

# Using higher-order Fourier analysis over general fields

Arnab Bhattacharyya\*  
 Department of Computer Science & Automation  
 Indian Institute of Science  
 arnabb@csa.iisc.ernet.in

Abhishek Bhowmick†  
 Department of Computer Science  
 The University of Texas at Austin  
 bhowmick@cs.utexas.edu

May 4, 2015

## Abstract

Higher-order Fourier analysis, developed over prime fields, has been recently used in different areas of computer science, including list decoding, algorithmic decomposition and testing. We extend the tools of higher-order Fourier analysis to analyze functions over general fields. Using these new tools, we revisit the results in the above areas.

- (i) For any fixed finite field  $\mathbb{K}$ , we show that the list decoding radius of the generalized Reed Muller code over  $\mathbb{K}$  equals the minimum distance of the code. Previously, this had been proved over prime fields [BL14] and for the case when  $|\mathbb{K}| - 1$  divides the order of the code [GKZ08].
- (ii) For any fixed finite field  $\mathbb{K}$ , we give a polynomial time algorithm to decide whether a given polynomial  $P : \mathbb{K}^n \rightarrow \mathbb{K}$  can be decomposed as a particular composition of lesser degree polynomials. This had been previously established over prime fields [Bha14, BHT15].
- (iii) For any fixed finite field  $\mathbb{K}$ , we prove that all locally characterized affine-invariant properties of functions  $f : \mathbb{K}^n \rightarrow \mathbb{K}$  are testable with one-sided error. The same result was known when  $\mathbb{K}$  is prime [BFH<sup>+</sup>13] and when the property is linear [KS08]. Moreover, we show that for any fixed finite field  $\mathbb{F}$ , an affine-invariant property of functions  $f : \mathbb{K}^n \rightarrow \mathbb{F}$ , where  $\mathbb{K}$  is a growing field extension over  $\mathbb{F}$ , is testable if it is locally characterized by constraints of bounded weight.

## 1 Introduction

Fourier analysis over finite groups has played a central role in the development of theoretical computer science. Examples of its applications are everywhere: analysis of random walks on graphs [CDG87], fast integer multiplication algorithms [SS71], learning algorithms [KM93a], the Kahn-Kalai-Linial theorem [KKL88], derandomization [NN93], tight inapproximability results using probabilistically checkable proofs [Has01], social choice theory [MOO10], and coding theory [NS05]. See the surveys of De Wolf [dW08] and Štefankovič [Šte00].

Higher-order Fourier analysis is a recent generalization of some aspects of Fourier analysis. Consider functions over the integers  $\mathbb{Z}$ . While classical Fourier analysis over  $\mathbb{Z}$  studies correlations of functions with linear phases  $e^{i\theta n}$ , higher-order Fourier analysis over  $\mathbb{Z}$  analyzes the correlation of

\*Supported in part by a DST Ramanujan Fellowship.

†Research supported in part by NSF Grant CCF-1218723.

functions with polynomial phases such as  $e^{i\theta n^2}$ . The modern<sup>1</sup> work on higher-order Fourier analysis over  $\mathbb{Z}$  began with the spectacular proof by Gowers of Szemerédi’s theorem [Gow98, Gow01], where the *Gowers norm* was introduced, and with the ergodic theory work of Host and Kra [HK05]. Subsequently, Green, Tao and Ziegler through several works [GT08, GT10, GTZ11, GTZ12] largely completed the research program of understanding the relationships between different aspects of the theory over  $\mathbb{Z}$ . This work was applied to solve several longstanding open problems in additive number theory, including the celebrated result showing the existence of arbitrarily long arithmetic progressions in the primes [GT10]. The book [Tao12] by Tao on the subject surveys the current state of knowledge.

In an influential article [Gre05], Green popularized the idea that it is useful to rephrase the problems arising in additive number theory into problems on vector spaces over fixed finite fields. The motivation was that many of the techniques in higher-order Fourier analysis over  $\mathbb{Z}$  simplify over finite fields, because of the presence of subspaces and of algebraic notions such as orthogonality and linear independence. However, it was soon realized that these questions over finite fields are also intrinsically interesting because of their connections to theoretical computer science. In particular, the Gowers norm for functions on  $\mathbb{F}^n$  for a finite prime field  $\mathbb{F}$  is directly related to low-degree testing, a problem intensely studied by computer scientists since the early 90’s.

Thanks to the sequence of works [GT09, KL08, TZ10, BTZ10, TZ12], the apparatus of higher-order Fourier analysis over  $\mathbb{F}^n$  for any fixed prime order field  $\mathbb{F}$  is also now largely complete. The theory has subsequently found several interesting applications in computer science that we detail below and has become part of the mainstream theorist toolkit. However, in all of these applications, the finite field in consideration was restricted to be a field of *prime* order, while the problems themselves are interesting over general finite fields. In this work, we show how the techniques of higher-order Fourier analysis continue to apply even when the underlying field is a non-trivial extension of a prime order field.

## 1.1 Applications

In this section, we describe three different problems involving a finite field  $\mathbb{K}$ , which previously had been solved only when  $|\mathbb{K}|$  was prime but which we can now solve for arbitrary finite  $\mathbb{K}$ .

Throughout, let  $\mathbb{F}$  be a fixed prime order field, and let  $\mathbb{K}$  be a finite field that extends  $\mathbb{F}$ . Let  $q = |\mathbb{K}|$ ,  $p = |\mathbb{F}|$  and  $q = p^r$  for  $r > 0$ .

### 1.1.1 List-decoding Reed-Muller codes

The notion of *list decoding* was introduced by Elias [Eli57] and Wozencraft [Woz58] to decode *error correcting codes* beyond half the minimum distance. The goal of a list decoding algorithm is to produce all the codewords within a specified distance from the received word. At the same time one has to find the right radius for which the number of such codewords is small, otherwise there is no hope for the algorithm to be efficient. After the seminal results of Goldreich and Levin [GL89] and Sudan [Sud97] which gave list decoding algorithms for the Hadamard code and the Reed-Solomon code respectively, there has been tremendous progress in designing list decodable codes. See the survey by Guruswami [Gur06, Gur04] and Sudan [Sud00].

List decoding has applications in many areas of computer science including hardness amplification in complexity theory [STV01, Tre03], derandomization [Vad12], construction of hard core

---

<sup>1</sup>In retrospect, Weyl’s results on equidistribution of polynomial phases [Wey14] laid the foundations of this theory.

predicates from one way functions [GL89, AGS03], construction of extractors and pseudorandom generators [TSZS01, SU05] and computational learning [KM93b, Jac97]. However, the largest radius up to which list decoding is tractable is still a fundamental open problem even for well studied codes like Reed-Solomon (univariate polynomials) and Reed-Muller codes (multivariate polynomials). The goal of this work is to analyse Reed-Muller codes over small fields (possibly non prime) and small degree.

Reed-Muller codes (RM codes) were discovered by Muller in 1954. Let  $d \in \mathbb{N}$ . The RM code  $\text{RM}_{\mathbb{K}}(n, d)$  is defined as follows. The message space consists of degree  $\leq d$  polynomials in  $n$  variables over  $\mathbb{K}$  and the codewords are evaluation of these polynomials on  $\mathbb{K}^n$ . Let  $\delta_q(d)$  denote the normalized distance of  $\text{RM}_{\mathbb{K}}(n, d)$ . Let  $d = a(q - 1) + b$  where  $0 \leq b < q - 1$ . We have

$$\delta_{\mathbb{K}}(d) = \frac{1}{q^a} \left( 1 - \frac{b}{q} \right).$$

RM codes are one of the most well studied error correcting codes. Many applications in computer science involve low degree polynomials over small fields, namely RM codes. Given a received word  $g : \mathbb{K}^n \rightarrow \mathbb{K}$  the objective is to output the list of codewords (e.g. low-degree polynomials) that lie within some distance of  $g$ . Typically we will be interested in regimes where list size is either independent of  $n$  or polynomial in the block length  $q^n$ .

Let  $\mathcal{P}_d(\mathbb{K}^n)$  denote the class of degree  $\leq d$  polynomials  $f : \mathbb{F}^n \rightarrow \mathbb{F}$ . Let  $\text{dist}$  denote the normalized Hamming distance. For  $\text{RM}_{\mathbb{K}}(n, d)$ ,  $\eta > 0$ , let

$$\ell_{\mathbb{F}}(n, d, \eta) := \max_{g: \mathbb{F}^n \rightarrow \mathbb{F}} |\{f \in \mathcal{P}_d(\mathbb{F}^n) : \text{dist}(f, g) \leq \eta\}|.$$

Let  $\text{LDR}_{\mathbb{K}}(n, d)$  (short for *list decoding radius*) be the maximum  $\rho$  for which  $\ell_{\mathbb{K}}(n, d, \rho - \varepsilon)$  is upper bounded by a constant depending only on  $\varepsilon, |\mathbb{K}|, d$  for all  $\varepsilon > 0$ .

It is easy to see that  $\text{LDR}_{\mathbb{K}}(n, d) \leq \delta_{\mathbb{K}}(d)$ . The difficulty lies in proving a matching lower bound. We review some previous work next. The first breakthrough result was the celebrated work of Goldreich and Levin [GL89] who showed that in the setting of  $d = 1$  over  $\mathbb{F}_2$  (Hadamard Codes)  $\text{LDR}_{\mathbb{F}_2}(n, 1) = \delta_{\mathbb{F}_2}(1) = 1/2$ . Later, Goldreich, Rubinfeld and Sudan [GRS00] generalized the field to obtain  $\text{LDR}_{\mathbb{K}}(n, 1) = \delta_{\mathbb{K}}(1) = 1 - 1/|\mathbb{K}|$ . In the setting of  $d < |\mathbb{K}|$ , Sudan, Trevisan and Vadhan [STV01] showed that  $\text{LDR}_{\mathbb{K}}(n, d) \geq 1 - \sqrt{2d/|\mathbb{K}|}$  improving previous work by Arora and Sudan [AS03], Goldreich *et al* [GRS00] and Pellikaan and Wu [PW04]. Note that this falls short of the upper bound which is  $\delta_{\mathbb{K}}(d)$ .

In 2008, Gopalan, Klivans and Zuckerman [GKZ08] showed that  $\text{LDR}_{\mathbb{F}_2}(n, d) = \delta_{\mathbb{F}_2}(d)$ . They posed the following conjecture.

**Conjecture 1.1** ([GKZ08]). *For fixed  $d$  and finite field  $\mathbb{K}$ ,  $\text{LDR}_{\mathbb{K}}(n, d) = \delta_{\mathbb{K}}(d)$ .*

It is believed [GKZ08, Gop10] that the hardest case is the setting of small  $d$ . An important step in this direction was taken in [Gop10] that considered quadratic polynomials and showed that  $\text{LDR}_{\mathbb{K}}(n, 2) = \delta_{\mathbb{K}}(2)$  for all fields  $\mathbb{K}$  and thus proved the conjecture for  $d = 2$ . Recently, Bhowmick and Lovett [BL14] resolved the conjecture for prime  $\mathbb{K}$ .

Our main result for list decoding is a resolution of Conjecture 1.1.

**Theorem 1.1.** *Let  $\mathbb{K}$  be a finite field. Let  $\varepsilon > 0$  and  $d, n \in \mathbb{N}$ . Then,*

$$\ell_{\mathbb{K}}(d, n, \delta_{\mathbb{K}}(d) - \varepsilon) \leq c_{|\mathbb{K}|, d, \varepsilon}.$$

Thus,

$$\text{LDR}_{\mathbb{K}}(n, d) = \delta_{\mathbb{K}}(d).$$

**Remark 1.2** (Algorithmic Implications). *Using the blackbox reduction of algorithmic list decoding to combinatorial list decoding in [GKZ08] along with Theorem 1.1, for fixed finite fields,  $d$  and  $\varepsilon > 0$ , we now have list decoding algorithms in both the global setting (running time polynomial in  $|\mathbb{K}|^n$ ) and the local setting (running time polynomial in  $n^d$ ).*

### 1.1.2 Algorithmic polynomial decomposition

Consider the following family of properties of functions over a finite field  $\mathbb{K}$ .

**Definition 1.3.** *Given a positive integer  $k$ , a vector of positive integers  $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_k)$  and a function  $\Gamma : \mathbb{K}^k \rightarrow \mathbb{K}$ , we say that a function  $P : \mathbb{K}^n \rightarrow \mathbb{K}$  is  $(k, \Delta, \Gamma)$ -structured if there exist polynomials  $P_1, P_2, \dots, P_k : \mathbb{K}^n \rightarrow \mathbb{K}$  with each  $\deg(P_i) \leq \Delta_i$  such that for all  $x \in \mathbb{K}^n$ ,*

$$P(x) = \Gamma(P_1(x), P_2(x), \dots, P_k(x)).$$

*The polynomials  $P_1, \dots, P_k$  are said to form a  $(k, \Delta, \Gamma)$ -decomposition.*

For instance, an  $n$ -variate polynomial over the field  $\mathbb{K}$  of total degree  $d$  factors nontrivially exactly when it is  $(2, (d-1, d-1), \text{prod})$ -structured where  $\text{prod}(a, b) = a \cdot b$ . We shall use the term *degree-structural property* to refer to a property from the family of  $(k, \Delta, \Gamma)$ -structured properties.

