2013 IEEE 54th Annual Symposium on Foundations of Computer Science

Estimating the distance from testable affine-invariant properties

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

Shachar Lovett
Computer Science and Engineering Department
University of California, San Diego
San Diego, USA
slovett@cs.ucsd.edu

***Abstract*—Let P be an affine invariant property of multivariate functions over a constant size finite field. We show that if P is locally testable with a constant number of queries, then one can estimate the distance of a function f from P with a constant number of queries. This was previously unknown even for simple properties such as cubic polynomials over the binary field.**

Our test is simple: take a restriction of f to a constant dimensional affine subspace, and measure its distance from P. We show that by choosing the dimension large enough, this approximates with high probability the global distance of f from P. The analysis combines the approach of Fischer and Newman [SIAM J. Comp 2007] who established a similar result for graph properties, with recently developed tools in higher order Fourier analysis, in particular those developed in Bhattacharyya et al. [STOC 2013].

***Keywords*-property testing; higher-order fourier analysis; affine invariant properties;**

I. INTRODUCTION

Blum, Luby, and Rubinfeld [1] observed that given a function $f : \mathbb{F}_p^n \to \mathbb{F}_p$, 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 [2] 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.

The focus of our work is on testing properties of multivariate functions over finite fields. Fix a prime $p \geqslant 2$ and an integer $R \geqslant 2$ throughout. Let $\mathbb{F} = \mathbb{F}_p$ be a prime field and $[R] = \{0, \dots, R-1\}$. We consider properties of functions $f : \mathbb{F}^n \to [R]$. We are interested in testing the *distance* of a function $f : \mathbb{F}^n \to [R]$ to a property. Here the distance corresponds to the minimum fraction of the points on which the function can be modified in order to satisfy the property. Fischer and Newman [3] showed that it is possible to estimate the distance from a graph to any given *testable* graph property. In this article we extend this result to the algebraic setting of affine-invariant properties on functions $f : \mathbb{F}^n \to [R]$. Furthermore we show that the Fischer-Newman test can be replaced by a more natural one: pick a sufficiently large subgraph H randomly and estimate the distance of H to the property. Analogously, in our setting, we pick a sufficiently large affine subspace of $\mathbb{F}^n$ randomly, and measure the distance of the restriction of the function to this subspace from the property.

A. Testability

Given a property $\mathcal{P}$ of functions in $\{\mathbb{F}^n \to [R] \mid n \in \mathbb{N}\}$, we say that $f : \mathbb{F}^n \to [R]$ is *ε-far* from $\mathcal{P}$ if

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

and we say that it is *ε-close* otherwise. We assume throughout the paper that the field $\mathbb{F}$ and range R are fixed, and n is going to infinity. In particular, any quantities (for example $q(\varepsilon)$ defined below) may implicitly depend on $\mathbb{F}, R$.

Definition I.1 (Testability). *A property $\mathcal{P}$ is said to be* testable *(with two-sided error) if there is a function $q : (0,1) \to \mathbb{N}$ and an algorithm T that, given as input a parameter $\varepsilon > 0$ and oracle access to a function $f : \mathbb{F}^n \to [R]$, makes at most $q(\varepsilon)$ queries to the oracle for f, accepts with probability at least $2/3$ if $f \in \mathcal{P}$ and rejects with probability at least $2/3$ if f is ε-far from $\mathcal{P}$.*

Note that if we do not require any restrictions on $\mathcal{P}$, then the algebraic structure of $\mathbb{F}^n$ becomes irrelevant, and $\mathbb{F}^n$ would be treated as a generic set of size $|\mathbb{F}|^n$. To take the algebraic structure into account, we have to require certain "invariance" conditions.

We say that a property $\mathcal{P} \subseteq \{\mathbb{F}^n \to [R] \mid n \in \mathbb{N}\}$ is *affine-invariant* if for any $f \in \mathcal{P}$ and any affine

0272-5428/13 $26.00 © 2013 IEEE
DOI 10.1109/FOCS.2013.33

transformation $A : \mathbb{F}^n \to \mathbb{F}^n$, we have $Af := f \circ A \in \mathcal{P}$ (an affine transformation A is of the form $L + c$ where L is linear and c is a constant vector in $\mathbb{F}^n$). Some well-studied examples of affine-invariant properties include Reed-Muller codes (in other words, bounded degree polynomials) [4], [5], [6], [7], [8] and Fourier sparsity [9]. In fact, affine invariance seems to be a common feature of most interesting properties that one would classify as "algebraic". Kaufman and Sudan in [10] made explicit note of this phenomenon and initiated a general study of the testability of affine-invariant properties (see also [11]). In particular, they asked for necessary and sufficient conditions for the testability of affine-invariant properties. This question initiated an active line of research, which have led to a near complete characterization of testable affine invariant properties over constant-sized fields, at least in the regime of one-sided error [12], [13], [14], [15], [16], [17].

