Correlation Testing for Affine Invariant Properties on $\mathbb{F}_p^n$ in the High Error Regime

[Extended Abstract] *

Hamed Hatami
School of Computer Science
McGill University
Montréal, Canada
hatami@cs.mcgill.ca

Shachar Lovett
School of Mathematics
Institute of Advanced Study
Princeton, USA
slovett@math.ias.edu

ABSTRACT

Recently there has been much interest in Gowers uniformity norms from the perspective of theoretical computer science. This is mainly due to the fact that these norms provide a method for testing whether the maximum correlation of a function $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$ with polynomials of degree at most $d \leq p$ is non-negligible, while making only a constant number of queries to the function. This is an instance of *correlation testing*. In this framework, a fixed test is applied to a function, and the acceptance probability of the test is dependent on the correlation of the function from the property. This is an analog of *proximity oblivious testing*, a notion coined by Goldreich and Ron, in the high error regime.

We study in this work general properties which are affine invariant and which are correlation testable using a constant number of queries. We show that any such property (as long as the field size is not too small) can in fact be tested by the Gowers uniformity test, and hence having correlation with the property is equivalent to having correlation with degree d polynomials for some fixed d . We stress that our result holds also for non-linear properties which are affine invariant. This completely classifies affine invariant properties which are correlation testable.

The proof is based on higher-order Fourier analysis, where we establish a new approximate orthogonality for structures defined by linear forms. In particular, this resolves an open problem posed by Gowers and Wolf. Another ingredient is a nontrivial extension of a graph theoretical theorem of Erdős, Lovász and Spencer to the context of additive number theory.

*A full version of this paper is available at <http://www.eccc.uni-trier.de/report/2011/029/>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

STOC'11, June 6–8, 2011, San Jose, California, USA.

Copyright 2011 ACM 978-1-4503-0691-1/11/06 ...\$10.00.

Categories and Subject Descriptors

F.2.m [ANALYSIS OF ALGORITHMS AND PROBLEM COMPLEXITY]: Miscellaneous

General Terms

Theory

Keywords

Gowers uniformity, property testing, higher-order Fourier analysis, true complexity

1. INTRODUCTION

Blum, Luby, and Rubinfeld [2] made a beautiful observation that given a function $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$, it is possible to inquire the value of f on a few random points, and accordingly probabilistically distinguish between the case that f is a linear function and the case that f has to be modified on at least $\varepsilon > 0$ fraction of points to become a linear function. Inspired by this observation, Rubinfeld and Sudan [14] defined the concept of property testing which is now a major area of research in theoretical computer science. Roughly speaking to test a function for a property means to examine the value of the function on a few random points, and accordingly (probabilistically) distinguish between the case that the function has the property and the case that it is not too close to any function with that property. Interestingly and to some extent surprisingly these tests exist for various basic properties. The first substantial investigation of property testing occurred in Goldreich, Goldwasser, and Ron [4] who showed that several natural combinatorial properties are testable. Since then there has been a significant amount of research on classifying the testable properties in combinatorial and algebraic settings.

Studying property testing in algebraic setting requires understanding the density of linear structures in subsets of Abelian groups. It is possible to express the density of those structures by certain analytical averages. Analyzing these averages and understanding the relations between them is the core of many problems and results in additive combinatorics and analytic number theory, and many tools and theories are developed for this purpose. The theory of uniformity, initiated by the proof of Szemerédi's theorem [15], plays an important role in this area, and it was a major breakthrough when Gowers [6] introduced a new notion of

uniformity in a Fourier-analytic proof for Szemerédi's theorem.

Gowers' work initiated an extension of Fourier analysis, called higher-order Fourier analysis. In the classical Fourier analysis of $\mathbb{F}_p^n$, a function is expressed as a linear combination of the characters of $\mathbb{F}_p^n$ which are exponentials of linear polynomials. In higher-order Fourier analysis, the linear polynomials are replaced by higher degree polynomials. Higher-order Fourier expansions are very useful in studying averages that are defined through linear structures. However for these expansions to be useful one needs some kind of orthogonality, or at least an approximation of it. The works of Green and Tao [9] and Kaufman and Lovett [12] provide an approximate orthogonality that can be used to analyze averages such as $\mathbb{E}_{X \in \mathbb{F}_p^n} [f_1(X) \dots f_m(X)]$ in a straightforward manner when proper higher-order Fourier expansions of $f_1, \dots, f_m$ are known. However it is not a priori clear that these results can be applied to analyze more general averages of the form

$$\mathbb{E}_{X_1, \dots, X_k \in \mathbb{F}_p^n} \left[f_1 \left(\sum_{i=1}^k \lambda_{1,i} X_i \right) \dots f_m \left(\sum_{i=1}^k \lambda_{m,i} X_i \right) \right], \quad (1)$$

where $\lambda_{i,j} \in \mathbb{F}_p$ are constants. Such averages arise naturally when one studies linear structures in subsets of Abelian groups.

Our first result is an extension of the results of Green and Tao [9] which can be applied to such general averages. Then, we apply these techniques to determine which Gowers norms of $f_1, \dots, f_m$ one needs to bound in order to guarantee that the average in (1) is small. Our result in particular improves a result of Gowers and Wolf [7] and settles a conjecture posed by them.

Gowers' notion of uniformity has an interesting implication in the context of property testing. The Gowers norm of a function can be expressed as an average of values of this function on a few random sample points. There are theorems which show that a bounded function has non-negligible Gowers uniformity norm if and only if it has a non-negligible correlation with a low degree polynomial. These two facts show that having correlation with low degree polynomials is testable. Our main result is that, roughly speaking, the only correlation testable families of functions are the ones that can be tested by Gowers uniformity norms.

The tests that we study in this article are slightly different in nature from the typical statements in the area of property testing. Typically in property testing the goal is to distinguish the functions that are in a set $\mathcal{D}$ (structured) from the functions that are in a non-negligible distance from every element in $\mathcal{D}$. For example linearity testing [2], probably the most celebrated result in this area, says that it is possible to test whether a function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ is linear or that it has to be modified on at least a non-negligible fraction of all points in order to become a linear function. In this article, we are interested in a different kind of tests. Here we seek a weaker structure in the function, and having correlation with an element in D replaces the usual condition of actually being in $\mathcal{D}$. This is in the spirit of proximity oblivious testing, a notion introduced by Goldreich and Ron [5] in the context of graph properties.

