Limits on the rate of locally testable affine-invariant codes

Eli Ben-Sasson* and Madhu Sudan**

Abstract. Despite its many applications, to program checking, probabilistically checkable proofs, locally testable and locally decodable codes, and cryptography, "algebraic property testing" is not well-understood. A significant obstacle to a better understanding, was a lack of a concrete definition that abstracted known testable algebraic properties and reflected their testability. This obstacle was removed by [Kaufman and Sudan, STOC 2008] who considered (linear) "affine-invariant properties", i.e., properties that are closed under summation, and under affine transformations of the domain. Kaufman and Sudan showed that these two features (linearity of the property and its affine-invariance) play a central role in the testability of many known algebraic properties. However their work does not give a complete characterization of the testability of affine-invariant properties, and several technical obstacles need to be overcome to obtain such a characterization. Indeed, their work left open the tantalizing possibility that locally testable codes of rate dramatically better than that of the family of Reed-Muller codes (the most popular form of locally testable codes, which also happen to be affine-invariant) could be found by systematically exploring the space of affine-invariant properties.

In this work we rule out this possibility and show that general (linear) affine-invariant properties are contained in Reed-Muller codes that are testable with a slightly larger query complexity. A central impediment to proving such results was the limited understanding of the structural restrictions on affine-invariant properties imposed by the existence of local tests. We manage to overcome this limitation and present a clean restriction satisfied by affine-invariant properties that exhibit local tests. We do so by relating the problem to that of studying the set of solutions of a certain nice class of (uniform, homogenous, diagonal) systems of multivariate polynomial equations. Our main technical result completely characterizes (combinatorially) the set of zeroes, or algebraic set, of such systems.

Keywords: Property testing, Symmetries, Direct sums, Error-correcting codes

* Department of Computer Science, Technion — Israel Institute of Technology, Haifa, Israel, eli@cs.technion.ac.il. Research funded partially by the European Community's Seventh Framework Programme (FP7/2007-2013) Grant 240258 and the US-Israel Binational Science Foundation Grant 2006104.

** Microsoft Research New England, Cambridge, Massachusetts, USA, madhu@mit.edu.

1 Introduction

In this work we consider an interesting subclass of "locally correctable" and "locally testable codes", namely "affine-invariant" codes, and prove upper bounds (limitations) on their rate. In the process we also introduce techniques of relevance to "algebraic property testing" and present a characterization of the set of solutions of a certain natural system of multivariate polynomials that lies at the core of our study.

1.1 Locally correctable and locally testable codes, affine-invariance, and main result

Locally correctable codes (LCCs) are error-correcting codes with the property that every entry of a corrupted codeword can be corrected, with high probability, by examining a small (random) subset of other entries of the corrupted codeword. Locally correctable codes are a special class of locally decodable codes (LDCs) studied in the work of [5, 24, 27] and formally defined by [21]. These codes are tightly connected to the construction of private information retrieval schemes [13] and we refer the reader to [30] for more information. One of the major open problems regarding LDCs is that of determining the minimal length n of a binary LDC by which messages of k bits can be encoded, and the current lower and upper bounds on n display an exponential gap. Namely, [28] showed n must be at least (roughly) $k^{1+\frac{2}{r}}$ whereas the best upper bounds of [29] show n is at most (roughly) $\exp(k^{1/\log\log k})$ (cf. [15]).

Locally testable codes are error-correcting codes for whom membership can be tested extremely efficiently, probabilistically. Specifically, a linear code $\mathcal{C} \subseteq \Sigma^N$ is k-locally testable if there exists an algorithm T that accesses a word $w \in \Sigma^N$ as an oracle, queries the value of w on k coordinates, and accepts with probability one if $w \in \mathcal{C}$ and rejects with constant probability if w is "far" from all codewords of $\mathcal{C}$. ("Far" here refers to the relativized Hamming distance between words.)

Locally testable codes have implicitly been a subject of active study ever since the work of [11] that showed that (effectively) the Hadamard code is 3-locally testable. They play a major role in the construction of PCPs [4, 3] from the early days of this theorem and continuing through the recent work of [14]. Their systematic investigation was started in [17] and yet most basic questions about their limits remain unanswered (e.g., is there an asymptotically good family of locally testable codes?).

A particularly interesting class of locally testable and locally correctable codes are the affine-invariant ones. Here the code is a linear code over some finite field $\mathbb{F}$ and the coordinates of the code are themselves vector spaces over some finite extension field $\mathbb{K}$ of $\mathbb{F}$. Thus words in such codes can be viewed as functions from $\mathbb{K}^m$ to $\mathbb{F}$ and the code is a subfamily of such functions. The code is said to be affine invariant if it is invariant under affine-transformations of the domain. Specifically if $A : \mathbb{K}^m \to \mathbb{K}^m$ is an affine transformation and $f : \mathbb{K}^m \to \mathbb{F}$ is a function in $\mathcal{C}$, then so is $f \circ A$ where $f \circ A(x) = f(A(x))$.

