Electronic Colloquium on Computational Complexity, Report No. 18 (2014)

Polynomial decompositions in polynomial time

Arnab Bhattacharyya

Indian Institute of Science
Bangalore, India
arnabb@csa.iisc.ernet.in

February 13, 2014

Abstract

Fix a prime p. Given a positive integer k, a vector of positive integers $\mathbf{\Delta} = (\Delta_1, \Delta_2, \ldots, \Delta_k)$ and a function $\Gamma : \mathbb{F}_p^k \to \mathbb{F}_p$, we say that a function $P : \mathbb{F}_p^n \to \mathbb{F}_p$ is $(k, \mathbf{\Delta}, \Gamma)$-*structured* if there exist polynomials $P_1, P_2, \ldots, P_k : \mathbb{F}_p^n \to \mathbb{F}_p$ with each $\deg(P_i) \leqslant \Delta_i$ such that for all $x \in \mathbb{F}_p^n$,

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

For instance, an n-variate polynomial over the field $\mathbb{F}_p$ of total degree d factors nontrivially exactly when it is $(2, (d-1, d-1), \mathsf{prod})$-structured where $\mathsf{prod}(a, b) = a \cdot b$.

We show that if $p > d$, then for **any** fixed $k, \mathbf{\Delta}, \Gamma$, we can decide whether a given polynomial $P(x_1, x_2, \ldots, x_n)$ of degree d is $(k, \mathbf{\Delta}, \Gamma)$-structured and if so, find a witnessing decomposition. The algorithm takes poly(n) time. Our approach is based on higher-order Fourier analysis.

1 Introduction

(Linear) Fourier analysis over a finite field $\mathbb{F}_p$ studies the structure of exponentials of linear functions, i.e. functions of the form $\omega^{\ell(x)}$ where $\ell : \mathbb{F}_p^n \to \mathbb{F}_p$ is a linear function and $\omega = e^{2\pi i/p}$ is the p'th root of unity. Fourier analysis over finite fields has, by now, a rich history of widespread success in theoretical computer science. Here is a sample of applications: coding theory, computational learning theory, influence of variables in boolean functions, probabilistically checkable proofs, cryptography, communication complexity, and quantum computing. For more, consult the lovely survey of de Wolf [dW08].

Higher-order Fourier analysis is a novel generalization of Fourier analysis. In higher-order Fourier analysis over finite fields, we study the structure of exponentials of low-degree polynomials, i.e. functions of the form $\omega^{Q(x)}$ where $Q : \mathbb{F}_p^n \to \mathbb{F}_p$ is a polynomial[1] of bounded degree. The theory (although conceptually originating with the classical equidistribution results of Weyl) really got its start from the spectacular proof by Gowers of Szemerédi's theorem [Gow98, Gow01], where the Gowers norm was introduced. Another significant influence was the work of Host and Kra [HK05] in ergodic theory. Subsequently, Green, Tao and Ziegler through several works [GT08, GT10, GTZ11, GTZ, TZ10, TZ12] largely completed the research program of understanding the relationships between different aspects of the theory. The book [Tao12] by Tao on the subject surveys the current state of knowledge.

Green, Tao and Ziegler applied higher-order Fourier analysis to find asymptotics for various linear patterns in the prime numbers. In theoretical computer science, low-degree polynomials over finite fields has long been under consideration due to the use of arithmetization. Specifically, there is a long history of testing whether a function is correlated with a low-degree polynomial, and higher-order Fourier analysis can be immediately phrased in this context. In fact, it was shown in [BCSX11, BGS10, BFL13, BFH$^+$13] that higher-order Fourier analysis can be used to analyze tests not only for low-degreeness but also for any locally

[1] Throughout, our functions are of n variables over $\mathbb{F}_p$, where n is growing but p is fixed.

ISSN 1433-8092

characterized affine-invariant property (see the cited papers for definitions). Besides property testing, the Gowers norm has also been used in computer science to show worst case to average case reductions for polynomials [KL08] and XOR lemmas for polynomials [VW08].

In this paper, we demonstrate a new algorithmic application of higher-order Fourier analysis. Consider the following family of properties of functions over a finite field $\mathbb{F}_p$ of fixed prime order p.

Definition 1.1. *Given a positive integer k, a vector of positive integers $\mathbf{\Delta} = (\Delta_1, \Delta_2, \ldots, \Delta_k)$ and a function $\Gamma : \mathbb{F}_p^k \to \mathbb{F}_p$, we say that a function $P : \mathbb{F}_p^n \to \mathbb{F}_p$ is* $(k, \mathbf{\Delta}, \Gamma)$-structured *if there exist polynomials $P_1, P_2, \ldots, P_k : \mathbb{F}_p^n \to \mathbb{F}_p$ with each $\deg(P_i) \leqslant \Delta_i$ such that for all $x \in \mathbb{F}_p^n$,*

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

The polynomials $P_1, \ldots, P_k$ are said to form a $(k, \mathbf{\Delta}, \Gamma)$-decomposition.

For instance, an n-variate polynomial over the field $\mathbb{F}_p$ of total degree d factors nontrivially exactly when it is $(2, (d-1, d-1), \mathsf{prod})$-structured where $\mathsf{prod}(a, b) = a \cdot b$. Informally, a *degree-structural property* refers to a property from the family of $(k, \mathbf{\Delta}, \Gamma)$-structured properties.

Our main result is that every degree-structural property can be decided in polynomial time:

Theorem 1.2. *For every positive integer k, every vector of positive integers $\mathbf{\Delta} = (\Delta_1, \Delta_2, \ldots, \Delta_k)$ and every function $\Gamma : \mathbb{F}_p^k \to \mathbb{F}_p$, there is a deterministic algorithm $\mathcal{A}_{k,\mathbf{\Delta},\Gamma}$ that takes as input a polynomial $P : \mathbb{F}_p^n \to \mathbb{F}_p$ of degree $d < p$, runs in time polynomial in n, and outputs a $(k, \mathbf{\Delta}, \Gamma)$-decomposition of P if one exists while otherwise returning* NO.

1.1 Discussion

The main result is surprisingly strong in that it holds for every k, $\mathbf{\Delta}$ and Γ. Thus, for instance, it immediately implies a (deterministic) poly(n)-time algorithm for factoring an n-variate polynomial of degree d over $\mathbb{F}_p$, as long as $p > d$ and p and d are fixed. Also, as we shall observe in Section 3, the proof of Theorem 1.2 implies a polynomial time algorithm for deciding whether a d-dimensional tensor over $\mathbb{F}_p$ has rank at most r, where d, p and r are constants and $d < p$.

