Optimal testing of multivariate polynomials over small prime fields

Elad Haramaty
Faculty of Computer Science
Technion — Israel Institute of Technology
Haifa, Israel
Email: eladh@cs.technion.ac.il

Amir Shpilka
Faculty of Computer Science
Technion — Israel Institute of Technology
Haifa, Israel
Email: shpilka@cs.technion.ac.il

Madhu Sudan
Microsoft Research
Cambridge, MA, USA
Email: madhu@mit.edu

Abstract— We consider the problem of testing if a given function $f : \mathbb{F}_q^n \to \mathbb{F}_q$ is close to a n-variate degree d polynomial over the finite field $\mathbb{F}_q$ of q elements. The natural, low-query, test for this property would be to pick the smallest dimension $t = t_{q,d} \approx d/q$ such that every function of degree greater than d reveals this aspect on *some* t-dimensional affine subspace of $\mathbb{F}_q^n$ and to test that f when restricted to a *random* t-dimensional affine subspace is a polynomial of degree at most d on this subspace. Such a test makes only q^t queries, independent of n.

Previous works, by Alon et al. [1], and Kaufman and Ron [7] and Jutla et al. [6], showed that this natural test rejected functions that were $\Omega(1)$-far from degree d-polynomials with probability at least $\Omega(q^{-t})$. (The initial work [1] considered only the case of $q = 2$, while the work [6] only considered the case of prime q. The results in [7] hold for all fields.) Thus to get a constant probability of detecting functions that are at constant distance from the space of degree d polynomials, the tests made q^{2t} queries. Kaufman and Ron also noted that when q is prime, then q^t queries are necessary. Thus these tests were off by at least a quadratic factor from known lower bounds. Bhattacharyya et al. [2] gave an optimal analysis of this test for the case of the binary field and showed that the natural test actually rejects functions that were $\Omega(1)$-far from degree d-polynomials with probability $\Omega(1)$.

In this work we extend this result for all fields showing that the natural test does indeed reject functions that are $\Omega(1)$-far from degree d polynomials with $\Omega(1)$-probability, where the constants depend only on q the field size. Thus our analysis thus shows that this test is optimal (matches known lower bounds) when q is prime. The main technical ingredient in our work is a tight analysis of the number of "hyperplanes" (affine subspaces of co-dimension 1) on which the restriction of a degree d polynomial has degree less than d. We show that the number of such hyperplanes is at most $O(q^{t_{q,d}})$ — which is tight to within constant factors.

Keywords-Low-degree testing; Property Testing; Reed-Muller codes

1. Introduction

Testing low-degree polynomials is one of the most basic problems in property testing. It is the prototypical problem in "algebraic property testing", and has seen many applications in probabilistic checking of proofs. In this work we focus on this basic problem and give optimal (to within large constant factors) results for the setting of degree d multivariate polynomials over fields of constant size. This setting has been considered before in [1], [7], [6], [2], but their results were off by a "quadratic factor". We remove this gap here, and in the process introduce some algebraic results about restrictions of low-degree polynomials to affine subspaces that may be of independent interest.

To describe our work and the previous work more precisely we start with some basic notation. For integer t, we let $[t]$ denote the set $\{1, \ldots, t\}$. We let $\mathbb{F}_q$ denote the finite field of cardinality q. We consider the task of testing functions mapping $\mathbb{F}_q^n$ to $\mathbb{F}_q$. Let $\mathcal{P}(n, d, q)$ denote the set of all n-variate polynomial functions over $\mathbb{F}_q$ of total degree at most d. We let $\delta(f, g) = \Pr_x[f(x) \neq g(x)]$ denote the distance between f and g, where the probability is over x chosen uniformly at random from $\mathbb{F}_q^n$. Let $\delta_d(f) = \min_{g \in \mathcal{P}(n,d,q)}\{\delta(f, g)\}$ denote the distance of f from the space of degree d polynomials. We say f is δ-far from g if $\delta(f, g) \geq \delta$ and δ-close otherwise. We say f is δ-far from the set of degree d polynomials if $\delta_d(f) \geq \delta$. The goal of low-degree testing is to design a test to distinguish the case where $\delta_d(f)$ is zero from the case where it is large.

A k-query *tester* (for $\mathcal{P}(n, d, q)$) is a probabilistic algorithm $T = T(n, d, q)$ that makes at most $k = k(d, q)$ queries to an oracle for the function $f : \mathbb{F}_q^n \to \mathbb{F}_q$ and accepts $f \in \mathcal{P}(n, d, q)$ with probability one. It has δ-soundness ϵ if it rejects every function f with $\delta_d(f) \geq \delta$ with probability at least ϵ. We say T is *absolutely sound* if for every q and $\delta > 0$ there exists $\epsilon > 0$ such that for every d and n, $T = T(n, d, q)$ has δ-soundness ϵ.

With the above definitions in place, we can now describe previous works. (We note that the testing problem was studied actively for large fields and small degrees starting with [10] and in the PCP literature, but we will not describe such works here.) The setting where the degree of the polynomial is larger than the field size was first studied by Alon et al. [1] who considered the setting of $q = 2$. They described a basic test that made $O(2^d)$ queries.[1] Their analysis showed that this test has δ-soundness $\Omega(\delta 2^{-d})$. Thus to get an absolutely sound test, they iterated this test $O(2^d)$ times, getting a query complexity of $O(4^d)$. They showed no test with $o(2^d)$ queries could test this family, thus giving a bound that was off by a quadratic factor.

The setting of general q was considered by Kaufman and

[1]Throughout this paper we think of q as a constant and so dependence on q may some times be suppressed. Dependence on d is crucial and complexity depending on n will be too large to be interesting.

Ron [7] and independently (for the case of prime q) by Jutla et al. [6]. They (in particular [7]) showed that there exists an integer $t = t_{q,d} \approx d/q$ (we will be more precise with this later) such that the natural test for low-degreeness makes $\Omega(q^t)$ queries. They also show that q^t is a lower bound on the number of queries if q is prime. Finally they analyzed this $O(q^t)$ query test, showing that the δ-soundness of this test is $\Omega(\delta q^{-t})$, again leading to an absolutely sound test with query complexity $O(q^{2t})$ which is off by a quadratic factor. The proof techniques of [1] and [7], [6] were similar and indeed the subsequent generalization of Kaufman and Sudan [8] shows how these results fall in the very general framework of "affine-invariant" property testing, where again all known tests are off by (at least) a quadratic factor.

In a recent work, Bhattacharyya et al. [2] raised the question of getting "optimal tests" for $\mathcal{P}(n, d, q)$. Again they restricted their attention to the case of $q = 2$ and came up with a new proof technique that allowed them to prove that the original $O(2^d)$-query test of [1] is absolutely sound. This also gave the first example of a linear-invariant property with tight bounds on query complexity.

