text{Form}_i$. We abbreviate $V^{\uparrow} = V_{\deg(V)}$, that is,

$$V^{\uparrow} = V \cap \text{Form}_{\deg(V)},$$

where we recall the notation $\deg(V) = \max_{X \in V} \deg(X)$.

For the proof of the rank-regularity lemma (Theorem 11) we will need the following lemma on rank-regularity of forms of top degree.

▶ **Lemma 12.** *Let $\mathbf{Y} \in \text{Form}^m$, and let $\mathbf{Y}^* \in \text{Form}^m$ have the same underlying set as $\mathbf{Y}^{\uparrow}$. For any t, if $\mathbf{Y}^*$ is t-rank-regular then $\mathbf{Y}$ is t-rank-regular.*

Proof. Put $V = \text{Sp}(\mathbf{Y})$, $V^* = \text{Sp}(\mathbf{Y}^*)$, and note $V^* = \text{Sp}(\mathbf{Y}^{\uparrow}) = (\text{Sp}\,\mathbf{Y})^{\uparrow} = V^{\uparrow}$. Let $R = tm$, and suppose $U_R(V) = V$. We need to show $U_R(V^*) = V^*$ (with the same R, as the tuples $\mathbf{Y}$ and $\mathbf{Y}^*$ have the same number m of polynomials), or equivalently, $U_R(V^{\uparrow}) = V^{\uparrow}$.

---

[15] For example, U might be $\text{Sp}(\mathbf{X}')$, provided $\mathbf{X}$ is linearly independent.

[16] Equivalently, this is about high-rank punctured spaces: if $V \leq \text{Poly}$ has a punctured space W $(= V \setminus U$, $U \lneq V)$ with $\text{rk}(P) \geq t \dim(V)$ for every $P \in W$, then any basis of forms for V is t-rank-regular (and vice versa).

[17] By convention, the zero space $\{0\}$ is homogeneous.

Let $P \in V^{\uparrow}$, that is, $P \in V$ is a form with $\deg(P) = \deg(V)$. Since $U_R(V) = V$, P is a linear combination of $P_i \in V$ with $\mathrm{rk}(P_i) < R$. Then P is a linear combination of their homogeneous parts of degree $\deg(V)$, which are:

- also in V (as V is homogeneous), and thus in $V^{\uparrow}$, and
- of rank $< R$ (by definition of rank for non-homogeneous polynomials).

Thus, $P \in U_R(V^{\uparrow})$. We deduce that $V^{\uparrow} \subseteq U_R(V^{\uparrow})$, and since the containment in the other direction trivially holds, we are done. ◀

For the rest of the proof of the weak regularity lemma (Theorem 8), we will need Lemma 14 and Lemma 15 below.

We begin with the following closure property of the subspaces $U_R(V)$.

▷ Claim 13. Let $V \leq \mathrm{Poly}$, $R > 0$. Every $X \in V$ with $\deg(X) < \deg(U_R(V))$ satisfies $X \in U_R(V)$.

Proof. Abbreviate $U = U_R(V)$. Let $X_0 \in U$ with $\mathrm{rk}(X_0) < R$ of maximal degree, $\deg(X_0) = \deg(U)$. We claim that $X_0 + X \in U$. Indeed, since $\deg(X) < \deg(U) = \deg(X_0)$, we have $\mathrm{rk}(X_0 + X) = \mathrm{rk}((X_0 + X)^{\uparrow}) = \mathrm{rk}(X_0^{\uparrow}) = \mathrm{rk}(X_0) < R$. Thus, as desired, $X \in U$, being the difference $X = (X_0 + X) - X_0$ of two members of the linear space U. ◁

▶ **Lemma 14.** *Let $V \leq \mathrm{Poly}$, $R > 0$. If V is homogeneous then so is $U_R(V)$.*

Proof. Put $U = U_R(V)$. For $X \in \mathrm{Poly}$, denote by $X^{(i)} \in \mathrm{Form}_i$ the degree-i part of X (so that $X = \sum_i X^{(i)}$). First, we prove another closure property of U (in a homogeneous V): for every $X \in V$, if $\mathrm{rk}(X) < R$ then $X^{(i)} \in U$ for every i. Let us prove this. If $i = \deg(X)$ (that is, $X^{(i)} = X^{\uparrow}$) then $\mathrm{rk}(X^{\uparrow}) = \mathrm{rk}(X) < R$; since V is homogeneous, $X^{\uparrow} \in V$, so $X^{\uparrow} \in U$, as needed. On the other hand, if $i < \deg(X)$ then $\deg(X^{(i)}) < \deg(X) \leq \deg(U)$, since $X \in U$, and since $X^{(i)} \in V$ by the homogeneity of V, we can apply Claim 13 on $X^{(i)}$ to deduce $X^{(i)} \in U$. This proves our claim.

Let $\mathbf{X} = \{X \in V \,|\, \mathrm{rk}(X) < R\}$, so that $U = \mathrm{Sp}(\mathbf{X})$ by definition. By the claim above, $\mathbf{X}^* := \{X^{(i)} \,|\, X \in \mathbf{X}, i \leq \deg(X)\}$ satisfies $\mathbf{X}^* \subseteq U$. Since $U = \mathrm{Sp}(\mathbf{X}) \subseteq \mathrm{Sp}(\mathbf{X}^*) \subseteq U$, we deduce that U has a spanning set of forms, which completes the proof. ◀

▶ **Lemma 15.** *Let $V \leq \mathrm{Poly}$, $R > 0$. If V is homogeneous, then $U_R(V) = V$ if and only if $U_R(V^{\uparrow}) = V^{\uparrow}$.*

Proof. The forward direction is equivalent to Lemma 12: if $U_R(V) = V$, then $\deg(U_R(V)) = \deg(V)$, which implies $V^{\uparrow} \subseteq U_R(V)^{\uparrow} \subseteq U_R(V^{\uparrow})$, as needed.

For the other direction, suppose $U_R(V^{\uparrow}) = V^{\uparrow}$. Then $V^{\uparrow} \subseteq U_R(V)$, as $U_R(V^{\uparrow}) \subseteq U_R(V)$, and in particular, $\deg(U_R(V)) = \deg(V)$. Let us denote $V' = \bigoplus_{i<\deg(V)} V_i$. Claim 13 says $V' \subseteq U_R(V)$. Then $V = V^{\uparrow} + V' \subseteq U_R(V) + U_R(V) = U_R(V) \subseteq V$, that is, $U_R(V) = V$, as needed. ◀

2.1.2 Proof of the rank-regularity lemma

Theorem 11 will be proved using a repeated application of the following decomposition lemma. It shows that if a tuple of forms $\mathbf{X}$, all of the same degree d, is not t-rank-regular, then we can decompose it into not-too-many forms of varying degrees, all strictly smaller than d. Crucially, the lemma considers not just individual low-rank forms, but the whole subspace they span.

▶ **Lemma 16.** *If* $\mathbf{X} \in \mathrm{Form}_d^r(\mathbb{F})$ *is not* t*-rank-regular, then* $\mathbf{X} \subseteq \mathbb{F}[\mathbf{Y}]$ *for some* $\mathbf{Y} \in \mathrm{Form}^R$ *with* $\deg(\mathbf{Y}) < d$ *and* $R \leq 2tr^2$.

Proof. Assume $d \geq 2$, as otherwise the hypothesis is vacuous. Put $V = \mathrm{Sp}\,\mathbf{X}$ ($\leq \mathrm{Form}_d$) and $k = \dim V$ ($\leq r$). Let

$$\mathbf{X}^- = \{X \in V \mid \mathrm{rk}(X) < tr\}.$$

Suppose $\mathbf{X}$ is not t-rank-regular, so that $\mathrm{Sp}(\mathbf{X}^-) = V$. Thus, $\mathbf{X}^-$ contains a basis $\mathbf{B} = (B_1, \ldots, B_k) \in \mathrm{Form}_d^k$ for V. This has the following implications:

- Since $\mathbf{X} \subseteq \mathrm{Sp}\,\mathbf{B}$ $(= V)$, there is a linear map $\mathbf{L}\colon \mathbb{F}^k \to \mathbb{F}^r$ such that $\mathbf{X} = \mathbf{L}(\mathbf{B})$.
- Since $\mathbf{B} \subseteq \mathbf{X}^- \cap \mathrm{Form}$, and by the definition of rk (recall Definition 9 and the subsequent remark), we have $\mathbf{B} = \mathbf{Q}(\mathbf{Y})$ for some quadratic $\mathbf{Q} \in \mathrm{Form}_2^k$ and some $\mathbf{Y} \in \mathrm{Form}^R$ with $\deg(\mathbf{Y}) < d$, where $R \leq \sum_{i\in[k]} 2\,\mathrm{rk}(B_i) < 2tr \cdot k$.