The problem here is, for arbitrary fixed  $k, \mathbb{K}, (\Delta), \Gamma$ , given a polynomial, decide efficiently if it is degree structural and if yes, output the decomposition. An efficient algorithm for the above would imply a (deterministic)  $\text{poly}(n)$ -time algorithm for factoring an  $n$ -variate polynomial of degree  $d$  over  $\mathbb{K}$ . Also, it implies a polynomial time algorithm for deciding whether a  $d$ -dimensional tensor over  $\mathbb{K}$  has rank at most  $r$ . Also, it would give polynomial time algorithms for a wide range of problems not known to have non-trivial solutions previously, such as whether a polynomial of degree  $d$  can be expressed as  $P_1 \cdot P_2 + P_3 \cdot P_4$  where each  $P_1, P_2, P_3, P_4$  are of degree  $d-1$  or less.

This problem was solved for prime  $\mathbb{K}$ , satisfying  $d < |\mathbb{F}|$  by Bhattacharyya [Bha14] and later for all  $d$  and prime  $|\mathbb{K}|$  by Bhattacharyya, Hatami and Tulsiani [BHT15].

Our main result in this line of work establishes this for all fixed finite fields.

**Theorem 1.4.** *For every finite field  $\mathbb{K}$ , positive integers  $k$  and  $d$ , every vector of positive integers  $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_k)$  and every function  $\Gamma : \mathbb{K}^k \rightarrow \mathbb{K}$ , there is a deterministic algorithm  $\mathcal{A}_{\mathbb{K}, d, k, \Delta, \Gamma}$  that takes as input a polynomial  $P : \mathbb{K}^n \rightarrow \mathbb{K}$  of degree  $d$  that runs in time polynomial in  $n$ , and outputs a  $(k, \Delta, \Gamma)$ -decomposition of  $P$  if one exists while otherwise returning NO.*

### 1.1.3 Testing affine-invariant properties

The goal of property testing, as initiated by [BLR93, BFL91] and defined formally by [RS96, GGR98], is to devise algorithms that query their input a very small number of times while correctly deciding whether the input satisfies a given property or is “far” from satisfying it. A property is called *testable* if the query complexity can be made independent of the size of the input.

More precisely, we use the following definitions. Let  $[R]$  denote the set  $\{1, \dots, R\}$ . Given a property  $\mathcal{P}$  of functions in  $\{\mathbb{K}^n \rightarrow [R] \mid n \in \mathbb{Z}_{\geq 0}\}$ , we say that  $f : \mathbb{K}^n \rightarrow [R]$  is  $\varepsilon$ -far from  $\mathcal{P}$  if

$$\min_{g \in \mathcal{P}} \Pr_{x \in \mathbb{K}^n} [f(x) \neq g(x)] > \varepsilon,$$

and we say that it is  $\varepsilon$ -close otherwise.

**Definition 1.5** (Testability). *A property  $\mathcal{P}$  is said to be testable (with one-sided error) if there are functions  $q : (0, 1) \rightarrow \mathbb{Z}_{>0}$ ,  $\delta : (0, 1) \rightarrow (0, 1)$ , and an algorithm  $T$  that, given as input a parameter  $\varepsilon > 0$  and oracle access to a function  $f : \mathbb{K}^n \rightarrow [R]$ , makes at most  $q(\varepsilon)$  queries to the oracle for  $f$ , always accepts if  $f \in \mathcal{P}$  and rejects with probability at least  $\delta(\varepsilon)$  if  $f$  is  $\varepsilon$ -far from  $\mathcal{P}$ . If, furthermore,  $q$  is a constant function, then  $\mathcal{P}$  is said to be proximity-obliviously testable (PO testable).*

The term proximity-oblivious testing is coined by Goldreich and Ron in [GR11]. As an example of a testable (in fact, PO testable) property, let us recall the famous result by Blum, Luby and Rubinfeld [BLR93] which initiated this line of research. They showed that linearity of a function  $f : \mathbb{K}^n \rightarrow \mathbb{K}$  is testable by a test which makes 3 queries. This test accepts if  $f$  is linear and rejects with probability  $\Omega(\varepsilon)$  if  $f$  is  $\varepsilon$ -far from linear.

Linearity, in addition to being testable, is also an example of a *linear-invariant* property. We say that a property  $\mathcal{P} \subseteq \{\mathbb{K}^n \rightarrow [R]\}$  is linear-invariant if it is the case that for any  $f \in \mathcal{P}$  and for any  $\mathbb{K}$ -linear transformation  $L : \mathbb{K}^n \rightarrow \mathbb{K}^n$ , it holds that  $f \circ L \in \mathcal{P}$ . Similarly, an *affine-invariant* property is closed under composition with affine transformations  $A : \mathbb{K}^n \rightarrow \mathbb{K}^n$  (an affine transformation  $A$  is of the form  $L + c$  where  $L$  is  $\mathbb{K}$ -linear and  $c \in \mathbb{K}$ ). The property of a function  $f : \mathbb{K}^n \rightarrow \mathbb{K}$  being affine is testable by a simple reduction to [BLR93], and is itself affine-invariant. Other well-studied examples of affine-invariant (and hence, linear-invariant) properties include Reed-Muller codes [BFL91, BFLS91, FGL<sup>+</sup>96, RS96, AKK<sup>+</sup>05] and Fourier sparsity [GOS<sup>+</sup>09]. In fact, affine invariance seems to be a common feature of most interesting properties that one would classify as “algebraic”. Kaufman and Sudan in [KS08] made explicit note of this phenomenon and initiated a general study of the testability of affine-invariant properties (see also [GK11]).

Our main theorem for testing is a very general positive result:

**Theorem 1.6** (Main testing result). *Let  $\mathcal{P} \subseteq \{\mathbb{K}^n \rightarrow [R]\}$  be an affine-invariant property that is  $t, w$ -lightly locally characterized, where  $t, R, w$ , and  $\text{char}(\mathbb{K})$  are fixed positive integers. Then,  $\mathcal{P}$  is PO testable with  $t$  queries.*

We are yet to define several terms in the above claim, but as we will see, the weight restriction is trivial when the field size is bounded. This yields the following characterization.

**Theorem 1.7** (Testing result for fixed fields). *Let  $\mathcal{P} \subseteq \{\mathbb{K}^n \rightarrow [R]\}$  be an affine-invariant property, where  $R \in \mathbb{Z}^+$  and field  $\mathbb{K}$  are fixed. Then,  $\mathcal{P}$  is PO testable with  $t$  queries if and only if  $\mathcal{P}$  is  $t$ -locally characterized.*

Previously, [BFH<sup>+</sup>13] (building on [BCSX11, BGS10, BFL13]) proved Theorem 1.6 in the case that  $\mathbb{K}$  is of fixed prime order using higher-order Fourier analytic techniques. We note that other recent results on 2-sided testability of affine-invariant properties over fixed prime-order fields [HL13, Yos14] can also be similarly extended to non-prime fields but we omit their description here.

**Local Characterizations** For a PO testable property  $\mathcal{P} \subseteq \{\mathbb{K}^n \rightarrow [R]\}$  of query complexity  $t$ , if a function  $f : \mathbb{K}^n \rightarrow [R]$  does not satisfy  $\mathcal{P}$ , then by Definition 1.5, the tester rejects  $f$  with positive probability. Since the test always accepts functions with the property, there must be  $t$  points  $a_1, \dots, a_t \in \mathbb{K}^n$  that form a witness for non-membership in  $\mathcal{P}$ . These are the queries that cause the tester to reject. Thus, denoting  $\sigma = (f(a_1), \dots, f(a_t)) \in [R]^t$ , we say that  $\mathcal{C} = (a_1, a_2, \dots, a_t; \sigma)$  forms a  *$t$ -local constraint* for  $\mathcal{P}$ . This means that whenever the constraint is violated by a function

$g$ , i.e.,  $(g(a_1), \dots, g(a_t)) = \sigma$ , we know that  $g$  is not in  $\mathcal{P}$ . A property  $\mathcal{P}$  is *t-locally characterized* if there exists a collection of  $t$ -local constraints  $\mathcal{C}_1, \dots, \mathcal{C}_m$  such that  $g \in \mathcal{P}$  if and only if none of the constraints  $\mathcal{C}_1, \dots, \mathcal{C}_m$  are violated. It follows from the above discussion that if  $\mathcal{P}$  is PO testable with  $q$  queries, then  $\mathcal{P}$  is  $t$ -locally characterized.

For an affine-invariant property, constraints can be defined in terms of affine forms, since the affine orbit of a constraint is also a constraint. So, we can describe each  $t$ -local constraint  $\mathcal{C}$  as  $(A_1, \dots, A_t; \sigma)$ , where for every  $i \in [t]$ ,  $A_i(X_1, \dots, X_t) = X_1 + \sum_{j=2}^t c_{i,j} X_j$  for some  $c_{i,j} \in \mathbb{K}$  is an affine form over  $\mathbb{K}$ . We define the *weight*  $\text{wt}$  of an element  $c \in \mathbb{K}$  as  $\sum_{k=1}^r |c_k|$ , where  $c$  is viewed as an  $r$ -dimensional vector  $(c_1, \dots, c_r)$  with each  $c_i$  in the base prime field<sup>2</sup>  $\mathbb{F}$  with respect to a fixed arbitrary basis. The *weight of an affine form*  $A_i$  to be  $\sum_{j=2}^m \text{wt}(c_{i,j})$  for  $c_{i,j}$  as above. A constraint is said to be of weight  $w$  if all its affine forms are of weight at most  $w$ , and a property  $\mathcal{P}$  is said to be  $t, w$ -lightly localyl characterized if there exist  $t$ -local constraints  $\mathcal{C}_1, \dots, \mathcal{C}_m$ , each of weight at most  $w$  that characterize  $\mathcal{P}$ .

Theorem 1.6 asserts that if  $\mathcal{P}$  has a light local characterization, then it is testable. There can exist many local characterizations of a property, and for the theorem to apply, it is only necessary that one such characterization be of bounded weight. Moreover, we can choose the basis with which to describe  $\mathbb{K}$  over  $\mathbb{F}$ . On the other hand, some restriction in addition to local characterization is needed, as Ben-Sasson et al. [BMSS11] show that there exist affine-invariant locally characterized properties of functions  $f : \mathbb{F}_{2^n} \rightarrow \mathbb{F}_2$  that require super-constant query complexity to test.

Another interesting observation is that if a property has a local characterization of bounded weight, then it has a local *single orbit characterization*, in the language of [KS08]. For linear<sup>3</sup> affine-invariant properties, [KS08] shows that any local single orbit characterized property is testable. Hence, our result is weaker than [KS08] in this aspect, though our Theorem 1.6 allows non-linear properties. It is an interesting open question as to whether dual-BCH codes and, more generally, sparse affine-invariant codes that were shown to be locally single orbit characterized in [KL05] and [GKS12] respectively also have local characterizations of bounded weight. It is also an open problem to describe a testable property  $\mathcal{P} \subseteq \{\mathbb{F}_{2^n} \rightarrow \mathbb{F}_2\}$  that does not have a local characterization of bounded weight.

## 1.2 Our Techniques

### 1.2.1 New Ingredients

Our starting point is the observation that  $\mathbb{K}$  is an  $r$ -dimensional vector space over  $\mathbb{F}$ . Thus, we can view a function  $Q : \mathbb{K}^n \rightarrow \mathbb{K}$  as determined by a collection of functions  $P_1, \dots, P_r : \mathbb{K}^n \rightarrow \mathbb{F}$  where  $\mathbb{K}^n$  is viewed as  $\mathbb{F}^{rn}$ . In view of this, we define the notion of an *additive polynomial*. A function<sup>4</sup>  $P : \mathbb{K}^n \rightarrow \mathbb{F}$  is said to have *additive degree*  $d$  if for all  $h_1, \dots, h_{d+1} \in \mathbb{K}^n$ ,  $D_{h_1} \cdots D_{h_{d+1}} P \equiv 0$ , where  $(D_h P)(x) = P(x + h) - P(x)$ . Additive polynomials are exactly the non-classical polynomials of [TZ12] when the domain is  $\mathbb{F}^{rn}$ . Moreover, if  $Q : \mathbb{K}^n \rightarrow \mathbb{K}$  has degree  $d$  (in the usual sense of having a monomial with degree  $d$ ), then  $\text{Tr}(\alpha Q)$  has additive degree  $\leq d$  for any  $\alpha \in \mathbb{K}$  where  $\text{Tr} : \mathbb{K} \rightarrow \mathbb{F}$  denotes the trace function.

<sup>2</sup>If  $x \in \mathbb{F}$ ,  $|x|$  is the obvious element of  $\{0, 1, \dots, |\mathbb{F}| - 1\}$ .