Affine-invariant codes form a natural class of algebraic codes and have been studied by the error-correcting-codes community since the late 1960's (cf. [20] and references therein). In the context of locally testable and locally correctable codes, affine-invariance facilitates natural local correctiong/testing procedures under minimal conditions. Specifically, it is well known that for a linear code to be testable it must have a low weight codeword in its dual, or equivalently a local "constraint" (see, for instance, [7]). In the notation used above for affine invariant codes, a k-local "constraint" is a collection of points $\alpha_1, \ldots, \alpha_k \in \mathbb{K}^m$ and values $\lambda_1, \ldots, \lambda_k \in \mathbb{F}$ such that for every function $f \in \mathcal{C}$, it is the case that $\sum_{i=1}^{k} \lambda_i f(\alpha_i) = 0$. For affine-invariant codes the presence of one local constraint immediately implies many local constraints by affine "rotations": For every affine map A, the set of points $A(\alpha_1), \ldots, A(\alpha_k)$ also define a constraint on $\mathcal{C}$. This abundance of constraints leads easily to a local-correction procedure and also raises the hope that affine-invariant codes may be locally testable, and indeed [23] show that if the code is characterized by the set of constraints derived from the affine rotations of a single constraint, then it is also locally testable. (The more optimistic hope, that all affine-invariant locally-characterized codes are also locally testable, has been recently refuted in [8] as a result of this work .)

We point out that it is the *abundance* of local constraints, not their mere existence, that seems to be essential for obtaining locally testable codes. In extreme cases where there is no abundance of local constraints, such as for low-density-parity-check (LDPC) codes based on random expanders, or for codes that have the very minimal number of local constraints needed to characterize them, there cannot be any hope for local testability [7, 6]. But, all things considered, abundance of local constraints should reduce the rate of the code, unless the constraints are carefully chosen in an algebraically consistent way. The class of affine invariant codes offered a promising approach to balance the need for abundance of local constraints with maintaining high rate.

This leads to the question: Which affine invariant codes have local constraints (and characterizations), and in particular how local can the constraints be, given other parameters of the code, most notably, its rate? One, somewhat optimistic hope, was that affine-invariance might lead to simpler constructions of locally testable codes matching the best known parameters (the current constructions are immensely complicated [9, 14]), or even improve on them, since the scope is wider than just the class of Reed-Muller codes. This question however, resisted attacks till now, since the question of determining when a low-weight constraint can exist in an affine-invariant code leads to questions about the zeroes of certain systems of multivariate polynomial equations and these are challenging to analyze.

Here we take some first steps in this direction, though unfortunately to give a negative answer to the optimistic hope above. Specifically, we give a full analysis of a certain class of polynomial equations that arise in this setting to get a rate upper bound on affine invariant codes. For simplicity of exposition we describe our result for the case of prime fields $\mathbb{F} = \mathbb{F}_p$. The statement for the case fields of size $p^r, r > 1$ is somewhat more technical but the rate bounds we get for

this case are similar to that of prime fields (cf. Theorem 2 and Corollary 1). Our main theorem (Theorem 2) shows that if $\mathbb{K}$ is an extension field of $\mathbb{F}$ and $\mathcal{C}$ is a k-locally testable/correctable code, then $\mathcal{C}$ is contained in a p^{k-1}-locally testable Reed-Muller code. If k and p are constants (which is the desired setting of parameters) then it says that going to general affine-invariance only buys (at best) a constant difference in the locality, when compared to the Reed-Muller codes. Since Reed-Muller codes with constant locality over constant field sizes are known to have exponentially low-rate, this rules out the hope described in the previous paragraph, by a long margin.

Notice there is an exponential gap between the query complexity of affine-invariant codes with a k-local constraint and the query complexity of the Reed-Muller code which we show contains them, which is p^{k-1}. Getting a full characterization of affine-invariant codes with a k-local constraint, even over specific fields (like $\mathbb{F}_{2^n}$ for prime n, a field which contains no subfields other than $\mathbb{F}_2$) seems to us like an interesting question for future research.

1.2 Algebraic property testing

Property testing considers the task of testing if a function f from a large domain D to a small range R satisfies some given property, where the property itself is given by the set of functions $\mathcal{F} \subseteq \{g : D \to R\}$ that satisfy the property. Again the interest here is in "quick and dirty" tests, i.e., probabilistic tests that query the given function f on few inputs, and accept if $f \in \mathcal{F}$ and reject with constant probability if f is far from $\mathcal{F}$. (Note that a locally testable code is just property testing where we view the set of functions $\mathcal{F}$ as an error-correcting code.)

Property testing also emerged in the work of [11], was formally defined by [25], and was systematically explored (in particular in non-algebraic contexts) by [16]. Subsequently the study of combinatorial property testing, and in particular, graph property testing has developed into a rich study and by now we have almost complete understanding (at least in the dense-graph model) of which graph properties are locally testable [1, 12].