1.1 Notations

The complex unit disk is denoted by $\mathbb{D} = \{z \in \mathbb{C} : |z| \leq 1\}$. We denote elements of $\mathbb{F}_p^n$ by x, y, z , where

$x = (x(1), \dots, x(n))$. We frequently need to work with the elements of $(\mathbb{F}_p^n)^k$ which we regard as vectors with k coordinates. We denote these by $\mathbf{x} = (x_1, \dots, x_k) \in (\mathbb{F}_p^n)^k$. Capital letters X, Y , etc are used to denote random variables. For an element $m \in \mathbb{F}_p$, let $e_p(m) := e^{\frac{2\pi i}{p} m}$. We denote by $f, g, \dots$ functions from $\mathbb{F}_p^n$ to $\mathbb{C}$, and by $\mathbf{f}, \mathbf{g}, \dots$ functions from $\mathbb{F}_p^n$ to $\mathbb{F}_p$. By $f := e_p(\mathbf{f})$ we mean $f(x) = e_p(\mathbf{f}(x))$. The *bias* of a function $f : \mathbb{F}_p^n \rightarrow \mathbb{C}$ is defined as

$$\text{bias}(f) := \left| \mathbb{E}_{X \in \mathbb{F}_p^n} [f(X)] \right|. \quad (2)$$

The *bias* of a function $\mathbf{f} : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$ is defined to be $\text{bias}(\mathbf{f}) := \text{bias}(e_p(\mathbf{f}))$. The inner product of two functions $f, g : \mathbb{F}_p^n \rightarrow \mathbb{C}$ is defined as

$$\langle f, g \rangle := \mathbb{E}_{X \in \mathbb{F}_p^n} [f(X) \overline{g(X)}]. \quad (3)$$

The *correlation* of a function $f : \mathbb{F}_p^n \rightarrow \mathbb{C}$ with a set D of functions from $\mathbb{F}_p^n$ to $\mathbb{C}$ is defined as

$$\|f\|_{u(D)} := \sup_{g \in D} |\langle f, g \rangle|. \quad (4)$$

By an abuse of notation, if D is a set of functions from $\mathbb{F}_p^n$ to $\mathbb{F}_p$, we define

$$\|f\|_{u(D)} := \sup_{g \in D} |\langle f, e_p(g) \rangle|.$$

Note that $\|\cdot\|_{u(D)}$ is always a semi-norm.

For two vector spaces V and W over $\mathbb{F}_p$, let $\text{Lin}(V, W)$ denote the set of all *linear transformations* from V to W . Let $\text{Aff}(n, \mathbb{F}_p)$ denote the group of all *invertible affine transformations* from $\mathbb{F}_p^n$ to itself. For a function $f : \mathbb{F}_p^n \rightarrow \mathbb{C}$, and an $A \in \text{Aff}(n, \mathbb{F}_p)$, we denote by Af the function that maps x to $f(Ax)$.

1.2 Gowers Uniformity Norms

Gowers uniformity norms are defined in the more general setting of arbitrary finite Abelian groups.

DEFINITION 1.1 (GOWERS UNIFORMITY NORMS). *Let G be a finite Abelian group and $f : G \rightarrow \mathbb{C}$. For an integer $k \geq 1$, the k -th Gowers norm of f , denoted $\|f\|_{U^k}$ is defined by*

$$\|f\|_{U^k}^{2^k} := \mathbb{E} \left[\prod_{S \subseteq [k]} c^{|S|} f \left(X + \sum_{i \in S} Y_i \right) \right], \quad (5)$$

where C denotes the complex conjugation operator, and $X, Y_1, \dots, Y_k$ are independent random variables taking values in G uniformly at random.

In this article we are only interested in the case where $G = \mathbb{F}_p^n$. These norms were first defined in [6] in the case where G is the group $\mathbb{Z}_N$. Note that $\|f\|_{U^1} = |\mathbb{E}[f(X)]|$, and thus $\|\cdot\|_{U^1}$ is a semi-norm rather than a norm. The facts that the right-hand side of (5) is always non-negative, and that for $k > 1$, $\|\cdot\|_{U^k}$ is actually a norm are easy to prove, but certainly not trivial (see [6] for a proof).

Let us explain the relevance of the uniformity norms to the area of property testing. The goal in property testing is to obtain certain information about a function by “reading” its values only on a small number of points. The simplest case is that of testing correlation with linear functions. For $a \in \mathbb{F}_p^n$, let $\ell_a : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$ be the corresponding linear function, defined as $\ell_a(x) = \sum_{i=1}^n a(i)x(i)$. Let $\text{Linear} = \{\ell_a : a \in$

$\mathbb{F}_p^n$ be the set of linear functions. Let $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$ be a function, and let $f = e_p(f)$. The correlation of f with linear functions is given by

$$\|f\|_{u(\text{Linear})} = \max_{\ell_a \in \text{Linear}} |\langle f, e_p(\ell_a) \rangle| = \max_{a \in \mathbb{F}_p^n} \text{bias}(f - \ell_a),$$

which is the same as the absolute value of the maximal Fourier coefficient of f . It is known [2, 6] that the correlation of f with linear functions is related to the U^2 norm of f . Specifically, for every $\varepsilon > 0$,

- *Direct Theorem:* If $\|f\|_{u(\text{Linear})} \geq \varepsilon$, then $\|f\|_{U^2} \geq \varepsilon$.
- *Inverse Theorem:* If $\|f\|_{U^2} \geq \varepsilon$, then $\|f\|_{u(\text{Linear})} \geq \varepsilon^2$.

These two facts together show that $\|f\|_{U^2}$ gives a rough estimate for the maximum correlation of f with linear functions. Recall that

$$\|f\|_{U^2}^4 = \mathbb{E}[e_p(f(X+Y+Z) - f(X+Y) - f(X+Z) + f(X))],$$

where $X, Y, Z \in \mathbb{F}_p^n$ are uniformly chosen. In other words, the joint distribution of $f(X+Y+Z), f(X+Y), f(X+Z), f(X)$ allows to distinguish between the case that f has correlation at least ε with linear functions, and the case where f has correlation at most $\delta = \varepsilon^2/2$ (say) with linear functions. Hence, correlation with linear functions is “testable with just 4 queries to f ” (for every $\varepsilon > 0$).