The proof of [2] was significantly more algebraic than those of [1], [7], [6]. (Indeed the work of [8] confirms that the central ingredient in the proofs in [1], [7], [6] are all the same and relies on very little algebra.) However, the proof of [2] seemed very carefully tailored to the case of $\mathbb{F}_2$ and extensions faced several obvious obstacles. In this work we manage to overcome these obstacles and show that the $O(q^t)$ query tester of [7] is also absolutely sound (though as it turns out, the dependence of the constant on q is terrible). En route of proving this we obtain several new results on the behavior of polynomials when restricted to lower dimensional affine spaces, that may be of independent interest. Below we explain our main theorem and some of the algebraic ingredients that we obtain along the way.

1.1. Our main results

To state the test of [1], [7] and our theorem we need a few more definitions. For an affine subspace A in $\mathbb{F}_q^n$, let $\dim(A)$ denote its dimension. For function $f : \mathbb{F}_q^n \to \mathbb{F}_q$ and affine subspace A, let $f|_A : A \to \mathbb{F}_q$ denote the restriction of f to A. For a function f, we let $\deg(f)$ denote its degree as a polynomial. We use the fact that $f|_A$ can be viewed as a $\dim(A)$-variate polynomial with $\deg(f|_A) \leq \deg(f)$. A special subclass of tests for $\mathcal{P}(n, d, q)$ would simply pick an affine subspace A of $\mathbb{F}_q^n$ and verify that $\deg(f|_A) \leq d$. We introduce the concept below of the testing dimension which attempts to explore the minimal dimension for which such a test has positive soundness.

Definition 1.1 (Testing dimension)**.** *For prime power q and non-negative d, the* testing dimension *of polynomials of degree d over $\mathbb{F}_q$ is the smallest integer t satisfying the following: For every positive integer n and every function $f : \mathbb{F}_q^n \to \mathbb{F}_q$ with $\deg(f) > d$, there exists an affine subspace A of dimension at most t such that $\deg(f|_A) > d$. We use $t_{q,d}$ to denote the testing dimension.*

This notion was studied in [7] who proved the following fact.

Proposition 1.2. *The testing dimension $t_{q,d} = \lceil \frac{d+1}{q-q/p} \rceil$.*

The test proposed by [7] is the following:

> t-dimensional (degree d) test:
>
> Given oracle access to $f : \mathbb{F}_q^n \to \mathbb{F}_q$, pick a random affine subspace A with $\dim(A) = t$ and accept if $\deg(f|_A) \leq d$.

[7] shows that the $t_{q,d}$-dimensional test, which has query complexity $q^{t_{q,d}}$ and accepts $f \in \mathcal{P}(n, d, q)$ with probability one, has δ-soundness roughly $\Omega(\delta q^{-t_{q,d}})$. We show that the test is absolutely sound (and in fact instead of losing a $q^{-t_{q,d}}$ factor we even gain it for small δ). Specifically, if we let $\rho_d(f, t)$ denote the probability which the t-dimensional test rejects a function f, then we show:

Theorem 1.3. *For every prime power q, there exist constants $\epsilon_1, \epsilon_2 > 0$ such that for every d and n and every function $f : \mathbb{F}_q^n \to \mathbb{F}_q$, it is the case that $\rho_d(f, t_{d,q}) \geq \min\{\epsilon_1 q^{t_{d,q}} \delta(f), \epsilon_2\}$. In other words the $t_{q,d}$-dimensional test rejects f with probability $\min\{\epsilon_1 q^{t_{q,d}} \delta(f), \epsilon_2\}$, where $t_{q,d}$ is the testing dimension for degree d polynomials over $\mathbb{F}_q$.*

Our analysis follows the approach of [2] who derive their analysis by first studying the behavior of functions that are not degree d polynomials, when restricted to affine subspaces of codimension one. Following their terminology we use the phrase *hyperplane* to refer to subspaces of $\mathbb{F}_q^n$ of codimension one (i.e., dimension $n-1$), and let $\mathrm{H}(q, n)$ denote the set of all hyperplanes in $\mathbb{F}_q^n$. We highlight two key quantities of interest to this approach. The first of these asks how often can a degree d polynomial drop in degree when restricted to hyperplanes. Formally:

Definition 1.4. *For prime power q and non-negative integer d, let $N = N_0(q, d)$ be the maximum over all n, and all functions $f \in \mathcal{P}(n, d, q)$ of the number of hyperplanes $A_1, \ldots, A_N$ such that $\deg(f|_{A_i}) < \deg(f)$. I.e., $N_0(q, d) = \max_{n, f \in \mathcal{P}(n,d,q)} |\{A \in \mathrm{H}(n, q) | \deg(f|_A) < \deg(f)\}|$.*

A priori it may not be clear that $N_0(d, q)$ is even bounded (i.e., is independent of n), but an easy argument from [2] shows this quantity is at most q^d. For our purposes we need a much tighter bound of roughly $q^{t_{q,d}}$ and our first main technical theorem (of two) shows that this is indeed the case.

Theorem 1.5. *For every q, d, $N_0(d, q) \leq q^{t_{q,d}+1}$. In other words, if $f \in \mathcal{P}(n, d, q)$, then $|\{A \in \mathrm{H}(q, n) | \deg(f) < \deg(f)\}| \leq N_0(d, q) \leq q^{t_{q,d}+1}$.*

(We note that it follows from the definition of N_0 and $t_{q,d}$ that $N_0(d,q) \geq q^{t_{q,d}}$.)

The above theorem gives a tight analysis (up to constant factors depending on the field size) of the number of hyperplanes where a degree d polynomial drops in degree. However for the analysis of the low-degree test, we need a similar theorem that talks about general functions. Extracting the correct quantity of interest (one that can be analyzed and is useful) turns out to be somewhat subtle. Rather than looking at general functions, or even functions that are far from polynomials, we look only at the restrictions of functions to hyperplanes and ask "when does pairwise consistency imply global consistency".

Definition 1.6. *For prime power q and non-negative integer d, let $N = N_1(q,d)$ be the largest integer such that the following holds: There exists n, and N hyperplanes $A_1, \ldots, A_N \in \mathrm{H}(n,q)$ and N polynomials $P_1, \ldots, P_N \in \mathcal{P}(n,d,q)$ such that the following hold:*

Pairwise consistency:

For every $i,j \in [N]$ it is the case that $P_i|_{A_i \cap A_j} = P_j|_{A_i \cap A_j}$.

Global inconsistency:

For every $Q \in \mathcal{P}(n,d,q)$, there exists $i \in [N]$ such that $Q|_{A_i} \neq P_i|_{A_i}$.

