On the Structure of Quintic Polynomials

Pooya Hatami*

Institute for Advanced Study, Princeton, NJ, USA
pooyahat@math.ias.edu

Abstract

We study the structure of bounded degree polynomials over finite fields. Haramaty and Shpilka [STOC 2010] showed that biased degree three or four polynomials admit a strong structural property. We confirm that this is the case for degree five polynomials also. Let $\mathbb{F} = \mathbb{F}_q$ be a prime field. Suppose $f : \mathbb{F}^n \to \mathbb{F}$ is a degree five polynomial with $\mathrm{bias}(f) = \delta$. We prove the following two structural properties for such f.

1. We have $f = \sum_{i=1}^{c} G_i H_i + Q$, where G_i and H_is are nonconstant polynomials satisfying $\deg(G_i) + \deg(H_i) \leqslant 5$ and Q is a degree $\leqslant 4$ polynomial. Moreover, c does not depend on n.
2. There exists an $\Omega_{\delta,q}(n)$ dimensional affine subspace $V \subseteq \mathbb{F}^n$ such that $f|_V$ is a constant.

Cohen and Tal [Random 2015] proved that biased polynomials of degree at most four are constant on a subspace of dimension $\Omega(n)$. Item [2.] extends this to degree five polynomials. A corollary to Item [2.] is that any degree five affine disperser for dimension k is also an affine extractor for dimension $O(k)$. We note that Item [2.] cannot hold for degrees six or higher.

We obtain our results for degree five polynomials as a special case of structure theorems that we prove for biased degree d polynomials when $d < |\mathbb{F}| + 4$. While the $d < |\mathbb{F}| + 4$ assumption seems very restrictive, we note that prior to our work such structure theorems were only known for $d < |\mathbb{F}|$ by Green and Tao [Contrib. Discrete Math. 2009] and Bhowmick and Lovett [arXiv:1506.02047]. Using algorithmic regularity lemmas for polynomials developed by Bhattacharyya, et. al. [SODA 2015], we show that whenever such a strong structure exists, it can be found algorithmically in time polynomial in n.

1998 ACM Subject Classification G.2.1 Combinatorics

Keywords and phrases Higher-order Fourier analysis, Structure Theorem, Polynomials, Regularity lemmas

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2016.33

1 Introduction

Let $\mathbb{F} = \mathbb{F}_q$ be a prime field. The bias of a function $f : \mathbb{F}^n \to \mathbb{F}$ is defined as

$$\mathrm{bias}(f) := \left| \mathop{\mathbf{E}}_{x \in \mathbb{F}^n} \left[\omega^{f(x)} \right] \right|,$$

where $\omega = e^{2\pi i/|\mathbb{F}|}$, is a complex primitive root of unity of order $|\mathbb{F}|$. The smaller the bias of a function, the more uniformly f is distributed over $\mathbb{F}$, thus a random function has negligible bias. This remains true, if f is a random degree d polynomial for a fixed degree $d > 0$. Thus bias can be thought of as a notion of pseudorandomness for polynomials, and as often lack of pseudorandomness implies structure, one may ask whether every biased degree d polynomial admits strong structural properties. Green and Tao [7] (in the case when $d < |\mathbb{F}|$) and later Kaufman and Lovett [11] (in the general case) proved this heuristic to be true by showing that

* Supported by the National Science Foundation grant No. CCF-1412958.

every biased degree d polynomial is determined by a few lower degree polynomials. Formally, these results state that for a degree d polynomial f, there is a constant $c \leqslant c(d, \mathrm{bias}(f), |\mathbb{F}|)$, degree $\leqslant d-1$ polynomials $Q_1, \ldots, Q_c$ and a function $\Gamma : \mathbb{F}^c \to \mathbb{F}$, such that

$$f = \Gamma(Q_1, \ldots, Q_c). \tag{1}$$

Note that here crucially c does not depend on the dimension n, meaning that for large n, it is very unlikely for a typical polynomial to be biased. Recently, Bhowmick and Lovett [3] proved that the dependence of the number of terms in (1) on $|\mathbb{F}|$ can be removed, in other words biased polynomials are very rare even when the field size is allowed to grow with n. These structure theorems for biased polynomials have had several important applications. For example they were used by Kaufman and Lovett [11] to give interesting worst case to average case reductions, and by Tao and Ziegler [16] in their proof of the inverse theorem for Gowers norms over finite fields. Such structure theorems have played an important role in determining the weight distribution and list decoding radius of Reed-Muller codes [12, 4, 3]. They were also used by Cohen and Tal [5] to show that any degree d affine disperser over a prime field is also an affine extractor with related parameters.

There are however two drawbacks to the structure theorems proved in [7, 11]. Firstly, the constant c has very bad dependence on δ which is due to the use of regularity lemmas for polynomials. Secondly, there is no restrictions on the function Γ obtained in (1), in particular there is nothing stopping it from being of degree c. In the special case of quadratic polynomials better bounds and structural properties follow from the following well-known theorem.

▶ **Theorem 1** (Structure of quadratic polynomials [13]). *For every quadratic polynomial $f : \mathbb{F}^n \to \mathbb{F}$ over a prime field $\mathbb{F}$, there exists an invertible linear map T, a linear polynomial ℓ, and field elements $\alpha_1, \ldots, \alpha_n$ such that*

- *If $|\mathbb{F}| = 2$, then $(f \circ T)(x) = \sum_{i=1}^{\lfloor n/2 \rfloor} \alpha_i x_{2i-1} x_{2i} + \ell(x)$.*
- *If $|\mathbb{F}|$ is odd, then $(f \circ T)(x) = \sum_{i=1}^{n} \alpha_i x_i^2 + \ell(x)$.*

It easily follows that every quadratic polynomial f, can be written in the form $\sum_{i=1}^{2\log(1/\mathrm{bias}(f))} \ell_i \ell_i' + \ell''$ where ℓ_i, ℓ_i's and ℓ'' are linear polynomials. This is a very strong structural property, moreover the dependence of the number of the terms on $\mathrm{bias}(f)$ is optimal. Haramaty and Shpilka [8] studied the structure of biased cubic and quartic polynomials and proved the following two theorems.

▶ **Theorem 2** (Biased cubic polynomials [8]). *Let $f : \mathbb{F}^n \to \mathbb{F}$ be a cubic polynomial such that $\mathrm{bias}(f) = \delta > 0$. Then there exist $c_1 = O(\log(1/\delta))$, $c_2 = O\left(\log^4(1/\delta)\right)$, quadratic polynomials $Q_1, ..., Q_{c_1} : \mathbb{F}^n \to \mathbb{F}$, linear functions $\ell_1, ..., \ell_{c_1}, \ell_1', ..., \ell_{c_2}' : \mathbb{F}^n \to \mathbb{F}$ and a cubic polynomial $\Gamma : \mathbb{F}^{c_2} \to \mathbb{F}$ such that*

$$f = \sum_{i=1}^{c_1} \ell_i Q_i + \Gamma\left(\ell_1', \ldots, \ell_{c_2}'\right).$$

▶ **Theorem 3** (Biased quartic polynomials [8]). *Let $f : \mathbb{F}^n \to \mathbb{F}$ be a quartic polynomial such that $\mathrm{bias}(f) = \delta$. There exist $c = \mathrm{poly}(|\mathbb{F}|/\delta)$ and polynomials $\{\ell_i, Q_i, Q_i', G_i\}_{i \in [c]}$, where the ℓ_is are linear, Q_i, Q_i's are quadratic, and G_i's are cubic polynomials, such that*

$$f = \sum_{i=1}^{c} \ell_i G_i + \sum_{i=1}^{c} Q_i Q_i'.$$

In the high characteristic regime when $d = \deg(f) < |\mathbb{F}|$, Green and Tao [7] showed that such a strong structure theorem holds, with a dependence that is really large in terms of bias. More precisely, if $d < |\mathbb{F}|$, then every degree d polynomial f, with $\text{bias}(f) \geqslant \delta$ can be written in the form $f = \sum_{i=1}^{c(\delta,\mathbb{F},d)} G_i H_i + Q$, where G_i and H_is are nonconstant polynomials satisfying $\deg(G_i) + \deg(H_i) \leqslant d$, and Q is a degree $\leqslant d-1$ polynomial. Recently, Bhowmick and Lovett [3] have proved that one can remove the dependence of c on $|\mathbb{F}|$. However, in the low characteristic case, i.e. when $|\mathbb{F}|$ can be smaller than d, the only general results before this work are Theorems 2 and 3.

1.1 Our results

Suppose that $\mathbb{F} = \mathbb{F}_q$ is a prime field. In this work we are interested in the case when q is a fixed prime, as the case of large q is addressed by a recent work of [3]. When the characteristic of $\mathbb{F}$ can be small, namely when $|\mathbb{F}| \leqslant 5$, it was not known whether a degree five biased polynomial admits a strong structure in the sense of Theorems 2 and 3. Moreover, the techniques from [8] seem to break down.

