Electronic Colloquium on Computational Complexity, Report No. 59 (2011)

Optimal testing of multivariate polynomials over small prime fields

Elad Haramaty* Amir Shpilka† Madhu Sudan‡

April 15, 2011

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 feature 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. [AKK+05], and Kaufman and Ron [KR06] and Jutla et al. [JPRZ04], showed that this natural test rejected functions that were $\Omega(1)$-far from degree d-polynomials with probability at least $\Omega(q^{-t})$ (the results of [KR06] hold for all fields $\mathbb{F}_q$, while the results of [JPRZ04] hold only for fields of prime order). Thus to get a constant probability of detecting functions that were 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. It was unclear if the soundness analysis of these tests were tight and this question relates closely to the task of understanding the behavior of the Gowers Norm. This motivated the work of Bhattacharyya et al. [BKS+10], who gave an optimal analysis 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 at least $\Omega(1)$.

In this work we give an optimal analysis of this test 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. Our analysis thus shows that this test is optimal (matches known lower bounds) when q is prime. (It is also potentially best possible for all fields.) Our approach extends the proof technique of Bhattacharyya et al., however it has to overcome many technical barriers in the process. The natural extension of their analysis leads to an $O(q^d)$ query complexity, which is worse than that of Kaufman and Ron for all q except 2! 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.

*Faculty of Computer Science, Technion — Israel Institute of Technology, Haifa, Israel, eladh@cs.technion.ac.il. This research was partially supported by the Israel Science Foundation (grant number 339/10).

†Faculty of Computer Science, Technion — Israel Institute of Technology, Haifa, Israel and Microsoft Research, Cambridge, MA, shpilka@cs.technion.ac.il. This research was partially supported by the Israel Science Foundation (grant number 339/10).

‡Microsoft Research, One Memorial Drive, Cambridge, MA 02142, USA, madhu@mit.edu

ISSN 1433-8092

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 [AKK+05, KR06, JPRZ04, BKS+10], 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 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 [RS96] 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 considered by Alon et al. [AKK+05] 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 Ron [KR06] and independently (for the case of prime q) by Jutla et al. [JPRZ04]. They (in particular [KR06]) 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 [AKK+05] and [KR06, JPRZ04] were similar and indeed the subsequent generalization of Kaufman and Sudan [KS08] 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.

[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.

In a recent work, Bhattacharyya et al. [BKS+10] 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 [AKK+05] is absolutely sound. This also gave the first example of a linear-invariant property with tight bounds on query complexity.

The proof of [BKS+10] was significantly more algebraic than those of [AKK+05, KR06, JPRZ04]. (Indeed the work of [KS08] confirms that the central ingredient in the proofs in [AKK+05, KR06, JPRZ04] are all the same and relies on very little algebra.) However, the proof of [BKS+10] 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 [KR06] 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 [AKK+05, KR06] 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 $\deg(f|_A)\le\deg(f)$. 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 $\deg(f|_A)\le d$. 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 degree d polynomials 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 [KR06] who proved the following fact. As it also follows easily from our results we give the proof in Section 4.3.

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

The test proposed by [KR06] 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 $\deg(f|_A)\le d$.

[KR06] 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 [BKS+10] 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 $\deg(f|_{A_i}) < d$. I.e.,*

$$N_0(q, d) = \max_{n, f \in \mathcal{P}(n,d,q)} |\{A \in \mathrm{H}(n, q) | \deg(f|_A) < d\}|.$$

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 [BKS+10] 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 $|\{A \in \mathrm{H}(q, n) | \deg(f) < d\}| \leq N_0(d, q) \leq q^{t_{q,d}+1}$.*

(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 [BKS+10]

While our proof outline does follow the same one as that of [BKS+10] 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 [BKS+10] — and we simply mimic their proofs. (We note that much of the novelty of [BKS+10] 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 [BKS+10] 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 Lemma 4.6) 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" [FK91, Pol09] 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 [KR06], 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 [BKS+10]. We give an affirmative answer to this question. Proving this turns out to be non-trivial and does not follow from either [KR06] or [BKS+10]. 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 $x_2,\ldots,x_m$ which 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 [KR06] 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 the 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.