Note that viewed contrapositively, the definition of N_1 says that if some arbitrary function f looks like a degree d polynomial on $N_1(q,d)+1$ hyperplanes, then its restriction to the union of these hyperplanes (which is typically an overwhelmingly large set) is a polynomial of degree d and hence f is *close* to a polynomial of degree d. Our second main technical theorem shows that N_1 is not much larger (in a technical sense) than $N_0(q,d)$.

Theorem 1.7. *For every q, there exists a constant λ_q such that for every d, $N_1(q,d) \leq q^{t_{q,d}+\lambda_q}$. In other words if $A_1, \ldots, A_K \in \mathrm{H}(n,q)$ and $P_1, \ldots, P_K \in \mathcal{P}(n,d,q)$ are such that $P_i|_{A_i \cap A_j} = P_j|_{A_i \cap A_j}$ for every $i,j \in [K]$ and $K > q^{t_{q,d}+\lambda_q}$, then there exists $Q \in \mathcal{P}(n,d,q)$ such that $Q|_{A_i} = P_i|_{A_i}$ for every $i \in [K]$.*

1.2. Comparison to [2]

While our proof outline does follow the same one as that of [2] the technical elements are much more complex and we point out the similarities and differences here. Both proofs work by induction on the number of variables. Key to this induction is an ability to understand how functions (that are not polynomials and are even far from them) behave on restrictions to hyperplanes. Once such an understanding is obtained, the proofs are immediate *given* the work of [2] — and we simply mimic their proofs. (We note that much of the novelty of [2] is in this part, but given their work their is no novelty in ours in this part.) Their proof roughly shows that for $\tilde{t} = \log_q N_1(q,d)$ the $\tilde{t}$-dimensional test is absolutely sound. To make this useful, one needs two more ingredients: (1) A good upper bound on $N_1(q,d)$ and (2) A (possibly weak) relationship between the soundness of a t-dimensional test and the soundness of the $(t-1)$-dimensional test (so that one can eventually analyze the $t_{q,d}$-dimensional test).

In [2] both of these elements turn out to be simple (once one has the right insights). $N_1(q,d)$ is at most q^d (by a simple linear algebra argument). And a t-dimensional test can be related to a $t-1$ also by similar linear algebra arguments for the case $q=2$. In our case it turns out both ingredients are non-trivial.

For (2) we prove (see Lemmas 4.6 and 4.7) that a $t-1$ dimensional test (as long as $t-1 \geq t_{q,d}$) has δ-soundness at least $1/q$ times the δ-soundness of the t-dimensional test. Even this step (though simple in comparison to the other part) is not immediate and requires a more algebraic view of restrictions than in previous works.

For (1), our task turns out to be much harder. We consider the simpler case of bounding $N_0(d,q)$ first and this ends up using several algebraic features of affine transformations and restrictions to hyperplanes (see Lemmas 4.4 and 4.8). This still leaves the question of bounding $N_1(d,q)$, for which we build an inductive proof, where each inductive step uses the bound on $N_0(d,q)$. The most problematic part however turns out to be the base case, where we need to show that the abundance of hyperplanes leads to a cover of most of $\mathbb{F}_q^n$ by q "near-parallel" hyperplanes. For this part we resort to the "density Hales-Jewett theorem" [4], [9] which says (for our purposes) that for every q and every $\epsilon > 0$ there is a $c = c_{q,\epsilon}$ such that $\epsilon \cdot q^c$ hyperplanes in c dimensions will contain q "near-parallel" ones. (Unfortunately this leads to a horrendous bound on $c_{q,\epsilon}$, but fortunately ϵ is independent of n and d and so this suffices for Theorem 1.3).

2. Overview of our proof

Here we give an overview of our proof and lead the reader through the technical parts of the paper. We start by listing ingredients in order of increasing "complexity" that we prove (each of which we argue is necessary), and then describe how these are put together to get our final analysis. All the novel technical ingredients talk about the behavior of some function f when restricted to hyperplanes.

Step 0: We start by considering an m-variate function f which is not a degree d polynomial, and ask: *Does there exist a single hyperplane on which f is not a degree d polynomial?* Obviously existence of such a hyperplane is a necessary condition for any $t < m$ dimensional test to work. By definition this question has an affirmative answer if $m > t_{q,d}$, the testing dimension. The testing dimension was already analyzed by Kaufman and Ron [7], but we end up reproving this result, since we need stronger versions of this analysis (as we describe next). Proposition 1.2 captures

this step. Its proof relies on Lemma 4.6 which is a central ingredient in our next step.

Step 1: Next we consider the same function f as above, but now ask: *Is the fraction of hyperplanes on which f has degree greater than d, a constant (independent of d)?* Such a statement is necessary to show that the q^{-m}-soundness of the $(m-1)$-dimensional test is an absolute constant (independent of d): the function f is q^{-m}-far from degree d polynomials and so the fraction of $(m-1)$-dimensional affine subspaces on which f is not of degree d better be a constant. Such a strong analysis is not implied by our theorem statement, but is essential to the proof approach of [2]. We give an affirmative answer to this question. Proving this turns out to be non-trivial and does not follow from either [7] or [2]. Indeed our proof is new even for the case of $q = 2$.

We manage to give a relatively clean proof of this statement by interpreting restrictions to hyperplanes algebraically. Since this style of analysis is central also to the next step, we give the essential details here (though formalizing some steps ends up requiring more work). For simplicity, assume we are restricting f to a hyperplane of the form $x_1 = \sum_{i=2}^{m} y_i x_i + y_0$. The restriction of the function f to this hyperplane is now given by the function $f_{y_2,\ldots,y_m,y_0}(x_2,\ldots,x_m) = f(\sum_{i=2}^{m} y_i x_i + y_0, x_2, \ldots, x_m)$, which can be viewed as a polynomial in $x_2,\ldots,x_m$ whose coefficients are themselves polynomials in $y_2,\ldots,y_m,y_0$. By the previous paragraph, it (roughly) follows that there exists a setting of $y_2,\ldots,y_m,y_0$ such that $f_{y_2,\ldots,y_m,y_0}$ is not a polynomial of degree d. In turn this implies that there is a monomial of degree greater than d in $x_2,\ldots,x_m$ whose coefficient is a non-zero function of $y_2,\ldots,y_m,y_0$. The key now is to notice that this coefficient is a polynomial in $y_2,\ldots,y_m,y_0$ of degree at most $q-1$ and so is non-zero with probability at least $1/q$ when $y_2,\ldots,y_m,y_0$ are assigned randomly.