It is not difficult to see that for affine-invariant properties testability has an equivalent "non-algorithmic" definition through the distribution of restrictions to affine subspaces. We will describe a restriction of $\mathbb{F}^n$ to an affine subspace of dimension k by an affine embedding $A : \mathbb{F}^k \to \mathbb{F}^n$ (an affine embedding is an injective affine transformation). The restriction of $f : \mathbb{F}^n \to [R]$ to the subspace is then given by $Af : \mathbb{F}^k \to [R]$.

Proposition I.2. *An affine-invariant property $\mathcal{P}$ is testable if and only if for every $\varepsilon > 0$, there exist a constant k and a set $\mathcal{H} \subseteq \{\mathbb{F}^k \to [R]\}$, such that for a function $f : \mathbb{F}^n \to [R]$ and a random affine embedding $A : \mathbb{F}^k \to \mathbb{F}^n$ the following holds. If $f \in \mathcal{P}$, then*

$$\mathbf{Pr}[Af \in \mathcal{H}] > 2/3,$$

and if f is ε-far from $\mathcal{P}$, then

$$\mathbf{Pr}[Af \notin \mathcal{H}] > 2/3.$$

B. Our contribution

For a property $\mathcal{P}$ and a positive real δ, let $\mathcal{P}_\delta$ denote the set of all functions that are δ-close to the property. Our main result is the following theorem.

Theorem I.3. *For every testable affine-invariant property $\mathcal{P}$ and every $\delta > 0$, the property $\mathcal{P}_\delta$ is testable.*

Theorem I.3 says that for every $\varepsilon, \delta > 0$ one can probabilistically distinguish between functions that are δ-close to the property and the functions that are $(\delta+\varepsilon)$-far from the property using only a constant number of queries (the constant is allowed to depend on the property $\mathcal{P}$ and on ε, δ). In fact the test is very natural. We show that there exists a constant $k_{\varepsilon,\delta,\mathcal{P}}$ such that for a random affine embedding $A : \mathbb{F}^k \to \mathbb{F}^n$, with probability at least $2/3$, $\mathrm{dist}(Af, \mathcal{P})$ provides a sufficiently accurate estimate of $\mathrm{dist}(f, \mathcal{P})$. Hence our test will be the following: Pick a random affine embedding $A : \mathbb{F}^k \to \mathbb{F}^n$. If $\mathrm{dist}(Af, \mathcal{P}) < \delta + \frac{\varepsilon}{2}$ accept, otherwise reject. This corresponds to taking $\mathcal{H} = \left\{h : \mathbb{F}^k \to [R] \mid \mathrm{dist}(h, \mathcal{P}) \leqslant \delta + \frac{\varepsilon}{2}\right\}$ in Proposition I.2.

We note that previously it was unknown if one can test distance to even simple properties, such as cubic polynomials over $\mathbb{F}_2$. The reason was that one specific natural test (the Gowers norm, or derivatives test) was shown not to perform well for such properties. Our work shows that a natural test indeed works, albeit the number of queries have to grow as a function of ε (a stronger possibility is to have a constant number of queries, and acceptance probability which depends on ε). We do not know if this is necessary for simple properties, such as cubic polynomials over $\mathbb{F}_2$, and leave this as an open problem.

On a technical level, our work combines two technologies developed in previous works. The first is the work of Fischer and Newman [3] which obtained similar results for graph properties. The second is higher order Fourier analysis, in particular a recent strong equidistribution theorem established in Bhattacharyya et al. [17]. From a high level, the approach for the graph case and the affine-invariant case are similar. One applies a regularization process, which allows to represent a graph (or a function) by a small structure. Then, one argues that a large enough random sample of the graph or function should have a similar small structure representing it. Hence, properties of the main object can be approximated by properties of a large enough sample of it. Fischer and Newman [3] implemented this idea in the graph case. We follow a similar approach in the algebraic case, which inevitably introduces some new challenges. One may see this result as an outcome of the large body of work on higher-order Fourier analysis developed in recent years. Once the machinery was developed, we can now apply it in various frameworks which were not accessible previously.