Similar to the case of linear functions, correlation with degree d polynomials can be tested by the U^{d+1} uniformity norm. The main ingredient is the inverse theorem for $\mathbb{F}_p^n$ which was proved in [1, 16]. Let $\text{Poly}_d(\mathbb{F}_p^n)$ be the set of polynomials of degree at most d over $\mathbb{F}_p^n$. These results show that, as long as $p > d$ (i.e. the field is not too small), for every $\varepsilon > 0$,

- *Direct Theorem:* If $\|f\|_{u(\text{Poly}_d(\mathbb{F}_p^n))} \geq \varepsilon$, then $\|f\|_{U^{d+1}} \geq \varepsilon$.
- *Inverse Theorem:* If $\|f\|_{U^{d+1}} \geq \varepsilon$, then $\|f\|_{u(\text{Poly}_d(\mathbb{F}_p^n))} \geq \delta(\varepsilon) > 0$.

The exact dependency of $\delta(\varepsilon)$ on ε is currently unknown, but crucially δ does not depend on n . Analogously to the case of linear functions and U^2 , we get that the joint distribution of $(f(X + \sum_{i \in I} Y_i) : I \subseteq [d+1])$ where $X, Y_1, \dots, Y_{d+1} \in \mathbb{F}_p^n$ are uniformly chosen, distinguishes the case where f has noticeable correlation ($\geq \varepsilon$) with degree d polynomials, from the case where f has negligible correlation ($\leq \delta(\varepsilon)/2$) with degree d polynomials, for every $\varepsilon > 0$. Hence, correlation with polynomials of total degree d is testable with just 2^{d+1} queries to f (for every $\varepsilon > 0$).

An equivalent qualitative formulation of the direct and inverse theorems stated above is as follows. Let $(f_n : \mathbb{F}_p^n \rightarrow \mathbb{F}_p)_{n \in \mathbb{N}}$ be a sequence of functions. Then as long as $d < p$ we have that

$$\lim_{n \rightarrow \infty} \|e_p(f_n)\|_{u(\text{Poly}_d(\mathbb{F}_p^n))} = 0 \iff \lim_{n \rightarrow \infty} \|e_p(f_n)\|_{U^{d+1}} = 0.$$

It will be convenient to express our results in such terms.

1.3 Correlation testable properties

The above discussion motivates the general definition of correlation testable properties. Let $\mathcal{D} = \{D_n\}_{n \in \mathbb{N}}$ be a family of sets, where each D_n is a set of functions from $\mathbb{F}_p^n$ to $\mathbb{F}_p$. Informally, $\mathcal{D}$ is correlation testable using q queries if

there exists a distribution over $x_1, \dots, x_q \in \mathbb{F}_p^n$, and for every $\varepsilon > 0$ there exists $\delta(\varepsilon) \in (0, \varepsilon)$, such that the following holds. The joint distribution of $(f(x_1), \dots, f(x_q))$ allows to distinguish between the case that f has noticeable correlation ($\geq \varepsilon$) with $\mathcal{D}$ and the case that f has negligible correlation ($\leq \delta(\varepsilon)$) with $\mathcal{D}$.

DEFINITION 1.2 (CORRELATION TESTABLE PROPERTIES). A family $\mathcal{D} = (D_n)$ is correlation testable with q queries, if there exists a distribution μ taking values in $(\mathbb{F}_p^n)^q$ and a mapping $\Gamma : \mathbb{F}_p^q \rightarrow \{0, 1\}$, such that the following holds. For every $\varepsilon > 0$, there exist $\delta \in (0, \varepsilon)$, $0 \leq \theta^- < \theta^+ \leq 1$ and $n_0 \in \mathbb{N}$, such that for every $n > n_0$ and $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$ we have:

- If $\|e_p(f)\|_{u(D_n)} \geq \varepsilon$ then $\Pr_{(X_1, \dots, X_q) \sim \mu}[\Gamma(f(X_1), \dots, f(X_q)) = 1] \geq \theta^+$.
- If $\|e_p(f)\|_{u(D_n)} \leq \delta$ then $\Pr_{(X_1, \dots, X_q) \sim \mu}[\Gamma(f(X_1), \dots, f(X_q)) = 1] \leq \theta^-$.

Following the discussion on Gowers uniformity norms, $\text{Poly}_d = \{\text{Poly}_d(\mathbb{F}_p^n)\}_{n \in \mathbb{N}}$ is correlation testable using $q = 2^{d+1}$ queries, as long as $d < p$.

Our goal is to study families $\mathcal{D} = \{D_n\}_{n \in \mathbb{N}}$ which are correlation testable using a constant number of queries. We shall require the sets D_n to be *consistent* with each other:

- **A1: (Consistency)** For positive integers $m > n$ and $g \in D_n$, the function $h : \mathbb{F}_p^m \rightarrow \mathbb{F}_p$ defined as $h(x_1, \dots, x_m) = g(x_1, \dots, x_n)$ belongs to D_m .

We need to make also a more crucial assumption. In this general setting, the algebraic structure of $\mathbb{F}_p^n$ is ignored, and we are treating $\mathbb{F}_p^n$ as a generic set of size p^n . In order to take the algebraic structure of $\mathbb{F}_p^n$ into account, we shall require D_n to be *affine invariant*:

- **A2: (Affine invariance)** For every positive integer n , if $g \in D_n$, then for every $A \in \text{Aff}(n, \mathbb{F}_p)$, we have $Ag \in D_n$.

Invariance plays a crucial role in the area of property testing. We refer to the work of Kaufman and Sudan [13] for the role of invariance in algebraic property testing. We stress that we **do not** require $\mathcal{D}$ to be linear; i.e. we do not require that if $f, g \in D_n$ then also $f + g \in D_n$.

The last condition relates to the size of D_n . We study families where one can distinguish functions with noticeable correlation from function with negligible correlation. In order not to make this meaningless, we would like to have functions with negligible correlation. For example, we would like a random function not to have correlation with $\mathcal{D}$ with high probability. Fix a function $g \in D_n$. The number of functions $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$ which have correlation at least δ with g is $p^{c(\delta) \cdot p^n}$ where $\lim_{\delta \rightarrow 0} c(\delta) = 1$. Thus, a sufficient condition for random functions not be correlated with D_n is that the size of D_n is $p^{o(p^n)}$. We thus add the following requirement.

- **A3: (Sparsity)** For every $\varepsilon > 0$ and large enough n , we have $|D_n| \leq p^{\varepsilon \cdot p^n}$.

We call every $\mathcal{D} = \{D_n\}_{n \in \mathbb{N}}$ satisfying assumptions **A1, A2, A3** a *proper dual*.