In contrast algebraic properties have not been understood as well, despite the overwhelming applications in complexity, and indeed till recently even an understanding of what makes a property algebraic was missing. The concept of affine-invariance was introduced by [23] to propose such a notion, and when the domain is a vector space over a small field $\mathbb{K}$ (of constant size) they manage to characterize locally testable properties completely. Such codes are constant-locally testable if and only if they admit a constant local constraint, and the size of the constraint can be related loosely to the highest degree of polynomials in the family.

This naturally leads to the question: What about affine invariant codes over large fields. (This family of properties includes, for instance, all sets of low-degree polynomials, i.e., the families of Hadamard, Reed-Solomon and Reed-Muller codes.) In particular for the extreme, and essentially most general case when $m = 1$ and the functions of interest map $\mathbb{K}$ to a prime subfield $\mathbb{F}_p$, there was no interesting relationships known between the degrees of the functions in

the family and the locality of the test. And such understanding is essential to get a characterization of affine-invariant locally testable codes that would be analogous to the characterizations of graph properties of [1, 12].

Our work takes a step in this direction by giving non-trivial lower bounds on the locality of tests for affine-invariant properties in the general case. Below we describe the main technical question resolved in this paper (which has a self-contained description).

2 Definitions and Main Results

2.1 Preliminaries — Locally testable, and Reed-Muller codes

Notation We use $[n]$ to denote the set $\{1, \ldots, n\}$. Throughout we let $\mathbb{F}, \mathbb{K}, \mathbb{L}$ denote fields. The q element finite field is denoted by $\mathbb{F}_q$. An $[N, K, D]_{\mathbb{F}}$-(linear) code is a K-dimensional subspace $\mathcal{C} \subseteq \mathbb{F}^N$ of Hamming distance D. Elements of $\mathcal{C}$ are referred to as codewords (of $\mathcal{C}$). Two vectors $u, w \in \mathbb{F}^N$ are said to be δ-close if they are within Hamming distance $\leq \delta N$ of each other, otherwise they are said to be δ-far. A vector u is said to be δ-close to $\mathcal{C}$ if it is δ-close to some codeword $w \in \mathcal{C}$, otherwise we say w is δ-far from $\mathcal{C}$. We define $\langle u, w \rangle \triangleq \sum_{i=1}^{N} u_i w_i$. Let $\mathcal{C}^{\perp} = \left\{ u \in \mathbb{F}^N \;\middle|\; \langle u, w \rangle = 0 \text{ for all } w \in \mathcal{C} \right\}$ denote the space that is dual to $\mathcal{C}$ (it is also known as the *dual code* of $\mathcal{C}$).

We recall the standard definitions of a tester for a linear code and a linear locally testable code. All codes considered in this paper are linear so from here on we drop further reference to this linearity (of testers, codes, etc.).

Definition 1 (Tester). *Suppose $\mathcal{C}$ is a $[N, K, D]_{\mathbb{F}}$-code. A* k-query tester *for $\mathcal{C}$ is a probabilistic oracle algorithm T that makes at most k oracle queries to a word $w \in \mathbb{F}^N$ and outputs an accept/reject verdict. The tester is said to have completeness c and ϵ-soundness s if it accepts every codeword of $\mathcal{C}$ with probability at least c and accepts words that are ϵ-far from $\mathcal{C}$ with probability at most s.*

Definition 2 (Locally Testable Code (LTC)). *An $[N, K, D]_{\mathbb{F}}$-code $\mathcal{C}$ is said to be a* (k, ϵ, ρ)-Locally Testable Code *(LTC) if there exists a k-query tester that has completeness c and ϵ-soundness $c - \rho$.*

We are typically interested in infinite family of codes. If an infinite family of codes is a (k, ϵ, ρ)-LTC for absolute constants k and $\epsilon, \rho > 0$, then we simply refer to this (family of) code(s) as an LTC. For linear LTCs the nature of tests can be simplified significantly, due to a result of [7], to get them to a canonical form, which has perfect completeness ($c = 1$), and is *non-adaptive* (while the soundness parameter ρ changes by a constant factor). This leads to the following definition.

Definition 3 (Canonical tester). *A* canonical k-query test *for $\mathcal{C}$ is given by an element $u \in \mathcal{C}^{\perp}$ that has support size at most k, i.e., $|\{i \mid u_i \neq 0\}| \leq k$, where the test accepts $w \in \mathbb{F}^n$ if and only if $\langle u, w \rangle = 0$. A* k-query canonical tester *T for $\mathcal{C}$ is defined by a distribution μ over canonical k-query tests. Invoking the tester T on a word $w \in \mathbb{F}^n$ is done by sampling a test u according to the distribution μ and outputting* accept *if the canonical test given by u accepts.*

The following proposition of [7] — stated as Theorem 3.3 there — shows that tests may always be assumed to be canonical (up to a constant factor change in soundness).