<sup>3</sup>These are properties of functions  $f : \mathbb{K}^n \rightarrow \mathbb{F}$ , where  $\mathbb{F}$  is a subfield of  $\mathbb{K}$ , for which  $f, g \in \mathcal{P}$  implies  $\alpha f + \beta g \in \mathcal{P}$  for any  $\alpha, \beta \in \mathbb{F}$ .

<sup>4</sup>To deal with low characteristics, we will actually use a slightly general definition valid for functions mapping to the torus  $\mathbb{R}/\mathbb{Z}$ .



Therefore, we can directly write any polynomial  $P : \mathbb{K}^n \rightarrow \mathbb{K}$  in terms of additive polynomials and then import all of the results shown in [TZ12] for non-classical polynomials to our setting! Unfortunately, we are not done. The reason is that our applications require, in addition to additive structure, some of the multiplicative structure of  $\mathbb{K}$ , which is lost when we view  $\mathbb{K}$  as  $\mathbb{F}^r$ .

To see why, recall the question of testing affine-invariant properties. When  $\mathbb{K}$  is of bounded order, we can view any one-sided test as examining the restriction of the input function on a random  $K$ -dimensional affine subspace of  $\mathbb{K}^n$ , for some constant integer  $K$ . In other words, the test will evaluate the input function at elements of the set  $H = \{x + \sum_{i=1}^K a_i y_i : a_1, \dots, a_K \in \mathbb{K}\}$  for some  $x, y_1, \dots, y_K \in \mathbb{K}$ . Clearly,  $H$  is not an affine subspace of  $\mathbb{F}^{rn}$ . An important component of the higher-order Fourier analytic approach is to show that any “sufficiently pseudorandom” collection of polynomials is equidistributed on  $H$ , and the proof of this fact in [BFH<sup>+</sup>13] crucially uses that  $H$  is a subspace of a vector space over a prime field. In our work, we show a strong equidistribution theorem (Theorem 3.3) that holds when  $H$  is an affine subspace of  $\mathbb{K}^n$ .

A different place where multiplicative structure rears its head is a key *Degree Preserving Lemma* of [BFH<sup>+</sup>13]. Informally, it states that if  $P_1, \dots, P_C$  form a “sufficiently pseudorandom” collection of polynomials and  $F(x) = \Gamma(P_1(x), \dots, P_C(x))$  is a polynomial of degree  $d$  where  $\Gamma$  is an arbitrary composition function, then for any other collection of polynomials  $Q_1, \dots, Q_C$  where  $\deg(Q_i) \leq \deg(P_i)$  for every  $i$ ,  $G(x) = \Gamma(Q_1(x), \dots, Q_C(x))$  also has degree  $\leq d$ . The lemma is crucially used for the analysis of the Reed-Muller list decoding bound in [BL14] and the polynomial decomposition algorithm in [Bha14, BHT15]. Its proof goes via showing that if all  $(d + 1)$  iterated derivatives of  $F : \mathbb{K}^n \rightarrow \mathbb{K}$  vanish, then so must all  $(d + 1)$  iterated derivatives of  $G : \mathbb{K}^n \rightarrow \mathbb{K}$ . However, when  $|\mathbb{K}|$  is non-prime, all  $(d + 1)$  iterated derivatives of a function  $G : \mathbb{K}^n \rightarrow \mathbb{K}$  may vanish without the degree being  $\leq d$ ; consider for example the polynomial  $x^p$  which vanishes after only 2 derivatives.

We resolve this issue by giving a different and more transparent proof of the Degree Preserving Lemma, which actually holds in a much more general setting (Theorem 3.4). Using the above notation, we prove that if  $F : \mathbb{K}^n \rightarrow \mathbb{K}$  satisfies some locally characterized property  $\mathcal{P}$ , then  $G : \mathbb{K}^n \rightarrow \mathbb{K}$  does also. Since due to a work of Kaufman and Ron [KR06], we know that degree is locally characterized, our desired result follows. Our new proof uses our strong equidistribution theorem on affine subspaces of  $\mathbb{K}^n$ .

An interesting point to note is that both the equidistribution theorem and the degree preserving lemma work only assuming that the field characteristic is constant and that the involved affine constraints are of bounded weight, without any assumption on the field size.

### 1.2.2 Reed-Muller codes

For a received word  $g : \mathbb{K}^n \rightarrow \mathbb{K}$  our goal is to upper bound  $|\{f \in \mathcal{P}_d : \text{dist}(f, g) \leq \eta\}|$ , where  $\eta = \delta_{\mathbb{K}}(d) - \varepsilon$  for some  $\eta > 0$  and  $\mathcal{P}_d$  is the class  $\{Q : \mathbb{K}^n \rightarrow \mathbb{K} : \deg(Q) \leq d\}$ . The proof technique is similar in structure as [BL14]. We apply the weak regularity lemma (Corollary 4.1) to the received word  $g : \mathbb{K}^n \rightarrow \mathbb{K}$  and reduce the problem to a structured word  $g' : \mathbb{K}^n \rightarrow \mathbb{K}$ . More specifically, whenever  $\text{dist}(f, g) \leq \eta$ , we have  $\text{dist}(f, g') \leq \eta + \varepsilon/2$ . From here, we first express each function  $f : \mathbb{K}^n \rightarrow \mathbb{K}$  as a linear combination of functions  $f' : \mathbb{K}^n \rightarrow \mathbb{F}$ . It can be then shown that the analysis in [BL14] works for functions  $f' : \mathbb{K}^n \rightarrow \mathbb{F}$ . A naive recombination of the  $f' : \mathbb{K}^n \rightarrow \mathbb{F}$  to  $f : \mathbb{K}^n \rightarrow \mathbb{K}$  gives us useful bounds only when  $d < \text{char}(|\mathbb{F}|)$ . To circumvent this problem, we use our improved degree preserving theorem. This is crucial to our analysis as the technique of [BL14] can be used only to analyze the additive degree of polynomials which is not enough for the argument to work for arbitrary  $d$  and  $|\mathbb{K}|$ .

### 1.2.3 Polynomial decomposition

The algorithm and its analysis follows the lines of [Bha14, BHT15]. Given a polynomial  $P : \mathbb{K}^n \rightarrow \mathbb{K}$  (where  $|\mathbb{K}|$  is bounded), we consider the collection of additive polynomials  $\{\text{Tr}(\alpha_1 P), \dots, \text{Tr}(\alpha_r P)\}$  where  $\alpha_1, \dots, \alpha_r \in \mathbb{K}$  are linearly independent. We regularize this collection into a pseudorandom additive polynomial factor and set one variable to 0 such that the degrees of the polynomials do not change. We then recursively solve the problem on  $n - 1$  variables and then apply a lifting procedure to get a decomposition for the original problem. A naive analysis of the lifting procedure over non-prime fields requires that  $\deg(P) < \text{char}(\mathbb{F})$ . In order to get around this, we use our improved degree preserving theorem which applies for arbitrary degrees.

### 1.2.4 Testing affine-invariant properties

Suppose  $\mathcal{P} \subseteq \{\mathbb{K}^n \rightarrow [R]\}$  is a locally characterized affine-invariant property (where  $R$  and  $\text{char}(\mathbb{K})$  are bounded but  $n|\mathbb{K}|$  is growing). Our proof follows the lines of [BGS10, BFL13, BFH<sup>+</sup>13]. Suppose  $f$  is far from  $\mathcal{P}$ . We first identify a low-rank function close to  $f$  in an appropriate Gowers norm which also contains the violation that  $f$  contains. Here, low rank is with respect to a collection  $\mathcal{B}$  of *additive* polynomials. We then investigate the distribution of  $\mathcal{B}$  on the affine constraint that  $f$  violates. Since these are affine with respect to  $\mathbb{K}^n$ , we need to use our strong equidistribution theorem. The rest of the proof proceeds along the same lines as [BFH<sup>+</sup>13].

Because the proof of Theorem 1.6 is very analogous to that in [BFH<sup>+</sup>13] (except for the use of additive polynomials and the new equidistribution theorem) and requires significant additional notation, we omit it here.

## 2 Preliminaries

Let  $\mathbb{N}$  denote the set of positive integers. For  $n \in \mathbb{N}$ , let  $[n] := \{1, 2, \dots, n\}$ . We use  $y = x \pm \varepsilon$  to denote  $y \in [x - \varepsilon, x + \varepsilon]$ . For  $n \in \mathbb{N}$ , and  $x, y \in \mathbb{C}^n$ , let  $\langle x, y \rangle := \sum_{i=1}^n x_i \bar{y}_i$  where  $\bar{a}$  is the conjugate of  $a$ . Let  $\|x\|_2 := \sqrt{\langle x, x \rangle}$ .

Let  $\mathbb{T}$  denote the torus  $\mathbb{R}/\mathbb{Z}$ . This is an abelian group under addition. Let  $e : \mathbb{T} \rightarrow \mathbb{C}$  be the function  $e(x) = e^{2\pi i x}$ . For an integer  $k \geq 0$ , let  $\mathbb{U}_k := \frac{1}{p^k} \mathbb{Z}/\mathbb{Z}$ . Note that  $\mathbb{U}_k$  is a subgroup of  $\mathbb{T}$ .

Let  $\iota : \mathbb{F} \rightarrow \mathbb{U}_1$  be the bijection  $\iota(a) = \frac{|a|}{p} \pmod{1}$ .

Fix a prime field  $\mathbb{F} = \mathbb{F}_p$ , and let  $\mathbb{K} = \mathbb{F}_q$  where  $q = p^r$  for a positive integer  $r$ . We denote by  $\text{Tr} : \mathbb{K} \rightarrow \mathbb{F}$  the trace function:

$$\text{Tr}(x) = x + x^p + x^{p^2} + \dots + x^{p^{r-1}}$$

Recall that  $\{x \rightarrow \text{Tr}(ax) : a \in \mathbb{K}\}$  is in bijection with the set of all linear maps from  $\mathbb{K}$  to  $\mathbb{F}$ . Also, we use  $|\cdot|$  to denote the obvious map from  $\mathbb{F}$  to  $\{0, 1, \dots, p-1\}$ . We will need the following useful fact.

**Proposition 2.1** (Dual basis). *For any  $r$  linearly independent elements  $\alpha_1, \dots, \alpha_r \in \mathbb{K}$ , there exist  $\beta_1, \beta_2, \dots, \beta_r$  in  $\mathbb{K}$  such that any  $x \in \mathbb{K}$  equals  $\sum_{i=1}^r \beta_i \text{Tr}(\alpha_i x)$ .*

Given a basis, i.e. collection of  $r$  linearly independent field elements,  $\alpha = (\alpha_1, \dots, \alpha_r)$ , we define  $\text{wt}_\alpha : \mathbb{K} \rightarrow \mathbb{Z}$  to be  $\text{wt}_\alpha(c) = \sum_{i=1}^r |\text{Tr}(\alpha_i c)|$ .



## 2.1 Affine forms and constraints

A *linear form on  $k$  variables* is a vector  $L = (w_1, w_2, \dots, w_k) \in \mathbb{K}^k$  that is interpreted as a function from  $(\mathbb{K}^n)^k$  to  $\mathbb{K}^n$  via the map  $(x_1, \dots, x_k) \mapsto w_1x_1 + w_2x_2 + \dots + w_kx_k$ . A linear form  $L = (w_1, w_2, \dots, w_k)$  is said to be *affine* if  $w_1 = 1$ . From now, linear forms will always be assumed to be affine. Given a basis  $\alpha = (\alpha_1, \dots, \alpha_r)$ , we define  $\text{wt}_\alpha$  of a linear form  $L = (w_1, \dots, w_k)$  to be  $\sum_{i=2}^k \text{wt}_\alpha(w_i)$ .

We specify a partial order  $\preceq$  among affine forms, with respect to a basis  $\alpha = (\alpha_1, \dots, \alpha_r)$ . We say  $(w_1, \dots, w_k) \preceq_\alpha (w'_1, \dots, w'_k)$  if  $|\text{Tr}(\alpha_j w_i)| \leq |\text{Tr}(\alpha_j w'_i)|$  for all  $i \in [k], j \in [r]$ .

**Definition 2.2** (Affine constraints). *An affine constraint of size  $m$  on  $k$  variables is a tuple  $A = (L_1, \dots, L_m)$  of  $m$  affine forms  $L_1, \dots, L_m$  over  $\mathbb{F}$  on  $k$  variables, where:  $L_1(x_1, \dots, x_k) = x_1$ . Moreover, it is said to be *weight-closed* if there exists a basis  $\alpha = (\alpha_1, \dots, \alpha_r)$  such that for any affine form  $L$  belonging to  $A$ , if  $L' \preceq_\alpha L$ , then  $L'$  also belongs to  $A$ .*

Observe that a weight-closed affine constraint is of bounded size if and only if all its affine forms are of bounded weight with respect to some  $\alpha$ .