Quintic polynomials

We combine ideas from [8] with arguments from polynomial regularity and prove such a structure theorem for quintic polynomials.

▶ **Theorem 4** (Biased quintic polynomials I). *Suppose $f : \mathbb{F}^n \to \mathbb{F}$ is a degree five polynomial with $\text{bias}(f) = \delta$. There exist $c_4 \leqslant c(\delta, |\mathbb{F}|)$, nonconstant polynomials $G_1, ..., G_c, H_1, ..., H_c$ and a polynomial Q such that the following holds.*

- $f = \sum_{i=1}^{c} G_i H_i + Q$.
- *For every $i \in [c]$, $\deg(G_i) + \deg(H_i) \leqslant 5$.*
- $\deg(Q) \leqslant 4$.

Note that c_4 has no dependence on n. We also prove that every biased quintic polynomial is constant on an affine subspace of dimension $\Omega(n)$.

▶ **Theorem 5** (Biased quintic polynomials II). *Suppose $f : \mathbb{F}^n \to \mathbb{F}$ is a degree five polynomial with $\text{bias}(f) = \delta$. There exists an affine subspace V of dimension $\Omega(n)$ such that $f|_V$ is constant.*

Theorem 5 was previously only known for degrees $\leqslant 4$. The case of quadratics when $\mathbb{F} = \mathbb{F}_2$ is Dickson's theorem [6], and the case of general $\mathbb{F}$ and $d \leqslant 4$ was proved recently by Cohen and Tal [5] building on Theorem 2 and Theorem 3. We also remark that the degree five is the largest degree that such a bound can hold. To see this, assume for example that $d = 6$ and $\mathbb{F} = \mathbb{F}_2$, and construct a degree 6 polynomial $f = G(x_1, ..., x_n) \cdot H(x_1, \ldots, x_n)$ by picking two random cubic polynomials G and H. One observes that f has large bias as $\mathbf{Pr}(f = 0) = 3/4 + o(1)$, however, f will not vanish over any subspace of dimension $\Omega(n^{1/2})$. Theorem 5 has the following immediate corollary.

▶ **Corollary 6.** *Suppose $f : \mathbb{F}^n \to \mathbb{F}$ is a degree five affine disperser for dimension k. Then f is also an affine extractor of dimension $O(k)$.*

We refer to [5] where affine dispersers and extractors and the relations between them are discussed.

Degree d polynomials, with $d < |\mathbb{F}| + 4$

We in fact prove a strong structure theorem for biased degree d polynomials when $d < |\mathbb{F}|+4$, from which Theorem 4 follows immediately.

▶ **Theorem 7** (Biased degree d polynomials I (when $d < |\mathbb{F}| + 4$))**.** *Suppose $0 < d$ and $\mathbb{F}$ is a prime field satisfying $d < |\mathbb{F}| + 4$. Let $f : \mathbb{F}^n \to \mathbb{F}$ be a degree d polynomial with $\mathrm{bias}(f) = \delta$. There exists $c_7 \leqslant c(\delta, d, |\mathbb{F}|)$, nonconstant polynomials $G_1, ..., G_c, H_1, ..., H_c$ and a polynomial Q such that the following hold.*

- $f = \sum_{i=1}^{c} G_i H_i + Q$.
- *For every $i \in [c]$,* $\deg(G_i) + \deg(H_i) \leqslant d$.
- $\deg(Q) \leqslant d - 1$.

We also prove a general version of Theorem 5 when $d < |\mathbb{F}| + 4$.

▶ **Theorem 8** (Biased degree d polynomials II (when $d < |\mathbb{F}| + 4$))**.** *Suppose $0 < d$ and $\mathbb{F}$ is a prime field satisfying $d < |\mathbb{F}| + 4$. Let $f : \mathbb{F}^n \to \mathbb{F}$ be a degree d polynomial with $\mathrm{bias}(f) = \delta$. There exists an affine subspace V of dimension $\Omega_{d,\delta}(n^{1/\lfloor \frac{d-2}{2} \rfloor})$ such that $f|_V$ is a constant.*

Cohen and Tal [5] recently showed that any degree d biased polynomial is constant on an $\Omega_\delta(n^{1/(d-1)})$ dimensional affine subspace. Theorem 8 improves on this by a quadratic factor, when $d < |\mathbb{F}| + 4$.

Our results for quintic polynomials follow immediately.

Proof of Theorem 4 and Theorem 5. Theorem 4 and Theorem 5 follow curiously as special cases of Theorem 7 and Theorem 8 as $|\mathbb{F}| \geqslant 2$ and $5 < 2 + 4$. ◀

Algorithmic aspects

Using a result of Bhattacharyya, et. al. [2] who gave an algorithm for finding prescribed decompositions of polynomials, we show that whenever such a strong structure exists, it can be found algorithmically in time polynomial in n. Combined with Theorem 7, we obtain the following algorithmic structure theorem.

▶ **Theorem 9.** *Suppose $\delta > 0$, $d > 0$ are given, and let $\mathbb{F}$ be a prime field satisfying $d < |\mathbb{F}|+4$. There is a deterministic algorithm that runs in time $O(n^{O(d)})$ and given as input a degree d polynomial $f : \mathbb{F}^n \to \mathbb{F}$ satisfying $\mathrm{bias}(f) = \delta$, outputs a number $c \leqslant c(\delta, |\mathbb{F}|, d)$, a collection of degree $\leqslant d - 1$ polynomials $G_1, ..., G_c, H_1, ..., H_c : \mathbb{F}^n \to \mathbb{F}$ and a polynomial $Q : \mathbb{F}^n \to \mathbb{F}$, such that*

- $f = \sum_{i=1}^{c} G_i H_i + Q$.
- *For every $i \in [c]$,* $\deg(G_i) + \deg(H_i) \leqslant d$.
- $\deg(Q) \leqslant d - 1$.

1.2 Overview of Proofs

As we saw above, Theorems 4 and 5 are immediate consequences of Theorems 7 and 8.

Proof overview of Theorem 7. Given a degree d biased polynomial $f : \mathbb{F}^n \to \mathbb{F}$, using an additive combinatorial argument we find a bounded index subspace restricted to which all the first degree partial derivatives of f are biased. We observe that since f was biased, it must be biased in some coset of this subspace also, and hence by a result of [11] it must be a function of a constant number of its derivatives. As each derivative of f is a degree $\leqslant d - 1$ and biased, we again invoke the "bias implies low-rank" result of [11] for these

lower-degree polynomials in order to rewrite f as a function of a constant number of degree $d-2$ polynomials. We finally show that under the assumption that $d < |\mathbb{F}| + 4$, we can "regularize" the resulting polynomials to a "regular" collection of polynomials from which the structure theorem can be deduced.

Proof overview of Theorem 8. Following the proof of Theorem 7, we pick an affine subspace W of constant codimension restricted to which f has the nice structure $f|_W = \sum G_i H_i + M$, where G_i, H_i, M are all of degrees $\leqslant d-2$. We moreover know that G_i, H_i, M are functions of a regular set of polynomials of degree $\leqslant d-2$. We argue by looking at the higher-order Fourier expansion of f that M must be a constant field element and since for each i, $\min\{\deg(G_i), \deg(H_i)\} \leqslant \lfloor \frac{d}{2} \rfloor$, using a result of Cohen and Tal [5] we can restrict to a subspace of dimension $\Omega(n^{1/\lfloor \frac{d-2}{2} \rfloor})$ making f a constant.

Organization

In Section 2 we present the basic tools from higher-order Fourier analysis. In Section 3 we discuss useful properties of a pseudorandom collection of polynomials. Theorem 7 is proved in Section 4.1, and Theorem 8 is proved in Section 4.2. We discuss the algorithmic aspects in Section 5. We end with a discussion of future directions in Section 6.

Notation

Let $\mathbb{D} = \{z \in \mathbb{C} : |z| \leqslant 1\}$ be the unit disk in the complex plane. Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Suppose that $\mathbb{F} = \mathbb{F}_q$ is a finite prime field, let $e_{\mathbb{F}} : \mathbb{F} \to \mathbb{D}$ denote the function $e_{\mathbb{F}}(x) := e^{\frac{2\pi i x}{|\mathbb{F}|}}$, and let $e : \mathbb{T} \to \mathbb{D}$ denote the function $e(x) := e^{2\pi i x}$. For functions $f, g : \mathbb{F}^n \to \mathbb{C}$, define

$$\langle f, g \rangle := \frac{1}{|\mathbb{F}|^n} \sum_{x \in \mathbb{F}^n} f(x)\overline{g(x)}.$$

For an integer a, denote by $[a] := \{1, \ldots, a\}$.