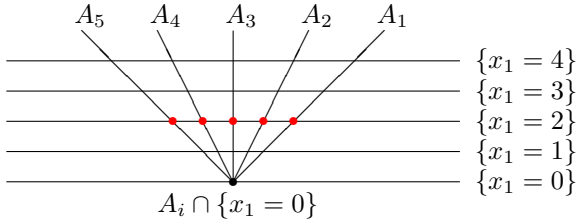
Step 3. This leads to the final step (which unfortunately ends up getting proved in two substeps) where we consider general functions that are $\Omega(q^{-t_{q,d}})$-far from degree d polynomials and show that even in this case (which subsumes the case of degree $d+1$ polynomials), the number of hyperplanes 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 (Claim 4.12) 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" [FK91, Pol09] — 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 [Pol09]). Nevertheless it does imply that if $\log N_1(q, d)$ is sufficiently large as a function of q (a constant we label $\lambda_{q,4}$), 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



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

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 $\log N_1(q, d) = t + \lambda_{q,4}$ 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 [BKS+10] 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. Section 3 contains some notations and basic facts regarding polynomial. 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 shows how to "rotate" the space in order to find one (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 give the proofs of our two main technical contributions. Theorem 1.5, in which we bound $N_0(q,d)$, is proved in Section 4.4 and Theorem 1.7 is proved in Section 4.5. Section 4.6 contains a strengthening of Theorem 1.7 (given as Theorem 4.16), that is proved in a relatively direct manner from Theorem 1.7. Finally, we analyze the $t_{q,d}$-dimensional test in Section 5, giving 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, $\binom{e}{d} \neq 0$ 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. [DK00]) that if $d = r(q-1)+s$ where $0 \leq s < q-1$ then the relative minimal distance is $(q-s)q^{-r-1}$. However, for completeness we provide an easy proof of a slightly weaker claim 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 $\delta(f,g) \geq q^{-d/(q-1)}$.*

Proof. By linearity it is enough to lower bound the distance of a non-zero f from the zero polynomial. In other words, we have to bound from below the number of non-zeros of f. We do so by induction on n. When $n = 1$, since f has degree at most $d < q$, it has at most d zeros and therefore $\delta(f,0) \geq (q-d)/q = 1 - d/q \geq q^{-d/(q-1)}$, where the last inequality follows from Claim 3.3 proved below. For the induction step, we express f as a polynomial in x_n

$$f(x_1, \ldots, x_n) = \sum_{e=0}^{q-1} x_n^e \cdot g_e(x_1, \ldots, x_{n-1}).$$

Let $e_{\max}$ be the degree of f as a polynomial in x_n. As $\deg(g_{e_{\max}}) \leq d - e_{\max}$, the induction hypothesis implies that the number of non-zeros of $g_{e_{\max}}$ is at least $q^{-(d-e_{\max})/(q-1)} \cdot q^{n-1}$. For any such non-zero $(a_1, \ldots, a_{n-1}) \in \mathbb{F}_q^{n-1}$ we get that $f(a_1, \ldots, a_{n-1}, x_n)$ is a non-zero polynomial in x_n of degree $e_{\max}$ and therefore has at least $q - e_{\max}$ non-zeros. Consequently,

$$\begin{aligned} \delta(f,0) &\geq (q - e_{\max}) \cdot q^{-(d-e_{\max})/(q-1)} \cdot q^{n-1}/q^n = (1 - e_{\max}/q) \cdot q^{-(d-e_{\max})/(q-1)} \\ &\geq^{(*)} q^{-e_{\max}/(q-1)} \cdot q^{-(d-e_{\max})/(q-1)} = q^{-d/(q-1)}, \end{aligned}$$

where inequality $(*)$ follows from Claim 3.3. □

Claim 3.3. *For any $0 \leq x \leq q-1$ it holds that $1 - x/q \geq q^{-x/(q-1)}$.*

Proof. Consider the function $F(x) = 1 - x/q - q^{-x/(q-1)}$. It is easy to see that $F(0) = F(q-1) = 0$ and that the second derivative of F is always negative. It immediately follows that $F \geq 0$ for $0 \leq x \leq q-1$. □

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 [FK91]. A more recent prove with explicit bounds on the density parameters was obtained in [Pol09].