II. Proof overview

For lack of space, we only give a high level proof overview in this extended abstract. The full paper is available online [18].

Let $R = 2$ for the simplicity of exposition, e.g. we consider functions $f : \mathbb{F}^n \to \{0, 1\}$. Let $\mathcal{P}$ be an affine invariant property of functions $\{\mathbb{F}^n \to \{0, 1\} : n \in \mathbb{N}\}$ which is locally testable, and fix $\varepsilon, \delta > 0$. We want to show that there exists an m (which depends only on

$\mathcal{P}, \varepsilon, \delta$) such that the following holds. Let $f : \mathbb{F}^n \to \{0,1\}$ be a function, and let $\tilde{f}$ be the restriction of the function to a random m-dimensional affine subspace of $\mathbb{F}^n$. Then

- Completeness: If f is δ-close to $\mathcal{P}$ then, with high probability, $\tilde{f}$ is $(\delta + \varepsilon/2)$-close to $\mathcal{P}$.
- Soundness: If f is $(\delta + \varepsilon)$-far from $\mathcal{P}$ then, with high probability, $\tilde{f}$ is $(\delta + \varepsilon/2)$-far from $\mathcal{P}$.

Once we show that we are done, as the local test computes the distance of $\tilde{f}$ from $\mathcal{P}$. If it is below $\delta+\varepsilon/2$ we declare that f is δ-close to $\mathcal{P}$; otherwise we declare it is $(\delta + \varepsilon)$-far from $\mathcal{P}$. The test correctness follows immediately from the completeness and soundness. We next argue why these hold.

Let us first fix notations. Let $A : \mathbb{F}^m \to \mathbb{F}^n$ be a random full rank affine transformation. Then, a restriction of f to a random m-dimensional affine subspace can be equivalently described by $\tilde{f} = Af$. The proof of the completeness is simple. If f is δ-close to a function $g : \mathbb{F}^n \to \{0,1\}$ which is in $\mathcal{P}$, then with high probability over a random restriction, the distance of Af and Ag is also at most $\delta + o_m(1)$. This is true because a random affine subspace is pairwise independent with regards to whether an element is contained in it. This, combined with Chebyshev's inequality implies the result. Then, by choosing m large enough we get the error term down to $\varepsilon/2$.

The main work (as in nearly all works in property testing) is to establish soundness. That is, we wish to show that if a function f is far from $\mathcal{P}$ then, with high probability, a random restriction of it is also from from the property. The main idea is to show that if for a typical restriction Af is δ-close to a function $h : \mathbb{F}^m \to \{0,1\}$ which is in $\mathcal{P}$, then h can be "pulled back" to a function $g : \mathbb{F}^n \to \{0,1\}$ which is both roughly δ-close to f and also very close to $\mathcal{P}$. This will contradict our initial assumption that f is $(\delta + \varepsilon)$-far from $\mathcal{P}$. In order to do so we apply the machinery of higher order Fourier analysis. The first description will hide various "cheats" but will present the correct general outline. We then note which steps need to be fixed to make this argument actually work.

First, we apply the assumption that $\mathcal{P}$ is locally testable to derive there exist a constant dimension $k = k(\mathcal{P}, \varepsilon)$ so that a random restriction to a k-dimensional subspace can distinguish functions in $\mathcal{P}$ from functions which are $\varepsilon/4$-far from $\mathcal{P}$. We want to decompose f to "structured" parts which we will study, and "pseudo-random" parts which do not affect the distribution of restrictions to k-dimensional subspaces. In order to do so, for a function $f : \mathbb{F}^n \to \{0,1\}$ define by $\mu_{f,k}$ the distribution of its restriction to k-dimensional subspaces. That is, for $v : \mathbb{F}^k \to \{0,1\}$ let

$$\mu_{f,k}[v] = \Pr_A[Af = v].$$

We need to slightly generalize this definition to functions where the output $f(x)$ can be random. In our context, a randomized function is a function $f : \mathbb{F}^n \to [0,1]$, which describes a distribution over functions $f' : \mathbb{F}^n \to \{0,1\}$, where for all x independently $\mathbf{Pr}[f'(x) = 1] = f(x)$. We extend the definition of $\mu_{f,k}$ to randomized functions by $\mu_{f,k}[v] = \mathbf{E}_{A,f'}\, \mu_{f',k}[v]$. By our definition, if two functions $f, g : \mathbb{F}^n \to [0,1]$ have distributions $\mu_{f,k}$ and $\mu_{g,k}$ close in statistical distance, then random restrictions to k-dimensional affine subspaces cannot distinguish f from g. This will be useful in the analysis of the soundness.