Proposition 1. *For every $\epsilon, \rho > 0$ and positive integer k, there exist $\rho' > 0$ such that every (k, ϵ, ρ)-LTC has a canonical k-query tester with perfect completeness and ϵ-soundness $1 - \rho'$.*

Our main theorem compares the performance of affine-invariant locally testable codes to that of Reed-Muller codes, which we define next.

Definition 4 (Reed-Muller codes). *For $\mathbb{F}$ a finite field of size q and m, k integers, the m-variate Reed-Muller code of degree k over $\mathbb{F}$, denoted $\mathrm{RM}[q, m, k]$ is the $[N = q^m, K = \binom{m+k}{k}, D = q^{m-k}]_{\mathbb{F}}$-code whose codewords are all evaluations of m-variate polynomials over $\mathbb{F}$ of degree at most k.*

These codes have also been studied for the testability properties (see, e.g., [25], [2], [26], [22], [19], and [10]) and the case most relevant to us is that of constant q and k and arbitrarily large m. For this choice of parameters the codes are known to be $(q^{O(k)}, \epsilon, \rho)$-locally testable for some constants $\epsilon, \rho > 0$ that may depend on q and k [2, 22].

2.2 Affine invariant codes

The main concept of interest to us is that of affine-invariance. We borrow some of the main definitions related to this concept from [23].

From here on we associate a code with a family of functions. Let p be a prime, $\mathbb{F} = \mathbb{F}_q$ for $q = p^r$ be a finite field and let $\mathbb{K} = \mathbb{F}_Q$ for $Q = q^n$ be an extension of $\mathbb{F}$. For integer m we can consider $[N = Q^m, k, d]_{\mathbb{F}}$-codes whose entries are indexed by elements of $\mathbb{K}^m$. In other words, from here on a code will be identified with an $\mathbb{F}$-linear subspace of $\{\mathbb{K}^m \to \mathbb{F}\}$, the space of all functions from $\mathbb{K}^m$ to $\mathbb{F}$.

Definition 5 (Affine invariant codes). *Let $\mathbb{K}$ be a finite degree extension of $\mathbb{F}$. A code $\mathcal{C} \subseteq \{\mathbb{K}^m \to \mathbb{F}\}$ is said to be* affine invariant *if it is invariant under the action of the affine monoid[1] over $\mathbb{K}^m$. In other words, for every $f \in \mathcal{C}$ and every affine transformation $A : \mathbb{K}^m \to \mathbb{K}^m$, the function $f \circ A$ defined by $(f \circ A)(x) = f(A(x))$ belongs to $\mathcal{C}$ as well.*

The work of [7] shows that in order for a linear property to be testable, it must have some "local constraints" (low-weight words in its dual). For affine invariant codes, [23] show that when $\mathbb{K}$ is small, then the existence of such constraints is also a sufficient condition. (Our main result will show that the existence of such constraints imposes a bound on the rate of a code, over any field $\mathbb{K}$, not just over fields of constant size.) We recall the following definition from [23].

[1] The set of all affine maps from $\mathbb{K}^m$ to itself forms a monoid under composition. If one restricted this set to full rank maps, then one gets a group.

Definition 6 (k-local constraint). *A k-local constraint is given by k distinct points in $\mathbb{K}^m$ $\alpha = (\alpha_1, \ldots, \alpha_k) \in (\mathbb{K}^m)^k$. We say that a code $\mathcal{C} \subseteq \{\mathbb{K}^m \to \mathbb{F}\}$ satisfies (or, has) a* k-local constraint *α if there exists nonzero $\Lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{F}^k$ such that $\sum_{i=1}^{k} \lambda_i f(\alpha_i) = 0$ for every $f \in \mathcal{C}$.*

The following statement is the main result of [23] regarding the local testability of affine invariant codes, and is stated as Theorem 2.10 there.

Theorem 1 (Affine invariant codes satisfying a k-local constraint over a small field are locally testable). *For fields $\mathbb{F} \subseteq \mathbb{K}$ with $|\mathbb{F}| = q$ and $|\mathbb{K}| = Q$, let $\mathcal{F} \subseteq \{\mathbb{K}^m \to \mathbb{F}\}$ be an affine-invariant code satisfying a k-local constraint. Then for any $\delta > 0$, the code $\mathcal{F}$ is*

$$\left(k' = (Q^2 k)^{Q^2}, \delta, \frac{\delta}{2(2k'+1)(k'+1)}\right) \text{-locally testable.}$$

Notice the above theorem implies local testability only when the field $\mathbb{K}$ is relatively small, and is of interest only when $m \to \infty$. When m is small (and $\mathbb{K}$ large) no general bounds were known on the locality of the tests. [18] show that it is possible to have affine invariant families with one 8-local constraint that cannot be characterized by $O(1)$-local constraints. And all this previous work leaves open the possibility that there may exist other affine-invariant families that are $O(1)$-locally characterized, perhaps even $O(1)$-testable (say, over fields of growing size and $m = 1$), and do have large rate. Our work rules this out.

We can now state our main theorem which bounds the rate of affine invariant codes containing a k-local constraint.