2 Preliminary results from higher-order Fourier analysis

2.1 Nonclassical Polynomials

Let $d \geqslant 0$ be an integer. It is well-known that for functions $P : \mathbb{F}^n \to \mathbb{F}$, a polynomial of degree $\leqslant d$ can be defined in two different ways. We say that P is a polynomial of degree $\leqslant d$ if it can be written as

$$P(x_1, ..., x_n) = \sum_{\substack{i_1, \ldots, i_n \geqslant 0 \\ i_1 + \cdots + i_n \leqslant d}} c_{i_1, \ldots, i_n} x_1^{i_1} \cdots x_n^{i_n},$$

with coefficients $c_{i_1, \ldots, i_n} \in \mathbb{F}$. This can be thought of as a global definition for polynomials in $\mathbb{F}[x_1, \ldots, x_n]$. The local definition of a polynomial uses the notion of additive directional derivatives.

▶ **Definition 10** (Polynomials over finite fields (local definition))**.** Suppose that G is an abelian group. For an integer $d \geqslant 0$, a function $P : \mathbb{F}^n \to G$ is said to be a *polynomial of degree* $\leqslant d$ if for all $y_1, \ldots, y_{d+1}, x \in \mathbb{F}^n$, it holds that

$$(D_{y_1} \cdots D_{y_{d+1}} P)(x) = 0,$$

where $D_yP(x) = P(x+y) - P(x)$ is the additive derivative of P with direction y evaluated at x. The *degree* of P is the smallest d for which the above holds.

It follows simply from the definition that for any direction $y \in \mathbb{F}^n$, $deg(D_yP) < deg(P)$. In the "classical" case of polynomials $P : \mathbb{F}^n \to \mathbb{F}$, i.e. $G = \mathbb{F}$, it is a well-known fact that the global and local definitions coincide. However, the situation is different when G is allowed to be other groups. For example when the range of P is $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, it turns out that the global definition must be refined to the "nonclassical polynomials". This phenomenon was noted by Tao and Ziegler [16] in the study of Gowers norms.

Nonclassical polynomials arise when studying functions $P : \mathbb{F}^n \to \mathbb{T}$ and their exponents $f = e(P) : \mathbb{F}^n \to \mathbb{C}$.

▶ **Definition 11** (Nonclassical Polynomials)**.** For an integer $d \geqslant 0$, a function $P : \mathbb{F}^n \to \mathbb{T}$ is said to be a *nonclassical polynomial of degree* $\leqslant d$ (or simply a *polynomial of degree* $\leqslant d$) if for all $y_1, \ldots, y_{d+1}, x \in \mathbb{F}^n$, it holds that

$$(D_{y_1} \cdots D_{y_{d+1}} P)(x) = 0. \tag{2}$$

The *degree* of P is the smallest d for which the above holds. A function $P : \mathbb{F}^n \to \mathbb{T}$ is said to be a *classical polynomial of degree* $\leqslant d$ if it is a nonclassical polynomial of degree $\leqslant d$ whose image is contained in $\frac{1}{q}\mathbb{Z}/\mathbb{Z}$.

Denote by $\mathrm{poly}(\mathbb{F}^n \to \mathbb{T})$, $\mathrm{poly}_d(\mathbb{F}^n \to \mathbb{T})$ and $\mathrm{poly}_{\leqslant d}(\mathbb{F}^n \to \mathbb{T})$, the set of all nonclassical polynomials over $\mathbb{F}^n$, all nonclassical polynomials of degree d and all nonclassical polynomials of degree $\leqslant d$ respectively.

From this point on by a polynomial we always mean a nonclassical polynomial, and we will make it clear when we talk about classical polynomials.

The following lemma of Tao and Ziegler [16] shows that a classical polynomial P of degree d must always be of the form $x \mapsto \frac{|Q(x)|}{q}$, where $Q : \mathbb{F}^n \to \mathbb{F}$ is a polynomial (in the usual sense) of degree d, and $|\cdot|$ is the standard map from $\mathbb{F}$ to $\{0, 1, \ldots, q-1\}$. This lemma also characterizes the structure of (nonclassical) polynomials.

▶ **Lemma 12** (Lemma 1.7 in [16])**.** *A function $P : \mathbb{F}^n \to \mathbb{T}$ is a polynomial of degree $\leqslant d$ if and only if P can be represented as*

$$P(x_1, \ldots, x_n) = \alpha + \sum_{\substack{0 \leqslant d_1, \ldots, d_n < q; k \geqslant 0: \\ 0 < \sum_i d_i \leqslant d - k(q-1)}} \frac{c_{d_1,\ldots,d_n,k} |x_1|^{d_1} \cdots |x_n|^{d_n}}{q^{k+1}} \mod 1,$$

for a unique choice of $c_{d_1,\ldots,d_n,k} \in \{0, 1, \ldots, q-1\}$ and $\alpha \in \mathbb{T}$. The element α is called the shift *of P, and the largest integer k such that there exist $d_1, \ldots, d_n$ for which $c_{d_1,\ldots,d_n,k} \neq 0$ is called the* depth *of P. A depth-k polynomial P takes values in an affine shift of the subgroup $\mathbb{U}_{k+1} := \frac{1}{q^{k+1}}\mathbb{Z}/\mathbb{Z}$. Classical polynomials correspond to polynomials with 0 shift and 0 depth.*

For convenience of exposition, henceforth we will assume that the shifts of all polynomials are zero. This can be done without affecting any of the results presented in this text. Under this assumption, all polynomials of depth k take values in $\mathbb{U}_{k+1}$.

2.2 Rank, Regularity, and Other Notions of Uniformity

The rank of a polynomial is a notion of its complexity according to lower degree polynomials.

▶ **Definition 13** (Rank of a polynomial)**.** Given a polynomial $P : \mathbb{F}^n \to \mathbb{T}$ and an integer $d \geqslant 1$, the *d-rank* of P, denoted $\mathsf{rank}_d(P)$, is defined to be the smallest integer r such that there exist polynomials $Q_1, \ldots, Q_r : \mathbb{F}^n \to \mathbb{T}$ of degree $\leqslant d-1$ and a function $\Gamma : \mathbb{T}^r \to \mathbb{T}$ satisfying $P(x) = \Gamma(Q_1(x), \ldots, Q_r(x))$. If $d = 1$, then 1-rank is defined to be ∞ if P is non-constant and 0 otherwise.

The *rank* of a polynomial $P : \mathbb{F}^n \to \mathbb{T}$ is its $\deg(P)$-rank. We say that P is r-regular if $\mathsf{rank}(P) \geqslant r$.

Note that for an integer $\lambda \in [1, q-1]$, $\mathrm{rank}(P) = \mathrm{rank}(\lambda P)$. In this article we are interested in obtaining a structure theorem for biased classical polynomials that does not involve nonclassical polynomials. Motivated by this, we define two other notions of rank.

▶ **Definition 14** (Classical rank of a polynomial)**.** Given a classical polynomial $P : \mathbb{F}^n \to \mathbb{F}$ and an integer $d \geqslant 1$, the classical d-rank of P, denoted by $\mathrm{crank}_d(P)$, is defined similarly to Definition 13 with the extra restriction that $Q_1, ..., Q_r : \mathbb{F}^n \to \mathbb{F}$ are classical polynomials.

The *classical rank* of a classical polynomial $P : \mathbb{F}^n \to \mathbb{F}$ is its classical $\deg(P)$-rank. We say that P is classical r-regular if $\mathsf{crank}(P) \geqslant r$.

▶ Remark. For a nonconstant affine-linear polynomial $P(x)$, $\mathrm{rank}(P) = \mathrm{crank}(P) = \infty$ and for a constant function $Q(x)$, $\mathrm{rank}(Q) = 0$.

▶ Remark. It is important to note that Definition 13 and Definition 14 are not equivalent. To see this, note that, as proved in [16] and [14], the degree 4 symmetric polynomial $S_4 := \sum_{i<j<k<\ell} x_i x_j x_k x_\ell$ has negligible correlation with any degree $\leqslant 3$ classical polynomial. A simple Fourier analytic argument implies that $\mathrm{crank}(S_4) = \omega(1)$, i.e. $\lim_{n\to\infty} \mathrm{crank}(S_4(x_1, ..., x_n)) = \infty$. However, S_4 turns out to have large *Gowers* U^4 *norm* and it follows by a theorem of Tao and Ziegler [16] that $\mathrm{rank}(S_4) \leqslant r(\mathbb{F})$ for some constant r.

In the above definitions of rank of a polynomial, we have allowed the function Γ to be arbitrary. It is interesting to ask whether a polynomial is structured in a stronger sense.