Our main result is the following theorem which roughly speaking says that the only correlation testable families of

functions are the ones that can be tested by Gowers uniformity norms provided that the field size p is not too small. Say a sequence of functions $(f_n)_{n \in \mathbb{N}}$ is *unbiased* if $\lim_{n \rightarrow \infty} \text{bias}(f_n) = 0$.

THEOREM 1.3 (MAIN THEOREM). *Consider a proper dual $\mathcal{D} = \{D_n\}_{n \in \mathbb{N}}$. If $\mathcal{D}$ is correlation testable with q queries, and if $p \geq q - 1$, then there exists $1 \leq k \leq q - 1$ such that the following holds. For every unbiased sequence of functions $(f_n : \mathbb{F}_p^n \rightarrow \mathbb{F}_p)_{n \in \mathbb{N}}$, we have*

$$\lim_{n \rightarrow \infty} \|e_p(f_n)\|_{u(D_n)} = 0 \iff \lim_{n \rightarrow \infty} \|e_p(f_n)\|_{U^{k+1}} = 0.$$

Equivalently, Theorem 1.3 can be phrased as follows: the sequence $(f_n)_{n \in \mathbb{N}}$ has correlation with $\mathcal{D}$ iff it has correlation with polynomials of degree at most k .

We remark on the seemingly odd requirement that f_n is unbiased. Let μ be some distribution over $\mathbb{F}_p$, and let $f_n : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$ be a random function, where each $f_n(x)$ is sampled independently according to μ . By condition **A3**, since $|D_n| = p^{o(p^n)}$, we have that with probability $1 - o_n(1)$,

$$\lim_{n \rightarrow \infty} \|e_p(f_n)\|_{u(D_n)} = 0.$$

However, if μ has a nonzero bias, then almost surely f_n will have correlation with constant functions. It turns out that ruling out sequences of functions which have correlation with constant functions is sufficient for establishing Theorem 1.3.

1.4 Paper organization

We give a short overview of the proof of our results in Section 2. For lack of space, the actual proofs are omitted, and can be found in the full version of this note [11]. We give some open problem in Section 3.

2. PROOF OVERVIEW

In this section we give an overview of the proofs of our main results. We skip most of the technicalities, and try to emphasize the essence of the proofs. The proof of Theorem 1.3 is based on studying averages of functions evaluated on linear forms.

2.1 Linear forms

A *linear form* in k variables is a vector $L = (\lambda_1, \dots, \lambda_k) \in \mathbb{F}_p^k$ regarded as a linear function from V^k to V , for every vector space V over $\mathbb{F}_p$. If $\mathbf{x} = (x_1, \dots, x_k) \in V^k$, then $L(\mathbf{x}) := \lambda_1 x_1 + \dots + \lambda_k x_k$. A *system of linear forms* in k variables is a finite set $\mathcal{L} = \{L_1, \dots, L_m\}$ of distinct linear forms L_i in k variables. For a function $f : \mathbb{F}_p^n \rightarrow \mathbb{C}$, and a system of linear forms $\mathcal{L} = \{L_1, \dots, L_m\}$ in k variables, define the average

$$t_{\mathcal{L}}(f) := \mathbb{E} \left[\prod_{i=1}^m f(L_i(\mathbf{X})) \right], \quad (6)$$

where $\mathbf{X}$ is a random variable taking values uniformly in $(\mathbb{F}_p^n)^k$. We define two generalizations of such averages. First, when $f : \mathbb{F}_p^n \rightarrow \mathbb{C}$, one may take conjugations. For $\alpha \in \{0, 1\}^m$, define

$$t_{\mathcal{L}, \alpha}(f) := \mathbb{E} \left[\prod_{i=1}^m C^{\alpha(i)} f(L_i(\mathbf{X})) \right], \quad (7)$$

where C is the conjugation operator. Second, if $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$, one may take coefficients in $\mathbb{F}_p$. For $\beta \in \mathbb{F}_p^n$, define

$$t_{\mathcal{L}, \beta}^*(f) := \mathbb{E} \left[e_p \left(\sum_{i=1}^m \beta(i) f(L_i(\mathbf{X})) \right) \right]. \quad (8)$$

We note that for functions $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$, (8) is more general than (7). Indeed, if $\alpha \in \{0, 1\}^m$, then setting $\beta(i) = (-1)^{\alpha(i)}$ gives $t_{\mathcal{L}, \alpha}(e_p(f)) = t_{\mathcal{L}, \beta}^*(f)$.

Let $\mathcal{D}$ be a proper dual family which is correlation testable. Then since $\mathcal{D}$ is affine invariant, it is easy to show that correlation with $\mathcal{D}$ can be expressed in terms of averages $t_{\mathcal{L}, \beta}^*$. That is, for every $\varepsilon > 0$ there exist $\delta \in (0, \varepsilon)$, systems of linear forms $\mathcal{L}_1, \dots, \mathcal{L}_\ell$ and corresponding coefficients $\beta_1, \dots, \beta_\ell$, and $n_0 \in \mathbb{N}$, such that the closures of the following two sets are disjoint:

$$T_\varepsilon := \{ (t_{\mathcal{L}_1, \beta_1}^*(f), \dots, t_{\mathcal{L}_\ell, \beta_\ell}^*(f)) \mid n > n_0, f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p, \|e_p(f)\|_{u(D_n)} \geq \varepsilon \},$$

and

$$S_\varepsilon := \{ (t_{\mathcal{L}_1, \beta_1}^*(f), \dots, t_{\mathcal{L}_\ell, \beta_\ell}^*(f)) \mid n > n_0, f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p, \|e_p(f)\|_{u(D_n)} \leq \delta \}.$$

Moreover, each of the systems $\mathcal{L}_1, \dots, \mathcal{L}_\ell$ has at most q linear forms.