Theorem 2 (Affine invariant families with a local constraint are contained in low-degree Reed-Muller codes). *Let p be a prime and r, n, m be positive integers and let $q = p^r$ and $Q = q^n$. For $\mathbb{F} = \mathbb{F}_q$ and $\mathbb{K} = \mathbb{F}_Q$ a degree-n extension of $\mathbb{F}$, let $\mathcal{C} \subseteq \{\mathbb{K}^m \to \mathbb{F}\}$ be an affine-invariant family that satisfies a k-local constraint for $k \geq 2$. Then*

1. *The dimension of $\mathcal{C}$ as a vector space over $\mathbb{F}_q$ is at most $(mrn)^{k-1}$. Since the blocklength of $\mathcal{C}$ is $Q^m = p^{rmn}$ we get*

$$\dim(\mathcal{C}) \leq (m \log_p Q)^{k-2}.$$

2. *$\mathcal{C}$ is isomorphic to a subcode[2] of $\mathrm{RM}[q, nm, (k-2)q/p]$. In particular, for $q = p$ we get that $\mathcal{C}$ is isomorphic to a subcode of $\mathrm{RM}[p, nm, k-2]$.*

Note that when $q = p$, Part (1) of the theorem above follows from Part (2), since the dimension of $\mathrm{RM}[p, nm, s]$ is at most $(mn)^s$. When $q = p^r$ for $r > 1$, this is not true, and the dimension of the code $\mathrm{RM}[q, nm, sq/p]$ is much larger. In this case Part (2) is a weak description of our understanding of $\mathcal{C}$. A

[2] In other words, there exists an isomorphism $\phi : \mathbb{F}^{nm} \to \mathbb{K}^m$ such that for every $f \in \mathcal{C}$, the function $(f \circ \phi) \in \{\mathbb{F}^{nm} \to \mathbb{F}\}$ defined by $(f \circ \phi)(x) = f(\phi(x))$ belongs to $\mathrm{RM}[q, nm, (k-2)q/p]$.

somewhat better understanding of affine-invariant codes over $\mathbb{F}_{p^r}$, for $r > 1$ can be obtained if we use a broader class of codes. In particular, by viewing a code over $\mathbb{F}_{p^r}$ as a code over the vector space $\mathbb{F}_p^r$, or as an r-tuple of codes over $\mathbb{F}_p$, one gets a more strict inclusion for such codes. Specifically, let $\mathrm{RM}[p, n, k-2]^r$ denote codes obtained by evaluations of $f = \langle f_1, \ldots, f_r \rangle : \mathbb{F}_p^n \to \mathbb{F}_p^r$, where each $f_i : \mathbb{F}_p^n \to \mathbb{F}_p$ is an n-variate polynomial over $\mathbb{F}_p$ of degree at most $k-2$. We then have the following Corollary of Theorem 2.

Corollary 1 (Affine invariant families with a local constraint over fields of prime powers). *Let p be a prime and r, n, m be positive integers and let $q = p^r$ and $Q = q^n$. For $\mathbb{F} = \mathbb{F}_q$ and $\mathbb{K} = \mathbb{F}_Q$ the degree-n extension of $\mathbb{F}$, let $\mathcal{C} \subseteq \{\mathbb{K}^m \to \mathbb{F}\}$ be an affine-invariant family that satisfies a k-local constraint. Then for every $\mathbb{F}_p$-linear bijection $\psi : \mathbb{F}_q \to \mathbb{F}_p^r$, the code $\mathcal{C}' = \{\psi \circ f | f \in \mathcal{C}\} \subseteq \{\mathbb{K}^m \to \mathbb{F}_p^r\}$ is isomorphic to a subcode of $\mathrm{RM}[p, nmr, k-2]^r$.*

Proof. For $i \in [r]$, let $\mathcal{C}_i$ be the projection of $\mathcal{C}'$ to the ith coordinate. Then $\mathcal{C}_i$ is a $\mathbb{F}_p$-linear, affine-invariant code (over the domain $\mathbb{F}_Q^m$). By Theorem 2 we get that it is isomorphic to a subcode of $\mathrm{RM}[p, nmr, k-2]$. It follows that $\mathcal{C}'$ is a subcode of $\mathrm{RM}[p, nmr, k-2]^r$.

3 Proof of Main Theorems

In this section we prove Theorem 2 modulo some technical lemmas. It is not hard to show that if Theorem 2 holds for the case $m = 1$ then it holds for all positive integers m. (Proof omitted due to space limitations.)

From now on we consider only univariate functions, i.e., $\mathcal{C} \subseteq \{\mathbb{K} \to \mathbb{F}\}$. Recall that every function from $\mathbb{K} \to \mathbb{K}$ and hence from $\mathbb{K} \to \mathbb{F}$ is the evaluation of a polynomial in $\mathbb{K}[x]$ of degree at most $q^n - 1$. For a polynomial $g \in \mathbb{K}[x]$ given by $g(x) = \sum_d c_d x^d$, let $\mathrm{supp}(g)$ denote its support, i.e., $\mathrm{supp}(g) = \{d | c_d \neq 0\}$. The set of degrees in the support of the functions in $\mathcal{C}$ turns out be a central ingredient in understanding the structure of $\mathcal{C}$, motivating the following definition.