Combining the above, we deduce that $\mathbf{X} = \mathbf{L}(\mathbf{Q}(\mathbf{Y})) = \mathbf{F}(\mathbf{Y})$, where $\mathbf{F} := \mathbf{L}(\mathbf{Q}) \in \mathrm{Form}_2^r$ and $\mathbf{Y} \in \mathrm{Form}^R$ with $\deg(\mathbf{Y}) < d$ and $R < 2tr \cdot k \leq 2tr^2$, which completes the proof. ◀

We prove the rank-regularity lemma by an iterative construction with at most d steps, maintaining at all times a tuple of forms of various degrees. At each step, we move to a minimal decomposition, and if it is not a t-rank-regular decomposition then, we show, its subset of forms of maximal degree is also not t-rank-regular. This last deduction requires care – as the rank requirement in Definition 10 ($\mathrm{rk}(X) \geq tr$) depends on the number of polynomials in the decomposition – and follows using Lemma 12. Having done that, we decompose the maximum-degree forms in terms of lower degree forms using Lemma 16, which strictly decreases the maximum degree of the forms in the decomposition, as needed.

Proof of Theorem 11. Assume $\mathbf{P}$ is non-constant, as otherwise there is nothing to prove. Let $\mathbf{X^0} \in \mathrm{Form}^{r_0}$ be obtained from $\mathbf{P} \in \mathrm{Poly}^m$ by replacing each component of $\mathbf{P}$ with its nonzero homogeneous parts of degree $1, 2, \ldots, d$. Note that $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X^0}]$ and $r_0 \leq dm$.

We claim that for some $s \geq 0$, there are s homogeneous polynomial maps $\mathbf{X^1}, \ldots, \mathbf{X^s}$, with $\mathbf{X^i} \in \mathrm{Form}^{r_i}$ for $i = 1, \ldots, s$, such that $\mathbf{P} \in \mathbb{F}[\mathbf{X^s}]$ is a t-rank-regular decomposition, and for every $i \geq 0$:

- $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X^i}]$,
- $\deg(\mathbf{X^i}) \leq d - i$, and
- $r_i \leq (2t+1)^{2^i-1}(dm)^{2^i}$.

The process terminates (that is, s is well defined); indeed, any $\mathbf{X} \subseteq \mathrm{Form}$ with $\deg(\mathbf{X}) = 1$ is t-rank-regular for any t, since $\mathrm{rk}(X) = \infty$ for every $X \in \mathrm{Sp}(\mathbf{X}) \setminus \{0\}$, so a minimal decomposition $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$ with $\deg(\mathbf{X}) = 1$ is a t-rank-regular decomposition for any t. Proving the above claim would complete the proof, since $1 \leq \deg(\mathbf{X^s}) \leq d - s$ implies $s < d$, and thus $\mathbf{X^s} \in \mathrm{Form}^{r_s}$ with $r_s \leq ((2t+1)dm)^{2^d}$, as desired.

For each i starting with $i = 0$, given $\mathbf{X^i}$ we proceed as follows:

- Consider a homogeneous subspace $W \leq \mathrm{Sp}(\mathbf{X^i})$ of minimal dimension satisfying $\mathbf{P} \subseteq \mathbb{F}[W]$, and let $\mathbf{Y} \in \mathrm{Form}^{r'}$ be a basis of W. Then $\mathbf{P} \subseteq \mathbb{F}[\mathbf{Y}]$ is a minimal decomposition, and moreover, $\deg(\mathbf{Y}) \leq \deg(W) \leq \deg(\mathbf{X^i})$, and $r' = \dim W \leq \dim \mathrm{Sp}(\mathbf{X^i}) \leq r_i$.
- Now, if $\mathbf{P} \subseteq \mathbb{F}[\mathbf{Y}]$ is t-rank-regular then set $s = i$, $\mathbf{X^s} = \mathbf{Y}$ and $r_s = r'$, finishing the construction. Otherwise, $\mathbf{Y}$ must not be t-rank-regular. We will use $\mathbf{Y}$ to construct $\mathbf{X^{i+1}}$ with $\deg(\mathbf{X^{i+1}}) < \deg(\mathbf{X^i})$. In order to apply Lemma 16, we need a tuple of forms satisfying two conditions: they are all of the same degree, and the tuple is not t-rank-regular.

Put $\ell = \deg(\mathbf{Y})$, and obtain $\mathbf{Y}^* \in \mathrm{Form}^{r'}$ from $\mathbf{Y}$ by replacing, in each component of $\mathbf{Y}$, a polynomial of degree smaller than ℓ by another polynomial from $\mathbf{Y}$ of degree ℓ.[18] By Lemma 12, $\mathbf{Y}^*$ is not t-rank-regular.

Apply Lemma 16 on $\mathbf{Y}^*$. Merge the resulting tuple of forms with the forms in $\mathbf{Y}$ of degree smaller than ℓ to obtain $\mathbf{X^{i+1}} \in \mathrm{Form}^{r_{i+1}}$ with $\mathbf{Y} \subseteq \mathbb{F}[\mathbf{X^{i+1}}]$, $\deg(\mathbf{X^{i+1}}) < \deg(\mathbf{X^i})$, and $r_{i+1} \le 2tr_i^2 + r_i \le (2t+1)r_i^2$.

To prove that this construction gives the claim above, we proceed by induction on $i \ge 0$. The induction base $i = 0$ follows trivially from the construction of $\mathbf{X^0}$, as $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X^0}]$ with $\deg(\mathbf{X^0}) \le d$, and $r_0 \le dm = (2t+1)^0(dm)^1$. For the induction step we have, by the construction of $\mathbf{X^{i+1}}$ and the induction hypothesis, that $\mathbf{P} \subseteq \mathbb{F}[\mathbf{Y}] \subseteq \mathbb{F}[\mathbf{X^{i+1}}]$, that $\deg(\mathbf{X^{i+1}}) \le \deg(\mathbf{X^i}) - 1 \le d - (i+1)$, and that $r_{i+1} \le (2t+1)r_i^2 \le (2t+1)((2t+1)^{2^i-1}(dm)^{2^i})^2 = (2t+1)^{2^{i+1}-1}(dm)^{2^{i+1}}$, as needed. This completes the proof. ◀

2.2 Bias and weak regularity

We will use the notion of bias of a polynomial, which is a standard analytic measure related to deviation from equidistribution. Henceforth, fix any nontrivial additive character[19] $\chi\colon \mathbb{F} \to \mathbb{C}$ (for concreteness, if $\mathbb{F} = \mathbb{F}_p$ is a prime field, take $\chi(y) = \omega^y$, where ω is the principal root of unity, $\omega = e^{2\pi i/p}$, and y on the right hand side is viewed as an element of $\{0, \dots, p-1\}$.) The *bias* of $P \in \mathrm{Poly}(\mathbb{F})$ is the complex number

$$\mathrm{bias}(P) = \mathop{\mathbb{E}}_{x \in \mathbb{F}^n} \chi(P(x)).$$

As we will see below, simple analysis relates bias and equidistribution as follows (see also e.g. [2], or Claim 33 in [4]):

$$\forall y \in \mathbb{F}\colon \left| \Pr[P = y] - \frac{1}{|\mathbb{F}|} \right| \le \max_{a \in \mathbb{F}^\times} |\,\mathrm{bias}(aP)|.$$

Towards proving that rank-regularity implies weak regularity, we will use the best bound known for the structure versus randomness problem for polynomials, which is almost linear.

▶ **Theorem 17** ([32], Corollary 8.5 and Remark 8.3)**.** *For any $P \in \mathrm{Poly}_d(\mathbb{F})$ with $\mathrm{char}(\mathbb{F}) > d$,*

$$\mathrm{rk}(P) \ge r \quad \Rightarrow \quad |\,\mathrm{bias}(P)| \le |\mathbb{F}|^{-c_d r/\log_2(r+1)}$$

and $c_d = \Omega_d(1)$, namely $c_d \ge 2^{-d^{1+o(1)}}$.

We remark that [32] even gives that the logarithmic term in Theorem 17 goes down with $|\mathbb{F}|$ (namely, denominator $\log_{|\mathbb{F}|}(r+1)+1$), but we do not need this. We also remark that one can replace the use of Theorem 17 in our proof by earlier bounds of Milićević [31] or Janzer [20], and still obtain reasonable, albeit somewhat worse bounds in Theorem 8.

Using bias, a Fourier-analytic argument will help us to deduce weak regularity.

▶ **Lemma 18.** *Let $U \lneq V \le \mathrm{Poly}(\mathbb{F}_q)$ with $\dim(V) = k$. Suppose $|\,\mathrm{bias}(X)| \le \epsilon q^{-k}$ for all $X \in V \setminus U$. Then any basis $\mathbf{X} \in \mathrm{Form}^k$ of V that contains a basis of U, with $X_1 \notin U$ and $\deg(X_1) = \deg(\mathbf{X})$, is weak ϵ-regular.*

[18] That is, $\mathbf{Y}^*$ is obtained from $\mathbf{Y}^\uparrow$ by arbitrarily duplicating components to achieve size r'.
[19] A nontrivial additive character of $\mathbb{F}$ is a homomorphism $\chi\colon \mathbb{F} \to \mathbb{C}^\times$ ($\chi(a+b) = \chi(a)\chi(b)$) with $\chi \not\equiv 1$. Equivalently, $\chi(y) = \omega^{\mathrm{Tr}(cy)}$, with $\omega = e^{2\pi i/p}$, $p = \mathrm{char}(\mathbb{F})$, $c \in \mathbb{F}^\times$, and Tr is the trace of $\mathbb{F}$ on $\mathbb{F}_p$.