Recall that $\mathbb{D}$ is the unit disk in $\mathbb{C}$. It turns out to be easier to first analyze sets with a stronger requirement which holds for all functions $f : \mathbb{F}_p^n \rightarrow \mathbb{D}$, and not just for functions of the form $f = e_p(f)$ for $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$. Say a family $\mathcal{D}$ is *strongly correlation testable* if the closures of the following two sets are disjoint:

$$T_\varepsilon := \{ (t_{\mathcal{L}_1, \alpha_1}(f), \dots, t_{\mathcal{L}_\ell, \alpha_\ell}(f)) \mid n > n_0, f : \mathbb{F}_p^n \rightarrow \mathbb{D}, \|f\|_{u(D_n)} \geq \varepsilon \},$$

and

$$S_\varepsilon := \{ (t_{\mathcal{L}_1, \alpha_1}(f), \dots, t_{\mathcal{L}_\ell, \alpha_\ell}(f)) \mid n > n_0, f : \mathbb{F}_p^n \rightarrow \mathbb{D}, \|f\|_{u(D_n)} \leq \delta \}.$$

2.2 Complexity of systems of linear forms

We turn to analyze averages of the form $t_{\mathcal{L}, \alpha}(f)$, and more generally averages of the form $\mathbb{E} \left[\prod_{i=1}^m C^{\alpha(i)} f_i(L_i(\mathbf{X})) \right]$ where $\mathbf{X} \in (\mathbb{F}_p^n)^k$ is uniform. A crucial ingredient is that we can “approximate” f_i by nice functions, and that these approximation are essentially undetected by averages of the above form.

Green and Tao [10] defined the notion of the *Cauchy-Schwarz complexity* of a system of linear forms. This gives a parameter $s \in \mathbb{N}$ for which the following lemma follows by a sequence of clever applications of the Cauchy-Schwarz inequality.

LEMMA 2.1 ([10]). *Let $\mathcal{L} = \{L_1, \dots, L_m\}$ be a system of linear forms of Cauchy-Schwarz complexity s . Let $f_i, g_i : \mathbb{F}_p^n \rightarrow \mathbb{D}$ be functions for $1 \leq i \leq m$. Assume that $\|f_i - g_i\|_{U^{s+1}} \leq \varepsilon/m$ for all $1 \leq i \leq m$. Then*

$$\left| \mathbb{E}_{\mathbf{X}} \left[\prod_{i=1}^m f_i(L_i(\mathbf{X})) \right] - \mathbb{E}_{\mathbf{X}} \left[\prod_{i=1}^m g_i(L_i(\mathbf{X})) \right] \right| \leq \varepsilon,$$

where $\mathbf{X} \in (\mathbb{F}_p^n)^k$ is uniform.

Gowers and Wolf [7, 8] asked for the minimal $d \in \mathbb{N}$ such that a bound on $\max_{i \in [m]} \|f_i - \mathbb{E}[f_i]\|_{U^{d+1}}$ will allow to approximate $\mathbb{E}[\prod_{i=1}^m f_i(L_i(\mathbf{X}))]$ by $\prod_{i=1}^m \mathbb{E}[f_i]$. They denoted this d as the *true complexity* of a system of linear forms. For a linear form $L = (\lambda_1, \dots, \lambda_k) \in \mathbb{F}_p^k$ and for $t \in \mathbb{N}$, define $L^t \in \mathbb{F}_p^{k^t}$ to be the t -tensor power of L , defined as

$$L^t = \left(\prod_{j=1}^t \lambda_{i_j} : i_1, \dots, i_t \in [k] \right).$$

Gowers and Wolf [7, 8] showed that if the field size is not too small, then the true complexity is the minimal d for which $L_1^{d+1}, \dots, L_m^{d+1}$ are linearly independent.

THEOREM 2.2 ([7, 8]). *Let $\mathcal{L} = \{L_1, \dots, L_m\}$ be a system of linear forms of Cauchy-Schwarz complexity s , and assume $p \geq s$. Let $d \geq 0$ be such that $L_1^{d+1}, \dots, L_m^{d+1}$ are linearly independent. Then for every $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds. Let $f_1, \dots, f_m : \mathbb{F}_p^n \rightarrow \mathbb{D}$ be functions where $\|f_i - \mathbb{E}[f_i]\|_{U^{d+1}} \leq \delta$. Then*

$$\left| \mathbb{E}_{\mathbf{X}} \left[\prod_{i=1}^m f_i(L_i(\mathbf{X})) \right] - \prod_{i=1}^m \mathbb{E}[f_i] \right| \leq \varepsilon,$$

where $\mathbf{X} \in (\mathbb{F}_p^n)^k$ is uniform.

We prove in this paper that Gowers and Wolf definition of true complexity extends also to the more general cases where g_i may be any functions, and not just constant functions.

THEOREM 2.3. *Let $\mathcal{L} = \{L_1, \dots, L_m\}$ be a system of linear forms of true complexity d and Cauchy-Schwarz complexity at most p . Then for every $\varepsilon > 0$, there exists $\delta > 0$ such that the following holds. Let $f_i, g_i : \mathbb{F}_p^n \rightarrow \mathbb{D}$ for $1 \leq i \leq m$ be functions such that $\|f_i - g_i\|_{U^{d+1}} \leq \delta$. Then*

$$\left| \mathbb{E}_{\mathbf{X}} \left[\prod_{i=1}^m f_i(L_i(\mathbf{X})) \right] - \mathbb{E}_{\mathbf{X}} \left[\prod_{i=1}^m g_i(L_i(\mathbf{X})) \right] \right| \leq \varepsilon,$$

where $\mathbf{X} \in (\mathbb{F}_p^n)^k$ is uniform.

In fact, Theorem 2.3 follows immediately from the following stronger theorem which we prove, which in particular resolves an open problem posed by Gowers and Wolf (Problem 7.6 in [7]).

THEOREM 2.4. *Let $\mathcal{L} = \{L_1, \dots, L_m\}$ be a system of linear forms of Cauchy-Schwarz complexity at most p . Let $d \geq 0$, and assume that L_1^{d+1} is not in the linear span of $L_2^{d+1}, \dots, L_m^{d+1}$. Then for every $\varepsilon > 0$, there exists $\delta > 0$ such that for any functions $f_1, \dots, f_m : \mathbb{F}_p \rightarrow \mathbb{D}$ where $\|f_1\|_{U^{d+1}} \leq \delta$ we have*

$$\left| \mathbb{E}_{\mathbf{X} \in (\mathbb{F}_p^n)^k} \left[\prod_{i=1}^m f_i(L_i(\mathbf{X})) \right] \right| \leq \varepsilon.$$

2.3 Approximation by averages over polynomial factors

We now define the notion of a “nice approximating function”, which will play the role of $g_1, \dots, g_m$ in the previous subsection. These functions will be averages of $f_1, \dots, f_m$ on polynomial factors.