We next decompose our function f based on the above intuition. The formal notion of pseud-randomness we use is that of Gowers uniformity. Informally, the d-th Gowers uniformity norm (denoted $\|\cdot\|_{U^d}$) measures correlation with polynomials of degree less than d. However, it turns to capture much more than that. For example, one can show that by choosing d large enough ($d = p^k$ suffices) then for any functions $f, g : \mathbb{F}^n \to [0,1]$, if $\|f - g\|_{U^d}$ is small enough then $\mu_{f,k}$ and $\mu_{g,k}$ are close in statistical distance. Thus, it makes sense to approximate f as

$$f = f_1 + f_2$$

where f_1 is structured (to be explained soon) and $\|f_2\|_{U^d}$ is small enough. This will allow us to replace f with f_1 for the purposes of analyzing its restrictions to k-dimensional subspaces. The structure of f_1 is as follows: it is a function of a constant number $C = C(\mathcal{P}, \varepsilon)$ of polynomials of degree less than d. That is,

$$f_1(x) = \Gamma(P_1(x), \ldots, P_C(x)),$$

where $P_1, \ldots, P_C$ are polynomials and $\Gamma : \mathbb{F}^C \to \{0,1\}$ is some function (not necessarily a low degree polynomial). The benefit of this decomposition is that f_1 is "dimension-less" in the sense that Γ does not depend on n; however, the polynomials $P_1, \ldots, P_C$ do depend on n. One can however "regularize" these polynomials in order to obtain "random-looking" (or high rank) polynomials. It can be shown that all properties of high rank polynomials are governed just by their degree (which is at most d), hence essentially the entire description of f_1 does not depend on n.

The next step is to show that the same type of decomposition can be applied to the restriction Af of

f. Clearly, $Af = Af_1 + Af_2$. We show that with high probability over the choice of A,

- $Af_1 = \Gamma(Q_1(x), \ldots, Q_C(x))$ where $Q_i = AP_i$ are the restrictions of $P_1, \ldots, P_C$; and $Q_1, \ldots, Q_C$ are still of "high enough rank" to behave like random polynomials.
- $\|Af_2\|_{U^d} \approx \|f_2\|_{U^d}$ so we can still approximate $Af \approx Af_1$ with respect to the distribution of their restrictions to random k-dimensional subspaces.

We next apply the same decomposition process to h, which we recall is the assumed function (in m variables) which is $(\delta + \varepsilon/2)$-close to Af. By choosing the conditions of regularity of h slightly weaker than those of f (but still strong enough), we get that we can decompose

$$h = h_1 + h_2$$

where

$$h_1(x) = \Gamma'(Q_1(x), \ldots, Q_{C'}(x))$$

for some $C' > C$ and $\|h_2\|_{U^d}$ is very small. The important aspect here is that, we can approximate h by the structured function h_1, and moreover that the polynomials $Q_1, \ldots, Q_C$ which compose Af_1 are part of the description of h_1. That is, both Af_1 and h_1 can be defined in terms of the same basic building blocks (high rank polynomials $Q_1, \ldots, Q_C$).

The next step is to "pull back" h to a function defined on $\mathbb{F}^n$. An easy first step is to pull back h_1. We need to define for $C < i \leqslant C'$ pullback polynomials $P_i : \mathbb{F}^n \to \{0, 1\}$ of $Q_i : \mathbb{F}^n \to \{0, 1\}$ such that both $Q_i = AP_i$; and such that $P_1, \ldots, P_{C'}$ are of high rank. This can be done for example by letting $P_i = DQ_i$ for any affine map $D : \mathbb{F}^n \to \mathbb{F}^m$ for which AD is the identity map on $\mathbb{F}^m$. This provides a pull-back ϕ of the "coarse" description of f_1 of h_1, but does not in general generate a function close to f (it makes sense, since we still haven't used the finer "pseudo-random" structure of f). Formally, we set $\phi(x) = \Gamma'(P_1(x), \ldots, P_{C'}(x))$. However, we can already show something about ϕ: it is very close to $\mathcal{P}$. More concretely, its distribution over restrictions to d-dimensional subspaces is very close to that of h. Hence, the tester which distinguishes function in $\mathcal{P}$ from those $(\varepsilon/4)$-far from $\mathcal{P}$ cannot distinguish ϕ from functions in $\mathcal{P}$, hence ϕ must be $(\varepsilon/4)$-close to $\mathcal{P}$.