That is, Lemma 18 requires of $\mathbf{X}$ that: $\mathbf{X}$ is a basis of V, $U \subseteq \mathrm{Sp}(\mathbf{X}')$, $X_1 \notin U$, and $\deg(X_1) = \deg(\mathbf{X})$.

Recall that $\mathbf{X}$ being weak ϵ-regular implies that for every $y \in \mathbb{F}^k$,

$$\big|\Pr[\mathbf{X} = \mathbf{y}] - |\mathbb{F}|^{-1}\Pr[\mathbf{X}' = \mathbf{y}']\big| \le \epsilon|\mathbb{F}|^{-k}.$$

To prove Lemma 18, we will bound the left hand side by $\max_{a_1\neq 0}|\operatorname{bias}(a\cdot\mathbf{X})|$.

It will be convenient to use a more geometric formulation of weak ϵ-regularity: for every line $\ell \subseteq \mathbb{F}^k$ in the direction of e_1 (so $\ell = \{te_1 + \mathbf{z} | t \in \mathbb{F}\}$ for some $\mathbf{z} \in \mathbb{F}^k$),[20]

$$\forall y \in \ell \colon |\Pr[\mathbf{X} = \mathbf{y}] - |\mathbb{F}|^{-1}\Pr[\mathbf{X} \in \ell]| \le \epsilon|\mathbb{F}|^{-k}. \tag{1}$$

Proof of Lemma 18. We will use the standard fact that the distribution $(\Pr[\mathbf{X} = y])_{y\in\mathbb{F}^k}$ can be expressed in terms of biases, by using simple Fourier analysis. Indeed, since[21]

$$\forall v \in \mathbb{F}^m\colon \qquad \mathop{\mathbb{E}}_{a\in\mathbb{F}^m}\chi(a\cdot v) = \delta[v=0] := \begin{cases}1 & v = 0\\ 0 & v\neq 0\end{cases}, \tag{2}$$

we have that for every $y \in \mathbb{F}^k$:

$$\Pr_x[\mathbf{X}(x) = y] = \mathop{\mathbb{E}}_x \delta[\mathbf{X}(x) - y = 0] = \mathop{\mathbb{E}}_x\mathop{\mathbb{E}}_{a\in\mathbb{F}^k}\chi(a\cdot(\mathbf{X}(x)-y)) = \mathop{\mathbb{E}}_{a\in\mathbb{F}^k}\operatorname{bias}(a\cdot(\mathbf{X}-y)). \tag{3}$$

Recall that $\mathbf{X}$ contains a basis for U, and $X_1 \notin U$. We claim that for every $a \in \mathbb{F}^k$ with $a_1 \neq 0$, we have that $a\cdot\mathbf{X} \notin U$. Indeed, write $a\cdot\mathbf{X} = a_1X_1 + X'$ with $X' \in \mathrm{Sp}(\mathbf{X}')$. Our assumption on $\mathbf{X}$ implies that $U \subseteq \mathrm{Sp}(\mathbf{X}')$, so if $a\cdot\mathbf{X} \in U$ then $a_1X_1 \in U - X' \subseteq \mathrm{Sp}(\mathbf{X}')$, a contradiction as $a_1 \neq 0$. Using our assumption on U, we deduce that

$$\forall a \in \mathbb{F}^k\colon \quad a_1 \neq 0 \quad\Rightarrow\quad |\operatorname{bias}(a\cdot\mathbf{X})| \le \epsilon|\mathbb{F}|^{-k}. \tag{4}$$

Put $v = e_1 \in \mathbb{F}^k$. Let $\ell \subseteq \mathbb{F}^k$ be a line in the direction of v, so that $\ell = \{tv + z | t \in \mathbb{F}\}$ for some $z \in \ell$. We claim that the following equality holds:

$$\Pr_x[\mathbf{X}(x) \in \ell] = \mathop{\mathbb{E}}_{\substack{a\in\mathbb{F}^k:\\ a\cdot v=0}}\operatorname{bias}(a\cdot(\mathbf{X}-z)). \tag{5}$$

Indeed,

$$\begin{aligned}\Pr_x[\mathbf{X}(x)\in\ell] &= \sum_{t\in\mathbb{F}}\Pr_x[\mathbf{X}(x) = tv+z] = \sum_{t\in\mathbb{F}}\mathop{\mathbb{E}}_{a\in\mathbb{F}^k}\mathop{\mathbb{E}}_x\chi(a\cdot(\mathbf{X}(x)-(tv+z)))\\ &= \mathop{\mathbb{E}}_{a\in\mathbb{F}^k}\Big(\sum_{t\in\mathbb{F}}\chi(t(a\cdot v))\Big)\mathop{\mathbb{E}}_x\chi(a\cdot(\mathbf{X}(x)-z))\\ &= \mathop{\mathbb{E}}_{a\in\mathbb{F}^k}\big(|\mathbb{F}|\delta[a\cdot v=0]\big)\operatorname{bias}(a\cdot(\mathbf{X}-z)) = \frac{1}{|\mathbb{F}|^{k-1}}\sum_{\substack{a\in\mathbb{F}^k:\\ a\cdot v=0}}\operatorname{bias}(a\cdot(\mathbf{X}-z))\end{aligned}$$

where (3) is used in the second equality, and (2) is used in the penultimate equality with $m = 1$. Combining the point equality (3) and the line equality (5), we deduce (for $y = z$)

$$\Pr_x[\mathbf{X}(x) = z] = |\mathbb{F}|^{-1}\Pr_x[\mathbf{X}(x)\in\ell] + |\mathbb{F}|^{-k}\sum_{\substack{a\in\mathbb{F}^k:\\ a\cdot v\neq 0}}\operatorname{bias}(a\cdot(\mathbf{X}-z)).$$

[20] Indeed, any $\mathbf{y} \in \ell$ is of the form $\mathbf{y} = (t, \mathbf{z}')$ for some $t \in \mathbb{F}$, so $\Pr[\mathbf{X} \in \ell] = \Pr[\mathbf{X}' = \mathbf{z}']$ and $\mathbf{z}' = \mathbf{y}'$.
[21] $\sum_{a\in\mathbb{F}^m}\chi(a\cdot v) = \sum_{a\in\mathbb{F}^m}\prod_{i=1}^m\chi(a_iv_i) = \prod_{i=1}^m\sum_{a_i\in\mathbb{F}}\chi(a_iv_i) = \prod_{i=1}^m \mathbf{1}(v_i=0)|\mathbb{F}| = \mathbf{1}(v=0)|\mathbb{F}|^m$.

Now, we bound the magnitude of the rightmost term as follows:

$$|\mathbb{F}|^{-k}\Big|\sum_{\substack{a\in\mathbb{F}^k:\\ a\cdot v\neq 0}} \mathrm{bias}(a\cdot(\mathbf{X}-z))\Big| \le \max_{\substack{a\in\mathbb{F}^k:\\ a\cdot v\neq 0}} |\,\mathrm{bias}(a\cdot(\mathbf{X}-z))| = \max_{\substack{a\in\mathbb{F}^k:\\ a\cdot v\neq 0}} |\,\mathrm{bias}(a\cdot\mathbf{X})|.$$

Therefore,

$$\Big|\Pr_x[\mathbf{X}(x)=y] - |\mathbb{F}|^{-1}\Pr_x[\mathbf{X}(x)\in\ell]\Big| \le \max_{\substack{a\in\mathbb{F}^k:\\ a\cdot v\neq 0}} |\,\mathrm{bias}(a\cdot\mathbf{X})| \le \epsilon|\mathbb{F}|^{-k},$$

using (4) in the last inequality. Thus, $\mathbf{X}$ is weak ϵ-regular, as desired. ◀

2.3 Putting everything together

The proof of the weak regularity lemma will rely on Theorem 11, Theorem 17, and Lemma 18. Careful work will be required to convert rank-regularity, which gives information only about the highest-degree part (recall Definition 9) of the polynomials in our linear space, to weak regularity, which takes into account polynomials of all degrees, while still retaining the homogeneity of the parts. Moreover, rank-regularity depends only on the linear space spanned by the parts, whereas weak regularity depends on the particular parts (and so does not allow the first part to be repeated, for example).

Proof of Theorem 8. Put $g(t,d,m) = ((2t+1)dm)^{2^d}$. We claim that the following holds for every polynomial $P \in \mathrm{Poly}_d(\mathbb{F})$ and every $1 \le k \le g(t,d,m)$:

$$|\,\mathrm{bias}(P)| \ge q^{-ck} \quad \Rightarrow \quad \mathrm{rk}(P) < tk \tag{6}$$

for some $t = 2^{d^{1+o(1)}}(c\log(2m))^{1+o(1)}$. Indeed, by Theorem 17, $\mathrm{rk}(P) \ge tk$ implies that $|\,\mathrm{bias}(P)| \le q^{-c_d tk/\log_2(tk+1)}$ with $c_d \ge 2^{-d^{1+o(1)}}$, that is,

$$\begin{aligned}(-\log_q|\,\mathrm{bias}(P)|)/k &\ge \frac{c_d t}{\log_2(tk+1)} \ge \frac{c_d t}{\log_2(tg(t,d,m)+1)}\\ &\ge \frac{c_d t}{(2^d+1)\log_2((2t+1)dm)} > c,\end{aligned}$$

that is, $|\,\mathrm{bias}(P)| < q^{-ck}$.