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 coordinates 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 ([FK91, Pol09])**.** *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.*

Proof. The proof is by induction on t. For $t = 1$, Theorem 3.4 implies that A contains a line and the claim follows.

Assume that we proved the statement for all $t' \leq t-1$ and consider $t' = t$. By the induction hypothesis we can find $t-1$ lines in linearly independent directions inside A. To simplify notations assume that those directions are $e_1, \ldots, e_{t-1}$ where $e_i \in \{0,1\}^n$ is zero everywhere except for the i'th coordinate (by applying an invertible linear transformation to A this can be assumed w.l.o.g.). By the pigeonhole principle there is some $u \in \mathbb{F}_q^{t-1}$ such that the number of elements $v \in A$ that identify with u on their first $t-1$ coordinates is large. Namely,

$$\#\{v \in A \mid (v_1, \ldots, v_{t-1}) = u\} \geq |A|/q^{t-1} \geq (q^n/q^c)/q^{t-1} = q^{n-t+1}/q^c.$$

In other words, the number of elements of A that belong to the $(n-t+1)$-dimensional flat

$$\mathcal{M} = \{v \in \mathbb{F}_q^n \mid (v_1, \ldots, v_{t-1}) = u\}$$

is at least $|\mathcal{M}|/q^c$. As the dimension of $\mathcal{M}$ is $n - t + 1 \geq \lambda_{q,c}$, we can apply Theorem 3.4 and get that $A \cap \mathcal{M}$ contains a line. It is immediate that the direction of this line is linearly independent of $e_1, \ldots, e_{t-1}$. □

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

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

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) $e_m < q$.*

Note that Property 3 implies Property 2, but for clarity we keep both.

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.*

Proof. Assume w.l.o.g. that $f(x_1, x_2) = \sum_{e:0 \leq e, d-e < q} \alpha_e x_1^e x_2^{d-e}$ (we can ignore monomials of degree smaller than d). Let $e_{\max}$ be the maximal degree of x_1 in f. If f already has a canonical monomial then we are done (i.e. we can take $\alpha = 0$). Otherwise, consider the monomial containing $x_1^{e_{\max}}$ and let i be such that $p^i \leq_p d - e_{\max}$ and $e_{\max} + p^i < q$. Consider the polynomial $f(x_1, x_2 + z x_1)$. By (1) it follows that

$$f(x_1, x_2 + z x_1) \equiv_p \sum_{e \leq d} \alpha_e x_1^e \sum_{r \leq_p d-e} \binom{d-e}{r} (z x_1)^r x_2^{d-e-r}.$$

The coefficient of $x_1^{e_{\max}+p^i} x_2^{d-(e_{\max}+p^i)}$ in the expression above is equal to

$$\sum_{r \le e_{\max}+p^i} \alpha_{e_{\max}+p^i-r} \binom{d-(e_{\max}+p^i-r)}{r} z^r =$$
$$\alpha_{e_{\max}} \binom{d-e_{\max}}{p^i} z^{p^i} + \sum_{\substack{r \le e_{\max}+p^i \\ r \neq p^i}} \alpha_{e_{\max}+p^i-r} \binom{d-(e_{\max}+p^i-r)}{r} z^r,$$

where some of the binomials $\binom{d-(e_{\max}+p^i-r)}{r}$ may be zero modulo p. However, by our choice of p^i it follows that the coefficient of z^{p^i} in the above expression is non-zero. Hence, since $e_{\max} + p^i < q$, the coefficient of $x_1^{e_{\max}+p^i} x_2^{d-(e_{\max}+p^i)}$ is a non-zero polynomial in z. It follows that there is some $\alpha \in \mathbb{F}_q$ such that the coefficient of $x_1^{e_{\max}+p^i} x_2^{d-(e_{\max}+p^i)}$ in $f(x_1, x_2 + \alpha x_1)$ is non-zero. In this way we can gradually increase the maximal degree of x_1 until we obtain a canonical monomial. Here we use the simple fact that composition of maps of the form $(x_1, x_2) \to (x_1, x_2 + \alpha x_1)$ has the same form. □

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 $d \le n(q-1)$ 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.