## 2.2 Polynomials, Degrees and Derivatives

A function  $P : \mathbb{K}^n \rightarrow \mathbb{K}$  is a *polynomial of degree  $d$*  if for all  $d_1, \dots, d_n \geq 0$  such that  $\sum_i d_i \leq d$ , there exists  $c_{d_1, \dots, d_n} \in \mathbb{K}$  such that:

$$P(x_1, \dots, x_n) = \sum_{\substack{d_1, \dots, d_n \in \mathbb{Z}^+ \\ d_1 + \dots + d_n \leq d}} c_{d_1, \dots, d_n} x_1^{d_1} x_2^{d_2} \dots x_n^{d_n}$$

We use the notion of *additive degree* for functions mapping to  $\mathbb{T}$ . Given a function  $f : \mathbb{K}^n \rightarrow \mathbb{T}$ , its *additive derivative in direction  $h \in \mathbb{K}^n$*  is  $D_h f : \mathbb{K}^n \rightarrow \mathbb{T}$ , given by

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

**Definition 2.3** (Additive Polynomials). *A function  $P : \mathbb{K}^n \rightarrow \mathbb{T}$  is a polynomial of additive degree  $d$  if for all  $x, h_1, h_2, \dots, h_{d+1} \in \mathbb{K}^n$ , we have*

$$D_{h_1} D_{h_2} \dots D_{h_{d+1}} P(x) = 0. \quad (1)$$

*A function of bounded additive degree is called an additive polynomial.*

For functions  $P$  mapping to  $\mathbb{T}$ ,  $\text{deg}(P)$  denotes its additive degree. Note that we can interpret  $P : \mathbb{K}^n \rightarrow \mathbb{T}$  as a function  $P' : \mathbb{F}^{nr} \rightarrow \mathbb{T}$  with the same additive degree by setting  $P(x_1, \dots, x_n) = P'(\text{Tr}(\alpha_1 x_1), \dots, \text{Tr}(\alpha_1 x_1), \dots, \text{Tr}(\alpha_1 x_n), \dots, \text{Tr}(\alpha_1 x_n))$ , using Proposition 2.1. By this identification, additive polynomials are exactly the same as the non-classical polynomials introduced by Tao and Ziegler [TZ12]. As a consequence, we have the following:

**Lemma 2.4** (Lemma 1.7 of [TZ12]).  *$P : \mathbb{K}^n \rightarrow \mathbb{T}$  is a polynomial of additive degree  $d$  if and only if it can be written in the form:*

$$P(x_1, \dots, x_n) = \alpha + \sum_{k \geq 0} \sum_{\substack{0 \leq d_{i,j} < p \quad \forall i \in [n], j \in [r] \\ 0 < \sum_{i=1}^n \sum_{j=1}^r d_{i,j} \leq d - k(p-1)}} \frac{c_{d_{1,1}, \dots, d_{n,r}, k} \prod_{i=1}^n \prod_{j=1}^r |\text{Tr}(\alpha_j x_i)|^{d_{i,j}}}{p^{k+1}} \pmod{1}$$

where  $\alpha \in \mathbb{T}$  and  $c_{d_1,1,\dots,d_n,r} \in \{0, 1, \dots, p-1\}$  are uniquely determined. The maximum  $k$  for which there is a nonzero  $c_{d_1,1,\dots,d_n,r,k}$  is the depth of  $P$ . Note that  $\text{depth}(P) \leq \left\lfloor \frac{d-1}{p-1} \right\rfloor$  and that  $P$  takes on at most  $p^{\text{depth}(P)+1}$  distinct values.

For a function  $f : \mathbb{K}^n \rightarrow \mathbb{C}$ , define the *multiplicative derivative in direction*  $h \in \mathbb{K}^n$  to be

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

### 2.3 Locally Characterized Properties

As described in the introduction, by a locally characterized property, we informally mean a property for which non-membership can be certified by a finite sized witness. Specifically for affine-invariant properties, we define:

**Definition 2.5** (Locally characterized properties).

An induced affine constraint of size  $m$  on  $\ell$  variables is a pair  $(A, \sigma)$  where  $A$  is an affine constraint of size  $m$  on  $\ell$  variables and  $\sigma \in [R]^m$ .

Given such an induced affine constraint  $(A, \sigma)$ , a function  $f : \mathbb{K}^n \rightarrow [R]$  is said to be  $(A, \sigma)$ -free if there exist no  $x_1, \dots, x_\ell \in \mathbb{K}^n$  such that  $(f(L_1(x_1, \dots, x_\ell)), \dots, f(L_m(x_1, \dots, x_\ell))) = \sigma$ . On the other hand, if such  $x_1, \dots, x_\ell$  exist, we say that  $f$  induces  $(A, \sigma)$  at  $x_1, \dots, x_\ell$ .

Given a (possibly infinite) collection  $\mathcal{A} = \{(A^1, \sigma^1), (A^2, \sigma^2), \dots, (A^i, \sigma^i), \dots\}$  of induced affine constraints, a function  $f : \mathbb{K}^n \rightarrow [R]$  is said to be  $\mathcal{A}$ -free if it is  $(A^i, \sigma^i)$ -free for every  $i \geq 1$ . The size of  $\mathcal{A}$  is the size of the largest induced affine constraint in  $\mathcal{A}$ .

Additionally,  $\mathcal{A} = \{(A^1, \sigma^1), (A^2, \sigma^2), \dots, (A^K, \sigma^K)\}$  is a  $W$ -light affine system if there exists a basis  $\alpha = (\alpha_1, \dots, \alpha_r)$  such that  $\text{wt}_\alpha(A^i) \leq W$  for all  $i \in [K]$ .

A property  $\mathcal{P} \subseteq \{\mathbb{K}^n \rightarrow [R]\}$  is said to be  $K, W$ -lightly locally characterized if it is equivalent to  $\mathcal{A}$ -freeness for some  $W$ -light affine system  $\mathcal{A}$  whose size is  $\leq K$ .

We recall that Kaufman and Ron [KR06] show that:

**Theorem 2.6** ([KR06]). The property  $\mathcal{P}_d = \{P : \mathbb{K}^n \rightarrow \mathbb{K} : \text{deg}(P) \leq d\}$  is  $q^{\lceil (d+1)/(q-q/p) \rceil}$ ,  $pr^{\lceil (d+1)/(q-q/p) \rceil}$ -lightly locally characterized.

### 2.4 Factors and Rank

Next, we define a polynomial factor which forms the basis for much of higher order Fourier analysis.

**Definition 2.7** (Factor). A polynomial factor  $\mathcal{B}$  is a sequence of additive polynomials  $P_1, \dots, P_C : \mathbb{K}^n \rightarrow \mathbb{T}$ . We also identify it with the function  $\mathcal{B} : \mathbb{K}^n \rightarrow \mathbb{T}^C$  mapping  $x$  to  $(P_1(x), \dots, P_C(x))$ . An atom of  $\mathcal{B}$  is a preimage  $\mathcal{B}^{-1}(y)$  for some  $y \in \mathbb{T}^C$ . When there is no ambiguity, we will in fact abuse notation and identify an atom of  $\mathcal{B}$  with the common value  $\mathcal{B}(x)$  of all  $x$  in the atom.

The partition induced by  $\mathcal{B}$  is the partition of  $\mathbb{K}^n$  given by  $\{\mathcal{B}^{-1}(y) : y \in \mathbb{T}^C\}$ . The complexity of  $\mathcal{B}$ , denoted  $|\mathcal{B}|$ , is the number of defining polynomials  $C$ . The order of  $\mathcal{B}$ , denoted  $\|\mathcal{B}\|$ , is the total number of atoms in  $\mathcal{B}$ . The degree of  $\mathcal{B}$  is the maximum additive degree among its defining polynomials  $P_1, \dots, P_C$ .

Note that due to Lemma 2.4, if  $\mathcal{B}$  is defined by polynomials  $P_1, \dots, P_C$ ,

$$\|\mathcal{B}\| = \prod_{i=1}^C p^{\text{depth}(P_i)+1}$$

**Definition 2.8** (Rank). *Let  $d \in \mathbb{N}$  and  $P : \mathbb{K}^n \rightarrow \mathbb{T}$ . Then  $\text{rank}_d(P)$  is defined as the smallest integer  $k$  such that there exist functions  $P_1, \dots, P_k : \mathbb{K}^n \rightarrow \mathbb{T}$  of additive degree  $\leq d - 1$  and a function  $\Gamma : \mathbb{T}^k \rightarrow \mathbb{T}$  such that  $P(x) = \Gamma(P_1(x), \dots, P_k(x))$ . If  $d = 1$ , then the rank is 0 if  $P$  is a constant function and is  $\infty$  otherwise. If  $P$  is a polynomial of additive degree  $d$ , then  $\text{rank}(P) = \text{rank}_d(P)$ .*

**Definition 2.9** (Rank and Regularity of Polynomial Factor). *Let  $\mathcal{B}$  be a polynomial factor defined by the sequence  $P_1, \dots, P_c : \mathbb{K}^n \rightarrow \mathbb{T}$  with respective depths  $k_1, \dots, k_c$ . Then, the rank of  $\mathcal{B}$  is  $\min_{(a_1, \dots, a_c)} \text{rank}(\sum_{i=1}^c a_i P_i)$  where the minimum is over  $(a_1, \dots, a_c) \in \mathbb{Z}^c$  such that  $(a_1 \bmod p^{k_1+1}, \dots, a_c \bmod p^{k_c+1}) \neq (0, \dots, 0)$ .*

*Given a polynomial factor  $\mathcal{B}$  and a non decreasing function  $r : \mathbb{Z}^+ \rightarrow \mathbb{Z}^+$ ,  $\mathcal{B}$  is  $r$ -regular if  $\mathcal{B}$  is of rank at least  $r(|\mathcal{B}|)$ .*

**Definition 2.10** (Semantic and Syntactic refinement). *Let  $\mathcal{B}$  and  $\mathcal{B}'$  be polynomial factors. A factor  $\mathcal{B}'$  is a syntactic refinement of  $\mathcal{B}$ , denoted by  $\mathcal{B}' \succeq_{\text{syn}} \mathcal{B}$  if the set of polynomials defining  $\mathcal{B}$  is a subset of the set of polynomials defining  $\mathcal{B}'$ . It is a semantic refinement, denoted by  $\mathcal{B}' \succeq_{\text{sem}} \mathcal{B}$  if for every  $x, y \in \mathbb{K}^n$ ,  $\mathcal{B}'(x) = \mathcal{B}'(y)$  implies  $\mathcal{B}(x) = \mathcal{B}(y)$ . Clearly, a syntactic refinement is also a semantic refinement.*

Our next lemma is the workhorse that allows us to convert any factor into a regular one.

**Lemma 2.11** (Polynomial Regularity Lemma). *Let  $r : \mathbb{Z}^+ \rightarrow \mathbb{Z}^+$  be a non-decreasing function and  $d > 0$  be an integer. Then, there is a function  $C_{2.11}^{(r,d)} : \mathbb{Z}^+ \rightarrow \mathbb{Z}^+$  such that the following is true. Suppose  $\mathcal{B}$  is a factor defined by polynomials  $P_1, \dots, P_C : \mathbb{K}^n \rightarrow \mathbb{T}$  of additive degree at most  $d$ . Then, there is an  $r$ -regular factor  $\mathcal{B}'$  consisting of polynomials  $Q_1, \dots, Q_{C'} : \mathbb{K}^n \rightarrow \mathbb{T}$  of additive degree  $\leq d$  such that  $\mathcal{B}' \succeq_{\text{sem}} \mathcal{B}$  and  $C' \leq C_{2.11}^{(r,d)}(C)$ .*

*Moreover, if  $\mathcal{B}$  is itself a refinement of some polynomial factor  $\hat{\mathcal{B}}$  that has rank  $> (r(C') + C')$ , then additionally  $\mathcal{B}'$  will be a syntactic refinement of  $\hat{\mathcal{B}}$ .*

*Proof.* Follows directly from Lemma 2.18 of [BFH<sup>+</sup>13] by identifying  $\mathbb{K}^n$  with  $\mathbb{F}^{rn}$ .  $\square$

In fact, the regularization process of Lemma 2.11 can be implemented in time  $O(n^{d+1})$  [BHT15].

## 2.5 Gowers norm and the inverse theorem

**Definition 2.12.** *The bias of a function  $f : \mathbb{K}^n \rightarrow \mathbb{C}$  is defined as  $\text{bias}(f) = |\mathbb{E}_{x \in \mathbb{K}^n} f(x)|$ . For  $P : \mathbb{K}^n \rightarrow \mathbb{T}$ , we use  $\text{bias}(P)$  to denote  $\text{bias}(e(P))$ .*

The Gowers norm of a function measures the bias of its iterated derivative. Precisely:

**Definition 2.13** (Gowers norm). *Given a function  $f : \mathbb{K}^n \rightarrow \mathbb{C}$  and an integer  $d \geq 1$ , the Gowers norm of order  $d$  for  $f$  is given by*

$$\|f\|_{U^d} = \left| \mathbb{E}_{h_1, \dots, h_d, x \in \mathbb{K}^n} [(\Delta_{h_1} \Delta_{h_2} \cdots \Delta_{h_d} f)(x)] \right|^{1/2^d}.$$

*If  $P : \mathbb{K}^n \rightarrow \mathbb{T}$ ,  $\|P\|_{U^d}$  denotes  $\|e(P)\|_{U^d}$ .*

Note that as  $\|f\|_{U^1} = \text{bias}(f)$  the Gowers norm of order 1 is only a semi-norm. However for  $d > 1$ , it is not difficult to show that  $\|\cdot\|_{U^d}$  is indeed a norm.