With t as above, apply Theorem 11 to find a t-rank-regular decomposition $\mathbf{Y} \subseteq \mathrm{Form}$ of $\mathbf{P}$ of size at most $g(t,d,m)$ so that $\mathbf{P} \subseteq \mathbb{F}[\mathbf{Y}]$ is a minimal decomposition and $\mathbf{Y}$ is t-rank-regular. For simplicity, assume without loss of generality that $\mathbf{Y}$ is linearly independent[22] (and nonempty, as otherwise there is nothing to prove). Put $V = \mathrm{Sp}\,\mathbf{Y}$ and $k = \dim V$ (> 0). To finish, we will use $\mathbf{Y}$ to obtain a weak ϵ-regular decomposition of $\mathbf{P}$ of size $k \le g(t,d,m) \le (2mc)^{2^{d(1+o(1))}}$.

Consider the subspace of V given by

$$U = \mathrm{Sp}\{X \in V \mid \mathrm{rk}(X) < tk\},$$

and the subset of V given by

$$U_\epsilon = \{X \in V \mid |\,\mathrm{bias}(X)| \ge q^{-ck}\}.$$

Then:

[22] This preserves $\mathbb{F}[\mathbf{Y}]$ and minimality, and $\mathrm{rk}(X) \ge t\cdot\dim\mathrm{Sp}(\mathbf{Y})$ outside of a strict subspace.

1. U is homogeneous. This is by Lemma 14 and since V is (it is spanned by $\mathbf{Y} \subseteq \text{Form}$).
2. $V^{\uparrow} \setminus U \neq \emptyset$. Indeed, by Lemma 15, the t-rank-regularity of $\mathbf{Y}$ implies $U_{tk}(V^{\uparrow}) \neq V^{\uparrow}$; thus, $V^{\uparrow} \setminus U = V^{\uparrow} \setminus (U \cap V^{\uparrow}) = V^{\uparrow} \setminus U_{tk}(V^{\uparrow}) \neq \emptyset$.
3. $U_{\epsilon} \subseteq U$. Indeed, for $X \in U_{\epsilon}$ we have $\text{rk}(X) < tk$ by (6), using that $k \leq g(t,d,m)$ and that $\deg(X) \leq \deg(\mathbf{Y}) \leq d$ (see Corollary 29).

Let $X_1 \in V^{\uparrow} \setminus U$ be nonzero, using 2 (as $V^{\uparrow} \setminus U \neq \emptyset, \{0\}$), add to it a homogeneous basis of U, using 1, and extend this independent set into a homogeneous basis $\mathbf{X} \in \text{Form}^k$ of V. Since U satisfies $(V \setminus U) \cap U_{\epsilon} = \emptyset$ by 3, we have $|\operatorname{bias}(X)| \leq q^{-ck}$ for every $X \in V \setminus U$. Therefore, we can apply Lemma 18 on $U \lneq V$, and $\mathbf{X}$ (with $\epsilon = q^{-(c-1)k}$), and deduce that $\mathbf{X}$ is weak ϵ-regular.

Finally, since $\operatorname{Sp}\mathbf{Y} = V = \operatorname{Sp}\mathbf{X}$, we have that $\mathbf{P} \subseteq \mathbb{F}[\mathbf{Y}] = \mathbb{F}[\mathbf{X}]$; moreover, by definition, since $\mathbf{P} \subseteq \mathbb{F}[\mathbf{Y}]$ is a minimal decomposition, so is $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$. In particular, $\mathbf{P} \nsubseteq \mathbb{F}[\mathbf{X}']$, since $\mathbf{X}' \subseteq \text{Form}$ satisfies $\operatorname{Sp}\mathbf{X}' \subsetneq \operatorname{Sp}\mathbf{X}$. Thus, $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$ is a weak ϵ-regular decomposition. The size of this decomposition is k, which, as discussed above, completes the proof. ◀

3 Applications

In this section we obtain several results along the theme that one can deduce strong bounds on the algebraic structure of polynomials, and polynomial maps, based solely on their image. The proofs use the weak regularity lemma.

We begin with the following standard quantification of algebraic structure, which constrains the degrees of the polynomials in the decomposition.

▶ **Definition 19** (t-rank). *Call a polynomial t-*reducible *if it is the product of polynomials of degree at most t. Denote, for $P \in \text{Poly}(\mathbb{F})$ with $\deg(P) = d$,*

$$\text{rk}_t(P) = \min\left\{r \,\middle|\, \exists R_1, \ldots, R_r \in \text{Poly}_d(\mathbb{F}) \text{ } t\text{-reducible} \colon P = R_1 + \cdots + R_r\right\}.$$

For completeness, set $\text{rk}_0(P) = 0$ for constant P.

This definition was used by Karam [21], and is closely related to an earlier definition of Green and Tao [15], $\text{rank}_t(P)$, which expresses P over a finite field as any function of $\text{rank}_t(P)$ polynomials of degree at most t. (Note that $\text{rank}_t(P) \leq O(\deg(P)/t)\,\text{rk}_t(P)$, as each t-reducible polynomial can be expressed in terms of $O(\deg(P)/t)$ polynomials of degree at most t; thus, proving an upper bound on $\text{rk}_t(P)$, as we will do below, would imply a similar upper bound on $\text{rank}_t(P)$.)

▶ Remark 20. If $\text{rk}_t(P)$ is small, for t bounded away from $\deg(P)$, then P is "correlated" with a polynomial of considerably smaller degree (unlike most polynomials of a given degree). More precisely, by a standard Fourier-analytic argument (see e.g. [20]), $P \in \text{Poly}(\mathbb{F}_q)$ has correlation $|\operatorname{bias}(P - Q)| \geq q^{-\text{rank}_t(P)}$ $(\geq q^{-\deg(P)\,\text{rk}_t(P)})$ with some polynomial Q of degree at most t.[23]

One can extend rk_t to polynomial tuples as follows. Define $\text{rk}_t(\mathbf{P})$ for $\mathbf{P} \in \text{Poly}^m$ to be the least number r of t-reducible polynomials $R_1, \ldots, R_r$ such that $\mathbf{P} \subseteq \operatorname{Sp}\{R_1, \ldots, R_r\}$. Note that for $m = 1$, this indeed reduces to Definition 19.

An application of the weak regularity lemma that is quite simple to state involves polynomial maps whose image is a line-free set. Recall that a line in $\mathbb{F}_q^m$ is a set of q elements of the form $\{ta + b \,|\, t \in \mathbb{F}_q\}$ for some $a, b \in \mathbb{F}_q^m$ $(a \neq 0)$. Henceforth, we denote

$$f(m,d) = (2m)^{2^{d(1+o(1))}}.$$

[23] In the context of the Gowers inverse theorem, this is denoted $\left\|\omega^P\right\|_{u^{t+1}} \geq q^{-\text{rank}_t(P)}$.

▶ **Theorem 21.** *Let* $\mathbf{P} \in \mathrm{Poly}_d^m(\mathbb{F})$ *with* $d < \mathrm{char}(\mathbb{F})$*. If* $\mathrm{Im}\,\mathbf{P} \subseteq \mathbb{F}^m$ *does not contain a line then* $\mathrm{rk}_{d/2}(\mathbf{P}) \le f(m,d)$*.*

In particular, the $m = 1$ case of Theorem 21 is about non-surjective polynomials.

Theorem 21 is a special case of a theorem about the degree of curves contained in the image of a polynomial map. We next recall the definition of the *univariate degree* of a polynomial map $\mathbf{P}$, which, roughly, measures the sparsity of the image of P, or how far P is from being surjective – the larger it is, the further P is from being surjective.

▶ **Definition 22** (Univariate degree)**.** *For* $\mathbf{P} \colon \mathbb{F}^n \to \mathbb{F}^m$*,*

$$\mathrm{udeg}(\mathbf{P}) = \min\{\deg(\mathbf{U}) \mid \mathbf{U} \colon \mathbb{F} \to \mathbb{F}^m \textit{ non-constant} \colon \mathrm{Im}\,\mathbf{U} \subseteq \mathrm{Im}\,\mathbf{P}\}.$$

By convention, if $\mathbf{P}$ *is constant then* $\mathrm{udeg}(\mathbf{P}) = \infty$*.*[24]

We include a brief list of properties of univariate degree in Section 3.1 below.

▶ Remark 23. The special case of Definition 22 where $(n, m, \mathrm{udeg}(P)) = (1, 1, 1)$ is already very interesting. Indeed, note that for any finite field $\mathbb{F}$, a *univariate* polynomial $P \colon \mathbb{F} \to \mathbb{F}$ satisfies $\mathrm{udeg}(P) = 1$ if and only if P is a permutation of $\mathbb{F}$; such polynomials are naturally called *permutation polynomials*, and have attracted much attention (see e.g. [18] and the references within).