This step is performed in Section 4.3. The heart of the proof is given by Lemma 4.6, which formalizes the above argument and extends it to general hyperplanes (which may not have support on x_1). An important ingredient of the general proof is that instead of trying to understand the function f we apply an invertible linear transformation to the space $\mathbb{F}_p^m$ and consider the function $f \circ A$. It is clearly enough to understand the restrictions of this function. The point is that we can pick A in such a way that $f \circ A$ contains a *canonical monomial* which is a monomial of a very special form (see Definition 4.1). Intuitively, a canonical monomial has its degree "squeezed" to a few variables. The notion of canonical-monomials did not appear in [7] and it makes our proofs considerably simpler. Roughly, having a canonical monomial in a polynomial enables us to focus almost entirely on this monomial instead of the whole polynomial. Furthermore, when restricting our attention to canonical monomials, the algebraic approach, hinted at the previous paragraph, becomes transparent and easy to use. For that reason canonical monomials will play an important role in all our proofs. Proving the existence of a transformation A such that $f \circ A$ has a canonical monomial, is done in Lemma 4.4. Basically, the proof shows that a canonical monomial for f can be found by taking the maximal monomial, in the graded lexicographical order, among all monomials in $\{f \circ B\}$, when we run over all invertible linear transformations B. We discuss canonical monomials in Section 4.1.

Step 2: We then move to the third in the series of questions. If previously we asked whether there exists a hyperplane, or even a noticeable fraction of hyperplanes where f has degree greater than d, we now ask: *Do an overwhelming number of hyperplanes reveal that f has degree greater than d?* We analyze this question when f is a polynomial of degree $d+1$, thus leading to an analysis of $N_0(q,d)$ (or $N_0(q,d+1)$ to be precise). We show that the number of hyperplanes on which f has degree d is $O(q^{t_{q,d}})$. So if the number of variables m is really large compared to q,d then the fraction of hyperplanes where f drops in degree is tiny.

This bound again views the restriction of f to hyperplanes of the form $x_1 = \sum_{i=2}^{m} y_i x_i + y_0$ as a polynomial in $x_2,\ldots,x_m$ and $y_2,\ldots,y_m,y_0$. We then perform an elementary, though somewhat non-obvious, algebraic analysis of this polynomial to show that there are few hyperplanes where f loses degree. Roughly, we show that when working with an appropriate basis for the space (i.e. when applying the linear transformation that guarantees the existence of a canonical monomial, found in the previous step) it is the case that for every fixing of $y_2,\ldots,y_t$, where $t = \log_q N_0(q,d) \approx t_{q,d}$, there is at most one setting of $y_{t+1},\ldots,y_m$ such that the degree of f decreases on the corresponding hyperplane. Canonical monomials again play a crucial role in the proof.

This step is captured by Theorem 1.5 that is proved in Section 4.4. Lemma 4.8 is the main step in which we give the analysis for hyperplanes of the form $x_1 = \sum_{i=2}^{m} y_i x_i + y_0$ that is described above.

Step 3: This leads to the final step (which unfortunately ends up getting proved in two substeps) where we consider general functions that are $\Omega(q^{-t_{q,d}})$-far from degree d polynomials and show that even in this case (which subsumes the case of degree $d+1$ polynomials), the number of hyperplanes on which f drops in degree is bounded by $O(q^{t_{q,d}})$, thus giving a bound on $N_1(q,d)$.

This part is itself proved by induction on the number of variables (with the base case being the hardest step; we will get to that later). And the inductive claim is somewhat different: instead of talking about functions that are far from polynomials (in some loose sense), we explicitly ignore a known small subset of the domain and argue f is a polynomial on the rest. Specifically, we assert that if a function f is a degree d polynomial on a large, $K > N_1(q,d)$, number of hyperplanes $A_1,\ldots,A_K$, then there is

a degree d polynomial Q that agrees with f on the union of $A_1, \ldots, A_K$. Since the union has large volume it follows that f is close to some degree d polynomial (specifically Q). The inductive claim is relatively easy when the number of variables is very large. In such case if we consider the restriction of f to some generic hyperplane A then all the intersections $A_i \cap A$ are distinct, and we can use the inductive claim to assert that $f|_{A \cap (\cup_i A_i)}$ is a degree d polynomial Q_0. Since this holds with overwhelmingly high probability over A, we can claim the same holds also for the $q-1$ parallel shifts of A, and since these cover $\mathbb{F}_q^m$, we can claim (by interpolation) that $f|_{\cup_i A_i}$ is a degree $d+q$ polynomial Q. Now, if $K > N_0(q, d+q)$, then this allows us to use the bound from the previous step (the low-degree polynomial Q cannot drop in degree too often) to claim that Q must be a degree d polynomial. This is the argument behind the induction step in the proof of Theorem 1.7, that is given in Section 4.5.

All this works fine when the number of variables is large. As the number of variables gets smaller, some things break down. $A \cap A_i$ starts coinciding with $A \cap A_j$ for some pairs etc., but careful counting makes sure we do not lose too much in this as long as the number of variables is sufficiently larger (by an additive constant) than $\log_q K$ (the number of given hyperplanes). This becomes our "base case", and we resort to a different argument at this stage.

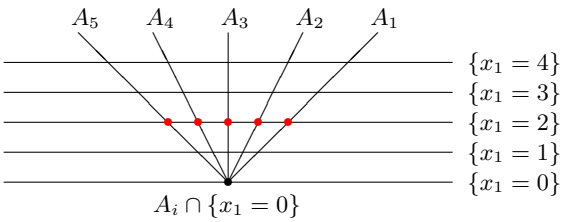
In the base case, we have that a constant fraction of all hyperplanes are "good" - i.e., f restricted to these form a degree d polynomial. It seems intuitive that at this stage f ought to be a degree d polynomial on the union of these (huge) number of hyperplanes, yet there seems to be no obvious way to conclude this intuitive fact. Furthermore, the density of hyperplanes is so high that restricting our attention to any lower dimensional hyperplane would not maintain the *number* of hyperplanes on the restriction (namely, for every hyperplane A there are $i, j \in [K]$ such that $A \cap A_i$ collides with $A \cap A_j$). However we now use the density in our favor by finding q hyperplanes, say $A_1, \ldots, A_q$, that have the same intersection. I.e., $A_i \cap A_j = A_j \cap A_k$ for every triple of distinct $i, j, k \in [q]$. To show that q such hyperplanes exist we use the "density Hales-Jewett theorem" [4], [9] — a somewhat heavy hammer with a high associated cost (see Theorem 3.4). The high cost is the base case dimension has to be lower bounded by a very large constant, albeit a constant — specifically it is some sort of Ackerman function of some polynomial in q (in the improved proof of the density Hales-Jewett theorem [9]). Nevertheless it does imply that if $\log N_1(q, d)$ is sufficiently large as a function of q (a constant we label $\lambda_{q,6}$), then this allows to conclude that q such "near-parallel" hyperplanes do exist. Now, with a linear change of basis, we can assume that the $A_i \cap A_j$ is contained in the hyperplane $x_1 = 0$, and that none of the hyperplanes $A_1, \ldots, A_q$ is equal to the hyperplane $x_1 = 0$. The crux of the idea is that now, on all the $q-1$ hyperplanes, $x_1 = \alpha$, $\alpha \in \mathbb{F}_q - \{0\}$, the hyperplanes $A_1 \cap \{x_1 = \alpha\}, \ldots, A_q \cap \{x_1 = \alpha\}$ are parallel. The situation is perhaps better explained by Figure 1 (for the case $q = 5$).