Definition 7 (Degree set of $\mathcal{C}$). *For a class of functions $\mathcal{C} \subseteq \{\mathbb{K} \to \mathbb{F}\}$, its degree set is the set $D(\mathcal{C}) = \cup_{g \in \mathcal{C}} \mathrm{supp}(g)$.*

It turns out that the representations of elements of $D(\mathcal{C})$ in base p play a central role in the structure of affine-invariant families over fields of characteristic p. To this end we introduce some terminology.

For integer d, let $[d]_p = \langle d_0, d_1, \ldots \rangle$ denotes its representation in base p (i.e., $0 \leq d_i < p$ and $d = \sum_{i=0}^{\infty} d_i p^i$). The *$p$-weight* of d, denoted $\mathrm{wt}_p(d)$, is the quantity $\sum_{i=0}^{\infty} d_i$. We say e is in the p-shadow of d, denoted $e \leq_p d$, if $[e]_p = \langle e_0, e_1, \ldots \rangle$ and $e_i \leq d_i$ for all i. The set $\{e | e \leq_p d\}$ is called the p-shadow of d. The following Lemma appears as Theorem 1 in [20].

Lemma 1. *For every affine invariant family $\mathcal{C} \subseteq \{\mathbb{K} \to \mathbb{F}\}$ where $\mathbb{F}, \mathbb{K}$ are fields of characteristic p, $D(\mathcal{C})$ is closed under p-shadow, i.e., if $d \in D(\mathcal{C})$ and $e \leq_p d$ the $e \in D(\mathcal{C})$.*

3.1 Uniform Homogenous Diagonal Systems of Polynomial Equations

The task of finding the set of zeroes of a system of multivariate polynomial equations is a central theme in mathematics. (Linear algebra considers the special case where all equations are linear/affine and understanding the "variety" of a given system of (higher-degree) equations is a central theme in algebraic geometry.) In general of course, the set of zeroes may be too complex, even for degree two polynomials. Nevertheless, our quest to understand the locality of constraints in an affine-invariant property leads to such a question, where the set of polynomials has a reasonably clean description. Somewhat surprisingly, we are even able to describe the set of zeroes in a fairly precise way. We describe the class of polynomial systems that we consider next.

Definition 8 (Uniform Homogenous Diagonal (UHD) System). *Fix a system of polynomials $P_1, \ldots, P_m \in \mathbb{F}[X_1, \ldots, X_k]$.*

- *We say the system is* homogenous *if every polynomial in the system is homogenous.*
- *We say that the system is* diagonal *if every monomial in the support of every polynomial is a power of a single variable. I.e, a homogenous system is diagonal if for every $j \in [m]$, it is the case that $P_j(X_1, \ldots, X_k) = \sum_{i=1}^{k} \lambda_{ji} \cdot X_i^{d_j}$.*
- *We say a homogenous diagonal system is* uniform *if the coefficients are the same for every polynomial, i.e., λ_{ji} is independent of j.*

We conclude that a uniform homogenous diagonal system is given by a sequence of coefficients $\Lambda = \langle \lambda_1, \ldots, \lambda_k \rangle \in \mathbb{F}^k$ and degrees $D = \{d_1, \ldots, d_m\}$ such that $P_j(X_1, \ldots, X_k) = \sum_{i=1}^{k} \lambda_i X_i^{d_j}$. We refer to such a system as the (D, Λ)-UHD system. We say that the (D, Λ)-system has a pairwise-distinct solution *over some field $\mathbb{K}$ if there exist* distinct *values $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ such that $P_j(\alpha_1, \ldots, \alpha_k) = 0$ for every $j \in [m]$.*

The following lemma motivates the study of UHD systems in our setting.

Lemma 2. *If an affine-invariant property $\mathcal{C} \subseteq \{\mathbb{K} \to \mathbb{F}\}$ has a k-local constraint, then there exists a non-zero vector $\Lambda \in \mathbb{F}^k$ such that the $(D(\mathcal{C}), \Lambda)$-UHD system has a pairwise-distinct solution over $\mathbb{K}$.*

Proof omitted due to space considerations. The following theorem is the main technical claim of this paper.

Theorem 3 (Shadow-closed UHD systems containing nontrivial solutions have bounded weight). *Let $\mathbb{F}$ be any field of characteristic p and let D be a p-shadow-closed set of integers containing an element d with $\mathrm{wt}_p(d) \geq k$. Then for every $\Lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{F}^k$ where not all λ_i's are zero, the (D, Λ)-UHD system has no pairwise-distinct solutions over $\mathbb{K}$ for any field $\mathbb{K}$ extending $\mathbb{F}$.*

3.2 Proof of Main Theorem

We are now ready to prove the main theorem assuming the lemmas claimed in the previous subsections.