Our main result in this section is as follows.

▶ **Theorem 24.** *For every* $\mathbf{P} \in \mathrm{Poly}_d^m(\mathbb{F})$ *with* $d < \mathrm{char}(\mathbb{F})$ *and with* $u = \mathrm{udeg}(\mathbf{P})$*, we have* $\mathrm{rk}_{d/u}(\mathbf{P}) \le f(m,d)$*.*

We highlight two extreme special cases. One is where $\mathrm{udeg}(\mathbf{P})$ is very large, namely $\mathrm{udeg}(\mathbf{P}) > \deg(\mathbf{P})/2$, in which case we directly obtain a small depth-3 (homogeneous) arithmetic formula for $\mathbf{P}$: each component P_i is expressed as a linear combination of (the same) few products of degree-1 polynomials.

▶ **Corollary 25.** *For* $\mathbf{P} \in \mathrm{Poly}_d^m(\mathbb{F})$ *with* $d < \mathrm{char}(\mathbb{F})$*, if* $\mathrm{udeg}(\mathbf{P}) > d/2$ *then* $\mathrm{rk}_1(\mathbf{P}) \le f(m,d)$*.*

The other extreme special case is Theorem 21, which is the special case where $\mathrm{udeg}(\mathbf{P}) > 1$. Indeed, $\mathrm{udeg}(\mathbf{P}) = 1$ means $\mathrm{Im}\,\mathbf{U} \subseteq \mathrm{Im}\,\mathbf{P}$ for some univariate $\mathbf{U} \colon \mathbb{F} \to \mathbb{F}^m$ with $\deg(\mathbf{U}) = 1$, and observe that $\deg(\mathbf{U}) = 1$ if and only if $\mathrm{Im}\,\mathbf{U}$ is a line. Indeed, $\deg(\mathbf{U}) = 1$ means $U_i(t) = ta_i + b_i$ with $a_i, b_i \in \mathbb{F}$, or equivalently, $\mathbf{U}(t) = ta + b$ with $a, b \in \mathbb{F}^m$, where $a \neq 0$. Thus, if $\mathrm{Im}\,\mathbf{P}$ does not contain a line then $\mathrm{udeg}(\mathbf{P}) \ge 2$, so Theorem 24 implies $\mathrm{rk}_{d/2}(\mathbf{P}) \le f(m,d)$ (that is, Theorem 21).

3.1 Properties of univariate degree

Recall the definition of univariate degree for a polynomial map $\mathbf{P} \colon \mathbb{F}^n \to \mathbb{F}^m$:

$$\mathrm{udeg}(\mathbf{P}) = \min\{\deg(\mathbf{U}) \mid \mathbf{U} \colon \mathbb{F} \to \mathbb{F}^m \text{ non-constant} \colon \mathrm{Im}\,\mathbf{U} \subseteq \mathrm{Im}\,\mathbf{P}\}.$$

We have $\mathrm{udeg}(Q^t) \ge \mathrm{udeg}(x^t)$ for any (multivariate) polynomial Q, with equality for any surjective Q (see Proposition 26 (3)), and moreover, it is not hard to show that $\mathrm{udeg}(x^t) = \gcd(t, p-1)$ assuming $\mathbb{F} = \mathbb{F}_p$ prime (as $\mathrm{Im}(x^t) = \mathrm{Im}(x^g)$ with $g = \gcd(t, p-1)$). Thus, for

[24] We require $\mathbf{U}$ to be non-constant as a polynomial. This is in fact equivalent to requiring $\mathbf{U}$ to be non-constant as a function, since we always have $\mathrm{udeg}(\mathbf{P}) < |\mathbb{F}|$ (for non-constant $\mathbf{P}$): if $a \neq b \in \mathrm{Im}\,\mathbf{P}$ then $\mathbf{U}(t) = a + (b-a)t^{|\mathbb{F}|-1}$ has $\mathrm{Im}\,\mathbf{U} \subseteq \mathrm{Im}\,\mathbf{P}$ and is non-constant.

$P = Q^t$ we have $\mathrm{udeg}(P) \geq t$ for any t dividing $p-1$. In this case, $\deg(P)/\mathrm{udeg}(P) \leq \deg(Q)$, with equality if Q is surjective. Now, if we take Q to be a random polynomial $Q\colon \mathbb{F}^n \to \mathbb{F}$, then we have that $\mathrm{rk}_{\deg(Q)-1}(P)$ is unbounded (as $n \to \infty$; see Example 1.4 in [21]). We conclude that one cannot hope to bound $\mathrm{rk}_t(P)$ for $t < \deg(P)/\mathrm{udeg}(P)$.

Below is a brief list of observations regarding the notion of univariate degree, which we include for the sake of completeness. For simplicity, we state them for the case of a single polynomial ($m = 1$). We will use the notation $\overline{P}$ for the unique *reduced* polynomial equal to P as a function, where reduced means that each variable has degree smaller than $|\mathbb{F}|$ (since $x^{|\mathbb{F}|} \equiv x$ as functions over $\mathbb{F}$).[25] Observe that $\mathrm{udeg}(P) = \mathrm{udeg}(\overline{P})$, since $\mathrm{Im}(P) = \mathrm{Im}(\overline{P})$.

▶ **Proposition 26.** *For non-constant $P \in \mathrm{Poly}(\mathbb{F})$, the univariate degree $\mathrm{udeg}(P)$ satisfies the following properties:*

1. $\mathrm{udeg}(P) \leq \min_i \deg_i(P)$ *($\leq \deg(P)$), over all i with $\deg_i(\overline{P}) > 0$.*
2. $\mathrm{udeg}(P) < |\mathbb{F}|$.
3. $\mathrm{udeg}(F(\mathbf{X})) \geq \mathrm{udeg}(F)$*, for any $F \in \mathrm{Poly}$, with equality if $\mathbf{X}$ is surjective.*
4. $\mathrm{udeg}(aP + b) = \mathrm{udeg}(P)$ *for any $a, b \in \mathbb{F}$ with $a \neq 0$.*
5. $\mathrm{udeg}(P^2) \geq 2$ *if $\mathrm{char}(\mathbb{F}) > 2$, with equality if P is surjective.*

Proof. In order:

1. It suffices to assume P is reduced, as $\mathrm{udeg}(P) = \mathrm{udeg}(\overline{P})$ and $\deg_i(\overline{P}) \leq \deg_i(P)$. Let i with $k := \deg_i(P) > 0$, and for concreteness assume $i = 1$. Write P as a univariate polynomial in x_1: $P = \sum_{j=0}^{k} x_1^j C_j(x_2, \ldots, x_n)$. Since C_k is both a nonzero polynomial and reduced (as P is), it is nonzero as a function as well: $C_k(c) \neq 0$ for some $c \in \mathbb{F}^{n-1}$. Thus, the univariate polynomial $U(t) = P(t, c) = \sum_{j=0}^{k} t^j C_j(c)$ is of degree k. Since $\mathrm{Im}\, U = \{P(t, c) \mid t \in \mathbb{F}\} \subseteq \mathrm{Im}\, P$ and $k > 0$, we have $\mathrm{udeg}(P) \leq k$.
2. $\mathrm{udeg}(P) = \mathrm{udeg}(\overline{P}) \leq \min_i \deg_i(P) < |\mathbb{F}|$ by the previous item.
3. We may assume F is non-constant. For U witnessing $\mathrm{udeg}(F(\mathbf{X}))$, we have $\mathrm{Im}(U) \subseteq \mathrm{Im}(F(\mathbf{X})) \subseteq \mathrm{Im}(F)$, so $\mathrm{udeg}(F) \leq \deg(U) = \mathrm{udeg}(F(\mathbf{X}))$. Moreover, if $\mathbf{X}$ is surjective then $\mathrm{Im}(F) = \mathrm{Im}(F(\mathbf{X}))$, and generally, $\mathrm{udeg}(P)$ depends only on $\mathrm{Im}(P)$.
4. If a univariate U satisfies $\mathrm{Im}(U) \subseteq \mathrm{Im}(P)$, then $\mathrm{Im}(aU + b) \subseteq \mathrm{Im}(aP + b)$, so $\mathrm{udeg}(aP + b) \leq \deg(aU + b) = \deg(U)$, which shows $\mathrm{udeg}(aP + b) \leq \mathrm{udeg}(P)$ if we choose U that witnesses $\mathrm{udeg}(P)$. The equality follows by symmetry.
5. $\mathrm{udeg}(P^2) \geq \mathrm{udeg}(x^2)$, and we claim that $\mathrm{udeg}(x^2) = 2$. It is not surjective as $\gcd(2, |\mathbb{F}| - 1) = 2$, using that $|\mathbb{F}|$ is odd. Thus, $2 \leq \mathrm{udeg}(x^2) \leq \deg(x^2) = 2$. Moreover, if P is surjective then $\mathrm{udeg}(P^2) = \mathrm{udeg}(x^2)$. ◀

3.2 Polynomial decompositions

First, we will need the following immediate property of weak ϵ-regular tuples. Note that for an arbitrary map $\mathbf{X}\colon \mathbb{F}^n \to \mathbb{F}^r$, if a line $\ell \subseteq \mathbb{F}^r$ in a given direction is chosen uniformly at random, then $\mathbb{E}_\ell \Pr[\mathbf{X} \in \ell] = q^{-(r-1)}$. We show that if $\mathbf{X}$ is weak regular, then lines that are somewhat popular (that is, close to being as popular as a random line) must be fully contained in the image of $\mathbf{X}$.