A *polynomial factor* $\mathcal{B}$ is a partition (sigma-algebra) of $\mathbb{F}_p^n$ defined by a collection of polynomials $P_1, \dots, P_C$. The atoms of the partitions are

$$\{x \in \mathbb{F}_p^n : P_1(x) = a(1), \dots, P_C(x) = a(C)\},$$

where $a(1), \dots, a(C) \in \mathbb{F}_p$. The *degree* of $\mathcal{B}$ is the maximal degree of $P_1, \dots, P_C$, and the *complexity* of $\mathcal{B}$ is C .

The *rank* of a polynomial P of degree k is the minimal number of lower degree polynomials required to compute it. That is, $\text{rank}(P) \leq r$ if there exists r polynomials $Q_1, \dots, Q_r$ of degree at most $k-1$ such that $P(x) = \Gamma(Q_1(x), \dots, Q_r(x))$ where $\Gamma : \mathbb{F}_p^r \rightarrow \mathbb{F}_p$ is some function. The rank of a set of polynomials $P_1, \dots, P_C$ is the minimal rank of a nonzero linear combination of them. The rank of a polynomial factor $\mathcal{B}$ is the rank of the set of polynomials defining it.

The average of a function $f : \mathbb{F}_p^n \rightarrow \mathbb{D}$ over a polynomial factor $\mathcal{B}$, denoted $\mathbb{E}(f|\mathcal{B}) : \mathbb{F}_p^n \rightarrow \mathbb{D}$, is

$$\mathbb{E}(f|\mathcal{B})(x) = \mathbb{E}_{\{y \in \mathbb{F}_p^n : P_1(y)=P_1(x), \dots, P_C(y)=P_C(x)\}} [f(y)].$$

That is, $\mathbb{E}(f|\mathcal{B})(x)$ is the average of f in the atom to which x belongs. The usefulness of these averages is that they may be used to approximate the function f . It is known that given any $d < p$ and $\delta > 0$, for every function $f : \mathbb{F}_p^n \rightarrow \mathbb{D}$, there exists a polynomial factor $\mathcal{B}$ of degree d and bounded complexity such that

$$\|f - \mathbb{E}(f|\mathcal{B})\|_{U^{d+1}} \leq \delta.$$

Combining this with the discussion in the previous subsection, we may replace any arbitrary function f with a structured function $\mathbb{E}(f|\mathcal{B})$ without changing averages $t_{\mathcal{L}, \alpha}$ by much. Note that $\mathbb{E}(f|\mathcal{B})$ is $\mathcal{B}$ -measurable, that is, it is constant on the atoms of $\mathcal{B}$. Thus, it will suffice to study averages $t_{\mathcal{L}, \alpha}(g)$ where g is a $\mathcal{B}$ -measurable function.

2.4 Strong orthogonality and an invariance theorem

Let $g : \mathbb{F}_p^n \rightarrow \mathbb{D}$ be a $\mathcal{B}$ -measurable function. Equivalently, it can be expressed as $g(x) = \Gamma(P_1(x), \dots, P_C(x))$, where $\Gamma : \mathbb{F}_p^C \rightarrow \mathbb{D}$ is some function. By the Fourier decomposition of Γ ,

$$g(x) = \sum_{\gamma \in \mathbb{F}_p^C} \hat{\Gamma}(\gamma) e_p \left(\sum_{j=1}^C \gamma(j) \cdot P_j(x) \right).$$

Thus, to analyze the average $t_{\mathcal{L}, \alpha}(g)$, we need to understand averages of the form

$$\mathbb{E}_{\mathbf{X}} \left[e_p \left(\sum_{i=1}^m \sum_{j=1}^C \lambda_{i,j} \cdot P_j(L_i(\mathbf{X})) \right) \right], \quad (9)$$

where $\lambda_{i,j} \in \mathbb{F}_p$ are coefficients.

It is known that polynomials which have large rank must be almost unbiased. That is, if $P_1, \dots, P_C$ have large enough rank, then

$$\left| \mathbb{E}_{\mathbf{X}} \left[e_p \left(\sum_{j=1}^C \gamma(j) \cdot P_j(\mathbf{X}) \right) \right] \right| \leq \varepsilon$$

for all nonzero $\gamma \in \mathbb{F}_p^C$. However, when studying averages of the form $\sum \lambda_{i,j} P_j(L_i(\mathbf{x}))$ there could be cancellations; for example, if P is a linear function then $P(x+y+z) - P(x+y) - P(x+z) + P(x) \equiv 0$.

We prove that this is the only bad case that can occur. If $P_1, \dots, P_C$ are polynomials of degree less than p , then for every $\varepsilon > 0$, if their rank is large enough, then for every setting of $\lambda_{i,j} \in \mathbb{F}_p$ one of the following two cases must hold:

$$\left| \mathbb{E}_{\mathbf{X}} \left[e_p \left(\sum_{i=1}^m \sum_{j=1}^C \lambda_{i,j} \cdot P_j(L_i(\mathbf{X})) \right) \right] \right| \leq \varepsilon,$$

or

$$\sum_{i=1}^m \sum_{j=1}^C \lambda_{i,j} \cdot P_j(L_i(\mathbf{x})) \equiv 0.$$

We study when the latter case may hold, and prove relevant conditions on the polynomials. Essentially, we show that the only important parameter is the degree of the polynomials. We use these to prove the following invariance theorem.

PROPOSITION 2.5. *Let $\mathcal{P} = \{P_1, \dots, P_C\}$, $\mathcal{Q} = \{Q_1, \dots, Q_C\}$ be two collections of polynomials over $\mathbb{F}_p^n$ of degree at most $d < p$ such that $\deg(P_i) = \deg(Q_i)$ for every $1 \leq i \leq C$. Let $L_1, \dots, L_m$ be linear forms, and $\Gamma : \mathbb{F}_p^C \rightarrow \mathbb{D}$ be an arbitrary function. Define $f, g : \mathbb{F}_p^n \rightarrow \mathbb{D}$ by*

$$f(x) = \Gamma(P_1(x), \dots, P_C(x))$$

and

$$g(x) = \Gamma(Q_1(x), \dots, Q_C(x)).$$

Then for every $\varepsilon > 0$, if $\text{rank}(\mathcal{P}), \text{rank}(\mathcal{Q})$ are large enough then for all $\alpha \in \{0, 1\}^m$,

$$|t_{\mathcal{L}, \alpha}(f) - t_{\mathcal{L}, \alpha}(g)| \leq \varepsilon,$$

if at least one of the following two conditions hold:

- (i) *The polynomials $P_1, \dots, P_C$ and $Q_1, \dots, Q_C$ are homogeneous.*
- (ii) *The system of linear forms $\{L_1, \dots, L_m\}$ is homogeneous.*

As we shall see, we may assume that all the systems of linear forms arising in our proofs are *homogeneous*. Informally, a system of linear forms is homogeneous if (possibly after a change of basis) we have $L_i(x_1, \dots, x_k) = L'_i(x_1, \dots, x_{k-1}) + x_k$.