Proof. Indeed, since composition with an invertible transformation does not affect the degree, the max-monomial of $f \circ A$ is of degree d. Assume that it is the monomial $M = \prod_{i=1}^m x_i^{e_i}$, where $e_1, \ldots, e_m > 0$. If M is not a canonical monomial then there must exist $i < m$ with $e_i < q - q/p$ (recall that we only consider monomials in which all individual degrees are smaller than q). Assume w.l.o..g. that[2] $i = m-1$. Consider the sum of all monomials of degree d in $f \circ A$ that involve only the variables $x_1, \ldots, x_m$ and that are divisible by $\prod_{i=1}^{m-2} x_i^{e_i}$. Clearly, the sum is a nonzero polynomial $\tilde{f}$ of the form

$$\tilde{f} = \prod_{i=1}^{m-2} x_i^{e_i} \cdot g(x_{m-1}, x_m).$$

[2] In fact, by the choice of A it must be the case that $i = m-1$.

Let $d' = e_{m-1} + e_m$. It follows that g is a nonzero bivariate polynomial of degree d'. Thus, by Lemma 4.2 there is $\alpha \in \mathbb{F}_q$ such that $g(x_{m-1}, x_m + \alpha x_{m-1})$ contains a canonical monomial of degree d'. It follows that the max-monomial of $\tilde{f}(x_1, \ldots, x_{m-1}, x_m + \alpha x_{m-1})$ is larger than M (since we 'pushed' degree from x_m to x_{m-1}). Let $A' = B \circ A$ where $B(v_1, \ldots, v_n) = (v_1, \ldots, v_{m-1}, v_m + \alpha v_{m-1}, v_{m+1}, \ldots, v_n)$. It is clear that A' is an invertible transformation and that the sum of all monomials of degree d in $f \circ A'$ that involve only the variables $x_1, \ldots, x_m$ and that are divisible by $\prod_{i=1}^{m-2} x_i^{e_i}$ is equal to $\tilde{f}(x_1, \ldots, x_{m-1}, x_m + \alpha x_{m-1})$. It is also clear that the max-monomial of $f \circ A'$ is equal to the max-monomial of $\tilde{f}$. This, however, contradicts the choice of A. Hence, it follows that the max-monomial in $f \circ A$ is a canonical monomial. □

A fact that we will use implicitly throughout our proofs, is that if $M = \prod_{i=1}^{m+1} x_i^{e_i}$ is a canonical monomial of degree $d = e_1 + \ldots + e_{m+1}$, then for any linear function $L(x_1, \ldots, x_{m+1})$, $\deg(M|_{L=0}) \leq d - e_{m+1}$. Indeed, assume for simplicity that $L(x) = x_m - \sum_{i=1}^{m} \alpha_i x_i - \alpha_0$ and let $e_{m+1} = \sum_j r_j p^j$ be the base p expansion of e_{m+1}. We have that

$$\begin{aligned} M|_{L=0} &= \left(\prod_{i=1}^{m} x_i^{e_i}\right) \cdot \left(\sum_{i=1}^{m} \alpha_i x_i + \alpha_0\right)^{e_{m+1}} \\ &= \left(\prod_{i=1}^{m} x_i^{e_i}\right) \cdot \left(\sum_{i=1}^{m} \alpha_i x_i + \alpha_0\right)^{\sum_j r_j p^j} \\ &= \left(\prod_{i=1}^{m} x_i^{e_i}\right) \cdot \left(\prod_j \left(\sum_{i=1}^{m} \alpha_i^{p^j} x_i^{p^j} + \alpha_0^{p^j}\right)^{r_j}\right). \end{aligned}$$