▶ **Definition 15** (Strong rank of a polynomial)**.** Given a classical polynomial $P : \mathbb{F}^n \to \mathbb{F}$ of degree d. The *strong rank* of P, denoted by $\text{strong-rank}_d(P)$, is the smallest $r \geqslant 0$, such that there exist nonconstant classical polynomials $G_1, ..., G_r, H_1, ..., H_r : \mathbb{F}^n \to \mathbb{F}$ and a classical polynomial Q such that

- $P(x) = \sum_{i=1}^{r} G_i H_i + Q$.
- For all $i \in [r]$, we have that $\deg(G_i) + \deg(H_i) \leqslant d$.
- $\deg(Q) \leqslant d-1$.

The *strong-rank* of a classical polynomial $P : \mathbb{F}^n \to \mathbb{F}$ is equal to $\text{strong-rank}_{\deg(P)}(P)$.

The above notion of rank is somewhat a stronger notion, in particular the following holds for any classical polynomial $P : \mathbb{F}^n \to \mathbb{F}$,

$$\mathrm{rank}(P) \leqslant \mathrm{crank}(P) \leqslant 2 \cdot \text{strong-rank}(P) + 1. \tag{3}$$

Due to the lack of multiplicative structure in $\frac{1}{p^k}\mathbb{Z}/\mathbb{Z}$ for $k > 1$, it is not clear how to define a similar structural notion to strong rank for nonclassical polynomials. Next, we will formalize the notion of a generic collection of polynomials. Intuitively, it should mean that there are no unexpected algebraic dependencies among the polynomials. First, we need to set up some notation.

▶ **Definition 16** (Factors)**.** If X is a finite set then by a *factor* $\mathcal{B}$ we simply mean a partition of X into finitely many pieces called *atoms*.

A finite collection of functions $\phi_1, \ldots, \phi_C$ from X to some other space Y naturally define a factor $\mathcal{B} = \mathcal{B}_{\phi_1,\ldots,\phi_C}$ whose atoms are sets of the form $\{x : (\phi_1(x), \ldots, \phi_C(x)) = (y_1, \ldots, y_C)\}$ for some $(y_1, \ldots, y_C) \in Y^C$. By an abuse of notation we also use $\mathcal{B}$ to denote the map $x \mapsto (\phi_1(x), \ldots, \phi_C(x))$, thus also identifying the atom containing x with $(\phi_1(x), \ldots, \phi_C(x))$.

▶ **Definition 17** (Polynomial factors)**.** If $P_1, \ldots, P_C : \mathbb{F}^n \to \mathbb{T}$ is a sequence of polynomials, then the factor $\mathcal{B}_{P_1,\ldots,P_C}$ is called a *polynomial factor*.

The *complexity* of $\mathcal{B}$, denoted $|\mathcal{B}| := C$, is the number of defining polynomials. The *degree* of $\mathcal{B}$ is the maximum degree among its defining polynomials $P_1, \ldots, P_C$. If $P_1, \ldots, P_C$ are of depths $k_1, \ldots, k_C$, respectively, then the number of atoms of $\mathcal{B}$ is at most $\prod_{i=1}^{C} q^{k_i+1}$ which we denote by $\|\mathcal{B}\|$.

The notions of rank discussed above can now be extended to quantify the structural complexity of a collection of (classical) polynomials.

▶ **Definition 18** (Rank, classical rank, and strong rank of a collection of polynomials)**.** A polynomial factor $\mathcal{B}$ defined by polynomials $P_1, \ldots, P_C : \mathbb{F}^n \to \mathbb{T}$ with respective depths $k_1, \ldots, k_C$ is said to have rank r if r is the least integer for which there exists $(\lambda_1, \ldots, \lambda_C) \in \mathbb{Z}^C$, with $(\lambda_1 \mod q^{k_1+1}, \ldots, \lambda_C \mod q^{k_C+1}) \neq 0^C$, such that $\operatorname{rank}_d(\sum_{i=1}^{C} \lambda_i P_i) \leqslant r$, where $d = \max_i \deg(\lambda_i P_i)$.

Given a collection of polynomials $\mathcal{P}$ and a function $r : \mathbb{N} \to \mathbb{N}$, we say that $\mathcal{P}$ is r-regular if $\mathcal{P}$ is of rank larger than $r(|\mathcal{P}|)$. We extend Definition 14 and Definition 15 to classical polynomial factors in a similar manner.

Notice that by the definition of rank, for a degree-d polynomial P of depth k we have

$$\operatorname{rank}(\{P\}) = \min\left\{\operatorname{rank}_d(P), \operatorname{rank}_{d-(q-1)}(qP), \ldots, \operatorname{rank}_{d-k(q-1)}(q^k P)\right\},$$

where $\{P\}$ is a polynomial factor consisting of one polynomial P.

In Section 3 we will see that regular collections of polynomials indeed do behave like a generic collection of polynomials in several manners.

Green and Tao [7] and Kaufman and Lovett [11] proved the following relation between bias and rank of a polynomial.

▶ **Theorem 19** ($d < |\mathbb{F}|$ [7], arbitrary $\mathbb{F}$ [11])**.** *For any $\varepsilon > 0$ and integer $d \geqslant 1$, there exists $r = r(d, \varepsilon, |\mathbb{F}|)$ such that the following is true. If $P : \mathbb{F}^n \to \mathbb{T}$ is a degree-d polynomial* $\operatorname{bias}(P) \geqslant \varepsilon$ *then crank*$(P) \leqslant r$.

More importantly, there are $y_1, \ldots, y_r \in \mathbb{F}^n$, and a function $\Gamma : \mathbb{F}^r \to \mathbb{F}$, such that

$$P = \Gamma(D_{y_1}P, \ldots, D_{y_r}P).$$

Kaufman and Lovett originally proved Theorem 19 for classical polynomials and classical rank. However, their proof extends to nonclassical polynomials without modification. Note that $r(d, \varepsilon, |\mathbb{F}|)$ does not depend on the dimension n. Motivated by Theorem 19 we define unbiasedness for polynomial factors.

▶ **Definition 20** (Unbiased collection of polynomials)**.** Let $\varepsilon : \mathbb{N} \to \mathbb{R}^+$ be a decreasing function. A polynomial factor $\mathcal{B}$ defined by polynomials $P_1, \ldots, P_C : \mathbb{F}^n \to \mathbb{T}$ with respective depths $k_1, \ldots, k_C$ is said to be ε-unbiased if for every collection $(\lambda_1, \ldots, \lambda_C) \in \mathbb{Z}^C$, with $(\lambda_1 \mod p^{k_1+1}, \ldots, \lambda_C \mod p^{k^C+1}) \neq 0^C$ it holds that

$$\left|\mathop{\mathbf{E}}_{x}\left[\mathsf{e}\left(\sum_i \lambda_i P_i(x)\right)\right]\right| < \varepsilon(|\mathcal{B}|).$$

2.3 Regularization of Polynomials

Due to the generic properties of regular factors, it is often useful to *refine* a collection of polynomials to a regular collection [16]. We will first formally define what we mean by refining a collection of polynomials.

▶ **Definition 21** (Refinement). A collection $\mathcal{P}'$ of polynomials is called a *refinement* of $\mathcal{P} = \{P_1, ..., P_m\}$, and denoted $\mathcal{B}' \succeq \mathcal{B}$, if the induced partition by $\mathcal{B}'$ is a combinatorial refinement of the partition induced by $\mathcal{B}$. In other words, if for every $x, y \in \mathbb{F}^n$, $\mathcal{B}'(x) = \mathcal{B}'(y)$ implies $\mathcal{B}(x) = \mathcal{B}(y)$.

Green and Tao [7], showed that given any nondecreasing function $r : \mathbb{N} \to \mathbb{N}$, any classical polynomial factor can be refined to an r classical-rank classical factor. The basic idea is simple; if some classical polynomial has low rank, decompose it to a few lower degree classical polynomials, and repeat. The formal proof uses a transfinite induction on the number of classical polynomials of each degree which defines the classical polynomial factor. The bounds on the number of classical polynomials obtained in the regularization process have Ackermann-type dependence on the degree d, even when the regularity parameter $r(\cdot)$ is a "reasonable" function. As such, it gives nontrivial results only for constant degrees. The extension of this regularity lemma to nonclassical polynomials is more involved, and was proved by Tao and Ziegler [16] as part of their proof of the inverse Gowers theorem.

▶ **Theorem 22** (Regularity lemma for (nonclassical) polynomials [16]). *Let $r : \mathbb{N} \to \mathbb{N}$ be a non-decreasing function and $d \geqslant 1$ be an integer. Then, there is a function $C_{\mathbb{F},r,d} : \mathbb{N} \to \mathbb{N}$ such that the following holds. Suppose $\mathcal{B}$ is a factor defined by polynomials $P_1, \ldots, P_C : \mathbb{F}^n \to \mathbb{T}$ of degree at most d. Then, there is an r-regular factor $\mathcal{B}'$ consisting of polynomials $Q_1, \ldots, Q_{C'} : \mathbb{F}^n \to \mathbb{T}$ of degree $\leqslant d$ such that $\mathcal{B}' \succeq \mathcal{B}$ and $C' \leqslant C_{22}^{(\mathbb{F},r,d)}(C)$.*

3 Properties of rank, crank, and strong-rank

A high-rank polynomial of degree d is, intuitively, a "generic" degree d polynomial; there are no unexpected ways to decompose it into lower degree polynomials. In this section we make precise this intuition.