We must remark that these results on factoring and tensor rank are not new, in the sense that there were already algorithms known for stronger versions of these two problems. Specifically, for deciding constant tensor rank, Karnin and Shpilka [KS09] showed a polynomial time algorithm for the more general problem of reconstructing multilinear $\Sigma\Pi\Sigma$ circuits with a constant number of multiplication gates. And for factoring multivariate polynomials over finite fields, it is known [vzGK85, Sud12] how to factor in time poly(n, d, p) deterministically and in time poly$(n, d, \log p)$ probabilistically.

However, Theorem 1.2 gives polynomial time algorithms for a whole host 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. Thus, these problems become useful targets for reductions in future. Our main result can be described as a *blackbox reconstruction algorithm* as in [KS09], in the sense that the algorithm is given blackbox query access to the polynomial and it runs in time linear in the dense representation of the polynomial (i.e., input size is measured as $\binom{n+d}{d}$). The property of having $(k, \mathbf{\Delta}, \Gamma)$-structure is also similar in spirit to a function having a *concise representation*, a notion introduced by Diakonikolas et al. in [DLM$^+$07]. We leave open as to whether there are formal connections here.

There are two main questions raised by Theorem 1.2:

1. Does Theorem 1.2 hold when $p \leqslant d$? The main difficulty here seems technical and stems from the fact that the proof of the Gowers inverse theorem for polynomials is currently very non-constructive [TZ12] when $p \leqslant d$, in contrast to the case of high characteristic [GT09].

2. Is there an analogous theorem when n is fixed and d and p are growing? Such questions are probably very difficult, because over $\mathbb{Z}_n$, we do not even know how to deterministically factorize the univariate polynomial $x^2 - a$, for a given $a \in \mathbb{Z}_n$. In fact, in recent work, Kopparty, Saraf and Shpilka [KSS14]

have shown an equivalence between deterministic factorization of multivariate polynomials and derandomization of polynomial identity testing, a long-standing challenge. Polynomial time *randomized* algorithms exist for factorization of course but are not known to exist for arbitrary degree-structural properties over large fields. In particular, Neeraj Kayal (personal communication) asks whether it is possible in randomized polynomial time to decompose a univariate polynomial $P : \mathbb{F}_p \to \mathbb{F}_p$ of degree $n < p$ as $P = P_1 \cdot P_2 + P_3 \cdot P_4$ where P_1, P_2, P_3, P_4 are of degree $< n$. Even an average-case algorithm would be interesting, meaning P is known to be formed out of random polynomials P_1, P_2, P_3, P_4, and the task is to recover them given access to P.

1.2 Techniques

The proof of Theorem 1.2 is actually a straightforward combination of ideas from [BFH+13] and [BHT13]. In [BFH+13], it was shown that any degree-structural property is constant query *testable*. That is, for all $k, \mathbf{\Delta}$, and Γ, one can decide correctly, with probability at least 2/3, whether a given function is $(k, \mathbf{\Delta}, \Gamma)$-structured or whether it is 1%-far from any $(k, \mathbf{\Delta}, \Gamma)$-structured function, by querying the input function's value on only a constant number of points. The main contribution of [BFH+13] is a reduction from the testability problem to the following combinatorial problem:

> Does there exist $s = s(k, \mathbf{\Delta}, \Gamma)$ such that a function is $(k, \mathbf{\Delta}, \Gamma)$-structured if and only if so is the restriction of the function to all affine subspaces of dimension s?

In other words, is it true that if a function f is not $(k, \mathbf{\Delta}, \Gamma)$-structured, there exists a subspace of dimension s on which the restriction of f is also not $(k, \mathbf{\Delta}, \Gamma)$-structured? [BFH+13] gave a positive answer to this problem (thus showing, by virtue of their main reduction, that degree-structure is testable). One can view their answer as a solution to the search problem of finding an s-dimensional subspace on which the function is not degree-structural. However, their proof is non-constructive, in the sense that no non-trivial algorithm is provided for finding the witnessing s-dimensional subspace.

The main reason for this non-constructivity is that the proof in [BFH+13] uses a notion called the *rank* of an n-variate polynomial over $\mathbb{F}_p$ (see Section 2.2.3) which we do not know how to compute in time polynomial in n. However, in [BHT13], it was noticed that when the polynomial degree is smaller than the field characteristic, instead of the rank of a polynomial, one could equally well work with the *Gowers uniformity norm* (see Section 2.2.2) of the polynomial, and the Gowers norm can be estimated upto constant additive error with good probability by evaluating the polynomial on a constant number of random samples. Via this approach, [BHT13] found an algorithmic *regularity lemma* for degree-d n-variate polynomials (see Section 2.3) that runs in time $O(n^d)$ when $d < p$.

The main contribution of this work is observing that a decision algorithm, not just a property test, follows if one plugs in the algorithmic regularity lemma into the analysis of [BFH+13]. The algorithm proceeds by restricting the input function to a carefully chosen constant dimensional subspace, solving the (finite) decomposition problem on that subspace, and then "lifting" the solution back to $\mathbb{F}_p^n$. At a high-level, this algorithm is very similar to Kaltofen's factorization algorithm [Kal95], where the polynomial is first restricted to a random two-dimensional subspace, then factored using bivariate factorization algorithms, and then lifted back to the original space. From this perspective, we show that the "restrict-solve-lift" paradigm can be used for any degree-structural decomposition problem, not just factorization (at least when the field order is a constant prime but larger than the degree of the input polynomial).

We hope that this work brings the techniques of higher-order Fourier analysis to the attention of a wider audience in computer science.

1.3 Organization

In Section 2, we lay out the results from higher-order Fourier analysis which will be needed in the proof of Theorem 1.2. We use this toolkit to finish the proof in Section 3.

2 Technical Preliminaries

From a bird's eye viewpoint, higher-order Fourier analysis is a study of how the analytic properties of a collection of polynomial relate to the collection's algebraic/combinatorial structure. We make precise all the needed notions in the subsections below.

2.1 Polynomial Factors

Given an integer $d \geqslant 0$, we say that a function $P : \mathbb{F}_p^n \to \mathbb{F}_p$ is a *polynomial of degree d* if $D_{h_1} D_{h_2} \cdots D_{h_d} P = 0$ for all $h_1, h_2, \ldots, h_d \in \mathbb{F}_p^n$, where $D_h P : \mathbb{F}_p^n \to \mathbb{F}_p$ is defined as:

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

It is easy to check that P is a polynomial of degree d exactly when it can be written as:

$$\sum_{\substack{0 \leqslant i_1, i_2, \ldots, i_n < p: \\ \sum i_j \leqslant d}} c_{i_1, i_2, \ldots, i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$$

Thus, polynomials of degree 1 correspond to all the linear polynomials over $\mathbb{F}_p^n$ and so on.