2.5 Interior of a set of averages of linear forms

Another ingredient in the proof of Theorem 1.3 is an extension of a graph theoretical result of Erdős, Lovász, and Spencer [3] to the setting of additive combinatorics. We show that if $\mathcal{L}_1, \dots, \mathcal{L}_\ell$ are systems of linear forms, then (excluding some trivial obstructions) the set

$$\{(t_{\mathcal{L}_1}(f), \dots, t_{\mathcal{L}_\ell}(f)) : f : \mathbb{F}_p^N \rightarrow [0, 1]\} \subset \mathbb{R}^\ell$$

has a nonempty interior for some finite $N \in \mathbb{N}$. Formally, for the theorem to hold, we require the systems of linear forms to be *non-isomorphic* and *connected*:

Two systems of linear forms $\mathcal{L}' = \{L'_1, \dots, L'_m\}$ and $\mathcal{L}'' = \{L''_1, \dots, L''_m\}$ are *isomorphic*, if the distributions $(L'_1(\mathbf{X}), \dots, L'_m(\mathbf{X}))$ and $(L''_1(\mathbf{X}), \dots, L''_m(\mathbf{X}))$ are identical. Note that if $\mathcal{L}', \mathcal{L}''$ are isomorphic then $t_{\mathcal{L}'}(f) = t_{\mathcal{L}''}(f)$ for all functions $f : \mathbb{F}_p^n \rightarrow \mathbb{D}$.

A system of linear forms $\mathcal{L} = \{L_1, \dots, L_m\}$ is *connected* if it cannot be partitioned as $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$ where $\text{span}(\mathcal{L}_1) \cap \text{span}(\mathcal{L}_2) = \{\vec{0}\}$. Note that if $\mathcal{L}$ is not connected, then $t_{\mathcal{L}}(f) = t_{\mathcal{L}_1}(f)t_{\mathcal{L}_2}(f)$ for all functions $f : \mathbb{F}_p^n \rightarrow \mathbb{D}$.

2.6 Proof for strongly correlation testable families

Let $\mathcal{D} = \{D_n\}_{n \in \mathbb{N}}$ be a family which is strongly correlation testable. Assume for simplicity of exposition in the proof overview that the following holds (in the actual proof we have generalized averages $t_{\mathcal{L}_i, \alpha_i}(f)$): there exists a system of linear forms $\mathcal{L}_1, \dots, \mathcal{L}_\ell$ such that for every $\varepsilon > 0$ there exists $\delta \in (0, \varepsilon)$ and $n_0 \in \mathbb{N}$, such that the following two sets are disjoint:

$$T_\varepsilon := \{(t_{\mathcal{L}_1}(f), \dots, t_{\mathcal{L}_\ell}(f)) \mid n > n_0, f : \mathbb{F}_p^n \rightarrow \mathbb{D}, \|f\|_{u(D_n)} \geq \varepsilon\},$$

and

$$S_\varepsilon := \{(t_{\mathcal{L}_1}(f), \dots, t_{\mathcal{L}_\ell}(f)) \mid n > n_0, f : \mathbb{F}_p^n \rightarrow \mathbb{D}, \|f\|_{u(D_n)} \leq \delta\}.$$

We can furthermore assume that these systems are homogeneous, and that their Cauchy-Schwarz complexity is $s < p$ (this follows from the assumption the the number of queries is $q \leq p$).

Let $t \in \mathbb{N}$ be maximal such the following holds. There exists polynomials Q_n of degree exactly t such that

- Q_n has noticeable correlation with D_n : $\liminf_{n \rightarrow \infty} \|e_p(Q_n)\|_{u(D_n)} > 0$.
- Q_n has “large enough rank”.

We can show that $t \leq s$: if Q_n are polynomials of degree $> s$ and large enough rank, then $\|e_p(Q_n)\|_{U^{s+1}} \approx 0$ and hence $t_{\mathcal{L}_i}(e_p(Q_n)) \approx 0$, for all $1 \leq i \leq \ell$. Thus we must have $\|e_p(Q_n)\|_{u(D_n)} \approx 0$.

Let $f_n : \mathbb{F}_p^n \rightarrow \mathbb{D}$ be a sequence of functions. We establish the main result (for strongly correlation testable families) by showing that:

$$\lim_{n \rightarrow \infty} \|f_n\|_{u(D_n)} = 0 \iff \lim_{n \rightarrow \infty} \|f_n\|_{U^{t+1}} = 0.$$

Proof sketch of $\Leftarrow$:

Assume by contradiction that $\lim_{n \rightarrow \infty} \|f_n\|_{U^{t+1}} = 0$ but that $\varepsilon := \liminf_{n \rightarrow \infty} \|f_n\|_{u(D_n)} > 0$. Let $\varepsilon' > 0$ be the L_∞ distance between T_ε and S_ε (here we use the fact that their closures are disjoint, hence there is positive distance between the sets). Let $\mathcal{B}_n$ be a polynomial factor of degree s such that $\|f_n - \mathbb{E}(f_n|\mathcal{B}_n)\|_{U^{s+1}} \leq \eta$. Choosing $\eta > 0$ small enough we can guarantee that $|t_{\mathcal{L}_i}(f_n) - t_{\mathcal{L}_i}(\mathbb{E}(f_n|\mathcal{B}_n))| \leq \varepsilon'/2$, hence we must have $\|\mathbb{E}(f_n|\mathcal{B}_n)\|_{u(D_n)} \geq \delta$.

Assume $\mathcal{B}_n$ is defined by polynomials $P_{n,1}, \dots, P_{n,C}$, which we can assume to be of high enough rank. For $\gamma \in \mathbb{F}_p^C$, define $P_{n,\gamma}(x) := \sum \gamma(i)P_{n,i}(x)$. We have $\mathbb{E}(f_n|\mathcal{B}_n)(x) = \sum_{\gamma \in \mathbb{F}_p^C} c_\gamma e_p(P_{n,\gamma}(x))$ where $|c_\gamma| \leq 1$. Using the assumption that $\lim_{n \rightarrow \infty} \|f_n\|_{U^{t+1}} = 0$, we show that if $\deg(P_\gamma) \leq t$, then its contribution to the sum is negligible, i.e. c_γ can be assumed to be arbitrarily small. Thus, there must exist a polynomial $P_{n,\gamma}$ of degree at least $t+1$ such that $\|e_p(P_{n,\gamma})\|_{u(D_n)} \geq \delta p^{-C}$. As this polynomial has “large enough rank”, this contradicts the maximality of t .