Using a standard observation that relates the bias of a function to its distribution on its range, Theorem 19 implies that high-rank polynomials behave like independent random variables. See [1, 10] for further discussion of stronger equidistribution properties of high-rank polynomials.

Another way that high-rank polynomials behave like generic polynomials is that their restriction to subspaces preserves degree and high rank. We refer to [1, 3] for a proof.

▶ **Lemma 23** (Degree and rank preservation). *Suppose $f : \mathbb{F}^n \to \mathbb{T}$ is a polynomial of degree d and rank $\geqslant r$, where $r > q + 1$. Let A be a hyperplane in $\mathbb{F}^n$. Then, $f|_A$ is a polynomial of degree d and rank $\geqslant \max\{r - d - 1, r - |\mathbb{F}| - 1\}$, unless $d = 1$ and f is constant on A.*

The following is a surprising and very useful property of high-rank polynomials that was proved by Bhattacharyya, et. al. [1].

▶ **Lemma 24** (Degree preservation, Lemma 2.13 of [1]). *Let $d > 0$ be given. There exists a nondecreasing function $r_{d,\mathbb{F}} : \mathbb{N} \to \mathbb{N}$ such that the following holds. Let $\mathcal{B}$ be a rank $\geqslant r_{d,\mathbb{F}}$ polynomial factor defined by degree $\leqslant d$ polynomials $P_1, ..., P_m : \mathbb{F}^n \to \mathbb{T}$. Let $\Gamma : \mathbb{T}^n \to \mathbb{T}$. Then*

$$\deg(\Gamma(Q_1(x), ..., Q_m(x))) \leqslant \deg(\Gamma(P_1(x), ..., P_m(x))),$$

for every collection of polynomial $Q_1, ..., Q_m : \mathbb{F}^n \to \mathbb{T}$, *with* $\deg(Q_i) \leqslant \deg(P_i)$ *and* $\text{depth}(Q_i) \leqslant \text{depth}(P_i)$.

We prove a lemma relating the strong-rank of a polynomial to its strong-rank over constant codimensional affine subspaces.

▶ **Lemma 25.** *Let* $f : \mathbb{F}^n \to \mathbb{F}$ *be a degree* d *classical polynomial and* V *be an affine subspace of* $\mathbb{F}^n$ *of dimension* $n - t$. *Then,*

$$\textit{strong-rank}(f) \leqslant \textit{strong-rank}(f|_V) + t.$$

Proof. It suffices to prove that for a hyperplane W, strong-rank$(f) \leqslant$ strong-rank$(f|_V) + 1$. The lemma then simply follows by induction on t, the codimension of V.

Suppose $W = \{x \in \mathbb{F}^n | \sum_{i=1}^n w_i x_i = a\}$, where $w \in \mathbb{F}^n$ and $a \in \mathbb{F}$. Applying an affine invertible projection, we can assume without loss of generality that $w = (1, 0, \ldots, 0)$ and $a = 0$, and thus $W = \{x \in \mathbb{F}^n | x_1 = 0\}$. Assume that strong-rank$(f|_W) = r$, hence there exist nonconstant classical polynomials $G_1, ..., G_r, H_1, ..., H_r : W \to \mathbb{F}$ where $\deg(G_i)+\deg(H_i) \leqslant d$ and a degree $\leqslant d-1$ classical polynomial $Q : W \to \mathbb{F}$ such that

$$f|_W = \sum_{i=1}^{r} G_i H_i + Q.$$

Now note that,

$$f(x_1, ..., x_n) = f|_W(0, x_2, \ldots, x_n) + x_1 R(x_1, ..., x_n),$$

where $\deg(R) \leqslant d - 1$. Thus

$$f = x_1 R + \sum_{i=1}^{r} G_i H_i + Q,$$

equivalently strong-rank$(f) \leqslant r + 1$. ◀

The above lemma shows that high strong-rank classical polynomials are generic in a strong sense. We finally observe that all the discussed notions of rank are subadditive.

▶ **Claim 26.** *For every fixed vectors* $a, b \in \mathbb{F}^n$,

(i) $\textit{strong-rank}(D_{a+b}f) \leqslant \textit{strong-rank}(D_a f) + \textit{strong-rank}(D_b f)$.
(ii) $\textit{crank}(D_{a+b}f) \leqslant \textit{crank}(D_a f) + \textit{crank}(D_b f)$.
(iii) $\textit{rank}(D_{a+b}f) \leqslant \textit{rank}(D_a f) + \textit{rank}(D_b f)$.

Proof. We compute $D_{a+b}f(x)$,

$$D_{a+b}f(x) = D_b f(x + a) + D_a f(x).$$

The claim follows since $\tau(D_b f(x+a)) \leqslant \tau(D_b f(x))$ for any choice of $\tau \in \{\text{strong-rank}, \text{crank}, \text{rank}\}$, as the degrees of polynomials are preserved under affine shifts. ◀

4 Structure of biased polynomials

Throughout the paper we will assume $\mathbb{F} = \mathbb{F}_q$ is a fixed prime field.

We will need the following theorem of Sanders [15] on the structure of sets with small doubling.

▶ **Theorem 27** ([15]). *Suppose $A \subseteq \mathbb{F}^n$ satisfies $\frac{|A|}{|G|} \geqslant \alpha$. Then $A+A+A = \{a_1+a_2+a_3 | a_1, a_2, a_3 \in A\}$ contains an affine subspace of codimension at most $O_{|\mathbb{F}|,\alpha}(1)$.*

The following lemma states that for a function $f : \mathbb{F}^n \to \mathbb{F}$ to be biased, there must be a positive set of directions y for which $D_y f$ is somewhat biased.

▶ **Lemma 28.** *Suppose $f : \mathbb{F}^n \to \mathbb{F}$ is such that $\mathrm{bias}(f) = \delta$. Then there exists a set $A \subseteq \mathbb{F}^n$, with $|A| \geqslant \frac{\delta^2}{2}|\mathbb{F}|^n$ such that for every $y \in A$, $\mathrm{bias}(D_y f) \geqslant \frac{\delta^2}{2}$.*

Proof. We compute the average bias of $D_y f$ for $y \in \mathbb{F}^n$ uniformly at random.

$$\mathop{\mathbf{E}}_{y\in\mathbb{F}^n}[\mathrm{bias}(D_y f)] = \mathop{\mathbf{E}}_{y\in\mathbb{F}^n}\left[\left|\mathop{\mathbf{E}}_{x\in\mathbb{F}^n} e_{\mathbb{F}}(f(x+y)-f(x))\right|\right] \geqslant \left|\mathop{\mathbf{E}}_{z,x\in\mathbb{F}^n}[e_{\mathbb{F}}(f(z))e_{\mathbb{F}}(-f(x))]\right| = \delta^2. \tag{4}$$

Thus, since $\mathrm{bias}(f) \leqslant 1$, we get

$$\mathop{\mathbf{Pr}}_{y\in\mathbb{F}^n}\left[\mathrm{bias}(D_y f) \geqslant \frac{\delta^2}{2}\right] \geqslant \frac{\delta^2}{2}. \tag{5}$$

The lemma follows by choosing $A := \{y \in \mathbb{F}^n | \mathrm{bias}(D_y f) \geqslant \frac{\delta^2}{2}\} \subseteq \mathbb{F}^n$. ◀

We will use this lemma along with Theorem 27 and Claim 26 to show that for every biased function f there exists a not too small subspace restricted to which all the derivatives of f are biased.

4.1 Structure of biased polynomials I, when $d < |\mathbb{F}| + 4$

In this section we prove that biased degree d classical polynomials are strongly structured when $d < |\mathbb{F}| + 4$.

▶ **Theorem 7** (restated – Biased degree d polynomials I (when $d < |\mathbb{F}|+4$))). *Suppose $d > 0$ and $\mathbb{F}$ be a prime field satisfying $d < |\mathbb{F}|+4$. Let $f : \mathbb{F}^n \to \mathbb{F}$ be a degree d classical polynomial with $\mathrm{bias}(f) = \delta$. Then strong-rank$(f) \leqslant c(\delta, d, q)$, namely there exists $c_7 \leqslant c(\delta, d, q)$, nonconstant classical polynomials $G_1, ..., G_c, H_1, ..., H_c : \mathbb{F}^n \to \mathbb{F}$ and a classical polynomial $Q : \mathbb{F}^n \to \mathbb{F}$ such that the following hold.*

- $f = \sum_{i=1}^{c} G_i H_i + Q$.
- *For every $i \in [c]$, $\deg(G_i) + \deg(H_i) \leqslant d$.*
- $\deg(Q) \leqslant d-1$.