Given a set X, a *factor* is a partition of X into finitely many subset called *atoms.* A finite collection of functions $\phi_1, \ldots, \phi_C$ from X to some other space Y naturally defines a factor $\mathcal{B} = \mathcal{B}_{\phi_1, \ldots, \phi_C}$ whose atoms are sets of the form $\{x : (\phi_1(x), \phi_2(x), \ldots, \phi_C(x)) = (y_1, y_2, \ldots, y_C)\}$ 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, the atom containing x is identified with $\mathcal{B}(x) = (\phi_1(x), \ldots, \phi_C(x))$.

Definition 2.1. *If $P_1, \ldots, P_C : \mathbb{F}_p^n \to \mathbb{F}$ is a sequence of polynomials, then the factor $\mathcal{B} = \mathcal{B}_{P_1, \ldots, P_C}$ is called a* polynomial factor*. The* complexity *of $\mathcal{B}$, denoted $|\mathcal{B}|$, is the number of defining polynomials, C. The* degree *of $\mathcal{B}$ is the maximum degree among its defining polynomials $P_1, \ldots, P_C$. Also, $\|\mathcal{B}\| = p^C$ is called the* order *of $\mathcal{B}$; the number of nonempty atoms of $\mathcal{B}$ is bounded by $\|\mathcal{B}\|$.*

We will often drop the adjective *polynomial*, and use "factors" to mean "polynomial factors".

2.2 Three notions of polynomial pseudorandomness

The main results of higher-order Fourier analysis revolve around three measures of pseudorandomness for polynomial factors. Each is a statistical test that is perfectly met by truly random polynomial factors, and the question is how well are they met by factors of degree d.

2.2.1 Bias

The first pseudorandomness measure is the familiar notion of bias, generalizing the definition of Naor and Naor [NN93] over $\mathbb{F}_2$.

Definition 2.2 (Unbiased)**.** *The* bias *of a function $F : \mathbb{F}_p^n \to \mathbb{F}_p$ is:*

$$\mathsf{bias}(F) = \left| \mathop{\mathbf{E}}_{x \in \mathbb{F}_p^n} [\mathsf{e}(F(x))] \right|$$

*Given a function $\beta : \mathbb{Z}^+ \to (0,1)$ and a polynomial factor $\mathcal{B}$ defined by a sequence of polynomials $P_1, \ldots, P_C : \mathbb{F}_p^n \to \mathbb{F}$, the factor $\mathcal{B}$ is said to be β-*unbiased *if for every $(a_1, \ldots, a_C) \in \{0, \ldots, p-1\}^C \setminus \{0^C\}$,*

$$\mathsf{bias}\left(\sum_{i=1}^{C} a_i P_i\right) < \beta(C).$$

The following facts are straightforward and folklore.

Lemma 2.3 (Equidistribution)**.** *Given $\beta : \mathbb{Z}^+ \to (0,1)$, let $\mathcal{B}$ be a β-unbiased polynomial factor of complexity C. For any $b \in \mathbb{F}^C$:*

$$\mathop{\mathbf{Pr}}_x[\mathcal{B}(x) = b] = \frac{1}{\|\mathcal{B}\|} \pm \beta(C).$$

Corollary 2.4 (Atom Dispersal)**.** *If $\beta(k) = \frac{1}{2p^k}$ and $\mathcal{B}$ is a β-unbiased polynomial factor, then all of the $\|\mathcal{B}\|$ atoms of $\mathcal{B}$ are nonempty.*

Moreover, if $\deg(\mathcal{B}) < p$, then polynomials defining an unbiased factor $\mathcal{B}$ can be composed with affine forms to yield another unbiased factor, unless the compositions cause a trivial linear dependency among the polynomials. Precisely:

Lemma 2.5 (Unbiased affine composition, Lemma 5.1 of [HL11])**.** *Suppose $d < p$. Let $\mathcal{B}$ be a polynomial factor of degree d defined by the sequence of polynomials $P_1, \ldots, P_C : \mathbb{F}_p^n \to \mathbb{F}_p$. Suppose $\mathcal{B}$ is β-unbiased for some $\beta : \mathbb{Z}^+ \to (0,1)$. For an integer $k \geqslant 1$, consider any set of coefficients $\Lambda = \{\lambda_{i,S} : i \in [C], S \subseteq [k]\}$, and define $P_\Lambda : (\mathbb{F}_p^n)^{k+1} \to \mathbb{F}_p$ as:*

$$P_\Lambda(x, y_1, \ldots, y_k) = \sum_{i=1}^{C} \sum_{S \subseteq [k]} \lambda_{i,S} P_i \left(x + \sum_{j \in S} y_j \right).$$

Then, either $P_\Lambda \equiv 0$ or $\mathsf{bias}(P_\Lambda) < \beta(C)$.

The following theorem[2], proved in [BFH+13], uses the above Lemma 2.5 to show that a function of an unbiased factor of degree d has the degree which one would expect from a generic collection of polynomials of degree d.

Theorem 2.6 (Degree Preservation, Theorem 4.1 of [BFH+13])**.** *For any positive integer $d < p$, there is a function $\alpha^d_{2.6} : \mathbb{Z}^+ \to (0,1)$ such that the following is true. Let $\mathcal{B}$ be any factor defined by polynomials $P_1, \ldots, P_C : \mathbb{F}_p^n \to \mathbb{F}_p$ of degree $\leqslant d$. Suppose $\mathcal{B}$ is $\alpha^d_{2.6}$-unbiased. Let $\Gamma : \mathbb{F}_p^C \to \mathbb{F}_p$ be an arbitrary function. Define the polynomial $F : \mathbb{F}_p^n \to \mathbb{F}_p$ by $F(x) = \Gamma(\mathcal{B}(x))$.*

Then, for any factor $\mathcal{B}'$ of degree $\leqslant d$, if $G : \mathbb{F}_p^n \to \mathbb{F}_p$ is the polynomial $G(x) = \Gamma(\mathcal{B}'(x))$, it holds that $\deg(G) \leqslant \deg(F)$.

2.2.2 Uniformity

Bias is often a very weak measure of pseudorandomness: the bias of any linear function is 0, even though it is clearly not a random function. We could strengthen low bias by additionally requiring that all the Fourier coefficients be small, which would ensure that the function is not (correlated with) a linear function. Continuing down this path leads us to the notion of uniformity, which measures the correlation of a function with polynomials of bounded degree.