There is a tight connection between additive polynomials and Gowers norms. In one direction, it is a straightforward consequence of the monotonicity of the Gowers norm ( $\|f\|_{U^d} \leq \|f\|_{U^{d+1}}$ ) and invariance of the Gowers norm with respect to modulation by lower degree polynomials ( $\|f\|_{U^{d+1}} = \|f \cdot e(P)\|_{U^{d+1}}$  for polynomials  $P$  of additive degree  $\leq d$ ) that if  $f : \mathbb{K}^n \rightarrow \mathbb{C}$  is  $\delta$ -correlated with a polynomial  $P$  of additive degree  $\leq d$ , meaning

$$\left| \mathbb{E}_x f(x) e(-P(x)) \right| \geq \delta$$

for some  $\delta > 0$ , then

$$\|f\|_{U^{d+1}} \geq \delta.$$

In the other direction, we have the following ‘‘Inverse theorem for the Gowers norm’’.

**Theorem 2.14** (Theorem 1.11 of [TZ12]). *Suppose  $\delta > 0$  and  $d \geq 1$  is an integer. There exists an  $\varepsilon = \varepsilon_{2.14}(\delta, d)$  such that the following holds. For every function  $f : \mathbb{K}^n \rightarrow \mathbb{C}$  with  $\|f\|_\infty \leq 1$  and  $\|f\|_{U^{d+1}} \geq \delta$ , there exists a polynomial  $P : \mathbb{K}^n \rightarrow \mathbb{T}$  of additive degree  $\leq d$  that is  $\varepsilon$ -correlated with  $f$ , meaning*

$$\left| \mathbb{E}_{x \in \mathbb{K}^n} f(x) e(-P(x)) \right| \geq \varepsilon.$$

We can be more explicit when  $f = e(P)$  for an additive polynomial  $P$ .

**Theorem 2.15** (Theorem 1.20 of [TZ12]). *Suppose  $\delta > 0$  and  $d \geq 1$  is an integer. There exists an  $r = r_{2.15}(\delta, d)$  such that the following holds. If a polynomial  $P : \mathbb{K}^n \rightarrow \mathbb{T}$  with additive degree  $d$  satisfies  $\|P\|_{U^d} \geq \delta$ , then  $\text{rank}(P) \leq r$ .*

## 3 New Tools

### 3.1 Equidistribution of regular factors

Our results in this section imply that a regular polynomial factor is ‘‘as random as possible’’, subject to the additive degree and depth bounds of its defining polynomials. Let us start with the following simple observation.

**Lemma 3.1.** *Given  $\varepsilon > 0$ , let  $\mathcal{B}$  be a polynomial factor of degree  $d > 0$ , complexity  $C$  and rank  $r_{3.1}(d, \varepsilon)$ , defined by a sequence of additive polynomials  $P_1, \dots, P_C : \mathbb{K}^n \rightarrow \mathbb{T}$  having respective depths  $k_1, \dots, k_C$ . Suppose  $\alpha = (\alpha_1, \dots, \alpha_C) \in \mathbb{U}_{k_1+1} \times \dots \times \mathbb{U}_{k_C+1}$ . Then:*

$$\Pr_x[\mathcal{B}(x) = \alpha] = \frac{1}{\|\mathcal{B}\|} \pm \varepsilon.$$

*Proof.* This is standard. See for example Lemma 3.2 of [BFH<sup>+</sup>13]. □

In our applications though, we will often need not just  $\mathcal{B}(x)$  to be nearly uniformly distributed but the tuple  $(\mathcal{B}(x) : x \in H)$  for a set  $H \subseteq \mathbb{K}^n$  to be nearly uniformly distributed. In particular, we consider the case when  $H$  is an affine subspace of  $\mathbb{K}^n$ . The following lemma is key.

**Lemma 3.2** (Near orthogonality). *Let  $A = (L_1, \dots, L_m)$  be a weight-closed affine constraint of bounded size on  $\ell$  variables. Suppose  $\mathcal{B}$  is a polynomial factor of degree  $d$  and rank  $\geq r$  (2.15)  $(d, \delta)$ , defined by the sequence of additive polynomials  $P_1, \dots, P_c : \mathbb{K}^n \rightarrow \mathbb{T}$ . Let  $\Lambda = (\lambda_{ij})_{i \in [c], j \in [m]}$  be a tuple of integers. Define:*

$$P_\Lambda(x_1, \dots, x_k) = \sum_{i \in [c], j \in [m]} \lambda_{ij} P_i(L_j(x_1, \dots, x_\ell)).$$

Then one of the following is true.

1. For every  $i \in [c]$ , it holds that  $\sum_{j \in [m]} \lambda_{ij} Q_i(L_j(\cdot)) \equiv 0$  for all polynomials  $Q_i : \mathbb{K}^n \rightarrow \mathbb{T}$  with the same additive degree and depth as  $P_i$ . Clearly, this implies  $P_\Lambda \equiv 0$ .
2.  $P_\Lambda \not\equiv 0$ . Moreover,  $\text{bias}(P_\Lambda) \leq \delta$ .

*Proof.* For  $j \in [m]$ , let  $(w_{j,1}, \dots, w_{j,\ell}) \in \mathbb{K}^\ell$  denote the affine form given by  $L_j$ . Note that  $w_{j,1} = 1$ .

Suppose  $\alpha = (\alpha_1, \dots, \alpha_r)$  is the basis with respect to which the affine forms are weight-closed. For each  $i$ , we do the following. If for some  $j$ , we have<sup>5</sup>  $\text{wt}_\alpha(L_j) > \deg(\lambda_{i,j} P_i)$ ,  $\lambda_{i,j} \neq 0$ , then using Proposition 2.1,  $L_j(x_1, \dots, x_\ell) = x_1 + \sum_{i=2}^\ell (\sum_{k=1}^r u_{i,k} \cdot \beta_k) x_i$  where  $\beta$  is the dual basis to  $\alpha$ , each  $u_{i,k} \in [0, p-1]$  and  $\sum_{i,k} u_{i,k} > \deg(\lambda_{i,j} P_i)$ . Using Equation (1), we can replace  $\lambda_{i,j} P_i(L_j)$  by a  $\mathbb{Z}$ -linear combination of  $P_i(L_{j'})$  where  $L_{j'} \preceq_\alpha L_j$  until no such  $j$  exists. This is where we use the fact that the affine constraint is weight-closed. Suppose the new coefficients are denoted by  $(\lambda'_{i,j})$ . If the  $\lambda'_{i,j}$  are all zero, then for every  $i \in [c]$  individually,  $\sum_{j \in [m]} P_i(L_j(x_1, \dots, x_\ell)) \equiv 0$ . Indeed,  $\sum_{j \in [m]} Q_i(L_j(x_1, \dots, x_\ell)) \equiv 0$  for any  $Q_i$  with the same additive degree and depth, as the transformation from  $\lambda_{i,j}$  to  $\lambda'_{i,j}$  did not use any other information about  $P_i$ .

Else some  $\lambda'_{i,j} \neq 0$ . Also,  $\text{wt}_\alpha(L_j) \leq \deg(\lambda'_{i,j} P_i)$ . Then we show the second part of the lemma, that is  $|\mathbb{E}[e(P_\Lambda(x_1, \dots, x_k))]| \leq \delta$ .

Suppose without loss of generality that the following is true.

$$\lambda'_{i,1} \neq 0 \text{ for some } i \in [C].$$

$L_1$  is maximal in the sense that for every  $j \neq 1$ , either  $\lambda'_{i,j} = 0$  for all  $i \in [C]$  or  $\text{wt}_\alpha(w_{j,s}) < \text{wt}_\alpha(w_{1,s})$  for some  $s \in [\ell]$ .

For  $a = (a_1, \dots, a_\ell) \in \mathbb{K}^\ell$  and  $y \in \mathbb{K}^n$  and  $P : \mathbb{K}^n \rightarrow \mathbb{T}$ , define

$$\overline{D}_{a,y} P(x_1, \dots, x_\ell) = P(x_1 + a_1 y, \dots, x_\ell + a_\ell y) - P(x_1, \dots, x_\ell).$$

Then

$$\overline{D}_{a,y} (P_i \circ L_j)(x_1, \dots, x_\ell) = (D_{L_j(a)y} P_i)(L_j(x_1, \dots, x_\ell)).$$

Let  $\Delta = \text{wt}_\alpha(L_1) \leq d$ . Define  $a_1, \dots, a_\Delta$  be the set of vectors of the form  $(-w, 0, \dots, 1, 0, \dots, 0)$  where 1 is in the  $i$ th coordinate for  $i \in [2, \ell]$  and for all  $w \in \mathbb{K}$  satisfying  $0 \leq \text{wt}_\alpha(w) < \text{wt}_\alpha(w_{1,i})$ . Note that  $\langle L_1, a_k \rangle \neq 0$  for  $k \in [\Delta]$  but for any  $j > 1$  there exists some  $k \in [\Delta]$  such that  $\langle L_j, a_k \rangle = 0$ . Thus,

$$\mathbb{E}_{y_1, \dots, y_\Delta, x_1, \dots, x_\ell} [e((\overline{D}_{a_\Delta, y_\Delta} \dots \overline{D}_{a_1, y_1} P_\Lambda)(x_1, \dots, x_\ell))] = \left\| \sum_{i=1}^C \lambda'_{i,1} P_i \right\|_{U^\Delta}^{2^\Delta}.$$

The rest of the analysis is same as Theorem 3.3 in [BFH<sup>+</sup>13] and we skip it here.  $\square$

<sup>5</sup>Here,  $\deg(\cdot)$  refers to the additive degree.

We can now use Lemma 3.2 to prove our result on equidistribution of regular factors over affine subspaces of  $\mathbb{K}^n$ .

**Theorem 3.3.** *Let  $\varepsilon > 0$ . Let  $\mathcal{B}$  be a polynomial factor defined by polynomials  $P_1, \dots, P_c : \mathbb{K}^n \rightarrow \mathbb{T}$  with respective additive degrees  $d_1, \dots, d_c \in \mathbb{Z}^+$  and depths  $k_1, \dots, k_c \in \mathbb{Z}^{\geq 0}$ . Suppose  $\mathcal{B}$  has rank at least  $r^{(2.15)}(d, \varepsilon)$  where  $d = \max(d_1, \dots, d_c)$ . Let  $A = (L_1, \dots, L_m)$  be a weight-closed affine constraint. For every  $i \in [c]$ , define  $\Lambda_i$  to be the set of tuples  $(\lambda_1, \dots, \lambda_m) \in [0, p^{k_i+1} - 1]$  such that  $\sum_{j=1}^m \lambda_j Q_i(L_j(\cdot)) \equiv 0$  for all polynomials  $Q_i$  with the same additive degree and depth as  $P_i$ . Consider  $(\alpha_{i,j} : i \in [c], j \in [m]) \in \mathbb{T}^{cm}$  such that for every  $i \in [c]$  and for every  $(\lambda_1, \dots, \lambda_m) \in \Lambda_i$ ,  $\sum_{j=1}^m \lambda_j \alpha_{i,j} = 0$ . Then:*

$$\Pr_{x_1, \dots, x_\ell \in \mathbb{K}^n} [\mathcal{B}(L_j(x_1, \dots, x_\ell)) = (\alpha_{1,j}, \dots, \alpha_{c,j}) \forall j \in [m]] = \frac{\prod_{i=1}^c |\Lambda_i|}{\|\mathcal{B}\|^m} \pm \varepsilon$$

*Proof.*

$$\begin{aligned} & \Pr_{x_1, \dots, x_\ell \in \mathbb{K}^n} [\mathcal{B}(L_j(x_1, \dots, x_\ell)) = (\alpha_{1,j}, \dots, \alpha_{c,j}) \forall j \in [m]] \\ &= \mathbb{E}_{x_1, \dots, x_\ell} \left[ \prod_{i,j} \frac{1}{p^{k_i+1}} \sum_{\lambda_{i,j}=0}^{p^{k_i+1}-1} e(\lambda_{i,j}(P_i(L_j(x_1, \dots, x_\ell)) - \alpha_{i,j})) \right] \\ &= \left( \prod_i p^{-(k_i+1)} \right)^m \sum_{(\lambda_{i,j})} e \left( - \sum_{i,j} \lambda_{i,j} \alpha_{i,j} \right) \mathbb{E} \left[ e \left( \sum_{i,j} \lambda_{i,j} P_i(L_j(x_1, \dots, x_\ell)) \right) \right] \\ &= p^{-m \sum_{i=1}^c (k_i+1)} \cdot \left( \prod_{i=1}^c |\Lambda_i| \pm \varepsilon p^{m \sum_{i=1}^c (k_i+1)} \right) \end{aligned}$$

The last line is due to the observation that from Lemma 3.2,  $\sum_{i=1}^c \sum_{j=1}^m \lambda_{i,j} P_i(L_j(x_1, \dots, x_\ell)) \equiv 0$  if and only if for every  $i \in [c]$ ,  $(\lambda_{i,1}, \dots, \lambda_{i,m}) \in \Lambda_i \pmod{p^{k_i+1}}$ . So,  $\sum_{i,j} \lambda_{i,j} P_i(L_j(\cdot))$  is identically 0 for  $\prod_i |\Lambda_i|$  many tuples  $(\lambda_{i,j})$  and for those tuples,  $\sum_{i,j} \lambda_{i,j} \alpha_{i,j} = 0$  also.  $\square$

Note that in Theorem 3.3, if  $\varepsilon$  is a constant,  $m$  needs to be bounded for the claim to be non-trivial, which in turn requires that the affine forms in  $L$  be of bounded weight.