Note that c_7 does not depend on n.

Proof. By Lemma 28 there exists a set $A \subseteq \mathbb{F}^n$, with $|A| \geqslant \frac{\delta^2}{2}|\mathbb{F}|^n$ such that for every $y \in A$,

$$\mathrm{bias}(D_y f) \geqslant \frac{\delta^2}{2}.$$

Thus by Theorem 19 for every $y \in A$,

$$\mathrm{crank}(D_y f) \leqslant r = r_{19}(d, |\mathbb{F}|, \delta).$$

Applying Theorem 27, there is a subspace $V \subset \mathbb{F}^n$ of co-dimension $t = O_{\delta,|\mathbb{F}|}(1)$ and $h_0 \in \mathbb{F}^n$ such that $V + h_0 \subseteq A + A + A$. By Claim 26 (ii), since $V + h_0 \subseteq A + A + A$ we have that for every $y \in V$,

$$\mathrm{crank}(D_y f) \leqslant c_1 \leqslant 3r.$$

By a simple averaging argument, there is an affine shift of V, $W := V + h$ such that $\mathrm{bias}(f|_W) \geqslant \delta$. Let us denote $\widetilde{f} := f|_W$. By Lemma 25, it is sufficient to prove that strong-rank$(\widetilde{f}) \leqslant c_1(|\mathbb{F}|, \delta)$. Since $\mathrm{bias}(\widetilde{f}) \geqslant \delta$, Theorem 19 implies $\mathrm{crank}(\widetilde{f}) \leqslant r_0 = r_0(\delta, |\mathbb{F}|)$. Moreover, there are $y_1, \ldots, y_{r_0} \in W$ and a $\Gamma : \mathbb{F}^{r_0} \to \mathbb{F}$ such that

$$\widetilde{f} = \Gamma(D_{y_1}\widetilde{f}, \ldots, D_{y_{r_0}}\widetilde{f}). \tag{6}$$

Note that for all $i \in [r_0]$,

$$\mathrm{crank}_{d-1}(D_{y_i}\widetilde{f}) \leqslant \mathrm{crank}(D_{y_i}f) \leqslant c_0 \tag{7}$$

This is due to the fact that an affine transformation can only decrease the degrees of classical polynomials and thus it can only decrease the crank of classical polynomials.

▶ Remark. We point out that the subscript $d-1$ in the LHS of (7) is necessary, as can be seen by the following example. Suppose $d - 1 = 4$, $m > 0$ and $n = 3m + 4$. Let $Q = x_{n-3}x_{n-2}x_{n-1}x_n + \sum_{i=1}^{m} x_{3i-2}x_{3i-1}x_{3i}$. Now note that

$$\mathrm{crank}(Q) \leqslant 3,$$

while

- $\mathrm{crank}(Q|_{x_n=0}) = \mathrm{crank}(\sum_{i=1}^{m} x_{3i-2}x_{3i-1}x_{3i}) = \omega_n(1)$, since $\|e_{\mathbb{F}}(\sum_{i=1}^{m} x_{3i-2}x_{3i-1}x_{3i})\|_{U^3} = o(1)$.
- $\mathrm{crank}_4(Q|_{x_n=0}) = 1$, since $\deg(Q|_{x_n=0}) < 4$.

By (7) there exist degree $\leqslant d-2$ classical polynomials $\left\{G_1^{(i)}, \ldots, G_{c_0}^{(i)}\right\}_{i=1}^{r_0}$ and a function $\Lambda : \mathbb{F}^{r_0 c_0} \to \mathbb{F}$ such that

$$\widetilde{f} = \Lambda\left((G_1^{(i)}, \ldots, G_{c_0}^{(i)})_{i=1}^{r_0}\right). \tag{8}$$

We would like to regularize this collection of classical polynomials, however we would like to avoid any appearance of nonclassical polynomials. The following observation allows us to do exactly that as long as $d < |\mathbb{F}| + 4$.

▶ **Claim 29** (Nonclassical regularity lemma over large characteristic). *Let $r : \mathbb{N} \to \mathbb{N}$ be a non-decreasing function. And d be such that $d < |\mathbb{F}| + 4$. Then, there is a function $C_{29}^{\mathbb{F},r} : \mathbb{N} \to \mathbb{N}$ such that the following holds. Suppose $\mathcal{B}$ is a factor defined by classical polynomials $P_1, \ldots, P_C : \mathbb{F}^n \to \mathbb{T}$ of degree at most $d-2$. Then, there is an r-regular factor $\mathcal{B}'$ consisting only of classical polynomials $Q_1, \ldots, Q_{C'} : \mathbb{F}^n \to \mathbb{T}$ of degree $\leqslant d-2$ such that $\mathcal{B}' \succeq_{sem} \mathcal{B}$ and $C' \leqslant C_{29}^{(\mathbb{F},r)}(C)$.*

▶ Remark. Note that the above claim does not hold for general degrees, as we require the obtained factor be high-rank as defined in Definition 13, which is complexity against (nonclassical) polynomials. To see this, we observe that in the case of quartic classical polynomials, the single classical polynomial $\{S_4\}$ cannot be refined to a high-rank polynomial factor defined by $O(1)$ *classical* polynomials. However, it can be refined to a high-rank nonclassical factor by Theorem 22. This is the barrier to extending our results to sextic and higher-degree classical polynomials. Starting with a biased sextic classical polynomial, dealing with non-classical polynomials seems to be unavoidable.

We postpone the proof of Claim 29 and show how it can be used to conclude Theorem 4. Fix $r_1 : \mathbb{N} \to \mathbb{N}$ a nondecreasing function as in Lemma 24 for degree $d-2$. Let $\mathcal{B}$ be the factor defined by degree $\leqslant d-2$ classical polynomials $\{G_1^{(i)}, \ldots, G_{c_0}^{(i)}\}_{i=1}^{r_0}$. Applying Claim 29 to $\mathcal{B}$ with regularity parameter r_1, we obtain a refinement $\mathcal{B}' \succeq_{sem} \mathcal{B}$, where $\mathcal{B}'$ is defined by

$c_2 := C_{29}^{(\mathbb{F},r_1)}(c_0 r_0)$ classical degree $\leqslant d-2$ polynomials $R_1, \ldots, R_{c_2} : \mathbb{F}^n \to \mathbb{F}$. Namely, there exists a function $\mathcal{K} : \mathbb{F}^{c_2} \to \mathbb{F}$, such that

$$\widetilde{f} = \mathcal{K}(R_1, \ldots, R_{c_2}).$$

Applying an affine transformation, assume without loss of generality that $W = \{x \in \mathbb{F}^n | x_1 = x_2 = \cdots = x_t = 0\}$. Moreover, we may assume that $n - t > c_2$, since otherwise, $\widetilde{f}$ has at most $d(n-t)^d = O(c_2^d)$ monomials, making the theorem statement trivial. For every $i \in [c_2]$, let $d_i := \deg(R_i)$, $s_i := \sum_{j=1}^{i} d_i$, and define $R'_i := x_{s_{i-1}+1} \cdots x_{s_i}$. We have that $\deg(R'_i) = \deg(R_i)$ and thus by Lemma 24,

$$\deg(\mathcal{K}(R'_1, \ldots, R'_{c_2})) \leqslant \deg(\mathcal{K}(R_1, \ldots, R_{c_2})) = \deg(\widetilde{f}) = d.$$

Note that $\mathcal{K} : \mathbb{F}^{c_2} \to \mathbb{F}$ is a classical polynomial, and $R'_1, ..., R'_{c_2}$ are monomials on disjoint variables, thus plugging in R'_is into $\mathcal{K}$'s variables, no cancelations can occur. In particular,

$$\mathcal{K}(y_1, \ldots, y_{c_2}) = \sum_{s \in \{0,\ldots,q-1\}^{c_2}, \sum_i s_i d_i \leqslant d} \alpha_s \prod_{i \in S} y_i^{s_i},$$

where $\alpha_S \in \mathbb{F}$ are coefficients of $\mathcal{K}$. Hence,

$$\widetilde{f} = \mathcal{K}(R_1, \ldots, R_{c_2}) = \sum_{s \in \{0,\ldots,q-1\}^{c_2}, \sum_i s_i d_i \leqslant d} \alpha_s \prod_{i \in S} R_i^{s_i}. \tag{9}$$

Namely, strong-rank$(\widetilde{f}) \leqslant dc_2^d$, and by Lemma 25 we deduce strong-rank$(f) \leqslant dc_2^d + t$ as desired. ◀

Proof of Claim 29. We observe that the iterative proof of Theorem 22 can be modified to include only classical polynomials. Theorem 22 is proved by a transfinite induction on the vector of number of (possibly nonclassical) polynomials of each degree and depth defining the polynomial factor. One then argues that a polynomial factor that is not of the desired rank, can always be refined to a polynomial factor where some polynomial is replaced by a collection of polynomials that are of either lower degree, or same degree with lower depth.