Definition 2.7 (Multiplicative Derivative)**.** *Given a function $f : \mathbb{F}_p^n \to \mathbb{C}$ and an element $h \in \mathbb{F}_p^n$, the* multiplicative derivative of f in direction h *is the function $\Delta_h f : \mathbb{F}_p^n \to \mathbb{C}$ satisfying $\Delta_h f(x) = f(x+h)\overline{f(x)}$ for all $x \in \mathbb{F}_p^n$.*

Definition 2.8 (Uniformity)**.** *Given a function $f : \mathbb{F}_p^n \to \mathbb{C}$ and an integer $d \geqslant 1$, the* Gowers uniformity norm of order d *for f is given by:*

$$\|f\|_{U^d} = \left| \mathop{\mathbf{E}}_{h_1, \ldots, h_d \in \mathbb{F}_p^n} \mathop{\mathbf{E}}_{x \in \mathbb{F}_p^n} [(\Delta_{h_1} \Delta_{h_2} \cdots \Delta_{h_d} f)(x)] \right|^{1/2^d}$$

[2] Versions of Lemma 2.5 and Theorem 2.6 are true when $p \leqslant d$ also, as shown in [BFH+13], but in that case, they require the stronger assumption of uniformity (see next section) instead of unbiasedness.

*Given a function $\gamma : \mathbb{Z}^+ \to (0,1)$ and a polynomial factor $\mathcal{B}$ defined by a sequence of polynomials $P_1, \ldots, P_C : \mathbb{F}_p^n \to \mathbb{F}$, the factor $\mathcal{B}$ is said to be γ-*uniform *if for every $(a_1, \ldots, a_C) \in \{0, \ldots, p-1\}^C \setminus \{0^C\}$,*

$$\left\| \mathsf{e}\left(\sum_{i=1}^{C} a_i P_i\right)\right\|_{U^d} < \gamma(C)$$

where $d = \max_i \deg(a_i P_i)$.

Note that $\mathsf{bias}(P) = \|\mathsf{e}(P)\|_{U^1}$ for any $P : \mathbb{F}_p^n \to \mathbb{F}$. Moreover, it holds that $\|f\|_{U^d} \leqslant \|f\|_{U^{d+1}}$ for any $f : \mathbb{F}_p^n \to \mathbb{C}$ and $d \geqslant 1$ [Gow98]. So:

Lemma 2.9 (Uniformity implies unbiased)**.** *If $\mathcal{B}$ is a polynomial factor that is γ-uniform for some function $\gamma : \mathbb{Z}^+ \to (0,1)$, then $\mathcal{B}$ is also γ-unbiased.*

2.2.3 Regularity

A third measure of pseudorandomness was introduced by Green and Tao [GT09] as a bridge between the algebraic structure of polynomials and the analytic notions of bias and uniformity.

Definition 2.10 (Regularity)**.** *Given a function $F : \mathbb{F}_p^n \to \mathbb{F}_p$ and an integer $d > 1$, the* d-rank of F, *denoted* $\mathsf{rank}_d(F)$, *is defined to be the smallest integer r such that there exist polynomials $Q_1, \ldots, Q_r : \mathbb{F}_p^n \to \mathbb{F}_p$ of degree $\leqslant d-1$ and a function $\Gamma : \mathbb{F}_p^r \to \mathbb{F}_p$ satisfying $P(x) = \Gamma(Q_1(x), \ldots, Q_r(x))$. If $d = 1$, the 1-rank is defined to be ∞ if F is non-constant and 0 otherwise.*

Given a function $R : \mathbb{Z}^+ \to \mathbb{Z}^+$ and a polynomial factor $\mathcal{B}$ defined by a sequence of polynomials $P_1, \ldots, P_C : \mathbb{F}_p^n \to \mathbb{F}_p$, the factor $\mathcal{B}$ is said to be R-regular *if for every $a_1, \ldots, a_C \in \{0, 1, \ldots, p-1\}^C \setminus \{0^C\}$,*

$$\mathsf{rank}_d\left(\sum_{i=1}^{C} a_i P_i\right) > R(C)$$

where $d = \max_i \deg(a_i P_i)$. Also, under these conditions, the rank of $\mathcal{B}$ *is at least $R(C)$.*

Regularity and uniformity turn out to be essentially equivalent, due to the following two remarkable theorems. The first theorem is folklore and essentially due to (linear) Fourier analysis.

Theorem 2.11 (Uniformity implies regularity)**.** *Suppose that $p > d$ and let $R : \mathbb{Z}^+ \to \mathbb{Z}^+$ be any non-decreasing function. Then, there is a function $\gamma_{2.11}^{d,R} : \mathbb{Z}^+ \to (0,1)$ such that the following holds. Any polynomial factor of degree d that is $\gamma_{2.11}^{d,R}$-uniform is also R-regular.*

The second theorem is really the key piece of higher-order Fourier analysis for polynomials.

Theorem 2.12 (Regularity implies uniformity, Proposition 6.1 of [GT09])**.** *Suppose that $p > d$, and let $\gamma : \mathbb{Z}^+ \to (0,1)$ be any non-increasing function. Then, there is a function $R_{2.12}^{d,\gamma} : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that the following holds. Any polynomial factor of degree d that is $R_{2.12}^{d,\gamma}$-regular is also γ-uniform.*

Remark 2.13. *Importantly, when $d < p$, $\gamma_{2.11}^{d,R}$ is explicitly known, given d and R. In other words, given access to an evaluation oracle for R, $\gamma_{2.11}^{d,R}$ is polynomial-time computable. Similarly, $R_{2.12}^{d,\gamma}$ is explicitly known.*

While unbiasedness and uniformity are analytic properties of a factor, regularity is an algebraic notion and is hence more amenable to algebraic operations on the function. For instance, we have:

Lemma 2.14 (Subspace Restriction, Lemma 2.13 of [BFH+13])**.** *Suppose $P : \mathbb{F}_p^n \to \mathbb{F}_p$ is a polynomial of degree d and rank r, where $r > p+1$. Let A be a hyperplane in $\mathbb{F}_p^n$, and denote by P' the restriction of P to A. Then, P' is a polynomial of degree d and rank $\geqslant r-p$, unless $d = 1$ and P is constant on A.*

2.3 Algorithmic Regularity Lemma

The celebrated Szemerédi graph regularity lemma [Sze78] permits the decomposition of an arbitrary graph into bipartite subgraphs which are regular (in the graph-theoretic sense). Actually, more strongly, given any partitioning of the vertices, Szemerédi's regularity lemma, roughly speaking, yields a refined partitioning such that most of the bipartite graphs between the new parts are regular.

One can carry out an analogous type of refinement for our notions of regularity also. First, let us specify what we mean by refinements of a factor.