### 3.2 Preservation of Locally Characterized Properties

**Theorem 3.4.** *Let  $\mathcal{P} \subset \{\mathbb{K}^n \rightarrow \mathbb{K}\}$  be a  $K, W$ -lightly locally characterized property. For an integer  $d$ , suppose  $P_1, \dots, P_c : \mathbb{K}^n \rightarrow \mathbb{T}$  are polynomials of additive degree  $\leq d$ , forming a factor of rank  $> r^{3.4}(d, K)$ , and  $\Gamma : \mathbb{T}^c \rightarrow \mathbb{K}$  is a function such that  $F : \mathbb{K}^n \rightarrow \mathbb{K}$  defined by  $F(x) = \Gamma(P_1(x), \dots, P_c(x))$  satisfies  $\mathcal{P}$ .*

*For every collection of additive polynomials  $Q_1, \dots, Q_c : \mathbb{K}^n \rightarrow \mathbb{T}$  with  $\deg(Q_i) \leq \deg(P_i)$  and  $\text{depth}(Q_i) \leq \text{depth}(P_i)$  for all  $i \in [c]$ , if  $G : \mathbb{K}^n \rightarrow \mathbb{K}$  is defined by  $G(x) = \Gamma(Q_1(x), \dots, Q_c(x))$ , then  $G \in \mathcal{P}$  too.*

*Proof.* For the sake of contradiction, suppose  $G \notin \mathcal{P}$ . Then, for a weight-closed affine constraint consisting of  $K'$  linear forms  $L_1, \dots, L_{K'}$ , there exist  $x_1, \dots, x_\ell$  such that  $(G(L_1(x_1, \dots, x_\ell)), \dots,$



$G(L_{K'}(x_1, \dots, x_\ell))$  which form a witness to  $G \notin \mathcal{P}$ . Note that  $K'$  is a function of only  $K$  and  $W$  because the affine forms characterizing  $\mathcal{P}$  can be made weight  $\leq W$  by a choice of basis for  $\mathbb{K}$  over  $\mathbb{F}$  and then completed into a weight-closed constraint. So, there exists  $x_1, \dots, x_\ell \in \mathbb{K}^n$  such that the tuple  $B = (Q_i(L_j(x_1, \dots, x_\ell)) : j \in [K'], i \in [c]) \in \mathbb{T}^{cK'}$  is a proof of the fact that  $G \notin \mathcal{P}$ .

Now we argue that there exist  $x'_1, \dots, x'_\ell$  such that  $(P_i(L_j(x'_1, \dots, x'_\ell)) : i \in [c], j \in [K])$  equals  $B$ , thus showing that  $F \notin \mathcal{P}$ , a contradiction. Notice that  $B$  satisfies the conditions required of  $\alpha$  in Theorem 3.3. So by Theorem 3.3,

$$\Pr_{x'_1, \dots, x'_\ell} [(P_i(L_j(x'_1, \dots, x'_\ell)) : i \in [c], j \in [K]) = B] > 0$$

if the rank of the factor formed by  $P_1, \dots, P_c$  is more than  $r^{(2.15)}\left(d, \frac{1}{2\|\mathcal{B}\|}\right)$ , where  $\|\mathcal{B}\| = p^{\sum_{i=1}^c (\text{depth}(P_i)+1)}$ .  $\square$

In our applications, we will use Theorem 3.4 for the property of having bounded degree, which is lightly locally characterized by Theorem 2.6.

## 4 List decoding of RM codes

We state the following corollary which we need in the proof to follow. We only state a special case of it which is enough.

**Corollary 4.1** (Corollary 3.3 of [BL14]). *Let  $g : K \rightarrow K$ ,  $\varepsilon > 0$ . Then there exist  $c \leq 1/\varepsilon^2$  functions  $h_1, h_2, \dots, h_c \in \text{RM}_{\mathbb{K}}(n, d)$  such that for every  $f \in \text{RM}_{\mathbb{K}}(n, d)$ , there is a function  $\Gamma_f : \mathbb{K}^c \rightarrow \mathbb{K}$  such that*

$$\Pr_x[\Gamma_f(h_1(x), \dots, h_c(x)) = f(x)] \geq \Pr_x[g(x) = f(x)] - \varepsilon.$$

**Theorem 1.1 (Restated).** Let  $\mathbb{K} = \mathbb{F}_q$  be an arbitrary finite field. Let  $\varepsilon > 0$  and  $d, n \in \mathbb{N}$ . Then,

$$\ell_{\mathbb{K}}(d, n, \delta_{\mathbb{K}}(d) - \varepsilon) \leq c_{q,d,\varepsilon}.$$

*Proof.* We follow the proof structure in [BL14]. Let  $g : \mathbb{K}^n \rightarrow \mathbb{K}$  be a received word. Suppose  $\Pr[g(x) = f(x)] \geq 1 - \delta_{\mathbb{K}}(d) + \varepsilon$ . Apply Corollary 4.1 with approximation parameter  $\varepsilon/2$  gives  $\mathcal{H}_0 = \{h_1, \dots, h_c\} \subseteq \text{RM}_{\mathbb{K}}(n, d)$ ,  $c \leq 4/\varepsilon^2$  such that, for every  $f \in \text{RM}_{\mathbb{K}}(n, d)$ , there is a function  $\Gamma_f : \mathbb{K}^c \rightarrow \mathbb{K}$  satisfying

$$\Pr[\Gamma_f(h_1(x), h_2(x), \dots, h_c(x)) = f(x)] \geq \Pr[g(x) = f(x)] - \varepsilon/2 \geq 1 - \delta_{\mathbb{K}}(d) + \varepsilon/2.$$

Let  $\alpha_1, \alpha_2, \dots, \alpha_r$  be an arbitrary basis for  $\mathbb{K}$  over  $\mathbb{F}$ . Let  $\delta(d) := \delta_{\mathbb{K}}(d)$ . By Proposition 2.1,

$$\Pr[\Gamma'_f(\text{Tr}(\alpha_i h_j(x)) : 1 \leq i \leq r, 1 \leq j \leq c) = F(\text{Tr}(\alpha_i f(x)) : 1 \leq i \leq r)] \geq 1 - \delta(d) + \varepsilon/2,$$

where  $\Gamma'_f : \mathbb{F}^{rc} \rightarrow \mathbb{K}$  and  $F : \mathbb{F}^r \rightarrow \mathbb{K}$ . From here onwards, we identify  $\mathbb{F}$  with  $\mathbb{U}_1$ . Let  $\mathcal{H} = \{\text{Tr}(\alpha_i h_j(x)) : 1 \leq i \leq r, 1 \leq j \leq c\}$  and  $\mathcal{H}_F = \{\text{Tr}(\alpha_i f(x)) : 1 \leq i \leq r\}$ .

Let  $r_1, r_2 : \mathbb{N} \rightarrow \mathbb{N}$  be two non decreasing functions to be specified later, and let  $C_{r,d}^{(2.11)}$  be as given in Lemma 2.11. We will require that for all  $m \geq 1$ ,

$$r_1(m) \geq r_2(C_{r_2,d}^{(2.11)}(m+1)) + C_{r_2,d}^{(2.11)}(m+1) + 1. \quad (2)$$

As a first step, we  $r_1$ -regularize  $\mathcal{H}$  by Lemma 2.11. This gives an  $r_1$ -regular factor  $\mathcal{B}'$  of degree at most  $d$ , defined by polynomials  $H_1, \dots, H_c : \mathbb{K}^n \rightarrow \mathbb{T}$ ,  $c' \leq C_{r_1, d}^{(2.11)}(cr)$  and  $\text{rank}(\mathcal{B}') \geq r_1(c')$ . We denote  $\mathcal{H}' = \{H_1, \dots, H_{c'}\}$ . Let  $\text{depth}(H_i) = k_i$  for  $i \in [c']$ . Let  $G_f : \otimes_{i=1}^{c'} \mathbb{U}_{k_i+1} \rightarrow \mathbb{U}_1$  be defined such that

$$\Gamma_f(h_1(x), \dots, h_c(x)) = G_f(h'_1(x), \dots, h'_{c'}(x)).$$

Next, we will show that  $f$  is measurable with respect to  $\mathcal{H}'$  and this would upper bound the number of such polynomials by  $c'(q, d, \varepsilon)$  independent on  $n$ .

Fix such a polynomial  $f$ . Call  $F_i = \text{Tr}(\alpha_i f)$ . Appealing again to Lemma 2.11, we  $r_2$ -regularize  $\mathcal{B}_f := \mathcal{B}' \cup \mathcal{H}_F$ . We get an  $r_2$ -regular factor  $\mathcal{B}'' \succeq_{\text{syn}} \mathcal{B}'$  defined by the collection  $\mathcal{H}'' = \{H_1, \dots, H_{c'}, H'_1, \dots, H'_{c''}\}$ . Note that it is a syntactic refinement of  $\mathcal{B}'$  as by our choice of  $r_1$ ,

$$\text{rank}(\mathcal{B}') \geq r_1(c') \geq r_2(C_{r_2, d}^{(2.11)}(c' + 1)) + C_{r_2, d}^{(2.11)}(c' + 1) + 1 \geq r_2(|\mathcal{B}''|) + |\mathcal{B}''| + 1.$$

We will choose  $r_2$  such that for all  $m \geq 1$ ,

$$r_2(m) = \max \left( r_d^{(3.1)} \left( \frac{\varepsilon/4}{\left( p^{\lfloor \frac{d-1}{p-1} \rfloor + 1} \right)^m} \right), r_d^{(3.4)}(m) \right). \quad (3)$$

Since each  $F_i$  is measurable with respect to  $\mathcal{B}''$ , there exists  $F' : S \rightarrow \mathbb{U}_1$  such that

$$f(x) = F'(H_1(x), \dots, H_{c'}(x), H'_1(x), \dots, H'_{c''}(x)).$$

Summing up, we have

$$\mathbf{Pr}[G(H_1(x), H_2(x), \dots, H_{c'}(x)) = F'(H_1(x), \dots, H_{c'}(x), H'_1(x), \dots, H'_{c''}(x))] \geq 1 - \delta(d) + \varepsilon/2.$$

We next show that we can have each polynomial in the factor have a disjoint set of inputs. This would simplify the analysis considerably.