We now make use of the fact that $d < |\mathbb{F}| + 4$. We observe that if we start with a polynomial factor defined by degree $\leqslant d-2$ classical polynomials, the only nonclassical polynomials that may arise are of degree $d - 3 \leqslant |\mathbb{F}|$ and thus of depth 1, this is due to the fact that any nonclassical polynomial of depth $\geqslant 2$ has degree $\geqslant 2|\mathbb{F}| - 1$. Now we use a known fact that polynomials of degree $|\mathbb{F}|$ that are not classical are unncessary in higher order Fourier analysis. More precisely in the inverse theorem for Gowers norms of [16] for the case of degree $|\mathbb{F}|$ polynomials, one can assume that the polynomial $P : \mathbb{F}^n \to \mathbb{T}$ in the statement of the theorem is a *classical* polynomial of degree at most $\leqslant |\mathbb{F}|$. More generally [9] showed a similar fact for higher depths.

▶ **Theorem 30** (Unnecessary depths [9])**.** *Let $k \geqslant 1$, and q the characteristic of $\mathbb{F}$. Every nonclassical polynomial $f : \mathbb{F}^n \to \mathbb{T}$ of degree $1 + k(q-1)$ and depth k, can be expressed as a function of three degree $\leqslant 1 + k(q-1)$ polynomials of depth $\leqslant k - 1$.*

By the above discussion we may assume that in our application of Theorem 22, $\mathcal{B}'$ is defined via only classical polynomials. ◀

4.2 Structure of biased polynomials II, when $d < |\mathbb{F}| + 4$

In this section we prove that a biased degree d classical polynomial is constant on a large subspace.

▶ **Theorem 8** (restated – Biased degree d polynomials II (when $d < |\mathbb{F}| + 4$))**.** *Suppose $d > 0$ and $\mathbb{F}$ be a prime field satisfying $d < |\mathbb{F}|+4$. Let $f : \mathbb{F}^n \to \mathbb{F}$ be a degree d classical polynomial with $\mathrm{bias}(f) = \delta$. There exists an affine subspace V of dimension $\Omega_{d,\delta}(n^{1/\lfloor \frac{d-2}{2} \rfloor})$ such that $f|_V$ is a constant.*

In the case of $d = 5$ we have $5 < 2 + 4 \leqslant |\mathbb{F}| + 4$ and $\lfloor (d-2)/2 \rfloor = 1$, hence we obtain a subspace of dimension $\Omega_\delta(n)$ as desired in Theorem 5.

We will need the following result of Cohen and Tal [5] on the structure of low degree polynomials.

▶ **Theorem 31** ([5], Theorem 3.5)**.** *Let q be a prime power. Let $f_1, \ldots, f_\ell : \mathbb{F}_q^n \to \mathbb{F}_q$ be (classical) polynomials of degree $d_1, \ldots, d_\ell$ respectively. Let k be the least integer such that*

$$n \leqslant k + \sum_{j=0}^{\ell} (d_i + 1) \sum_{j=0}^{d_i - 1} (d_i - j) \cdot \binom{k+j-1}{j}.$$

Then, for every $u_0 \in \mathbb{F}_q^n$ there exists a subspace $U \subseteq \mathbb{F}_q^n$ of dimension k, such that for all $i \in [\ell]$, f_i restricted to $u_0 + U$ is a constant function.

In particular, if $d_1, ..., d_\ell \leqslant d$, then the above holds for $k = \Omega((n/\ell)^{\frac{1}{d-1}})$.

Proof of Theorem 8. Following the proof of Theorem 7, there exists an affine subspace W of dimension $n - t$ for $t = poly(\log(\frac{1}{\delta^2}))$, for which (9) holds. By Theorem 19, choosing a proper regularity parameter in the application of Claim 29, we can further assume that the factor defined by $R_1, ..., R_{c_2}$ is $\frac{\delta}{2} q^{-c_2}$-unbiased in the sense of Definition 20. We may rewrite (9) in the form

$$f|_W = \sum_{i=1}^{C} \alpha_i G_i H_i + M,$$

where $C \leqslant c_2^d$, α_i are field elements, M is a degree $\leqslant d-2$ classical polynomial, G_is and H_is are nonconstant degree $\leqslant d-2$ classical polynomials satisfying $\deg(G_i) + \deg(H_i) \leqslant d$. Moreover, every G_i and H_i is product of a subset of $\{R_1, ..., R_{c_2}\}$. We crucially observe that M can be taken to be of the form

$$M = \sigma_0 + \sum_{i=1}^{c_2} \sigma_i R_i,$$

where σ_i are field elements, such that $\sigma_i \neq 0$ implies that R_i does not appear in $\sum_{i=1}^{C} \alpha_i G_i H_i$.

▶ **Claim 32.** *Let f, W, $R_1, ...R_{c_2}$ and M be as above. Then M is a constant.*

Proof. Assume for contradiction that M is nonconstant. By the above discussion, letting

$$S := \{j \in [c_2] : R_j \text{ appears in } \sum_i \alpha_i G_i H_i\},$$

we have

$$f|_W = \Lambda(R_j)_{j \in S} + \sum_{i \in [c_2] \setminus S} \sigma_j R_j,$$

for some function $\Lambda : \mathbb{F}^{|S|} \to \mathbb{F}$. Writing the Fourier expansion of $e_{\mathbb{F}}(\Lambda)$, we have

$$e_{\mathbb{F}}(f|_W) = \sum_{\gamma \in \mathbb{F}^{|\S|}} \widehat{\Lambda}(\gamma) e_{\mathbb{F}}(\sum_{j \in S} \gamma_j R_j + M).$$

Note that W was chosen such that $\mathrm{bias}(f|_W) \geqslant \delta$. Thus,

$$\begin{aligned}
\mathrm{bias}(f|_W) &= |\mathop{\mathbf{E}}_{x \in \mathbb{F}^n} e_{\mathbb{F}}(\Lambda(R_j)_{j \in S} + M)| \\
&= |\mathop{\mathbf{E}}_{x \in \mathbb{F}^n} \sum_{\gamma \in \mathbb{F}^{|S|}} \widehat{\Lambda}(\gamma) e_{\mathbb{F}}(M + \sum_{j \in S} \gamma_j R_j)| \\
&\leqslant \sum_{\gamma \in \mathbb{F}^{|S|}} |\widehat{\Lambda}(\gamma)| \cdot \mathrm{bias}(M + \sum_{j \in S} \gamma_j R_j) \\
&\leqslant q^{c_2} \cdot \frac{\delta}{2} q^{-c_2} < \delta,
\end{aligned}$$

contradicting $\mathrm{bias}(f|_W) = \delta$, where the last inequality uses the fact that the factor defined by $R_1, ..., R_{c_2}$ is $\frac{\delta}{2} q^{-c_2}$-unbiased. ◀

By the above claim M is a constant, and thus

$$f|_W = \sigma_0 + \sum_{i=1}^{C} \alpha_i G_i H_i.$$

Recall that $\deg(G_i) + \deg(H_i) \leqslant d$, hence for every i, $\min\{\deg(G_i), \deg(H_i)\} \leqslant \lfloor \frac{d}{2} \rfloor$. Thus by Theorem 31, there is an $\Omega_C((n-t)^{1/\lfloor \frac{d-2}{2} \rfloor}) = \Omega_{\delta,\mathbb{F},d}(n^{1/\lfloor \frac{d-2}{2} \rfloor})$ dimensional affine subspace W' such that $f|_{W'}$ is constant. ◀

5 Algorithmic Aspects

In this section we show that the strong structures implied by Theorem 4 and Theorem 7 can be found by a deterministic algorithm that runs in time polynomial in n.

▶ **Theorem 9** (restated). *Suppose $\delta > 0$, $d > 0$ are given, and let $\mathbb{F}$ be a prime field satisfying $d < |\mathbb{F}| + 4$. There is a deterministic algorithm that runs in time $O(n^{O(d)})$ and given as input a degree d classical polynomial $f : \mathbb{F}^n \to \mathbb{F}$ satisfying $\mathrm{bias}(f) = \delta$, outputs a number $c \leqslant c(\delta, |\mathbb{F}|, d)$, a collection of degree $\leqslant d-1$ classical polynomials $G_1, ..., G_c, H_1, ..., H_c : \mathbb{F}^n \to \mathbb{F}$ and a classical polynomial $Q : \mathbb{F}^n \to \mathbb{F}$, such that*

- $f = \sum_{i=1}^{c} G_i H_i + Q$.
- *For every $i \in [c]$, $\deg(G_i) + \deg(H_i) \leqslant d$.*
- $\deg(Q) \leqslant d - 1$.