▶ **Lemma 27.** *Let $\mathbf{X} \subseteq \mathrm{Form}^r(\mathbb{F}_q)$ be weak ϵ-regular. For every line $\ell \subseteq \mathbb{F}^r$ in direction e_1, if $\Pr[\mathbf{X} \in \ell] > \epsilon q^{-(r-1)}$ then $\ell \subseteq \mathrm{Im}\, \mathbf{X}$.*

[25] That is, $\overline{P} = P \pmod{x_1^{|\mathbb{F}|} - x_1, \ldots, x_n^{|\mathbb{F}|} - x_n}$.

Proof. By the definition of weak ϵ-regularity (recall (1)), for every $y \in \ell$ we have

$$\Pr[\mathbf{X} = y] \geq \Pr[\mathbf{X} \in \ell]/q - \epsilon q^{-r} > 0.$$

Thus $y \in \operatorname{Im} \mathbf{X}$, which completes the proof. ◀

Motivated by Lemma 27, we proceed to prove results about popular lines in the image of polynomial decompositions (equivalently, of polynomials in not necessarily independent variables). These results do not assume weak regularity.

We say that a polynomial decomposition $P = F(\mathbf{X})$ is *irredundant* if there is no polynomial F' with $\operatorname{Supp} F' \subsetneq \operatorname{Supp} F$ such that $F'(\mathbf{X}) = F(\mathbf{X})$.[26] Note that the definition of a weak ϵ-regular decomposition (Definition 7) is independent of the choice of F, so there is no loss of generality in assuming it is irredundant, as we will later do.

First, we show that a decomposition of a polynomial P into forms need not involve terms of higher degree than that of P.

▶ **Lemma 28** (irredundant degree). *For $P \in \mathrm{Poly}$, let $P = F(\mathbf{X})$, with $\mathbf{X} \subseteq \mathrm{Form}$, be an irredundant decomposition. For every monomial M in the support of F, $\deg M(\mathbf{X}) \leq \deg P$.*

Proof. Since every X_i is homogeneous, for every monomial $M \in \operatorname{Supp}(F)$, all the monomials that result from expanding $M(\mathbf{X})$ are of the same degree. Thus, for each k, the degree-k part of $F(\mathbf{X})$ is equal to $\sum M(\mathbf{X})$ with the sum over all monomials $M \in \operatorname{Supp}(F)$ with $\deg M(\mathbf{X}) = k$. Therefore, if F' is obtained from F by removing from its support all monomials M with $\deg M(\mathbf{X}) = k$, for $k > \deg F(\mathbf{X})$, then we still have $F'(\mathbf{X}) = F(\mathbf{X})$. We deduce that if there exists a monomial $M \in \operatorname{Supp}(F)$ with $\deg M(\mathbf{X}) > \deg F(\mathbf{X})$, then $P = F(\mathbf{X})$ is not an irredundant decomposition, completing the proof. ◀

▶ **Corollary 29.** *For any decomposition $P \subseteq \mathbb{F}[\mathbf{X}]$, with $P \in \mathrm{Poly}$ and $\mathbf{X} \subseteq \mathrm{Form}$, if $P \nsubseteq \mathbb{F}[\mathbf{X} \setminus \{X_i\}]$ then $\deg(X_i) \leq \deg(P)$.*

Proof. Without loss of generality, $P = F(\mathbf{X})$ is an irredundant decomposition. We have $\deg_i(F) \geq 1$, since $P \nsubseteq \mathbb{F}[\mathbf{X} \setminus \{X_i\}]$. Apply Lemma 28 on any monomial M in the support of F with $\deg_i(M) \geq 1$. Then $\deg(X_i) \leq \deg M(\mathbf{X}) \leq \deg(P)$, as needed. ◀

We next prove that in the image of any polynomial map $\mathbf{X}$ one can find a popular line that avoids a given low-degree zero set.

▶ **Lemma 30** (Schwartz-Zippel for decompositions). *Let $P \in \mathrm{Poly}_d(\mathbb{F}_q)$, with $d < q$, be nonzero. If $P = F(\mathbf{X})$ with $\mathbf{X} \in \mathrm{Poly}^k$ (and $F \in \mathrm{Poly}$), then there exists $y \in \mathbb{F}_q^k$ such that $F(y) \neq 0$ and*

$$\Pr[F \neq 0] \cdot \Pr[\mathbf{X} = y] \geq (1 - d/q)q^{-k}.$$

In particular, $\Pr[\mathbf{X} = y] \geq (1 - d/q)q^{-k}$. Note that if $\mathbf{X}$ is perfectly equidistributed (e.g., independent variables), we retrieve the usual Schwartz-Zippel lemma: $\Pr[F \neq 0] \geq 1 - d/q$.

Proof. Consider $\{y \in \mathbb{F}^k \mid F(y) \neq 0\}$, the complement of the zero set $\mathrm{Z}(F)$ of F. We have

$$\Pr_x[P(x) \neq 0] = \Pr_x[\mathbf{X} \notin \mathrm{Z}(F)] = \sum_{y \notin \mathrm{Z}(F)} \Pr[\mathbf{X} = y] \leq q^k \Pr[F \neq 0] \max_{y \notin \mathrm{Z}(F)} \Pr[\mathbf{X} = y].$$

On the other hand, as P is nonzero, the Schwartz-Zippel lemma gives $\Pr[P \neq 0] \geq 1 - d/q$. Combining the above, we get $\Pr[F \neq 0] \max_{y \notin \mathrm{Z}(F)} \Pr[\mathbf{X} = y] \geq (1 - d/q)q^{-k}$, as desired. ◀

[26] As usual, this is an equality of polynomials. Recall that $\operatorname{Supp} F$ denotes the support of F, meaning the set of monomials appearing with a nonzero coefficient in F.

We combine the above lemmas to find a popular line that the decomposition nontrivially depends on. Denote by $\deg_1(F)$ the degree of $F \in \mathbb{F}[x_1, \dots, x_r]$ in its first variable (that is, the largest exponent of x_1 in the monomials in the support of F).

▶ **Corollary 31.** *Let $P \in \mathrm{Poly}_d(\mathbb{F}_q)$ with $d < q$. If $P = F(\mathbf{X})$ with $\mathbf{X} \in \mathrm{Form}^r$ is an irredundant decomposition with $\deg_1 F > 0$, then there exists a line $\ell \subseteq \mathbb{F}_q^r$ in direction e_1 such that $F|_\ell$ is non-constant and $\Pr[\mathbf{X} \in \ell] > (1 - d/q)q^{-(r-1)}$.*

Proof. Put $\mathbb{F} = \mathbb{F}_q$. Write $F(t, \mathbf{y}) = \sum_{i \geq 0} t^i C_i(\mathbf{y})$ with $C_i \in \mathrm{Poly}$, the unique representation of F as a linear combination of the monomials in its support. We claim that the following holds for some $j \geq 1$:

- $C_j(\mathbf{X}')$ is a nonzero polynomial; indeed, otherwise we have $P = F(\mathbf{X}) = F(X_1, \mathbf{X}') = C_0(\mathbf{X}')$; moreover, $\mathrm{Supp}\, C_0 \subseteq \mathrm{Supp}\, F$, and $\mathrm{Supp}\, C_0 \neq \mathrm{Supp}\, F$ by the assumption $\deg_1(F) > 0$. This contradicts the assumption that $P = F(\mathbf{X})$ is an irredundant decomposition.
- $C_j(\mathbf{X}')$ is of degree smaller than d; indeed, since $\mathbf{X} \subseteq \mathrm{Form}$ and the decomposition $P = F(\mathbf{X})$ is irredundant, we have that $\deg(C_j(\mathbf{X}')) < \deg(X_1^j C_j(\mathbf{X}')) \leq \deg(P) \leq d$, using Lemma 28 in the second inequality.

Apply Lemma 30 on the nonzero $C_j(\mathbf{X}') \in \mathrm{Poly}_{d-1}$ to obtain $y_0 \in \mathbb{F}_q^{r-1}$ such that $C_j(y_0) \neq 0$ and $\Pr[\mathbf{X}' = y_0] > (1 - d/q)q^{-(r-1)}$. Observe that $\deg(F(t, y_0)) \geq j$ (≥ 1), so in particular, $F(t, y_0)$ is a non-constant univariate polynomial. Thus, the restriction $F|_\ell$ of F to the line $\ell = \{(t, y_0) \mid t \in \mathbb{F}\} \subseteq \mathbb{F}^r$ is non-constant. Moreover, observe that $\Pr[\mathbf{X} \in \ell] = \Pr[\mathbf{X}' = y_0]$. This completes the proof. ◀

Finally, we use the following easy (and tight) binomial coefficient bound.