Definition 2.15 (Semantic and syntactic refinements)**.** *$\mathcal{B}'$ is called a* semantic refinement *(or simply, a* refinement*) if the partition induced by $\mathcal{B}'$ is a combinatorial refinement of the partition induced by $\mathcal{B}$. In other words, if for every $x, y \in \mathbb{F}_p^n$, $\mathcal{B}'(x) = \mathcal{B}'(y)$ implies $\mathcal{B}(x) = \mathcal{B}(y)$. $\mathcal{B}'$ is called a* syntactic refinement *of $\mathcal{B}$ if the sequence of polynomials defining $\mathcal{B}'$ extends that of $\mathcal{B}$. A syntactic refinement is clearly a semantic refinement but not necessarily, vice versa.*

The algorithmic regularity lemma of [BHT13] (analogous to the algorithmic version [ADL+94] of Szemerédi's regularity lemma) is as follows:

Theorem 2.16 (Uniform refinement, Lemma 4.1 of [BHT13])**.** *Suppose $d < p$ is a positive integer, $\rho \in (0, 1)$, and $\gamma : \mathbb{Z}^+ \to (0, 1)$ is a non-increasing function. There is a function $C_{2.16}^{\gamma,d} : \mathbb{Z}^+ \to \mathbb{Z}^+$ and an algorithm that takes as input a factor $\mathcal{B}$ of $\mathbb{F}_p^n$ of degree d, runs in time $O(n^d)$ and with probability $1 - \rho$, outputs a γ-uniform factor $\tilde{\mathcal{B}}$ where $\tilde{\mathcal{B}}$ is a refinement of $\mathcal{B}$, is of degree d, and $|\tilde{\mathcal{B}}| \leqslant C_{2.16}^{\gamma,d}(|\mathcal{B}|)$.*

Combining with Theorem 2.11 immediately implies:

Corollary 2.17 (Regular refinement)**.** *Suppose $d < p$ is a positive integer, $\rho \in (0, 1)$ and $R : \mathbb{Z}^+ \to \mathbb{Z}^+$ is a non-decreasing function. There is a function $C_{2.17}^{R,d} : Z^+ \to \mathbb{Z}^+$ and an algorithm that takes as input a factor $\mathcal{B}$ of $\mathbb{F}_p^n$ of degree d, runs in time $O(n^d)$ and with probability $1 - \rho$, outputs a R-regular factor $\tilde{\mathcal{B}}$ where $\tilde{\mathcal{B}}$ is a refinement of $\mathcal{B}$, is of degree d, and $|\tilde{\mathcal{B}}| \leqslant C_{2.17}^{R,d}(|\mathcal{B}|)$. Additionally, if $\mathcal{B}$ is defined by polynomials $P_1, P_2, \ldots, P_m$, then we can find functions $\Gamma_1, \ldots, \Gamma_m : \mathbb{F}_p^{|\tilde{\mathcal{B}}|} \to \mathbb{F}_p$ such that $P_i(x) = \Gamma_i(\tilde{\mathcal{B}}(x))$ for every $i \in [m]$.*

Moreover, if $\mathcal{B}$ is itself a syntactic refinement of some $\mathcal{B}'$ that is of rank at least $R(|\mathcal{B}|) + 1$, then $\tilde{\mathcal{B}}$ will also be a syntactic refinement of $\mathcal{B}'$.

The second-to-last sentence of Corollary 2.17 comes from observing that the proof of Lemma 4.1 in [BHT13] explicitly constructs the functions Γ_i. The last sentence of Corollary 2.17 follows from Lemma 3.17 of [BFL13].

3 The main proof

First, we prove that every degree-structural property is in randomized polynomial time.

Theorem 3.1. *If $p > d$, then for any fixed $k, \boldsymbol{\Delta}$ and Γ, there is a randomized algorithm which given a polynomial $P : \mathbb{F}_p^n \to \mathbb{F}_p$ of degree d runs in time $O(n^{d+1})$ and has the following behavior:*

1. *If P is $(k, \boldsymbol{\Delta}, \Gamma)$-structured, with probability $2/3$, it finds a $(k, \boldsymbol{\Delta}, \Gamma)$-decomposition of P.*
2. *Otherwise, it always outputs NO.*

Proof. Let $R : \mathbb{Z}^+ \to \mathbb{Z}^+$ be defined as $R(m) = r(C_{2.17}^{r,d}(m+k)) + C_{2.17}^{r,d}(m+k) + p$ for a function $r : \mathbb{Z}^+ \to \mathbb{Z}^+$ to be fixed later. First, we apply Corollary 2.17 to the factor defined by $\{P\}$ so that with probability 9/10, we find an R-regular polynomial factor $\mathcal{B}$ of degree d defined by polynomials $P_1, P_2, \ldots, P_C : \mathbb{F}_p^n \to \mathbb{F}_p$ such that $P(x) = G(\mathcal{B}(x))$ for some $G : \mathbb{F}_p^C \to \mathbb{F}_p$. Here, $C \leqslant C^{R,d}(1) = O(1)$.

If $n \leqslant Cd$, then we can decide whether f is $(k, \boldsymbol{\Delta}, \Gamma)$-structured by brute force in $O(1)$ time.

Otherwise, we are in the case $n > Cd$. From each P_i, pick a monomial m_i with degree equal to $\deg(P_i)$. Since $n > Cd$, there exists $i_0 \in [n]$ such that x_{i_0} does not appear in any of the m_i's. Let $P'_1, P'_2, \ldots, P'_C$ be $P_1|_{x_{i_0}=0}, P_2|_{x_{i_0}=0}, \ldots, P_C|_{x_{i_0}=0}$ respectively, and let $\mathcal{B}'$ be the factor defined by these polynomials. Clearly, $\deg(P'_i) = \deg(P_i)$ for each $i \in [C]$. Moreover, by Subspace Restriction Lemma 2.14, $\mathcal{B}'$ is $(R-p)$-regular.

Recursively, decide $(k, \mathbf{\Delta}, \Gamma)$-structure for the polynomial $P' \overset{\text{def}}{=} P|_{x_{i_0}=0}$ on $n-1$ variables. Note that:

$$P'(x) = G(P'_1(x), P'_2(x), \ldots, P'_C(x)).$$

If P' is not $(k, \mathbf{\Delta}, \Gamma)$-structured, then clearly P cannot be, and the algorithm can output NO. Otherwise, suppose that:

$$P'(x) = \Gamma(S_1(x), S_2(x), \ldots, S_k(x))$$

where $\deg(S_1), \ldots, \deg(S_k)$ are at most $\Delta_1, \ldots, \Delta_k$ respectively. We need to show how to extract $(k, \mathbf{\Delta}, \Gamma)$-structure for P from this decomposition for P'.