The next step is to define a more refined pull-back of f. Define an atom as a subset $\{x \in \mathbb{F}^n : P_1(x) = a_1, \ldots, P_{C'}(x) = a_{C'}\}$ for values $a_1, \ldots, a_{C'} \in \mathbb{F}$. Note that the functions f_1, h_1 are constant over atoms. We next define $\psi : \mathbb{F}^n \to [0, 1]$ by redefining ϕ inside each atom, so that the average over the atoms of ϕ, ψ is the same, but such that ψ is as close as possible to f given this constraint. For example, if in an atom the average of f is higher than the value ϕ assigns to this atom (and so it needs to be reduced to match ϕ), we set for all x in this atom $\psi(x) = 0$ if $\phi(x) = 0$ and $\psi(x) = \alpha$ if $f(x) = 1$, where α is appropriately chosen so that the averages match. We then show that ψ is a proper pull-back of h in the sense that

- The distance between f, ψ is very close to the distance between Af, h, which we recall is at most $\delta + \varepsilon/2$.
- ψ is nearly $\varepsilon + 4$ close to $\mathcal{P}$ in the distributional sense.

To finalize, we show that sampling a function $g : \mathbb{F}^n \to \{0, 1\}$ based on ψ has the same properties, which shows that f is not $(\delta + \varepsilon)$-far from $\mathcal{P}$.

Let us remark on a few technical points overlooked in the above description. First, there are the exact notions of "high rank polynomials". It turns that in order to make this entire argument work, one needs to consider more general objects, called non-classical polynomials. We rely on a series of results on the distributional properties of high-rank non-classical polynomials, in particular these recently established in [17]. Also, the decomposition theorems are actually to three parts,

$$f = f_1 + f_2 + f_3,$$

where f_1 is structured as before, $\|f_2\|_2$ is somewhat small (but not very small) and $\|f_3\|_{U^d}$ is very small. This requires a somewhat more refined analysis to make the argument work, but does not create any significant change in the proof outline as described above.

III. Comparison with graph property testing

The main outline of our proof follows closely that of Fischer and Newman [3]. They study graph properties, where decompositions are given by the Szemerédi regularity lemma. Their test, in the notation above, can be described as measuring the distance between Γ and all potential Γ' which can be achieved from graphs that have the property. Our argument (when applied to graph properties instead of affine invariant properties) shows that a much more natural test achieves the same behaviour: choose a random small subgraph and measure its distance from the property. In quantitative terms it is hard to compare the two results, as both get outrageous bounds coming from the bounds in the regularity lemma. So, we view this part of our work as having contribution in the simplicity of the test, and not in terms of the simplicity of the proof or the quantitative bounds (which are both very similar).

The more challenging aspect of our work is to take this approach and carry it out in the affine invariant settings. The main reason is that in the affine invariant setup the structural parts have more structure in them than in the graph setting. In the graph setup, the structure of a graph can be represented by a constant size graph with weighted edges. In the affine invariant case, the structured part is a constant size function applied to polynomials. However there will be *no constant bound* on the number of variables, and they can grow as n grows. So, at first glance, these "compact descriptions" have sizes which grow with the input size; this is very different from the graph case. The reason these compact descriptions are useful is because, as long as the polynomials participating in them are "random enough", then their exact definitions do not matter, just a few simple properties of them (their degree, and "depth" for non-classical polynomials). This is fueled by the recent advances on higher-order Fourier analysis. In essence, the state of the art has reached a stage where these tools are powerful enough to simulate the counterpart arguments which were initially developed in the context of graph properties.

IV. Open problems

As we mentioned, the ultimate goal of this line of work is to achieve a complete understanding of algebraic property testing, analogously to the complete understanding we have for (dense) graph property testing, at least in qualitative terms. Still, the affine invariant case is more complex, and there are some problems which we do not know yet how to handle. For example,

- A complete classification of one-sided testable properties (e.g. can properties of "infinite complexity" be locally testable?); See [17]
- A complete classification of two-sided testable properties.
- Properties where any non-trivial distance from them can be witnessed by a constant number of queries (also called correlation testing [19]). For example, can one test correlation to cubics over $\mathbb{F}_2$ using a constant number of queries?