▷ Claim 32. $\binom{n+t}{t} \leq \frac{3}{2} n^t$ for every $n \geq 2$, $t \geq 1$.

Proof. By induction on t, with base case $\binom{n+1}{1} = n+1 \leq \frac{3}{2}n$ as $n \geq 2$. For the induction step, $\binom{n+t+1}{t+1} = \binom{n+t}{t}\frac{n+t+1}{t+1} \leq \frac{3}{2}n^t \cdot n$, using the bound $\frac{n+t+1}{t+1} \leq n$, or equivalently, $1 \leq t(n-1)$. ◁

3.3 Proofs of the applications

Proof of Theorem 24. Put $\mathbb{F} = \mathbb{F}_q$. Assume $\mathbf{P}$ is non-constant, as otherwise $\mathrm{rk}_t(\mathbf{P}) = 0$ by convention and we are done. Put $\epsilon = 1 - d/q$. By Theorem 8 (with $c = 2$, so $q^{-(c-1)k} \leq 1/q \leq \epsilon$), $\mathbf{P}$ has a weak ϵ-regular decomposition $\mathbf{P} = \mathbf{F}(\mathbf{X})$, with $\mathbf{X} \in \mathrm{Form}^r$, of size $r \leq (2m)^{2^{d(1+o(1))}}$. We will prove that $\deg(\mathbf{X}) \leq \deg(\mathbf{P}) / \mathrm{udeg}(\mathbf{P})$.

We assume without loss of generality that each decomposition $P_i = F_i(\mathbf{X})$ is irredundant (by removing monomials from the support of F_i if necessary). Since $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$ but $\mathbf{P} \nsubseteq \mathbb{F}[\mathbf{X}']$, there is a component P_j of $\mathbf{P}$ such that $P_j \in \mathbb{F}[\mathbf{X}] \setminus \mathbb{F}[\mathbf{X}']$, and therefore $\deg_1(F_j) > 0$. Apply Corollary 31 on $P_j \in \mathrm{Poly}_d(\mathbb{F})$ to obtain a line $\ell \subseteq \mathbb{F}^r$ in direction e_1 such that $F_j|_\ell$ is non-constant and $\Pr[\mathbf{X} \in \ell] > (1 - d/q)q^{-(r-1)}$. Apply Lemma 27 on ℓ, using the weak ϵ-regularity of $\mathbf{X}$ and our choice of ϵ, to deduce that

$$\ell \subseteq \mathrm{Im}\,\mathbf{X} \text{ and } \mathbf{F}|_\ell \text{ is non-constant.} \tag{7}$$

Let $\mathbf{U} \colon \mathbb{F} \to \mathbb{F}^m$ be the curve $\mathbf{U} = \mathbf{F}(\mathbf{L})$, where $\mathbf{L} \colon \mathbb{F} \to \mathbb{F}^r$ satisfies $\mathrm{Im}(\mathbf{L}) = \ell$. (Explicitly, if $\ell = \{te_1 + y \mid t \in \mathbb{F}\}$, then $\mathbf{L}(t) = te_1 + y$, so $\mathbf{U}(t) = \mathbf{F}(\mathbf{L}(t)) = \mathbf{F}(t + y_1, y_2, \dots, y_r)$.) Then $\mathrm{Im}\,\mathbf{U} = \mathbf{F}(\mathrm{Im}(\mathbf{L})) = \mathbf{F}(\ell)$, which implies that

$$\mathrm{Im}\,\mathbf{U} \subseteq \mathrm{Im}\,\mathbf{P} \text{ and } \mathbf{U} \text{ is non-constant,}$$

since $\mathbf{F}(\ell) \subseteq \mathbf{F}(\mathrm{Im}\,\mathbf{X}) = \mathrm{Im}\,\mathbf{P}$ by (7). Thus, by definition, $\mathrm{udeg}(\mathbf{P}) \leq \deg(\mathbf{U})$.

Let $i \in [m]$; abbreviate $P = P_i$, $F = F_i$, and $U = U_i$, so that $P = F(\mathbf{X})$ and $U = F(\mathbf{L})$. We claim that the following chain of inequalities holds:

$$\deg(U) \le \deg_1(F) \le \deg(P)/\deg(X_1) = \deg(P)/\deg(\mathbf{X}).$$

The first inequality is immediate from the construction of U. The second inequality follows using Lemma 28: as $P = F(\mathbf{X})$, for every monomial $M = \prod_{j=1}^{r} x_j^{m_j(M)}$ in the support of F, we have $m_1(M)\deg(X_1) \le \deg(M(\mathbf{X})) \le \deg(P)$; since $\deg_1(F) = \max_M m_1(M)$ over $M \in \operatorname{Supp} F$, we indeed obtain that $\deg_1(F)\deg(X_1) \le \deg(P)$. The final equality follows from the construction of $\mathbf{X}$. We obtain the following degree bound for $\mathbf{X}$, using i satisfying $\deg(U_i) = \deg(\mathbf{U})$:

$$\deg(\mathbf{X}) \le \deg(P)/\deg(U) \le \deg(\mathbf{P})/\deg(\mathbf{U}) \le \deg(\mathbf{P})/\operatorname{udeg}(\mathbf{P}).$$

Finally, as every component P_i of $\mathbf{P}$ satisfies $P_i = F_i(\mathbf{X})$ with $F_i \colon \mathbb{F}^r \to \mathbb{F}$ and $\deg(F_i) \le \deg(P_i) \le d$, by Lemma 28 (and as $\deg(X_i) \ge 1$), we deduce that $P_1, \dots, P_m$ lie in the span of all products of at most d members of $\mathbf{X}$. We are therefore done (using Claim 32, assuming $r \ge 2$ without loss of generality):

$$\operatorname{rk}_{d/u}(\mathbf{P}) \le \binom{r+d}{d} \le \frac{3}{2} r^d \le (2m)^{2^{d(1+o(1))}}. \qquad \blacktriangleleft$$

We are now ready to deduce Theorems 3 and 4.

Proof of Theorem 3. Let P be an n-variate polynomial with $\deg(P) \le d$ and $\operatorname{udeg}(P) \ge u$. Assume P is non-constant, as otherwise there is nothing to prove. Put $t = \lfloor d/u \rfloor$ (≥ 1). Then $P = R_1 + \dots + R_{\operatorname{rk}_t(P)}$ where each $R_i \in \operatorname{Poly}_d$ is a product of polynomials of degree at most t, that is, $R_i = \prod_{j \le d_i} F_{i,j}$ with $\deg(F_{i,j}) \le t$ and $d_i \le d$, by Definition 19. Since, trivially, $L(F_{i,j}) \le O(\binom{n+t}{t}) \le O(n^t)$ (Claim 32, assuming $n \ge 2$), we deduce that

$$L(P) \le O\big(\sum_{i,j} L(F_{i,j})\big) \le O(d \cdot \operatorname{rk}_t(P) \cdot n^t) \le n^{t+o(1)},$$

where the last inequality uses Theorem 24, and the statement's assumption on d, to bound

$$\operatorname{rk}_t(P) \le f(1,d) = 2^{2^{d(1+o(1))}} \le 2^{(\log_2(n))^{\frac{1}{2}+o(1)}} = n^{(\log_2(n))^{-(\frac{1}{2}+o(1))}} \le n^{o(1)}. \qquad \blacktriangleleft$$

Proof of Theorem 4. Let $P \subseteq \mathbb{F}[\mathbf{P}]$ with $\operatorname{udeg}(\mathbf{P}) = u$. As in the proof of Theorem 24, using the weak regularity lemma we obtain a decomposition $\mathbf{P} \subseteq \mathbb{F}[\mathbf{X}]$ with $\mathbf{X} \in \operatorname{Form}^s$ of size $s \le (2m)^{2^{d(1+o(1))}}$ such that $\deg(\mathbf{X}) \le d/u$. Observe that $P \in \mathbb{F}[\mathbf{P}] \subseteq \mathbb{F}[\mathbf{X}]$. Using Lemma 28 with an irredundant representation (by removing monomials if necessary), it follows that we can write

$$P = \sum_{i=1}^{r} c_i \prod_j Q_{i,j} \tag{8}$$

with $Q_{i,j} \in \mathbf{X}$, $c_i \in \mathbb{F}$, $\deg(\prod_j Q_{i,j}) \le d$, and $r \le \sum_{i=0}^{d} s^i \le (2m)^{2^{d(1+o(1))}}$. Note that $\deg(Q_{i,j}) \le \deg(\mathbf{X}) \le d/u$.