Use Corollary 2.17 to find, with probability at least 9/10, an r-regular refinement $\mathcal{B}'$ of the factor defined by $\{P'_1, \ldots, P'_C, S_1, \ldots, S_k\}$. Note that the rank of $\mathcal{B}'$ is at least $r(|\mathcal{B}'|)$, while the rank of the factor defined by $\{P'_1, \ldots, P'_C\}$ is at least $R(C) - p = r(C^{r,d}_{2.17}(C+k)) + C^{r,d}_{2.17}(C+k) \geqslant r(|\mathcal{B}'|) + |\mathcal{B}'|$. Because of the last part of Corollary 2.17, $\mathcal{B}'$ is a syntactic refinement of of $\{P'_1, \ldots, P'_C\}$. That is, we obtain a polynomial factor $\mathcal{B}' = \{P'_1, \ldots, P'_C, S'_1, \ldots, S'_D\}$ which has degree d and rank $> r(C+D)$, where $C + D \leqslant C^{r,d}_{2.17}(C+k)$ and where $S_i(x) = G_i(P'_1(x), \ldots, P'_C(x), S'_1(x), \ldots, S'_D(x))$ for some function $G_i : \mathbb{F}_p^{C+D} \to \mathbb{F}_p$. Thus, we have that for all x:

$$\begin{aligned} & G(P'_1(x), \ldots, P'_C(x)) \\ = \Gamma(G_1(P'_1(x), \ldots, P'_C(x), S'_1(x), \ldots, S'_D(x)), \ldots, & G_k(P'_1(x), \ldots, P'_C(x), S'_1(x), \ldots, S'_D(x)) \end{aligned}$$

Note that by Corollary 2.17, we find the functions $G_1, \ldots, G_k$ explicitly.

Let $\gamma(m) = \frac{1}{2p^m}$, and suppose $r(m) > R^{d,\gamma}_{2.12}(m)$. Then, by Theorem 2.12, Lemma 2.9 and Corollary 2.4, we see that $\mathcal{B}'(x)$ acquires every possible value in its range. Thus, we have the identity:

$$G(a_1, \ldots, a_C) = \Gamma(G_1(a_1, \ldots, a_C, b_1, \ldots, b_D), \ldots, G_k(a_1, \ldots, a_C, b_1, \ldots, b_D))$$

for *all* $a_1, \ldots, a_C, b_1, \ldots, b_D \in \mathbb{F}_p$. In particular:

$$\begin{aligned} P(x) = G(P_1(x), \ldots, & P_C(x)) \\ & = \Gamma(G_1(P_1(x), \ldots, P_C(x), 0, \ldots, 0), \ldots, G_k(P_1(x), \ldots, P_C(x), 0, \ldots, 0)) \end{aligned}$$

Define $Q_i(x) = G_i(P_1(x), \ldots, P_C(x), 0, \ldots, 0)$ for each $i \in [k]$. Now, suppose $r(m) > R^{d,\alpha^d_{2.6}}_{2.12}(m)$. By Lemma 2.9 and Theorem 2.6, since $\deg(P_i) = \deg(P'_i)$, it follows that $\deg(Q_i) \leqslant \deg(S_i) \leqslant \Delta_i$ for each $i \in [k]$. Then, our $(k, \mathbf{\Delta}, \Gamma)$-decomposition is given by:

$$P(x) = \Gamma(Q_1(x), \ldots, Q_k(x))$$

Hence, set $r = \max(R^{d,\gamma}_{2.12}, R^{d,\alpha^d_{2.6}}_{2.12})$.

In order to see the guarantees in the theorem statement, consider repeating the above algorithm infinitely until a $(k, \mathbf{\Delta}, \Gamma)$-decomposition is discovered for P. If P is not $(k, \mathbf{\Delta}, \Gamma)$-structured, then any candidate $(k, \mathbf{\Delta}, \Gamma)$-decomposition discovered (due to the error probability in Corollary 2.17) can be ruled out in $O(n^d)$ time. Otherwise, if P is $(k, \mathbf{\Delta}, \Gamma)$-structured the expected time before a valid $(k, \mathbf{\Delta}, \Gamma)$-decomposition is discovered will be the expected time for discovering a decomposition for P' plus expected $O(n^d)$ time for finding valid regular refinements. Thus, the expected time to find a $(k, \mathbf{\Delta}, \Gamma)$-decomposition for P is $O(n^{d+1})$. Therefore, if we stop repeating the algorithm after $O(n^{d+1})$ time steps, our desired result is true by Markov's theorem. □

Theorem 3.1 can be derandomized using existing pseudorandom generators for low-degree polynomials [Vio09] to yield Theorem 1.2. This idea was suggested by Shachar Lovett.

Proof of Theorem 1.2. If we unravel the proof of Theorem 3.1, we find that the only uses of randomness are in choosing a constant number of random points uniformly from $\mathbb{F}_p^n$ and querying the value of degree-d polynomials on them. In particular, there are three instantiations of sampling:

1. In Lemma 2.1 (Algorithmic Bogdanov-Viola Lemma) of [BHT13], for a degree-d polynomial $P : \mathbb{F}_p^n \to \mathbb{F}_p$, the distribution of $P(x)$ is estimated upto constant statistical distance by sampling P at a constant number of points.

2. In Lemma 2.1 (Algorithmic Bogdanov-Viola Lemma) of [BHT13], for a degree-d polynomial P, it is shown that with probability $\geqslant 1/3$, a tuple $\boldsymbol{h} = (h_1, h_2, \ldots, h_C)$ of $C = O(1)$ many uniformly chosen random points from $\mathbb{F}_p^n$ has the property that for at least $(1-\sigma)p^n$ many $x \in \mathbb{F}_p^n$, it holds that for every $t \in \mathbb{F}_p$:
$$\left| \frac{1}{C} \sum_{i=1}^{C} \mathbf{1}_{P(x+h_i)=t} - \mathop{\mathbf{Pr}}_{y} \left[P(y) = t \right] \right| < \frac{\delta}{p}$$

3. In Lemma 4.1 (Uniform Refinement) of [BHT13], for a degree-d polynomial $Q : \mathbb{F}_p^n \to \mathbb{F}_p$, the quantity $\|\mathsf{e}\left(Q(x)\right)\|_{U^k}$ is estimated upto constant additive error by sampling a constant number of points.

Our key tool to derandomize these sampling steps will be existing pseudorandom generators that fool polynomials over finite fields [BV07, Lov09, Vio09]. First, we formalize what we require of these PRGs:

Definition 3.2. *A distribution Z on $\mathbb{F}_p^n$ ε-*fools *degree-d polynomials in n variables over $\mathbb{F}_p$ if for every such polynomial P, we have:*
$$|\mathop{\mathbf{E}}_{Z}[\mathsf{e}\left(P(Z)\right)] - \mathop{\mathbf{E}}_{U}[\mathsf{e}\left(P(U)\right)]| < \varepsilon$$
where U is the uniform distribution over $\mathbb{F}_p^n$.