Thus, $M|_{L=0}$ contains the monomial $\prod_{i=1}^{m} x_i^{e_i}$ which is of degree $d - e_{m+1}$ and any other monomial contains a variable x_i of degree at least $e_i + r_j p^j$. However, by the definition of canonical monomials, it must be the case that if $r_j \neq 0$ then $e_i + p^j \geq q$ so, when after reducing modulo $x_k^q - x = 0$, the degree of the monomial drops by $q - 1 \geq e_{m+1}$.

In fact, we will usually apply this simple observation for a monomial M that is both a canonical monomial and the max-monomial in some polynomial f (i.e. the max-monomial that is found in Lemma 4.4). Then, by the maximality of M, it follows that the same conclusion will be true for any linear function L, and not just L that is supported on x_{m+1}.

4.2 Monotonicity

Here we prove that $\rho_d(f, k)$ is monotone in k. This is a simple fact that has an easy proof.

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

Proof. Consider the following way to randomly sample a k'-dimensional flat: Choose uniformly at random a k dimension flat $A \subseteq \mathbb{F}_q^n$. Then, choose uniformly at random a k'-dimensional flat $B \subseteq A$.

We have that

$$\begin{aligned}
\rho_d(f,k') &= \Pr_{B:\dim(B)=k'}[\deg(f|_B) > d] \\
&= \Pr_{A:\dim(A)=k}[\deg(f|_A) > d] \cdot \Pr_{B\subseteq A:\dim(B)=k'}[\deg(f|_B) > d \mid \deg(f|_A) > d] \\
&= \rho_d(f,k) \cdot \Pr_{B\subseteq A:\dim(B)=k'}[\deg(f|_B) > d \mid \deg(f|_A) > d] \\
&\leq \rho_d(f,k).
\end{aligned}$$

□

4.3 Relating Different Dimensions

The first lemma in this section shows 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 $\rho_d(f,k) \geq 1/q$.*

Proof. Let A be the invertible linear transformation guaranteed by Lemma 4.4. To simplify notations, assume w.l.o.g. that A is the identity transformation. Let M be the max-monomial of f. By the choice of A, M is a canonical monomial. Denote, $M = \prod_{i=1}^m x_i^{e_i}$, where $\sum_{i=1}^m e_i = \deg(f) > d$. Roughly, we will show that in every linear function L, we can either tweak the coefficient of x_{k+1}, or the free term, so that $\deg(f|_{L=0}) = \deg(f)$. This will prove the claim as it will map at most q different functions to one 'good' function. Formally, we analyze two cases.

Case $m \leq k$. Notice that if $L(x_{m+1},\ldots,x_{k+1})$ is a linear function then $\deg(f|_{L=0}) = \sum_{i=1}^m e_i > d$. Indeed, M is still a canonical monomial in $f|_{L=0}$ as L does not involve $x_1,\ldots,x_m$. Any other linear transformation has the form (after a possible rescaling) $L = x_i - (\sum_{j=i+1}^{k+1} \alpha_j x_j + \alpha_0)$, where $1 \leq i \leq m$. Given $\bar{\alpha} = (\alpha_{i+1},\ldots,\alpha_k,\alpha_0)$ consider the function $L_{\bar{\alpha},z}(x_i,\ldots,x_{k+1}) = x_i - (\sum_{j=i+1}^{k} \alpha_j x_j + z x_{k+1} + \alpha_0)$. Note, that L and $L_{\bar{\alpha},z}$ only differ in the coefficient of x_{k+1}. We will show that for any $\bar{\alpha}$ there is $\beta \in \mathbb{F}_q$ such that $\deg(f|_{L_{\bar{\alpha},\beta}=0}) > d$, which is sufficient to establish the claim. To ease notations and w.l.o.g., assume that $i = 1$. Namely, $L_{\bar{\alpha},z}(x_1,\ldots,x_{k+1}) = x_1 - (\sum_{j=2}^{k} \alpha_j x_j + z x_{k+1} + \alpha_0)$. Observe that the function $f|_{L_{\bar{\alpha},z}=0}$ has the same degree as $f(\sum_{j=2}^{k} \alpha_j x_j + z x_{k+1} + \alpha_0, x_2,\ldots,x_{k+1})$, when both are considered as polynomials in $x_2,\ldots,x_{k+1}$.