Acknowledgment

The authors thank the anonymous reviewers for helpful comments and suggestions.

References

[1] M. Blum, M. Luby, and R. Rubinfeld, "Self-testing/correcting with applications to numerical problems," *J. Comp. Sys. Sci.*, vol. 47, pp. 549–595, 1993, earlier version in STOC'90. 1

[2] R. Rubinfeld and M. Sudan, "Robust characterizations of polynomials and their applications to program testing," Ithaca, NY, USA, Tech. Rep., 1993. 1

[3] E. Fischer and I. Newman, "Testing versus estimation of graph properties," *SIAM J. Comput.*, vol. 37, no. 2, pp. 482–501 (electronic), 2007. 1, 2, 4

[4] L. Babai, L. Fortnow, and C. Lund, "Non-deterministic exponential time has two-prover interactive protocols," *Computational Complexity*, vol. 1, no. 1, pp. 3–40, 1991. 2

[5] L. Babai, L. Fortnow, L. A. Levin, and M. Szegedy, "Checking computations in polylogarithmic time," in *Proc. 23rd Annual ACM Symposium on the Theory of Computing*. New York: ACM Press, 1991, pp. 21–32. 2

[6] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy, "Interactive proofs and the hardness of approximating cliques," *Journal of the ACM*, vol. 43, no. 2, pp. 268–292, 1996. 2

[7] R. Rubinfeld and M. Sudan, "Robust characterizations of polynomials with applications to program testing," *SIAM J. on Comput.*, vol. 25, pp. 252–271, 1996. 2

[8] N. Alon, T. Kaufman, M. Krivelevich, S. Litsyn, and D. Ron, "Testing Reed-Muller codes," *IEEE Transactions on Information Theory*, vol. 51, no. 11, pp. 4032–4039, 2005. 2

[9] P. Gopalan, R. O'Donnell, R. A. Servedio, A. Shpilka, and K. Wimmer, "Testing Fourier dimensionality and sparsity," in *Proceedings of the 36th International Colloquium on Automata, Languages and Programming (ICALP '09)*, 2009, pp. 500–512. 2

[10] T. Kaufman and M. Sudan, "Algebraic property testing: the role of invariance," in *Proc. 40th Annual ACM Symposium on the Theory of Computing*, 2008, pp. 403–412. 2

[11] O. Goldreich and T. Kaufman, "Proximity oblivious testing and the role of invariances," in *Approximation, randomization, and combinatorial optimization*, ser. Lecture Notes in Comput. Sci. Heidelberg: Springer, 2011, vol. 6845, pp. 579–592. 2

[12] A. Bhattacharyya, V. Chen, M. Sudan, and N. Xie, "Testing linear-invariant non-linear properties," *Theory of Computing*, vol. 7, no. 1, pp. 75–99, 2011. 2

[13] D. Král, O. Serra, and L. Vena, "A removal lemma for systems of linear equations over finite fields," *Israel Journal of Mathematics*, pp. 1–15, 2012. 2

[14] A. Shapira, "Green's conjecture and testing linear-invariant properties," in *Proc. 41st Annual ACM Symposium on the Theory of Computing*, 2009, pp. 159–166. 2

[15] A. Bhattacharyya, E. Grigorescu, and A. Shapira, "A unified framework for testing linear-invariant properties," in *Proc. 51st Annual IEEE Symposium on Foundations of Computer Science*, 2010, pp. 478–487. 2

[16] A. Bhattacharyya, E. Fischer, and S. Lovett, "Testing low complexity affine-invariant properties," in *Proc. 24th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2013, pp. 1337-1355. 2

[17] A. Bhattacharyya, E. Fische, H. Hatami, , P. Hatami, and S. Lovett, "Every locally characterized afine-invariant property is testable," in *Proc. 45th Annual ACM Symposium on the Theory of Computing*, 2013, pp. 429–436. 2, 4, 5

[18] H. Hatami and S. Lovett, "Estimating the distance from testable affine-invariant properties," in *Electronic Colloquium in Computational Complexity*, vol. 20, no. 87, 2013. 2

[19] ——, "Correlation testing for affine invariant properties on $\mathbb{F}_p^n$ in the high error regime," in *Proc. 43rd Annual ACM Symposium on the Theory of Computing*, 2011, pp. 187–194. 5