This allows us to prove (using arguments similar to the inductive step) that f on these hyperplanes is a degree d polynomial, and roughly tells us what $Q \mod (x_1^{q-1} - 1)$ is (where Q is the desired polynomial of degree d that agrees with f on the union $\cup_{i \in [K]} A_i$). Pushing our luck further, we note that if $\log N_1(q, d) = t + \lambda_{q,6}$ then we can find t independent variables $x_1, \ldots, x_t$ such that we know the polynomial $Q \mod \prod_{i=1}^{t} (x_i^{q-1} - 1)$. If $t > d/(q-1)$ this should tell us exactly what Q is, and with some careful examination we confirm this intuition, and show that this polynomial Q agrees with f on every one of the given hyperplanes, thus concluding the analysis in the base case. The base case is given in Lemma 4.11.

Putting things together: Once we have the upper bound on $N_1(q, d)$ (tight to within constant factors that depends only on q), it is straightforward to mimic the work of [2] to derive an analysis of the (roughly) $\log_q N_1(q.d)$-dimensional test, which shows that this test is absolutely sound. We then use the fact from Step 2 (for every $m > t_{q,d}$ an m-dimensional non-degree d polynomial f is of degree greater than d on at least $1/q$ fraction of the hyperplanes) to conclude that the soundness of the $(m-1)$-dimensional test is at least a $1/q$-fraction of the soundness of the m-dimensional test, as long as $m > t_{q,d}$. After a constant number of such steps, we end up with a soundness analysis of $t_{q,d}$-dimensional test also!

Organization of this paper: In what follows we present a skeleton of the proof in the form of the statements of principal lemmas and theorems that are proved in the path to the proof of Theorem 1.3 (and were mentioned earlier in this section). Proofs of the lemmas themselves are omitted, but can be found in the full version [5]. To assist the reader, we use the same numbering for sections, lemmas and theorems here as in the full version (and as a result the numbering here is not consecutive).

Section 3 contains some notations and basic facts regarding polynomials. We discuss the density Hales-Jewett theorem in Section 3.2. The main body of the paper is Section 4. The section is organized as follows. In Section 4.1 we give the definition of canonical monomials and assert how (for any given polynomial) the space $\mathbb{F}_q^n$ can be "rotated" to find a canonical monomial (Lemma 4.4). Section 4.2 shows the basic and simple fact that the rejection probability of the ℓ-dimensional test is monotone in ℓ and in Section 4.3 we prove that although the rejection probability is monotone, it does not decrease too fast when we go from ℓ to $\ell - 1$ (Lemma 4.6). We then present our two main technical contributions. Theorem 1.5, in which we bound $N_0(q, d)$, is described in Section 4.4 and we give some details on how this leads to the proof of Theorem 1.7 in Section 4.5. Section 4.6 contains a strengthening of Theorem 1.7 (given



We think of the horizontal line $\{x_1 = \alpha\}$ as a description of the corresponding hyperplane. The diagonal lines correspond to the different A_i's. The intersection of all A_i's, depicted by the black circle, is an affine subspace of co-dimension 1 inside $\{x_1 = 0\}$. Focusing on the subspace $\{x_1 = 2\}$, we have the q different intersections $A_i \cap \{x_1 = 2\}$, captured by the red circles. Furthermore, all those subspaces are parallel to each other and cover all of $\{x_1 = 2\}$ (which is not evident from the picture).

Figure 1. Near parallel hyperplanes

as Theorem 4.16). Finally, we give lemmas analyzing the $t_{q,d}$-dimensional test in Section 5, leading to a proof of Theorem 1.3 – our main theorem.

3. Preliminaries

Throughout the paper $q = p^k$ is a power of a prime number p and $\mathbb{F}_q$ is the field of characteristic p with q elements. We denote by $\equiv_p$ equality modulo p. Recall that for every $0 \neq \alpha \in \mathbb{F}_q$ it holds that $\alpha^{q-1} \equiv_p 1$. For an integer t we denote $[t] = \{1, \ldots, t\}$.

Recall that $H(q, n)$ is the set of hyperplanes in $\mathbb{F}_q^n$. Similarly, we denote $\mathrm{Aff}(q, n)$ the set of affine linear functions in $\mathbb{F}_q^n$. We will often use the fact that every hyperplane is the set of zeros of an affine linear function. We will also use the term *flat* to denote an affine subspace (of dimension possibly lower than $n-1$). When $L = \sum_{i=1}^n \alpha_i x_i + \alpha_0$ is a linear function, we call α_0 the *free term* of L.

Let $d, e \in \mathbb{N}$ be integers and denote by $d = \sum_i d_i p^i$ and $e = \sum_i e_i p^i$ their base p expansion. Namely, $\forall i\ 0 \leq d_i, e_i < p$. We denote $d \leq e$ if d is not larger than e as integers and $d \leq_p e$ if for every i it holds that $d_i \leq e_i$. We recall Lucas' theorem.

Theorem 3.1 (Lucas' theorem)**.** *In the notations above,* $\binom{e}{d} \equiv_p \prod_i \binom{e_i}{d_i}$*, where* $\binom{e_i}{d_i} = 0$ *if* $e_i < d_i$.

In particular, $\binom{e}{d} \not\equiv_p 0$ if and only if $d \leq_p e$. It follows that for $e < q$ the expansion of $(y+z)^e$ in $\mathbb{F}_q$ has the form

$$(y+z)^e \equiv_p \sum_{d \leq_p e} \binom{e}{d} y^{e-d} z^d. \tag{1}$$

We will represent functions $f : \mathbb{F}_q^n \to \mathbb{F}_q$ as n-variate polynomials, with individual degrees at most $q-1$. Whenever we have a polynomial that has a variable of degree larger than $q-1$ we will use the identity $x^q - x \equiv_p 0$ to reduce its degree.

3.1. The Distance Between Polynomials

A basic fact that is required for understanding the testing dimension for polynomials of degree d is the minimal distance between any two such polynomials. It is well known (cf. [3]) that if $d = r(q-1)+s$ where $0 \leq s < q-1$ then the relative minimal distance is $(q-s)q^{-r-1}$. We state a slightly weaker version below that still suffices for our needs.