Proof (Theorem 2). We know that it suffices to prove the theorem for the univariate case (i.e., $m = 1$). Let $D(\mathcal{C})$ be the degree set of $\mathcal{C}$. By Lemma 1, we know that $D(\mathcal{C})$ is p-shadow closed. Furthermore if $\mathcal{C}$ has a k-local constraint then, by Lemma 2 there exists a non-zero vector $\Lambda \in \mathbb{F}^{k'}$ such that the $(D(\mathcal{C}), \Lambda)$-UHD system has a pairwise-distinct solution. But then, by Theorem 3, we have that the weight of every element $d \in D(\mathcal{C})$ must be at most $k-2$.

The dimension of $\mathcal{C}$, which is at most $|D(\mathcal{C})|$, can now be bounded from above by the number of integers $d \in \{0, \ldots, q^n - 1\}$ of p-weight less than $k-1$ which is (crudely) at most $(rn)^{k-2}$ (where $q = p^r$). It follows that $\mathcal{C}$ is isomorphic to a subcode of $\mathrm{RM}[q, nm, (k-2)q/p]$, thus concluding the proof of the theorem.

3.3 Proof of Theorem 3

We now prove our main technical theorem, Theorem 3. The proof below is a simplification of our original proof, due to Shachar Lovett. We start by introducing notation that will help us in the proof. The *p-weight* of a set of integers D, denoted $\mathrm{wt}_p(D)$, is the maximal p-weight of an element in D. For the purposes of this proof, the *length* of the UHD system defined by $(D, \Lambda = (\lambda_1, \ldots, \lambda_k)$ is k. We assume throughout the proof that $\Lambda \neq 0$. We say that $\alpha = (\alpha_1, \ldots, \alpha_k)$ is a *solution* of the UHD system if $\alpha_i \neq \alpha_j$ for all $i \neq j$ and α is a root of the UHD system. Using this terminology, Theorem 3 says

> If the $(D, (\lambda_1, \ldots, \lambda_k))$-UHD system has a solution then $k > \mathrm{wt}_p(D) + 1$.

Notice that it suffices to prove the theorem for the case where $D = \mathrm{shadow}_p(d)$ for some integer d. (Else we can simply take the element d of largest weight in D and work with the set $D' = \mathrm{shadow}_p(d)$.)

We prove this by induction on $\mathrm{wt}_p(d)$. The base case of $\mathrm{wt}_p(d) = 0$ says that if the $(\mathrm{shadow}_p(d), \Lambda)$-UHD system is of p-weight 0 and length 1 then it has no solution. The proof in this case is immediate because $\mathrm{shadow}_p(d) = \{0\}$, so if $\Lambda = (\lambda_1)$ is nonzero then there is no solution to the system $\lambda X^0 = 0$ (recall $0^0 = 1$).

For the inductive case we have $\mathrm{wt}_p(d) > 0$. Let $[d]_p = \langle d_0, d_1, \ldots \rangle$ be the base-p representation of d and suppose $d_j > 0$. Assume by way of contradiction that $\alpha = (\alpha_1, \ldots, \alpha_k)$ is a solution to the $(\mathrm{shadow}_p(d), \Lambda)$-UHD system of weight k. The base case of the induction shows $k > 1$ because α is also a solution to the $(\{0\}, \Lambda)$-UHD system, so we may assume without loss of generality that $\alpha_k \neq 0$ because all α_i are distinct. Our proof goes by showing that $\alpha' = (\alpha_1, \ldots, \alpha_{k-1})$ is a solution of a UHD-system of p-weight $w = \mathrm{wt}_p(d) - 1$. By the inductive hypothesis we have $k - 1 > w$ which implies $k > \mathrm{wt}_p(d) + 1$ and completes the proof.

To construct the said UHD system set $e = d - p^j$, noticing $\mathrm{wt}_p(e) = \mathrm{wt}_p(d) - 1$ and let $E = \mathrm{shadow}_p(e)$. Construct $\Lambda' = (\lambda'_1, \ldots, \lambda'_{k-1})$ by setting

$$\lambda'_i = \lambda_i(\alpha_i^{p^j} - \alpha_k^{p^j}).$$

Notice $\lambda'_i \neq 0$ because $\lambda_i \neq 0$ and $\alpha_i \neq \alpha_k$ and the transformation $\alpha \mapsto \alpha^{p^j}$ is a bijection on $\mathbb{K}$. We shall now show that $(\alpha_1, \ldots, \alpha_{k-1})$ is a solution of length $k-1$ to the (E, Λ')-UHD system of p-weight $\mathrm{wt}_p(d) - 1$ thereby completing the proof of the theorem. To show that $(\alpha_1, \ldots, \alpha_{k-1})$ is a solution we argue that for all $r \in \mathrm{shadow}_p(e)$ we have $\sum_{i=1}^{k-1} \lambda'_i \alpha_i^r = 0$. Indeed

$$\begin{aligned}
\sum_{i=1}^{k-1} \lambda'_i \alpha_i^r &= \sum_{i=1}^{k-1} \lambda_i(\alpha_i^{p^j} - \alpha_k^{p^j})\alpha_i^r \\
&= \sum_{i=1}^{k} \lambda_i(\alpha_i^{p^j} - \alpha_k^{p^j})\alpha_i^r \\
&= \sum_{i=1}^{k} \lambda_i \alpha_i^{r+p^j} - \alpha_k^{p^j} \sum_{i=1}^{k} \lambda_i \alpha_i^r = 0
\end{aligned}$$

The last equality follows because $r + p^j \in \mathrm{shadow}_p(d)$ for all $r \in \mathrm{shadow}_p(e)$. This completes the proof of the theorem.