We claim that we can guarantee that the number of multiplicands in each product $\prod_j Q_{i,j}$ is at most $2u$. Note that we can express each product as a product of homogeneous polynomials $\prod_{j=1}^{a_i} Q'_{i,j}$ all of whose degrees, except at most one of them, satisfy $\frac{1}{2}(d/u) < \deg(Q'_{i,j}) \le d/u$ (note that $Q'_{i,j}$ are not necessarily members of $\mathbf{X}$); indeed, this follows by repeatedly using the fact that if $\deg(Q_{i,j_1}), \deg(Q_{i,j_2}) \leq \frac{1}{2}(d/u)$ then $\deg(Q_{i,j_1}Q_{i,j_2}) \leq d/u$. So, $d \geq \sum_{j=1}^{a_i} \deg(Q'_{i,j}) > (a_i - 1)\frac{1}{2}(d/u)$, hence $a_i \leq 2u$ as claimed. Thus, (8) gives a $\sum^{[r]}\prod^{[2u]}\sum\prod^{[d/u]}$-formula for P, as needed.

Finally, if P is homogeneous then, since every member of $\mathbf{X}$ is homogeneous, without loss of generality the products in the formula satisfy $\deg(\prod_j Q_{i,j}) = \deg(P)$ for every i, and therefore the formula is homogeneous, as desired. ◀

References

1. Tigran Ananyan and Melvin Hochster. Small subalgebras of polynomial rings and Stillman's Conjecture. *Journal of the American Mathematical Society*, 33(1):291–309, 2020. doi:10.1090/jams/932.
2. Paul Beame, Shayan Oveis Gharan, and Xin Yang. On the bias of Reed–Muller codes over odd prime fields. *SIAM Journal on Discrete Mathematics*, 34(2):1232–1247, 2020. doi:10.1137/18M1215104.
3. Abhishek Bhowmick and Shachar Lovett. Bias vs structure of polynomials in large fields, and applications in information theory. *IEEE Transactions on Information Theory*, 69(2):963–977, 2023. doi:10.1109/TIT.2022.3214372.
4. Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. *SIAM Journal on Computing*, 39(6):2464–2486, 2010. doi:10.1137/070712109.
5. Alex Cohen and Guy Moshkovitz. Partition and analytic rank are equivalent over large fields. *Duke Mathematical Journal*, 172(12):2433–2470, 2023. doi:10.1215/00127094-2022-0086.
6. Zeev Dvir, János Kollár, and Shachar Lovett. Variety evasive sets. *Computational Complexity*, 23(4):509–529, 2014. doi:10.1007/s00037-013-0073-9.
7. Zeev Dvir and Shachar Lovett. Subspace evasive sets. In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing*, STOC 2012, pages 351–358. ACM, 2012. doi:10.1145/2213977.2214010.
8. Christian Elsholtz, Zach Hunter, Laura Proske, and Lisa Sauermann. Improving Behrend's construction: Sets without arithmetic progressions in integers and over finite fields, 2024. arXiv:2406.12290.
9. Daniel Erman, Steven V. Sam, and Andrew Snowden. Cubics in 10 variables vs. cubics in 1000 variables: uniformity phenomena for bounded degree polynomials. *Bulletin of the American Mathematical Society*, 56(1):87–114, 2019. doi:10.1090/bull/1652.
10. Jacob Fox and Huy Tuan Pham. Popular progression differences in vector spaces II. *Discrete Analysis*, 16:39 pp., 2019. doi:10.19086/da.11002.
11. Alan Frieze and Ravi Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999. doi:10.1007/s004930050052.
12. Abhibhav Garg, Rafael Oliveira, and Akash Kumar Sengupta. Rank bounds and PIT for depth-4 circuits with top fan-in 3 and constant bottom fan-in via a non-linear Edelstein–Kelly theorem. In *Proceedings of the 66th IEEE Annual Symposium on Foundations of Computer Science*, FOCS 2025, pages 2143–2169. IEEE, 2025. doi:10.1109/FOCS63196.2025.00114.
13. W. T. Gowers and Julia Wolf. Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$. *Geometric and Functional Analysis*, 21(1):36–69, 2011. doi:10.1007/s00039-010-0106-3.
14. Ben Green. Montreal lecture notes on quadratic Fourier analysis. In Andrew Granville, Melvyn B. Nathanson, and József Solymosi, editors, *Additive Combinatorics*, volume 43 of *CRM Proceedings and Lecture Notes*, pages 69–102. American Mathematical Society, 2007. arXiv:math/0604089.
15. Ben Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms. *Contributions to Discrete Mathematics*, 4(2):1–36, 2009. doi:10.55016/ojs/cdm.v4i2.62086.
16. Jaime Gutiérrez, Rosario Rubio, and David Sevilla. On multivariate rational function decomposition. *Journal of Symbolic Computation*, 33(5):545–562, 2002. doi:10.1006/jsco.2000.0529.

17 Hamed Hatami and Shachar Lovett. Higher-order Fourier analysis of $\mathbb{F}_p^n$ and the complexity of systems of linear forms. *Geometric and Functional Analysis*, 21(6):1331–1357, 2011. doi:10.1007/s00039-011-0141-8.

18 Xiang-dong Hou. Permutation polynomials over finite fields—a survey of recent advances. *Finite Fields and Their Applications*, 32:82–119, 2015. doi:10.1016/j.ffa.2014.10.001.

19 Pavel Hrubeš and Amir Yehudayoff. Arithmetic complexity in ring extensions. *Theory of Computing*, 7(8):119–129, 2011. doi:10.4086/toc.2011.v007a008.

20 Oliver Janzer. Polynomial bound for the partition rank vs the analytic rank of tensors. *Discrete Analysis*, 7:18 pp., 2020. doi:10.19086/da.12935.

21 Thomas Karam. Ranges of polynomials control degree ranks of Green and Tao over finite prime fields. *Proceedings of the Edinburgh Mathematical Society*, pages 1–28, 2026. doi:10.1017/S0013091526101382.

22 Tali Kaufman and Shachar Lovett. Worst case to average case reductions for polynomials. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, FOCS 2008, pages 166–175. IEEE, 2008. doi:10.1109/FOCS.2008.17.

23 Tali Kaufman, Shachar Lovett, and Ely Porat. Weight distribution and list-decoding size of Reed–Muller codes. *IEEE Transactions on Information Theory*, 58(5):2689–2696, 2012. doi:10.1109/TIT.2012.2184841.

24 Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC 2014, pages 146–153. ACM, 2014. doi:10.1145/2591796.2591847.

25 David Kazhdan, Amichai Lampert, and Alexander Polishchuk. Schmidt rank and singularities. *Ukrainian Mathematical Journal*, 75(9):1420–1442, 2024. doi:10.1007/s11253-024-02270-6.

26 Mrinal Kumar and Shubhangi Saraf. The limits of depth reduction for arithmetic formulas: it's all about the top fan-in. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC 2014, pages 136–145. ACM, 2014. doi:10.1145/2591796.2591827.

27 Amichai Lampert. Small ideals in polynomial rings and applications, 2023. arXiv:2309.16847.

28 Amichai Lampert and Tamar Ziegler. On rank in algebraic closure. *Selecta Mathematica. New Series*, 30(1), 2024. doi:10.1007/s00029-023-00903-5.

29 Amichai Lampert and Tamar Ziegler. Relative rank and regularization. *Forum of Mathematics, Sigma*, 12, 2024. doi:10.1017/fms.2024.15.

30 Shachar Lovett. Computing polynomials with few multiplications. *Theory of Computing*, 7(13):185–188, 2011. doi:10.4086/toc.2011.v007a013.

31 Luka Milićević. Polynomial bound for partition rank in terms of analytic rank. *Geometric and Functional Analysis*, 29(5):1503–1530, 2019. doi:10.1007/s00039-019-00505-4.

32 Guy Moshkovitz and Daniel G. Zhu. Quasi-linear relation between partition and analytic rank, 2022. arXiv:2211.05780.

33 Eric Naslund. The partition rank of a tensor and k-right corners in $\mathbb{F}_q^n$. *Journal of Combinatorial Theory, Series A*, 174, 2020. doi:10.1016/j.jcta.2019.105190.

34 Rafael Oliveira and Akash Kumar Sengupta. Strong algebras and radical Sylvester–Gallai configurations. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, pages 95–105. ACM, 2024. doi:10.1145/3618260.3649617.

35 Ran Raz. Elusive functions and lower bounds for arithmetic circuits. *Theory of Computing*, 6(7):135–177, 2010. doi:10.4086/toc.2010.v006a007.

36 Wolfgang M. Schmidt. The density of integer points on homogeneous varieties. *Acta Mathematica*, 154:243–296, 1985. doi:10.1007/BF02392473.

37 Hans-Jörg Stoß. Lower bounds for the complexity of polynomials. *Theoretical Computer Science*, 64(1):15–23, 1989. doi:10.1016/0304-3975(89)90094-7.

38 Terence Tao and Tamar Ziegler. The inverse conjecture for the Gowers norm over finite fields in low characteristic. *Annals of Combinatorics*, 16(1):121–188, 2012. doi:10.1007/s00026-011-0124-3.

39 Joachim von zur Gathen and Konstantin Ziegler. Survey on counting special types of polynomials. In Jaime Gutiérrez, Josef Schicho, and Martin Weimann, editors, *Computer Algebra and Polynomials*, volume 8942 of *Lecture Notes in Computer Science*, pages 50–75. Springer Cham, 2015. doi:10.1007/978-3-319-15081-9_3.