Lemma 3.2. *Let $q = p^k$, where p is a prime number. Let $f \neq g \in \mathbb{F}_q[x_1, \ldots, x_n]$ be two distinct polynomials of degree at most d and individual degrees at most $q-1$. Then $\delta(f, g) \geq q^{-d/(q-1)}$.*

3.2. Density Hales-Jewett Theorem

We will need to use the following version of the density Hales-Jewett theorem. The theorem was first proved by Furstenberg and Katznelson [4]. A more recent prove with explicit bounds on the density parameters was obtained in [9].

Before stating the theorem we need to define the notion of a combinatorial line. Let $\Sigma = \{a_1, \ldots, a_q\}$ be an alphabet of size q. E.g., one can think of Σ as being $\mathbb{F}_q$. A set $\mathcal{L} = \{v_1, \ldots, v_q\} \subset \Sigma^n$ is a *combinatorial line* if we can partition the coordinates $[n]$ to two disjoint sets $[n] = I \cup J$, $I \cap J = \emptyset$ such that: (1) For all $i \in I$ and $k, k' \in [q]$, $(v_k)_i = (v_{k'})_i$. Namely, for all $i \in I$, the i'th coordinate of all elements in $\mathcal{L}$ is fixed. (2) For $j \in J$ and $k \in [q]$, $(v_k)_j = a_k$. I.e., the j'th coordinate advances with k.

It is not hard to see that if we set $\Sigma = \mathbb{F}_q$ then a combinatorial line in $\mathbb{F}_q^n$ corresponds to a set of the form $\{v + tu \mid t \in \mathbb{F}_q\}$ where $v \in \mathbb{F}_q^n$, $u \in \{0,1\}^n \setminus \{\bar{0}\}$ and v, u have disjoint supports. In particular, a combinatorial line in $\mathbb{F}_q^n$ is a line in the geometric sense.

Theorem 3.4 ([4], [9])**.** *For any integer q and any $0 < c \in \mathbb{R}$ there exists an integer $\lambda_{q,c}$, such that if $n \geq \lambda_{q,c}$ then any set $A \subseteq \mathbb{F}_q^n$, of size $|A| \geq q^n/q^c$, contains a combinatorial line.*

We now state an easy corollary of the theorem. We say that u is the *direction* of the line $\{v + tu \mid t \in \mathbb{F}_q\}$. Notice that, say, $2u$ is also the direction of the line but since u and $2u$ are linearly dependent we ignore this small issue.

Corollary 3.5. *Let $1 \leq t$ be an integer. If $n \geq \lambda(q,c)+t-1$ then any set $A \subseteq \mathbb{F}_q^n$, of size $|A| \geq q^n/q^c$, contains t combinatorial lines whose directions are linearly independent.*

4. Restrictions to Hyperplanes

In this section we will study the behavior of polynomials when restricted to hyperplanes. Recall that a hyperplane $A \subset \mathbb{F}_q^n$ is an $(n-1)$-dimensional affine subspace. For each hyperplane there is a linear function L such that

$$A = \{x \mid L(x) = 0\}.$$

It will be convenient to express L as $L(x) = x_k - \sum_{i=k+1}^n \alpha_i x_i - \alpha_0$, where k is the first non-zero coefficient in L (the coefficient of x_k is not necessarily 1, but scaling L by a constant does not change the definition of A so we can assume this w.l.o.g.). For such an L we will express the restriction of f to A as

$$\begin{aligned} f|_A &= f(x_1,\ldots,x_n)|_{L=0} \\ &= f(x_1,\ldots,x_{k-1},\sum_{i=k+1}^n \alpha_i x_i + \alpha_0, x_{k+1},\ldots,x_n), \end{aligned}$$

since setting $L = 0$ is equivalent to substituting $\sum_{i=k+1}^n \alpha_i x_i + \alpha_0$ to x_k.

4.1. Canonical Monomials

The notion of canonical monomial will play an important role in our proofs. Intuitively, the reason for defining canonical monomials is because they decrease in degree on any hyperplane, and thus give an extremal example that is useful to study.

Definition 4.1. *A canonical monomial of degree d in $m \leq n$ variables over $\mathbb{F}_q$ is a monomial $\prod_{i=1}^m x_i^{e_i}$ such that (1) $\sum_{i=1}^m e_i = d$. (2) For all $1 \leq i < m$, $q - q/p \leq e_i < q$. (3) If $p^i \leq_p e_m$ then for every $j < m$, $p^i + e_j > q-1$. (4) $e_m < q$.*

The following simple lemma shows that whenever we have a bivariate polynomial over $\mathbb{F}_q$ there exists an invertible linear transformation $A : \mathbb{F}_q^2 \to \mathbb{F}_q^n$, such that $f \circ A$ contains a canonical monomial of maximal degree.

Lemma 4.2. *Let $f(x_1,x_2)$ be a degree $d \leq 2(q-1)$ polynomial over $\mathbb{F}_q$. Then, there exists $\alpha \in \mathbb{F}_q$ such that $f(x_1, x_2 + \alpha x_1)$ contains a canonical monomial of degree d.*

Theorem 3.1 is the main ingredient in the proof of the lemma above. The next lemma generalizes the above claim to n-variate polynomials. In fact, we will prove a slightly stronger property. For that end we will need the following definition.

Definition 4.3 (Graded Lexicographical Order)**.** *We denote $\prod_{i=1}^n x_i^{e_i} >_m \prod_{i=1}^n x_i^{r_i}$ if $\sum_{i=1}^n e_i > \sum_{i=1}^n r_i$ or if $\sum_{i=1}^n e_i = \sum_{i=1}^n r_i$ and the first i for which $e_i \neq r_i$ satisfies $e_i > r_i$. Note that we only consider monomials in which all individual degrees are smaller than q (we can reduce the degree of other monomials). The* max-monomial *of a polynomial g is the maximal monomial appearing in g (with a non-zero coefficient of course).*

Lemma 4.4. *Let $f(x_1,\ldots,x_n)$ be a degree $d \leq n(q-1)$ polynomial over $\mathbb{F}_q$. Let*

$$A = \operatorname*{argmax}_{\text{invertible } B} \textit{max-monomial of } (f \circ B)(x_1,\ldots,x_n).$$

In words, $A : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is an invertible linear transformation such that the max-monomial of $(f \circ A)$ is maximal, in the graded lexicographical order, among all monomials of all polynomials of the form $f \circ B$, for invertible B. Then, the max-monomial of $f \circ A$ is a canonical monomial of degree d.

4.2. Monotonicity

Next we note a monotonicity property of the rejection probability, namely that $\rho_d(f,k)$ is monotone in k. This turns out to be useful in our eventual analysis.