Acknowledgements

We thank the anonymous referees for helpful comments. We thank Shachar Lovett for allowing us to include the simplified proof of Theorem 3 in this paper.

References

1. N. Alon, E. Fischer, I. Newman, and A. Shapira. A combinatorial characterization of the testable graph properties: it's all about regularity. In *STOC 2006*, pages 251–260.
2. N. Alon, T. Kaufman, M. Krivelevich, S. Litsyn, and D. Ron. Testing Reed-Muller codes. *IEEE Transactions on Information Theory*, 51(11):4032–4039, 2005.
3. S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, May 1998.
4. S. Arora and S. Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM*, 45(1):70–122, Jan. 1998.
5. L. Babai, L. Fortnow, L. A. Levin, and M. Szegedy. Checking computations in polylogarithmic time. In *STOC 1991*, pages 21–32.
6. E. Ben-Sasson, V. Guruswami, T. Kaufman, M. Sudan, and M. Viderman. Locally testable codes require redundant testers. In *CCC 2009*, pages 52–61.
7. E. Ben-Sasson, P. Harsha, and S. Raskhodnikova. Some 3CNF properties are hard to test. *SIAM J. Comput*, 35(1):1–21, 2005.

8. E. Ben-Sasson, G. Maatouk, A. Shpilka, and M. Sudan. Symmetric LDPC codes are not necessarily locally testable. In *CCC*, 2011.
9. E. Ben-Sasson and M. Sudan. Simple PCPs with poly-log rate and query complexity. In *STOC 2005*, pages 266–275.
10. A. Bhattacharyya, S. Kopparty, G. Schoenebeck, M. Sudan, and D. Zuckerman. Optimal testing of Reed-Muller codes. In *FOCS 2010*, pages 488–497.
11. M. Blum, M. Luby, and R. Rubinfeld. Self-testing/correcting with applications to numerical problems. In *STOC 1982*, pages 73–83.
12. C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, B. Szegedy, and K. Vesztergombi. Graph limits and parameter testing. In *STOC 2006*, pages 261–270.
13. B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan. Private information retrieval. *Journal of the ACM*, 45:965–981, 1998.
14. I. Dinur. The PCP theorem by gap amplification. *Journal of the ACM*, 54(3):12:1–12:44, June 2007.
15. K. Efremenko. 3-query locally decodable codes of subexponential length. *STOC 2009*, pages 39–44.
16. O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, 1998.
17. O. Goldreich and M. Sudan. Locally testable codes and PCPs of almost-linear length. *J. ACM*, 53(4):558–655, 2006.
18. E. Grigorescu, T. Kaufman, and M. Sudan. 2-transitivity is insufficient for local testability. In *CCC 2008*, pages 259–267.
19. C. S. Jutla, A. C. Patthak, A. Rudra, and D. Zuckerman. Testing low-degree polynomials over prime fields. In *FOCS 2004*, pages 423–432.
20. T. Kasami, S. Lin, and W. Peterson. Some results on cyclic codes which are invariant under the affine group and their applications. *Information and Control*, 11(5-6):475 – 496, 1967.
21. J. Katz and L. Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *STOC 2000*, pages 80–86.
22. T. Kaufman and D. Ron. Testing polynomials over general fields. In *FOCS 2004*, pages 413–422.
23. T. Kaufman and M. Sudan. Algebraic property testing: the role of invariance. In *STOC 2008*, pages 403–412.
24. A. Polishchuk and D. A. Spielman. Nearly-linear size holographic proofs. In *STOC 1994*, pages 194–203.
25. R. Rubinfeld and M. Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM J. Comput.*, 25(2):252–271, 1996.
26. A. Samorodnitsky. Low-degree tests at large distances. In *STOC 2007*, pages 506–515.
27. M. Sudan. *Efficient checking of polynomials and proofs and the hardness of approximation problems*. PhD thesis, UC Berkeley, 1992.
28. D. Woodruff. New lower bounds for general locally decodable codes. *Electronic Colloquium on Computational Complexity (ECCC)*, (006), 2007.
29. S. Yekhanin. Towards 3-query locally decodable codes of subexponential length. *Journal of the ACM*, 55:1–16, 2008.
30. S. Yekhanin. Locally decodable codes. *Foundations and Trends in Theoretical Computer Science*, 2011.