The lemma below showing that such a distribution Z can fool multiple polynomials simultaneously will be useful:

Lemma 3.3. *If Z ε-fools degree-d polynomials, then for any k degree-d polynomials $P_1, \ldots, P_k$:*
$$\sum_{t_1,\ldots,t_k \in \mathbb{F}_p} \left| \mathop{\mathbf{Pr}}_{Z}[P_1(Z) = t_1 \wedge \cdots \wedge P_k(Z) = t_k] - \mathop{\mathbf{Pr}}_{U}[P_1(U) = t_1 \wedge \cdots \wedge P_k(U) = t_k] \right| < p^k \varepsilon$$

Proof. Observe that:
$$\begin{aligned}\mathop{\mathbf{Pr}}_{Z}[P_1(Z) = t_1 \wedge \cdots \wedge P_k(Z) = t_k] &= \mathop{\mathbf{E}}_{Z}\left[\frac{1}{p}\sum_{\lambda_1 \in \mathbb{F}_p} \mathsf{e}\left(\lambda_1(P_1(Z) - t_1)\right) \cdots \frac{1}{p}\sum_{\lambda_k \in \mathbb{F}_p} \mathsf{e}\left(\lambda_k(P_k(Z) - t_k)\right)\right] \\ &= \frac{1}{p^k} \sum_{\lambda_1,\ldots,\lambda_k \in \mathbb{F}_p} \mathsf{e}\left(-\sum_{i=1}^{k} \lambda_i t_i\right) \mathop{\mathbf{E}}_{Z}\left[\mathsf{e}\left(\sum_{i=1}^{k} \lambda_i P_i(Z)\right)\right]\end{aligned}$$

Thus:
$$\begin{aligned}&\sum_{t_1,\ldots,t_k \in \mathbb{F}_p} \left| \mathop{\mathbf{Pr}}_{Z}[P_1(Z) = t_1 \wedge \cdots \wedge P_k(Z) = t_k] - \mathop{\mathbf{Pr}}_{U}[P_1(U) = t_1 \wedge \cdots \wedge P_k(U) = t_k] \right| \\ &= \frac{1}{p^k} \sum_{t_1,\ldots,t_k \in \mathbb{F}_p} \left| \sum_{\lambda_1,\ldots,\lambda_k \in \mathbb{F}_p} \mathsf{e}\left(-\sum_{i=1}^{k} \lambda_i t_i\right) \left(\mathop{\mathbf{E}}_{Z}\left[\mathsf{e}\left(\sum_{i=1}^{k} \lambda_i P_i(Z)\right)\right] - \mathop{\mathbf{E}}_{U}\left[\mathsf{e}\left(\sum_{i=1}^{k} \lambda_i P_i(U)\right)\right]\right)\right| \\ &< \frac{1}{p^k} \cdot p^k \cdot p^k \varepsilon = p^k \varepsilon\end{aligned}$$

where we used the fact that linear combinations of degree-d polynomials are also degree-d polynomials. □

We can now invoke the following positive result of Viola.

Theorem 3.4 ([Vio09])**.** *There is an explicit generator $g : \mathbb{F}_p^s \to \mathbb{F}_p^n$ whose output distribution ε-fools degree-d polynomials with seed length $s = d\log_p n + O(d \cdot 2^d \cdot \log(1/\varepsilon))$.*

Let us see now how to derandomize the three sampling steps listed above, while blowing up the running time only polynomially.

1. This step is easy to handle. For the generator g from Theorem 3.4, look at $(P(g(x)) : x \in \mathbb{F}_p^s)$. By definition of g, we obtain the required approximation of $P(U)$. This takes time $O(p^s) = O(n^d)$.

2. The proof of Lemma 2.1 in [BHT13] shows that:

$$\Pr_{x,h_1,\dots,h_C \leftarrow U}\left[\left|\frac{1}{C}\sum_{i=1}^{C}\mathbf{1}_{P(x+h_i)=t} - \Pr_y[P(y)=t]\right| > \frac{\delta}{m}\right] < \frac{2\sigma}{3}$$

   Now, for any given x, consider the k polynomials $Q_1, \dots, Q_k : \mathbb{F}_p^{Cn} \to \mathbb{F}_p$ defined as: $Q_i(h_1,\dots,h_C) = P(x+h_i)$. Invoking Lemma 3.3 on the degree-d polynomials $Q_1,\dots,Q_C$, we have that if Z ε-fools degree-d polynomials on $\mathbb{F}_p^{Cn}$, then the L_1 distance between the distributions $(Q_1(Z),\dots,Q_k(Z))$ and $(Q_1(U),\dots,Q_k(U))$ is at most $p^C\varepsilon$. Since this is true for every x, we have that:

$$\Pr_{x\leftarrow U,h_1,\dots,h_C \leftarrow Z}\left[\left|\frac{1}{C}\sum_{i=1}^{C}\mathbf{1}_{P(x+h_i)=t} - \Pr_y[P(y)=t]\right| > \frac{\delta}{m}\right] < \frac{2\sigma}{3} + p^C\varepsilon$$

   By choosing ε sufficiently small, we can therefore ensure there exist some $h_1,\dots,h_C$ in the support of Z such that

$$\Pr_{x\leftarrow U}\left[\left|\frac{1}{C}\sum_{i=1}^{C}\mathbf{1}_{P(x+h_i)=t} - \Pr_y[P(y)=t]\right| > \frac{\delta}{m}\right] < \sigma \tag{1}$$

   To search for $h_1,\dots,h_C$, we can simply iterate through all possible selections of C elements from the support of Z, where $Z = g(s)$ from Theorem 3.4 and $s = d\log_p Cn + O(1)$. The above argument guarantees that there will be one selection of $h_1,\dots,h_C$ where (1) holds. This iteration multiplies the running time of the algorithm by an overhead factor of $(p^s)^C = \mathrm{poly}(n)$.

3. For each $I \subseteq [k]$, let $Q_I : \mathbb{F}_p^{(k+1)n} \to \mathbb{F}_p$ be defined as $Q_I(x, y_1, y_2, \dots, y_k) = Q(x + \sum_{i\in I} y_i)$. Invoking Lemma 3.3, letting U be the uniform distribution on $\mathbb{F}_p^{(k+1)n}$, and letting Z be a distribution that ε-fools degree-d polynomials on $\mathbb{F}_p^{(k+1)n}$, we have that $(Q_I(x,y_1,\dots,y_k) : x, y_1,\dots,y_k \leftarrow Z)$ and $(Q_I(x,y_1,\dots,y_k) : x,y_1,\dots,y_k \leftarrow U)$ are $p^{2^k}\varepsilon$-close in L_1-distance. Thus:

$$\mathop{\mathbf{E}}_{x,y_1,\dots,y_k\leftarrow Z}\prod_{I\subseteq[k]}\mathsf{e}\left((-1)^{|I|}Q_I(x,y_1,\dots,y_k)\right)$$

   is away from $\|Q\|_{U^k}^{2^k}$ only by an additive constant controlled by ε. Using Z from Theorem 3.4, this expectation can be computed in poly(n) time.