Lemma 4.5. *Let $k > k'$ be two integers and $f : \mathbb{F}_q^n \to \mathbb{F}_q$ a function. Then $\rho_d(f,k) \geq \rho_d(f,k')$.*

4.3. Relating Different Dimensions

The first lemma in this section asserts that if a $(k+1)$-variate function f has degree larger than d (when k is not too small relatively to d) then $\rho_d(f,k) \geq 1/q$. Notice that we need to lower bound k as, for example, when $k = d/(q-q/p)$, the degree of $x_1^{q-q/p} \cdot \ldots \cdot x_k^{q-q/p}$ decreases by $q - q/p$ on any subspace. Proposition 1.2 is an (almost) immediate consequence of this lemma.

Lemma 4.6. *Let $k \geq (d+1)/(q-q/p)$ and let $f : \mathbb{F}_q^{k+1} \to \mathbb{F}_q$ have degree larger than d. Then $\rho_d(f,k) \geq 1/q$.*

Applying the above lemma iteratively we obtain the following.

Lemma 4.7. *Let $n \geq k \geq (d+1)/(q-q/p)$ and let $f : \mathbb{F}_q^n \to \mathbb{F}_q$ have degree larger than d. Then $\rho_d(f,k) \geq q^{-(n-k)}$. Moreover, if $n \geq k' \geq k$ then $\rho_d(f,k) \geq \rho_d(f,k') \cdot q^{-(k'-k)}$.*

4.4. The Case of Polynomials of Degree $d+1$

In this section we show that $N_0(q,d)$, the number of hyperplanes on which a degree d polynomial has degree at most $d-1$ (see Definition 1.4), is not too large. Specifically, we show that $N_0(q,d) \leq \widehat{N}_0(q,d) = q^{\left\lfloor \frac{d-q/p}{q-q/p} \right\rfloor + 1}$. Observe that

$$q^{t_{q,d}-1} \leq \widehat{N}_0(q,d) = q^{\left\lfloor \frac{d-q/p}{q-q/p} \right\rfloor + 1} < q^{t_{q,d-1}+1} \leq q^{t_{q,d}+1}.$$

As a first step we bound the number of such hyperplanes that 'depend' on x_1.

Lemma 4.8. *Let f be a polynomial of degree d. Assume that f has a monomial of degree d that contains x_1 and at most $t-1$ other variables. Then there are at most $(q-1)q^{t-1}$ linear functions L of the form $L(x_1,\ldots,x_n) = x_1 + \sum_{i=2}^{n} \alpha_i x_i + \alpha_0$ such that $\deg(f|_{L=0}) \leq d-1$.*

In words, if the minimal number of variables that appear with x_1 in a monomial of degree d in f is $t-1$, then there are at most $(q-1)q^{t-1}$ linear functions, that depend on x_1, such that the degree of f decreases on the hyperplanes defined by them. The proof is similar in spirit to the proof of Lemma 4.6. We basically show that after fixing some coefficients in a linear function, the number of completions to linear functions L that have those fixed coefficients and such that $\deg(f|_{L=0}) < \deg(f)$ is small.
The following lemma extends the above to functions that do not necessarily depend on x_1.

Lemma 4.9. *Let f be a polynomial that has a max-monomial containing only t variables. Then there are at most q^t linear functions L such that $\deg(f|_{L=0}) \leq \deg(f) - 1$.*

The above lemmas immediately give a proof of Theorem 1.5. We repeat the statement here (in a slightly different form).
Theorem 1.5 restated. *Let $f : \mathbb{F}_q^n \to \mathbb{F}_q$ be a polynomial of degree d. Then the number of linear functions L such that $\deg(f|_{L=0}) < d$ is at most $\widehat{N}_0(q,d) = q^{\left\lfloor \frac{d-q/p}{q-q/p} \right\rfloor + 1}$. In particular $N_0(q,d) \leq \widehat{N}_0(q,d)$.*

Corollary 4.10. *Let n, d, q, K be integers such that $K > \widehat{N}_0(q,d)$. Let f be an n-variate polynomial of degree at most d over F_q. If there exist K hyperplanes $A_1, ..., A_K$, such that for all $i \in [K]$ $\deg f|_{A_i} \leq d' < d$, then $\deg f \leq d'$.*

4.5. Interpolating from Exact Agreement

Next we turn to the proof of Theorem 1.7 that shows that if we have enough 'pairwise consistent' polynomials then it is possible to obtain 'global' consistency. We first restate the theorem in more explicit terms.
Theorem 1.7 restated. *Let $A_1,\ldots,A_K$ be distinct hyperplanes in $\mathbb{F}_q^n$ and $P_1,\ldots,P_K$ be polynomials of degree d satisfying $P_i|_{A_i \cap A_j} = P_j|_{A_i \cap A_j}$ for every pair $i,j \in [K]$. If*

$$K \geq \widehat{N}_1(q,d) = 2\widehat{N}_0(q,d+q) \cdot q^{\lambda_{q,6}} = 2q^{\left\lfloor \frac{d}{q-q/p} \right\rfloor + 2 + \lambda_{q,6}},$$

where $\lambda_{q,6}$ is the constant $\lambda_{q,c}$ from Theorem 3.4 for $c=6$, then there exists a polynomial Q, of degree d, such that $Q|_{A_i} = P_i|_{A_i}$ for every $i \in [K]$. In other words $N_1(q,d) \leq \widehat{N}_1(q,d)$.
In fact, we prove a slightly stronger statement. Specifically, we show that the conclusion holds when

$$K \geq \widetilde{N_1}(q,d,n) \triangleq \frac{\widehat{N}_1(q,d)}{2\prod_{i=1}^{n-\log_q \widehat{N}_1(q,d)-3}\left(1-\frac{\widehat{N}_1(q,d)}{q^{n-i-1}}\right)}.$$