**Claim 4.2.** *Let  $x^i, y^j$ ,  $i \in [c'], j \in [c'']$  be pairwise disjoint sets of  $n \in \mathbb{N}$  variables each. Let  $n' = n(c' + c'')$ . Let  $\tilde{f} : \mathbb{K}^{n'} \rightarrow \mathbb{K}$  and  $\tilde{g} : \mathbb{K}^{n'} \rightarrow \mathbb{K}$  be defined as*

$$\tilde{f}(x) = F(H_1(x^1), \dots, H_{c'}(x^{c'}), H'_1(y^1), \dots, H'_{c''}(y^{c''}))$$

and

$$\tilde{g}(x) = G(H'_1(x^1), \dots, H_{c'}(x^{c'})).$$

Then  $\text{deg}(\tilde{f}) \leq d$  and

$$\left| \mathbf{Pr}_{x \in \mathbb{F}^{n'}}[\tilde{f}(x) = \tilde{g}(x)] - \mathbf{Pr}_{x \in \mathbb{F}^n}[f(x) = G_f(h'_1(x), h'_2(x), \dots, h'_c(x))] \right| \leq \varepsilon/4.$$

*Proof.* The bound  $\text{deg}(\tilde{f}) \leq \text{deg}(f) \leq d$  follows from Lemma 3.4 since  $r_2(|\mathcal{H}''|) \geq r_d^{(3.4)}(|\mathcal{H}''|)$ . To establish the bound on  $\mathbf{Pr}[\tilde{f} = \tilde{g}]$ , for each  $s \in S$  let

$$p_1(s) = \mathbf{Pr}_{x \in \mathbb{F}^n}[(h'_1(x), \dots, h'_{c'}(x), h''_1(x), \dots, h''_{c''}(x)) = s].$$

Applying Lemma 3.1 and since our choice of  $r_2$  satisfies  $\text{rank}(\mathcal{H}'') \geq r_d^{(3.1)}(\varepsilon/4|S|)$ , we have that  $p_1$  is nearly uniform over  $S$ ,

$$p_1(s) = \frac{1 \pm \varepsilon/4}{|S|}.$$

Similarly, let

$$p_2(s) = \Pr_{x^1, \dots, x^{c'}, y^1, \dots, y^{c''} \in \mathbb{F}^n} [(h'_1(x^1), \dots, h'_{c'}(x^{c'}), h''_1(y^1), \dots, h''_{c''}(y^{c''})) = s].$$

Note that the rank of the collection of polynomials  $\{h'_1(x^1), \dots, h'_{c'}(x^{c'}), h''_1(y^1), \dots, h''_{c''}(y^{c''})\}$  defined over  $\mathbb{F}^{n'}$  cannot be lower than that of  $\mathcal{H}''$ . Applying Lemma 3.1 again gives

$$p_2(s) = \frac{1 \pm \varepsilon/4}{|S|}.$$

For  $s \in S$ , let  $s' \in \otimes_{i=1}^{c'} \mathbb{U}_{k_i+1}$  be the restriction of  $s$  to first  $c'$  coordinates, that is,  $s' = (s_1, \dots, s_{c'})$ . Thus

$$\begin{aligned} \Pr_{x \in \mathbb{F}^{n'}} [\tilde{f}(x) = \tilde{g}(x)] &= \sum_{s \in S} p_2(s) \mathbf{1}_{F(s)=G_f(s')} \\ &= \sum_{s \in S} p_1(s) \mathbf{1}_{F(s)=G_f(s')} \pm \varepsilon/4 \\ &= \Pr_{x \in \mathbb{F}^n} [f(x) = G_f(h'_1(x), h'_2(x), \dots, h'_c(x))] \pm \varepsilon/4. \end{aligned}$$

□

So, we obtain that

$$\Pr_{x \in \mathbb{F}^{n'}} [\tilde{f}(x) = \tilde{g}(x)] \geq \Pr_{x \in \mathbb{F}^n} [f(x) = G_f(h'_1(x), \dots, h'_c(x))] - \varepsilon/4 \geq 1 - \delta(d) + \varepsilon/4.$$

Next, we need the following variant of the Schwartz-Zippel lemma from [BL14].

**Claim 4.3.** *Let  $d, n_1, n_2 \in \mathbb{N}$ . Let  $f_1 : \mathbb{K}^{n_1+n_2} \rightarrow \mathbb{K}$  and  $f_2 : \mathbb{K}^{n_1} \rightarrow \mathbb{K}$  be such that  $\deg(f_1) \leq d$  and*

$$\Pr[f_1(x_1, \dots, x_{n_1+n_2}) = f_2(x_1, \dots, x_{n_1})] > 1 - \delta(d)$$

*Then,  $f_1$  does not depend on  $x_{n_1+1}, \dots, x_{n_1+n_2}$ .*

With claim 4.3 applied to  $f_1 = \tilde{f}$ ,  $f_2 = \tilde{g}$ ,  $n_1 = nc'$ ,  $n_2 = nc''$ . We obtain that  $\tilde{f}$  does not depend on  $y^1, \dots, y^{c''}$ . Hence,

$$\tilde{f}(x^1, \dots, x^{c'}, y^1, \dots, y^{c''}) = F(H'_1(x^1), \dots, H'_{c'}(x^{c'}), C_1, \dots, C_{c''})$$

where  $C_j = H''_j(0)$  for  $j \in [c'']$ . If we substitute  $x^1 = \dots = x^{c'} = x$  we get that

$$f(x) = F(H'_1(x), \dots, H'_{c'}(x), H''_1(x), \dots, H''_{c''}(x)) = F(H'_1(x), \dots, H'_{c'}(x), C_1, \dots, C_{c''}),$$

which shows that  $f$  is measurable with respect to  $\mathcal{H}'$ , as claimed.

□

## 5 Polynomial decomposition

**Definition 5.1.** Given  $k \in \mathbb{N}$  and  $\Delta = (\Delta_1, \dots, \Delta_k) \in \mathbb{N}^k$  and a function  $\Gamma : \mathbb{K}^k \rightarrow \mathbb{K}$ , a function  $P : \mathbb{K}^n \rightarrow \mathbb{K}$  is  $(k, \Delta, \Gamma)$ -structured if there exist polynomials  $P_1, \dots, P_k : \mathbb{K}^n \rightarrow \mathbb{K}$  with  $\deg(P_i) \leq \Delta_i$  such that for  $x \in \mathbb{K}^n$ , we have

$$P(x) = \Gamma(P_1(x), \dots, P_k(x)).$$

The polynomials  $P_1, \dots, P_k$  form a  $(k, \Delta, \Gamma)$ -decomposition.

The main result we prove is the following.

**Theorem 5.2.** Let  $k \in \mathbb{N}$ . For every  $\Delta = (\Delta_1, \dots, \Delta_k) \in \mathbb{N}^k$  and every function  $\Gamma : \mathbb{K}^k \rightarrow \mathbb{K}$ , there is a randomized algorithm  $A$  that on input  $P : \mathbb{K}^n \rightarrow \mathbb{K}$  of degree  $d$ , runs in time  $\text{poly}_{q,k,\Delta}(n^{d+1})$  and outputs a  $(k, \Delta, \Gamma)$ -decomposition of  $P$  if one exists while otherwise returning NO.

We first show that the notion of rank is robust to hyperplane restrictions over nonprime fields. More precisely, we have the following.

**Lemma 5.3.** Let  $P : \mathbb{K}^n \rightarrow \mathbb{T}$  be an additive polynomial such that  $\text{rank}(P) \geq r$ . Let  $H$  be a hyperplane in  $\mathbb{K}^n$ . Then the restriction of  $P$  to  $H$  has rank at least  $r - q$ .

*Proof.* Without loss of generality, let  $H$  be defined by  $x_1 = 0$ . Let  $P' : \mathbb{K}^{n-1} \rightarrow \mathbb{T}$  be the restriction of  $P$  defined by  $P'(y) = P(0y)$ . Let  $\pi : \mathbb{K}^n \rightarrow \mathbb{K}^{n-1}$  be the map  $\pi(x_1 x_2 \dots x_n) = x_2 \dots x_n$ . Let  $P'' : \mathbb{K}^n \rightarrow \mathbb{T}$  be defined by  $P''(x) = P(x) - P' \circ \pi$ . Then  $P''(x) = 0$  for  $x \in H$ . For  $i \in \mathbb{K} \setminus \{0\}$ , let  $h_i = (i, 0, \dots, 0)$ . Then, for  $y \in H$ , define  $R_j : \mathbb{K}^n \rightarrow \mathbb{T}$  by

$$R_j(y) = P''(y + h_j) = (D_{h_j} P'')(y).$$

Note that  $\deg(R_j) \leq d - 1$ . Now, since  $P(x) = P''(x) + P' \circ \pi(x)$ , we have

$$P(x) = \Gamma(P' \circ \pi, x_1, \{R_y(x) : y \in \mathbb{F}\}).$$

Now, if  $\text{rank}(P') \leq r$ , then  $\text{rank}(P' \circ \pi) \leq r$  and hence  $\text{rank}(P) \leq r + q$ . This finishes the proof.  $\square$

We now start with the proof of Theorem 5.2.

*Proof.* Let  $R_1 : \mathbb{N} \rightarrow \mathbb{N}$  be defined as  $R_1(m) = R_2(c_{2.11}^{(R_1, d)}(m + k)) + c_{2.11}^{(R_1, d)}(m + k) + q$  where  $R_2 : \mathbb{N} \rightarrow \mathbb{N}$  will be specified later.

Let  $\alpha_1, \dots, \alpha_r$  be an arbitrary basis for  $\mathbb{K}$  over  $\mathbb{F}$ . By Proposition 2.1,  $P(x) = \sum_i \beta_i \text{Tr}(\alpha_i P(x))$  for the dual basis  $\beta_1, \dots, \beta_r$ . Set  $f_i(x) = \text{Tr}(\alpha_i P(x))$ . Identifying  $\mathbb{F}$  with  $\mathbb{U}_1$  we treat  $f_i : \mathbb{K}^n \rightarrow \mathbb{T}$ . Regularize  $\{f_1, \dots, f_r\}$  using the algorithm of [BHT15] to find  $R_1$ -regular  $\mathcal{B} = \{g_1, \dots, g_C : \mathbb{K}^n \rightarrow \mathbb{T}\}$  where  $C \leq c_{2.11}^{(R_1, d)}(r)$ . So,  $f_i(x) = G_i(g_1(x), \dots, g_C(x))$  and  $P(x) = \sum_i \alpha_i G_i(g_1(x), \dots, g_C(x))$ . Thus, if  $n \leq Cd$ , then we are done by a brute force search.

Else,  $n > Cd$ . For each  $g_i$ , pick a monomial  $m_i$  with degree  $\deg(P_i)$ . Then there is  $i_0 \in [n]$  such that  $x_{i_0}$  does not appear in any  $g_i$ . Set  $g'_i := g_i|_{x_{i_0} = 0}$ . Let  $\mathcal{B}'$  be the factor defined by the  $g'_i$ 's. Note that  $\deg(g'_i) = \deg(g_i)$  and  $\text{depth}(g'_i) = \text{depth}(g_i)$ . Also, by Lemma 5.3,  $\mathcal{B}'$  is  $R_1 - q$ -regular.

Now, using recursion, we solve the problem on  $n - 1$  variables. That is, decide if for  $P' := P|_{x_{i_0} = 0}$  is  $(k, \Delta, \Gamma)$ -structured. If  $P'$  is not, then  $P$  is not either, so we are done. Else, suppose the algorithm does not output NO.

Say

$$P'(x) = \Gamma(S_1(x), \dots, S_k(x)) = \Gamma'(\text{Tr}(\alpha_j S_i(x)) : i \in [k], j \in [r]),$$

where

$$\Gamma'(a_{ij} : i \in [k], j \in [r]) = \Gamma\left(\sum_j \beta_i a_{ij} : i \in [k]\right).$$

Note that while  $\Gamma : \mathbb{K}^k \rightarrow \mathbb{K}$ , we have  $\Gamma' : \mathbb{F}^{kr} \rightarrow \mathbb{K}$ . Let  $\mathcal{B}_1$  be the factor formed by  $\{\text{Tr}(\alpha_j S_i)\}$ . Via the algorithm of [BHT15], regularize  $\mathcal{B}' \cup \mathcal{B}_1$  using  $R_2 : \mathbb{N} \rightarrow \mathbb{N}$  and we get a syntactic refinement  $\mathcal{B}' \cup \mathcal{B}'_1$  by the choice of  $R_1$ . Let  $\mathcal{B}'_1 = \{s'_1, \dots, s'_D\}$ . where

$$\text{Tr}(\alpha_j S_i) = G_{ij}(g'_i, s'_j : i \in [C], j \in [D]).$$

Choose  $R_2$  large enough such that the map induced by  $\mathcal{B}' \cup \mathcal{B}'_1$  is surjective. Now, fix any  $\ell \in [r]$ . Then,

$$\text{Tr}(\alpha_\ell P') = G_\ell(g'_1, \dots, g'_C) = F_\ell(G_{ij}(g'_i, s'_j)),$$

where  $F_\ell = \text{Tr}(\alpha_\ell \Gamma')$ . Thus, for  $a_1, \dots, a_C, b_1, \dots, b_D \in \mathbb{F}$ ,

$$G_\ell(a_1, \dots, a_C) = F_\ell(G_{ij}(a_1, \dots, b_D) : i \in [C], j \in [D]).$$

Substituting,  $a_i = g_i(x)$  and  $b_j = 0$  we have

$$\text{Tr}(\alpha_\ell P) = G_\ell(g_1, \dots, g_C) = F_\ell(G_{ij}(g_i, 0)).$$

Now,

$$\text{Tr}(\alpha_\ell P) = \text{Tr}(\alpha_\ell \Gamma(Q_i : i \in [k])),$$

where  $Q_i(x) = \sum_{j=1}^r \alpha_j G_{ij}(g'_i, \dots, 0)$ .