□

Tensor Rank Although tensor rank is not degree-structural by definition, it has a similar flavor and indeed can be decided by the same algorithm presented above. Given a tensor $A : [n]^r \to \mathbb{F}_p$ and r variables $X^{(1)},\dots,X^{(r)}$ over $\mathbb{F}_P^n$, consider the natural set-multilinear polynomial $f_A : \mathbb{F}_p^{rn} \to \mathbb{F}_p$:

$$f_A(X^{(1)},\dots,X^{(r)}) = \sum_{i_1,i_2,\dots,i_r\in[n]} A(i_1,\dots,i_r)\prod_{j=1}^{r} X_{i_j}^{(j)}$$

Then A has tensor rank s exactly when there are rs linear forms $\ell_{1,1},\ldots,\ell_{1,s},\ldots,\ell_{r,1},\ldots,\ell_{r,s} : \mathbb{F}_p^n \to \mathbb{F}_p$ such that:

$$f_A(X^{(1)},\ldots,X^{(r)}) = \sum_{j=1}^{s}\prod_{i=1}^{r}\ell_{i,j}(X^{(i)})$$

This is close to being degree-structural except that every $\ell_{i,j}$ is only allowed to depend on $X^{(i)}$ and not on the rest of the input. However, upon examining the proof of Theorem 3.1, it can be checked that if in the base case, the decomposition respects this partitioning of variables, then so will be the case at the end of the algorithm.

Acknowledgments

Part of this work was done when visiting the Simons Institute and MIT during December, 2013. Many thanks to Ronitt Rubinfeld for useful discussions, to Pablo Parrilo for asking about tensor rank and to Neeraj Kayal, Amir Shpilka and Madhu Sudan for pointers to previous work.

References

[ADL$^+$94] Noga Alon, Richard A. Duke, Hanno Lefmann, Vojtech Rödl, and Raphael Yuster. The algorithmic aspects of the regularity lemma. *J. Algorithms*, 16(1):80–109, 1994.

[BCSX11] Arnab Bhattacharyya, Victor Chen, Madhu Sudan, and Ning Xie. Testing linear-invariant nonlinear properties. *Theory Comput.*, 7(1):75–99, 2011.

[BFH$^+$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.

[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.

[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.

[BHT13] Arnab Bhattacharyya, Pooya Hatami, and Madhur Tulsiani. Algorithmic regularity for polynomials and applications. Technical report, November 2013. http://arxiv.org/abs/1311.5090.

[BV07] Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. In *Proc. 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 41–51, Washington, DC, USA, 2007. IEEE Computer Society.

[DLM$^+$07] Ilias Diakonikolas, Homin K. Lee, Kevin Matulef, Krzysztof Onak, Ronitt Rubinfeld, Rocco A. Servedio, and Andrew Wan. Testing for concise representations. In *Proc. 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 549–558, 2007.

[dW08] Ronald de Wolf. *A Brief Introduction to Fourier Analysis on the Boolean Cube.* Number 1 in Graduate Surveys. Theory of Computing Library, 2008.

[Gow98] William T. Gowers. A new proof of Szeméredi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998.

[Gow01] William T. Gowers. A new proof of Szeméredi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.

[GT08] Ben Green and Terence Tao. An inverse theorem for the Gowers U^3-norm. *Proc. Edin. Math. Soc.*, 51:73–153, 2008.

[GT09] Ben Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms. *Contrib. Discrete Math.*, 4(2), 2009.

[GT10] Ben Green and Terence Tao. Linear equations in primes. *Ann. of Math.*, 171:1753–1850, 2010.

[GTZ] Ben Green, Terence Tao, and Tamar Ziegler. An inverse theorem for the Gowers U^{s+1}-norm. *Ann. of Math.*, to appear. http://arxiv.org/abs/1009.3998.

[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.

[HK05] Bernard Host and Bryna Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.*, 161(1):397–488, 2005.

[HL11] Hamed Hatami and Shachar Lovett. Higher-order Fourier analysis of $\mathbb{F}_p^n$ and the complexity of systems of linear forms. *Geom. Funct. Anal.*, 21:1331–1357, 2011.

[Kal95] Erich Kaltofen. Effective Noether irreducibility forms and applications. *J. Comp. Sys. Sci.*, 50(2):274 – 295, 1995.

[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.

[KS09] Zohar Shay Karnin and Amir Shpilka. Reconstruction of generalized depth-3 arithmetic circuits with bounded top fan-in. In *Proc. 24th Annual IEEE Conference on Computational Complexity*, pages 274–285, 2009.

[KSS14] Swastik Kopparty, Shubhangi Saraf, and Amir Shpilka. Equivalence of polynomial identity testing and deterministic multivariate polynomial factorization. Technical Report 001, Electronic Colloquium on Computational Complexity, January 2014. http://eccc.hpi-web.de/report/2014/001/.

[Lov09] Shachar Lovett. Unconditional pseudorandom generators for low degree polynomials. *Theory Comput.*, 5(1):69–82, 2009.

[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.

[Sud12] Madhu Sudan. Lecture 11 of "Algebra and Computation". http://people.csail.mit.edu/madhu/ST12/scribe/lect11.pdf, March 2012. Lecture notes.

[Sze78] Endre Szemerédi. Regular partitions of graphs. In J.C. Bremond, J.C. Fournier, M. Las Vergnas, and D. Sotteau, editors, *Proc. Colloque Internationaux CNRS 260 – Problèmes Combinatoires et Théorie des Graphes*, pages 399–401, 1978.

[Tao12] Terence Tao. *Higher Order Fourier Analysis*, volume 142 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012.

[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.

[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.

[Vio09] Emmanuele Viola. The sum of D small-bias generators fools polynomials of degree D. *Computational Complexity*, 18(2):209–217, 2009.

[VW08] Emanuele Viola and Avi Wigderson. Norms, XOR lemmas, and lower bounds for polynomials and protocols. *Theory of Computing*, 4(7):137–168, 2008.

[vzGK85] Joachim von zur Gathen and Erich Kaltofen. Factorization of multivariate polynomials over finite fields. *Mathematics of Computation*, 45(171):251–261, 1985.