This is indeed a stronger statement as the denominator above

$$\begin{aligned}
& 2\prod_{i=1}^{n-\log_q \widehat{N}_1(q,d)-3}\left(1-\frac{\widehat{N}_1(q,d)}{q^{n-i-1}}\right) \\
\geq\ & 2\left(1-\sum_{i=1}^{n-\log_q \widehat{N}_1(q,d)-3}\frac{\widehat{N}_1(q,d)}{q^{n-i-1}}\right) \\
=\ & 2-\frac{2\widehat{N}_1(q,d)}{q^{n-1}}\sum_{i=1}^{n-\log_q \widehat{N}_1(q,d)-3} q^i \\
>\ & 2-\frac{2\widehat{N}_1(q,d)}{q^{n-1}} q^{n-\log_q \widehat{N}_1(q,d)-2} \\
=\ & 2-2q^{-1} \geq 1,
\end{aligned}$$

namely, $\widetilde{N_1}(q,d,n) < \widehat{N}_1(q,d)$ for all n, and so the requirement on K is weaker.
We first set some notation. Let $L_i \in \mathrm{Aff}_q^n$ be an affine linear function such that $A_i = \{u \in \mathbb{F}_q^n \mid L_i(u) = 0\}$. For the rest of the proof we denote $\mathcal{L} = \{L_1,\ldots,L_K\}$. We will abuse notations and denote, for $L \in \mathcal{L}$, $P_L = P_i$ and $A_L = A_i$ when $L = L_i$. Another important notation is the following. For $L \in \mathrm{Aff}_q^n$ and $\gamma \in \mathbb{F}_q$ we denote

$$B_{L,\gamma} \stackrel{\text{def}}{=} \{v \in \mathbb{F}_q^n \mid L(v) = \gamma\} \quad \text{and} \quad A_{i,L,\gamma} \stackrel{\text{def}}{=} A_i \cap B_{L,\gamma}.$$

Note that for γ_1, γ_2, the hyperplanes B_{L,γ_1} and B_{L,γ_2} are shifts of each other (they can also be empty sets if L is a constant function).
The proof is by induction on the number of variables n. The idea of the proof is to find a linear function L and restrict our attention to the different hyperplanes $B_{L,\gamma}$. We show that we can find an L such that the induction assumption holds for every $B_{L,\gamma}$. By the induction hypothesis, for each $B_{L,\gamma}$ there is a polynomial P_γ, of degree d, that is defined over $B_{L,\gamma}$ and is consistent there with the P_i's. Then we 'glue' the P_γ's together to get a polynomial of degree at most $d+q$ that is consistent with all the P_i's. But now, we can use Theorem 1.5 to claim that this resulting polynomial must have degree at most d.
This is indeed the idea, but what is swept under the rug here is the base case which is technically challenging. The base of the induction for us is the case $n < \log_q \widehat{N}_1(q,d)+4$. For such n it holds that $\widetilde{N_1}(q,d,n) = \frac{1}{2}\widehat{N}_1(q,d)$. The analysis of this case, which is the technical heart of the proof, is given in the next lemma. While we omit the proof of this lemma also, we note that it is this proof that invokes the "density-Hales-Jewett" result (Theorem 3.4). (See the discussion in Section 2.)

Lemma 4.11 (Main Lemma)**.** *Let $n < \log_q \widehat{N}_1(q,d) + 4$ and $K \geq \widetilde{N_1}(q,d,n) = \widehat{N}_1(q,d)/2$. Let $A_1, ..., A_K$ be distinct hyperplanes in $\mathbb{F}_q^n$ and let $P_1, ..., P_K$ be polynomials of degree d satisfying $P_i|_{A_i \cap A_j} = P_j|_{A_i \cap A_j}$ for every*

$i, j \in [K]$. *Then there exists a degree d polynomial P such that for every $i \in [K]$, $P|_{A_i} = P_i$.*

We omit the proof of Lemma 4.11 here. Theorem 1.7 follows immediately from this lemma.

4.6. Interpolating from Approximate Agreement

We use Theorem 1.7 to prove a version which applies to functions which are *close* to degree d polynomials. Specifically, we consider a function f whose restriction on many hyperplanes is close to some degree d polynomial, and show that such a function is close to a degree d polynomial. This proof essentially follows [2] (using Theorem 1.7 of course) and we omit it here.

Theorem 4.16. *Let $\delta_1 < \frac{1}{2}q^{-(1+(d/(q-1))}$ and $K \geq N_1(q,d)$. If the function $f : \mathbb{F}_q^n \to \mathbb{F}_q$ and hyperplanes $A_1, \ldots, A_K$ are such that $\delta_d(f|_{A_i}) \leq \delta_1$ for every $i \in [K]$, then $\delta_d(f) \leq 2\delta_1 + 4(q-1)/K$.*

5. Analysis of the low-degree tests

The following lemmas are now natural extensions of analogous ones from [2]. The proof of Lemma 5.1 is straightforward, and the proof of Lemma 5.2 uses Theorem 4.16 as a central ingredient. Theorem 1.3 follows immediately from these lemmas. We skip all proofs here.

Lemma 5.1. *Let $t \geq d/(q-1)$ be an integer. Then, if $\delta_d(f) \leq \frac{1}{2}q^{-d/(q-1)}$ then $\rho_d(f,t) \geq \min\{\frac{1}{4q}, \frac{1}{2} \cdot q^t \cdot \delta_d(f)\}$.*

Lemma 5.2. *For every q, there exists $\epsilon > 0$ and c such that for every d, $t \geq t_{q,d} + c$ and n, the following hold: Let $f : \mathbb{F}_q^n \to \mathbb{F}_q$ be a function with $\delta_d(f) \geq q^{-t}$. Then $\rho_d(f,t) \geq \epsilon + \frac{1}{8}q^t \cdot \sum_{i=n+1}^{\infty} q^{-i}$.*

6. Acknowledgement

Research of E.H and A.S was partially supported by the Israel Science Foundation (grant number 339/10).

References

[1] 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] A. Bhattacharyya, S. Kopparty, G. Schoenebeck, M. Sudan, and D. Zuckerman, "Optimal testing of reed-muller codes," in *Proceedings of the 51th Annual FOCS*, 2010, pp. 488–497.

[3] P. Ding and J. D. Key, "Minimum-weight codewords as generators of generalized reed-muller codes," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2152–2158, 2000.

[4] H. Furstenberg and Y. Katznelson, "A density version of the Hales-Jewett theorem," *J. d'Analyse Math.*, vol. 57, pp. 64–119, 1991.

[5] E. Haramaty, A. Shpilka, and M. Sudan, "Optimal testing of multivariate polynomials over small prime fields," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 18, no. 59, 2011.

[6] C. S. Jutla, A. C. Patthak, A. Rudra, and D. Zuckerman, "Testing low-degree polynomials over prime fields," in *Proceedings of the 45th Annual FOCS*, 2004, pp. 423–432.

[7] T. Kaufman and D. Ron, "Testing polynomials over general fields," *SIAM J. Comput.*, vol. 36, no. 3, pp. 779–802, 2006.

[8] T. Kaufman and M. Sudan, "Algebraic property testing: the role of invariance," in *Proceedings of the 40th Annual STOC*, 2008, pp. 403–412.

[9] D. H. J. Polymath, "A new proof of the density Hales-Jewett theorem," *CoRR*, vol. arxiv.org/abs/0910.3926, 2009.

[10] R. Rubinfeld and M. Sudan, "Robust characterizations of polynomials with applications to program testing," *SIAM J. Comput.*, vol. 25, no. 2, pp. 252–271, 1996.