Since, this is true for all  $\ell \in [r]$ , we have

$$P(x) = \Gamma(Q_1(x), \dots, Q_k(x)).$$

where  $Q_i$  is defined as above. This finishes the proof.  $\square$

## References

- [AGS03] A. Akavia, S. Goldwasser, and S. Safra. Proving hard-core predicates using list decoding. In *Proc. 44<sup>th</sup> IEEE Symposium on Foundations of Computer Science (FOCS'03)*, 2003. [3](#)
- [AKK<sup>+</sup>05] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing Reed-Muller codes. *IEEE Trans. Inform. Theory*, 51(11):4032–4039, 2005. [5](#)
- [AS03] S. Arora and M. Sudan. Improved low-degree testing and its applications. *Combinatorica*, 23(3):365–426, 2003. [3](#)
- [BCSX11] Arnab Bhattacharyya, Victor Chen, Madhu Sudan, and Ning Xie. Testing linear-invariant non-linear properties. *Theory Comput.*, 7(1):75–99, 2011. [5](#)

- [BFH<sup>+</sup>13] Arnab Bhattacharyya, Eldar Fischer, Hamed Hatami, Pooya Hatami, and Shachar Lovett. Every locally characterized affine-invariant property is testable. In *Proc. 45th Annual ACM Symposium on the Theory of Computing*, pages 429–436, 2013. [1](#), [5](#), [7](#), [8](#), [11](#), [12](#), [13](#)
- [BFL91] László Babai, Lance Fortnow, and Carsten Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1(1):3–40, 1991. [4](#), [5](#)
- [BFL13] Arnab Bhattacharyya, Eldar Fischer, and Shachar Lovett. Testing low complexity affine-invariant properties. In *Proc. 24th ACM-SIAM Symposium on Discrete Algorithms*, pages 1337–1355, 2013. <http://arxiv.org/abs/1201.0330v2>. [5](#), [8](#)
- [BFLS91] László Babai, Lance Fortnow, Leonid A. Levin, and Mario Szegedy. Checking computations in polylogarithmic time. In *Proc. 23rd Annual ACM Symposium on the Theory of Computing*, pages 21–32, New York, 1991. ACM Press. [5](#)
- [BGS10] Arnab Bhattacharyya, Elena Grigorescu, and Asaf Shapira. A unified framework for testing linear-invariant properties. In *Proc. 51st Annual IEEE Symposium on Foundations of Computer Science*, pages 478–487, 2010. [5](#), [8](#)
- [Bha14] Arnab Bhattacharyya. Polynomial decompositions in polynomial time. In *Proc. 22nd Annual European Symposium on Algorithms*, pages 125–136, 2014. [1](#), [4](#), [7](#), [8](#)
- [BHT15] Arnab Bhattacharyya, Pooya Hatami, and Madhur Tulsiani. Algorithmic regularity for polynomials and applications. In *Proc. 26th ACM-SIAM Symposium on Discrete Algorithms*, pages 1870–1889, 2015. [1](#), [4](#), [7](#), [8](#), [11](#), [18](#), [19](#)
- [BL14] Abhishek Bhowmick and Shachar Lovett. List decoding Reed-Muller codes over small fields. Technical report, July 2014. Preprint at <http://arxiv.org/abs/1407.3433>. To appear in STOC ‘15. [1](#), [3](#), [7](#), [15](#), [17](#)
- [BLR93] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *J. Comp. Sys. Sci.*, 47:549–595, 1993. Earlier version in STOC’90. [4](#), [5](#)
- [BMSS11] Eli Ben-Sasson, Ghid Maatouk, Amir Shpilka, and Madhu Sudan. Symmetric LDPC codes are not necessarily locally testable. In *Proc. 26th Annual Conference on Computational Complexity (CCC)*, pages 55–65. IEEE, 2011. [6](#)
- [BTZ10] Vitaly Bergelson, Terence Tao, and Tamar Ziegler. An inverse theorem for the uniformity seminorms associated with the action of  $\mathbb{F}^\omega$ . *Geom. Funct. Anal.*, 19(6):1539–1596, 2010. [2](#)
- [CDG87] Fan R. K. Chung, Persi Diaconis, and Ronald L. Graham. Random walks arising in random number generation. *Ann. Probab.*, 15(3):1148–1165, 07 1987. [1](#)
- [dW08] Ronald de Wolf. *A Brief Introduction to Fourier Analysis on the Boolean Cube*. Number 1 in Graduate Surveys. Theory of Computing Library, 2008. [1](#)
- [Eli57] P. Elias. List decoding for noisy channels. Technical Report 335, Research Laboratory of Electronics, MIT, 1957. [2](#)



- [FGL<sup>+</sup>96] Uriel Feige, Shafi Goldwasser, László Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996. [5](#)
- [GGR98] Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45:653–750, 1998. [4](#)
- [GK11] Oded Goldreich and Tali Kaufman. Proximity oblivious testing and the role of invariances. In *Studies in Complexity and Cryptography*, pages 173–190. 2011. [5](#)
- [GKS12] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. Succinct representation of codes with applications to testing. *SIAM Journal on Discrete Mathematics*, 26(4):1618–1634, 2012. [6](#)
- [GKZ08] P. Gopalan, A. Klivans, and D. Zuckerman. List decoding Reed-Muller codes over small fields. In *Proc. 40<sup>th</sup> ACM Symposium on the Theory of Computing (STOC’08)*, pages 265–274, 2008. [1](#), [3](#), [4](#)
- [GL89] O. Goldreich and L. Levin. A hard-core predicate for all one-way functions. In *Proc. 21<sup>st</sup> ACM Symposium on the Theory of Computing*, pages 25–32, 1989. [2](#), [3](#)
- [Gop10] P. Gopalan. A Fourier-analytic approach to Reed-Muller decoding. In *Proc. 51<sup>st</sup> IEEE Symp. on Foundations of Computer Science (FOCS’10)*, pages 685–694, 2010. [3](#)
- [GOS<sup>+</sup>09] Parikshit Gopalan, Ryan O’Donnell, Rocco A. Servedio, Amir Shpilka, and Karl Wimmer. Testing Fourier dimensionality and sparsity. In *Proc. 36th Annual International Conference on Automata, Languages, and Programming*, pages 500–512, 2009. [5](#)
- [Gow98] William T. Gowers. A new proof of Szemerédi’s theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998. [2](#)
- [Gow01] William T. Gowers. A new proof of Szemerédi’s theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001. [2](#)
- [GR11] Oded Goldreich and Dana Ron. On proximity oblivious testing. *SIAM J. Comput.*, 40(2):534–566, 2011. [5](#)
- [Gre05] Ben Green. Finite field models in additive combinatorics. In Bridget S Webb, editor, *Surveys in combinatorics 2005*, pages 1–27. Cambridge Univ. Press, 2005. [2](#)
- [GRS00] O. Goldreich, R. Rubinfeld, and M. Sudan. Learning polynomials with queries: The highly noisy case. *SIAM J. Discrete Math.*, 13(4):535–570, 2000. [3](#)
- [GT08] Ben Green and Terence Tao. An inverse theorem for the Gowers  $U^3$ -norm. *Proc. Edin. Math. Soc.*, 51:73–153, 2008. [2](#)
- [GT09] B. Green and T. Tao. The distribution of polynomials over finite fields, with applications to the gowers norms. *Contrib. Discrete Math*, 4(2):1–36, 2009. [2](#)
- [GT10] Ben Green and Terence Tao. Linear equations in primes. *Ann. of Math.*, 171:1753–1850, 2010. [2](#)

- [GTZ11] Ben Green, Terence Tao, and Tamar Ziegler. An inverse theorem for the Gowers  $U^4$ -norm. *Glasgow Math. J.*, 53(1):1–50, 2011. <http://arxiv.org/abs/0911.5681>. 2
- [GTZ12] Ben Green, Terence Tao, and Tamar Ziegler. An inverse theorem for the Gowers  $U^{s+1}$ -norm. *Ann. of Math.*, 176(2):1231–1372, 2012. 2
- [Gur04] V. Guruswami. *List Decoding of Error-Correcting Codes*, volume 3282 of *Lecture Notes in Computer Science*. Springer, 2004. 2
- [Gur06] V. Guruswami. *Algorithmic Results in List Decoding*, volume 2 of *Foundations and Trends in Theoretical Computer Science*. Now Publishers, 2006. 2
- [Has01] Johan Hastad. Some optimal inapproximability results. *Journal of the ACM (JACM)*, 48(4):798–859, 2001. 1
- [HK05] Bernard Host and Bryna Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.*, 161(1):397–488, 2005. 2
- [HL13] Hamed Hatami and Shachar Lovett. Estimating the distance from testable affine-invariant properties. In *Proc. 54th Annual IEEE Symposium on Foundations of Computer Science*, pages 237–242. IEEE, 2013. 5
- [Jac97] J. Jackson. An efficient membership-query algorithm for learning DNF with respect to the uniform distribution. *Journal of Computer and System Sciences*, 55:414–440, 1997. 3
- [KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions. In *Proc. 29th Annual IEEE Symposium on Foundations of Computer Science*, pages 68–80, 1988. 1
- [KL05] Tali Kaufman and Simon Litsyn. Almost orthogonal linear codes are locally testable. In *Proc. 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 317–326. IEEE, 2005. 6
- [KL08] Tali Kaufman and Shachar Lovett. Worst case to average case reductions for polynomials. In *Proc. 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 166–175, 2008. 2
- [KM93a] D. Koller and N. Megiddo. Constructing small sample spaces satisfying given constraints. In *Proc. 25th Annual ACM Symposium on the Theory of Computing*, pages 268–277, 1993. 1
- [KM93b] E. Kushilevitz and Y. Mansour. Learning decision trees using the Fourier spectrum. *SIAM Journal of Computing*, 22(6):1331–1348, 1993. 3
- [KR06] Tali Kaufman and Dana Ron. Testing polynomials over general fields. *SIAM J. on Comput.*, 36(3):779–802, 2006. 7, 10
- [KS08] Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In *Proc. 40th Annual ACM Symposium on the Theory of Computing*, pages 403–412, 2008. 1, 5, 6

- [MOO10] Elchanan Mossel, Ryan O’Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. of Math.*, 171(1), 2010. [1](#)
- [NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: efficient constructions and applications. *SIAM J. on Comput.*, 22(4):838–856, 1993. Earlier version in STOC’90. [1](#)
- [NS05] Michael Navon and Alex Samorodnitsky. On delarte’s linear programming bounds for binary codes. In *Proc. 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 327–338, 2005. [1](#)
- [PW04] R. Pellikaan and X. Wu. List decoding of q-ary Reed-Muller codes. *IEEE Transactions on Information Theory*, 50(4):679–682, 2004. [3](#)
- [RS96] Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM J. on Comput.*, 25:252–271, 1996. [4](#), [5](#)
- [SS71] Arnold Schönhage and Volker Strassen. Schnelle multiplikation grosser zahlen. *Computing*, 7:281–292, 1971. [1](#)
- [Šte00] Daniel Štefankovič. Fourier transform in computer science. Master’s thesis, University of Chicago, 2000. [1](#)
- [STV01] M. Sudan, L. Trevisan, and S. P. Vadhan. Pseudorandom generators without the XOR lemma. *J. Comput. Syst. Sci.*, 62(2):236–266, 2001. [2](#), [3](#)
- [SU05] Ronen Shaltiel and Christopher Umans. Simple extractors for all min-entropies and a new pseudorandom generator. *J. ACM*, 52(2):172–216, 2005. [3](#)
- [Sud97] M. Sudan. Decoding of Reed-Solomon codes beyond the error-correction bound. *Journal of Complexity*, 13(1):180–193, 1997. [2](#)
- [Sud00] M. Sudan. List decoding: Algorithms and applications. *SIGACT News*, 31(1):16–27, 2000. [2](#)
- [Tao12] Terence Tao. *Higher Order Fourier Analysis*, volume 142 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012. [2](#)
- [Tre03] L. Trevisan. List-decoding using the XOR lemma. In *Proc. 44<sup>th</sup> IEEE Symposium on Foundations of Computer Science (FOCS’03)*, page 126, 2003. [2](#)
- [TSZS01] A. Ta-Shma, D. Zuckerman, and S. Safra. Extractors from Reed-Muller codes. In *Proc. 42<sup>nd</sup> IEEE Symp. on Foundations of Computer Science (FOCS’01)*, pages 638–647, 2001. [3](#)
- [TZ10] Terence Tao and Tamar Ziegler. The inverse conjecture for the Gowers norm over finite fields via the correspondence principle. *Analysis & PDE*, 3(1):1–20, 2010. [2](#)
- [TZ12] 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. [2](#), [6](#), [7](#), [9](#), [12](#)

- [Vad12] Salil P. Vadhan. Pseudorandomness. *Foundations and Trends in Theoretical Computer Science*, 7(1-3):1–336, 2012. [2](#)
- [Wey14] Hermann Weyl. Über ein problem aus dem gebiete der diophantischen approximationen. *Nachr. Ges. Wiss. Gttingen*, pages 234–244, 1914. [2](#)
- [Woz58] J. Wozencraft. List decoding. Technical Report 48:90-95, Quarterly Progress Report, Research Laboratory of Electronics, MIT, 1958. [2](#)
- [Yos14] Yuichi Yoshida. A characterization of locally testable affine-invariant properties via decomposition theorems. In *Proc. 46th Annual ACM Symposium on the Theory of Computing*, pages 154–163, 2014. [5](#)