Proof. We will use the following result of Bhattacharyya, et. al. [2] who proved several algorithmic regularity lemmas for polynomials.

▶ **Theorem 33** ([2], Theorem 1.6). *For every finite field $\mathcal{F}$ of fixed prime order, positive integers d, k, every vector of positive integers $\Delta = (\Delta_1, ..., \Delta_k)$, and every function $\Gamma : \mathbb{F}^k \to \mathbb{F}$, there is a deterministic algorithm that takes as input a classical polynomial $f : \mathbb{F}^n \to \mathbb{F}$ of degree d, runs in time polynomial in n, and outputs classical polynomials $Q_1, ..., Q_k$ of degrees respectively at most $\Delta_1, ..., \Delta_k$ such that*

$$f = \Gamma(Q_1, ..., Q_k),$$

if such a decomposition exists, while otherwise accurately returning NO.

By Theorem 7, we know that there is $c \leqslant C(\delta, |\mathbb{F}|, d)$ such that there exist a collection of nonconstant classical polynomials $G_1, ..., G_c, H_1, ..., H_c : \mathbb{F}^n \to \mathbb{F}$, and a classical polynomial $Q : \mathbb{F}^n \to \mathbb{F}$, such that

$$f = \sum_{i=1}^{c} G_i H_i + Q, \tag{10}$$

for every $i \in [c]$, $\deg(G_i) + \deg(H_i) \leqslant d$, and $\deg(Q) \leqslant d - 1$. The algorithm is now straight-forward.

1 Iterate through all choices for $c \leqslant C(\delta, |\mathbb{F}|, d)$. This is our guess for the number of terms in the summation in (10).

1.1 Iterate through all choices of $d_1, \ldots, d_c, d'_1, \ldots, d'_c \leqslant d-1$ and $d'' \leqslant d-1$ such that $d_i + d'_i \leqslant d$. These are our guesses for degree sequences for $G_1, ..., G_c, H_1, ..., H_c$ and Q. Note that this step does not depend on n.

1.1.1 Define $\Gamma : \mathbb{F}^{2c+1} \to \mathbb{F}$ as

$$\Gamma(x_1, \ldots, x_c, y_1, \ldots, y_c, z) := \sum_{i=1}^{c} x_i y_i + z.$$

1.1.2 Run Theorem 33 on the classical polynomial f, with $\Delta = (d_1, \ldots, d_c, d'_1, \ldots, d'_c, d'')$ and Γ as inputs.

1.1.2.a If the algorithm outputs NO, then continue.

1.1.2.b If the algorithm outputs a collection of classical polynomials satisfying the decomposition, halt and output the desired decomposition.

By Theorem 7 and Theorem 33 the above algorithm will always halt with a decomposition of desired form. The number of possible choices in **1** and **1.1** do not depend on n, and step **1.1.2** runs in polynomial time in n, as a result making the algorithm polynomial time in n. ◀

6 Conclusions

Green and Tao [7] and Kaufman and Lovett [11] proved that every degree d classical polynomial f with $\text{bias}(f) = \delta$ can be written in the form

$$f = \Gamma(P_1, ..., P_c), \tag{11}$$

for $c \leqslant c(\delta, d, \mathbb{F})$ and degree $\leqslant d-1$ classical polynomials $P_1, ..., P_c$. However, nothing is known on the structure of the function Γ in (11). In this work we showed that in the case of degree five polynomials we can say more about the structure of f. More generally for degree d classical polynomials when $d < |\mathbb{F}| + 4$, we can write

$$f = \sum_{i=1}^{C} G_i H_i + Q,$$

for nontrivial classical polynomials G_i, H_i satisfying $\deg(G_i) + \deg(H_i) \leqslant d$, and $\deg(Q) \leqslant d-1$. It is a fascinating question whether similar structure theorems hold in the case of $d \geqslant |\mathbb{F}| + 4$, more specifically we suspect that answering this question for degree 6 classical polynomials and $\mathbb{F} = \mathbb{F}_2$ will suffice resolve the question for all degrees and characteristics.

▶ **Open Problem 34.** *Can every biased degree six classical polynomial* $f : \mathbb{F}_2^n \to \mathbb{F}_2$ *be written in the form*

$$f = \sum_{i=1}^{C} G_i H_i + Q,$$

for $C \leqslant C(\mathrm{bias}(f))$*, nontrivial classical polynomials* Q*,* G_i, H_i *satisfying* $\deg(G_i)+\deg(H_i) \leqslant 6$*, and* $\deg(Q) \leqslant 5$*?*

A somewhat weaker question that also remains open is whether we can bound the degree of Γ in (11) in terms of d only.

▶ **Open Problem 35.** *Suppose that* $\mathbb{F}$ *is a prime field. Can every degree* d *classical polynomial* $f : \mathbb{F}^n \to \mathbb{F}$ *be written in the form*

$$f = \Gamma(P_1, ..., P_{C_1}),$$

where $C \leqslant C(\mathrm{bias}(f), \mathbb{F}, d)$*,* $P_1, ..., P_C$ *are degree* $\leqslant d-1$ *classical polynomials, and* $\deg(\Gamma) \leqslant O_d(1)$?

Finally, we note that the constants obtained in Theorems 4, 5, 7 and 8, unlike Theorems 2 and 3, have very bad dependence on δ and d. In particular, in the case of degree five polynomials, an interesting problem that remains unaddressed is to find out what the optimum constant achievable in Theorem 4 is.

Acknowledgements. We thank Avishay Tal and Avi Wigderson for helpful discussions. We also thank the referees for their useful comments.

References

1 Arnab Bhattacharyya, Eldar Fischer, Hamed Hatami, Pooya Hatami, and Shachar Lovett. Every locally characterized affine-invariant property is testable. In *Proceedings of the 45th annual ACM symposium on Symposium on theory of computing*, STOC'13, pages 429–436, New York, NY, USA, 2013. ACM. doi:10.1145/2488608.2488662.

2 Arnab Bhattacharyya, Pooya Hatami, and Madhur Tulsiani. Algorithmic regularity for polynomials and applications. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 1870–1889, 2015. doi:10.1137/1.9781611973730.125.

3 Abhishek Bhowmick and Shachar Lovett. Bias vs structure of polynomials in large fields, and applications in effective algebraic geometry and coding theory. *CoRR*, abs/1506.02047, 2015. URL: http://arxiv.org/abs/1506.02047.

4 Abhishek Bhowmick and Shachar Lovett. The list decoding radius of reed-muller codes over small fields. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC'15, pages 277–285, New York, NY, USA, 2015. ACM. doi:10.1145/2746539.2746543.

5 Gil Cohen and Avishay Tal. Two structural results for low degree polynomials and applications. *RANDOM*, 2015.

6 Leonard Eugene Dickson. *Linear groups: With an exposition of the Galois field theory.* with an introduction by W. Magnus. Dover Publications, Inc., New York, 1958.

7 Ben Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms. *Contrib. Discrete Math.*, 4(2):1–36, 2009.

8 Elad Haramaty and Amir Shpilka. On the structure of cubic and quartic polynomials. In *STOC'10 – Proceedings of the 2010 ACM International Symposium on Theory of Computing*, pages 331–340. ACM, New York, 2010.

9 Hamed Hatami, Pooya Hatami, and James Hirst. Limits of Boolean functions on $\mathbb{F}_p^n$. *Electron. J. Combin.*, 21(4):Paper 4.2, 15, 2014.

10 Hamed Hatami, Pooya Hatami, and Shachar Lovett. General systems of linear forms: equidistribution and true complexity. *Advances in Mathematics*, 292:446–477, 2016.

11 Tali Kaufman and Shachar Lovett. Worst case to average case reductions for polynomials. *Foundations of Computer Science, IEEE Annual Symposium on*, 0:166–175, 2008. doi:10.1109/FOCS.2008.17.

12 Tali Kaufman, Shachar Lovett, and Ely Porat. Weight distribution and list-decoding size of reed–muller codes. *Information Theory, IEEE Transactions on*, 58(5):2689–2696, 2012.

13 Rudolf Lidl and Harald Niederreiter. *Introduction to finite fields and their applications*. Cambridge university press, 1994.

14 Shachar Lovett, Roy Meshulam, and Alex Samorodnitsky. Inverse conjecture for the Gowers norm is false. *Theory Comput.*, 7:131–145, 2011. doi:10.4086/toc.2011.v007a009.

15 Tom Sanders. Additive structures in sumsets. *Math. Proc. Cambridge Philos. Soc.*, 144(2):289–316, 2008. doi:10.1017/S030500410700093X.

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