Proof sketch of $\Rightarrow$:

Assume that $\varepsilon := \lim_{n \rightarrow \infty} \|f_n\|_{U^{t+1}} > 0$ (actually we have $\liminf$, which we can replace by an actual limit by Condition **A1**). Thus, there exist polynomials P_n of degree at most t such that $|\langle f_n, P_n \rangle| \geq \varepsilon$. We can assume these polynomials to have “large enough rank”. Assume first that these polynomials are of degree exactly t . By the definition of t , there exist polynomials Q_n of degree t and “large enough rank” such that $\|Q_n\|_{u(D_n)} \geq \varepsilon'$. We use the invariance theorem to construct a new sequence of functions $f'_n : \mathbb{F}_p^n \rightarrow \mathbb{D}$ (essentially replacing P_n with Q_n) such that

- f'_n has correlation with Q_n . In fact, this correlation is strong enough to guarantee that $\|f'_n\|_{u(D_n)} \geq \varepsilon''$ for some $\varepsilon'' > 0$ independent of n .
- Averages $t_{\mathcal{L}_i}$ cannot distinguish f_n from f'_n . So, we must also have $\|f_n\|_{u(D_n)} \geq \delta'' > 0$.

The case where P_n have degrees less than t is reduced to the case of degree t . We use the theorem on the interior of a set of averages of linear forms to argue that we can (essentially) tweak f_n slightly, without changing averages $t_{\mathcal{L}_i}$ by much, but such that f_n will have correlation with polynomials of degree exactly t and large enough rank.

2.7 Proof for correlation testable families

The proof for the correlation testable families (i.e. with guarantees only for functions $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$) follows similar steps, albeit slightly more involved. Let $P(\mathbb{F}_p) \subset \mathbb{R}^p$ denote the convex set of probability distributions defined over $\mathbb{F}_p$. Then correlation testable families can in fact test randomized functions as well, i.e. functions $f : \mathbb{F}_p^n \rightarrow P(\mathbb{F}_p)$. Once this is established the remainder of the proof follows identical lines to the case of strongly correlation testable families, where we employ the fact that $P(\mathbb{F}_p)$ is a convex set to appropriately define averages such as $\mathbb{E}(f|\mathcal{B})$.

3. CONCLUDING REMARKS

In this paper we study affine invariant properties which are testable. We show that essentially every such property can be tested by an appropriate Gowers uniformity norm. One technical limitation of our techniques is that they hold only if the field size is not too small (i.e. if the Cauchy-Schwarz complexity is smaller than the field size). The main reason for this obstacle is that the inverse theorem for Gowers norm, which fully established for large fields, is only partially understood for small fields. Thus, we do not have corresponding decomposition theorems for small fields which are suitable for our needs.

Even if the conjectured form of the inverse theorem for the Gowers norm for small fields was proved, it would still not answer the following problem: is it possible to test, using a constant number of queries, whether a function $f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p$ is correlated to a polynomial of degree d , where $d > p$? We know that the Gowers norm test fails, as it actually tests distance to a larger set of functions (non-classical polynomials). The simplest case which is unknown is the following:

PROBLEM 3.1. *Let $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$. Does there exist a test which queries f on a constant number of positions, and which can distinguish whether f has noticeable or negligible correlation with cubic polynomials?*

In fact, even the following simpler problem is unknown:

PROBLEM 3.2. *Let $\varepsilon > 0$ and $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$. Does there exist a test which queries f on $q(\varepsilon)$ positions, and which can distinguish whether f has correlation at least ε , or at most $\delta(\varepsilon)$, with cubic polynomials?*

4. REFERENCES

- [1] V. Bergelson, T. Tao, and T. Ziegler. An inverse theorem for the uniformity seminorms associated with the action of $\mathbb{F}_p^\infty$. *Geom. Funct. Anal.*, 19(6):1539–1596, 2010.
- [2] M. Blum, M. Luby, and R. Rubinfeld. Self-testing/correcting with applications to numerical problems. In *STOC '90: Proceedings of the twenty-second annual ACM symposium on Theory of computing*, pages 73–83, New York, NY, USA, 1990.
- [3] P. Erdős, L. Lovász, and J. Spencer. Strong independence of graphcopy functions. In *Graph theory and related topics (Proc. Conf., Univ. Waterloo, Waterloo, Ont., 1977)*, pages 165–172. Academic Press, New York, 1979.
- [4] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, 1998.
- [5] O. Goldreich and D. Ron. On proximity oblivious testing. In *STOC '09: Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 141–150, New York, NY, USA, 2009. ACM.
- [6] W. T. Gowers. A new proof of Szemerédi’s theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.
- [7] W. T. Gowers and J. Wolf. Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$. *Geom. Funct. Anal.*, to appear, 2010.
- [8] W. T. Gowers and J. Wolf. The true complexity of a system of linear equations. *Proc. Lond. Math. Soc.* (3), 100(1):155–176, 2010.
- [9] 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.
- [10] B. Green and T. Tao. Linear equations in primes. *Ann. of Math.* (2), 171(3):1753–1850, 2010.
- [11] H. Hatami and S. Lovett. Correlation testing for affine invariant properties on $\mathbb{F}_p^n$ in the high error regime. *manuscript*.
- [12] T. Kaufman and S. Lovett. Worst case to average case reductions for polynomials. In *FOCS '08: Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 166–175, Washington, DC, USA, 2008. IEEE Computer Society.
- [13] T. Kaufman and M. Sudan. Algebraic property testing: the role of invariance. In *STOC '08: Proceedings of the 40th annual ACM symposium on Theory of computing*, pages 403–412. ACM, 2008.
- [14] R. Rubinfeld and M. Sudan. Robust characterizations of polynomials and their applications to program testing. Technical report, Ithaca, NY, USA, 1993.
- [15] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975. Collection of articles in memory of Juriui Vladimirovich Linnik.
- [16] T. Tao and T. Ziegler. The inverse conjecture for the Gowers norm over finite fields via the correspondence principle. *Anal. PDE*, 3(1):1–20, 2010.