Let $\tilde{f}$ be the sum of all monomials, of maximal degree in f, that involve only the variables $x_1,\ldots,x_m$. Clearly M is such a monomial and therefore $\tilde{f}$ is not zero. Let $e_{\max}$ be the maximal degree of x_1 in $\tilde{f}$. As M is a max-monomial we have that $e_{\max} = e_1$. We can express $\tilde{f}$ as

$$\tilde{f} = x_1^{e_{\max}} \cdot h_{e_{\max}}(x_2,\ldots,x_m) + \sum_{e<e_{\max}} x_1^e \cdot h_e(x_2,\ldots,x_m),$$

where $h_{e_{\max}} \neq 0$. Let $\hat{f}$ be such that $f = \tilde{f} + \hat{f}$. Hence, $f(\sum_{j=2}^{k} \alpha_j x_j + z x_{k+1} + \alpha_0, x_2, \ldots, x_{k+1}) = \tilde{f}(\sum_{j=2}^{k} \alpha_j x_j + z x_{k+1} + \alpha_0, x_2, \ldots, x_{k+1}) + \hat{f}(\sum_{j=2}^{k} \alpha_j x_j + z x_{k+1} + \alpha_0, x_2, \ldots, x_{k+1})$. Consider all monomials of degree $\deg(f)$ in $f(\sum_{j=2}^{k} \alpha_j x_j + z x_{k+1} + \alpha_0, x_2, \ldots, x_{k+1})$ that have degree exactly $e_{\max}$ in both z and x_{k+1} and that only involve, besides z and x_{k+1}, the variables $x_2, \ldots, x_m$. Notice that the sum of those monomials is *exactly* $z^{e_{\max}} x_{k+1}^{e_{\max}} h_{e_{\max}}(x_2, \ldots, x_m)$. Furthermore,

$$\deg(x_{k+1}^{e_{\max}} h_{e_{\max}}(x_2, \ldots, x_m)) = \deg(x_1^{e_{\max}} h_{e_{\max}}(x_2, \ldots, x_m)) = \deg(\tilde{f}) = \deg(f).$$

Therefore, if we look at all monomials (in $x_2, \ldots, x_{k+1}$) of maximal degree in $f(\sum_{j=2}^{k} \alpha_j x_j + z x_{k+1} + \alpha_0, x_2, \ldots, x_{k+1})$, and think of their coefficients as polynomials in z, then at least one of those monomials, call it M', has a coefficient which is a non-zero polynomial in z. Hence, there is some value $\beta \in \mathbb{F}_q$ such that if we substitute $z = \beta$ then the coefficient of M' will not be zero. In particular $\deg(f|_{L_{\bar{\alpha},\beta}=0}) = \deg(f)$ as required. This completes the proof of this case.

Case $m = k + 1$. The analysis of this case is of a similar spirit to the previous case, only now we show that, with high probability, the degree cannot go down by too much. Again we consider $M = \prod_{i=1}^{k+1} x_i^{e_i}$. By the choice of A it follows that $e_1 \geq e_2 \geq \ldots \geq e_{k+1}$. For this case we will only focus on linear functions that are supported on x_{k+1}. Given $\bar{\alpha} = (\alpha_1, \ldots, \alpha_k)$ consider the linear function $L_{\bar{\alpha},z} = \sum_{i=1}^{k} \alpha_i x_i - x_{k+1} + z$ (we consider the case that the coefficient of x_{k+1} is -1, but the analysis of other cases is the same). Consider the coefficient of $\prod_{i=1}^{k} x_i^{e_i}$ in $f(x_1, \ldots, x_k, \sum_{i=1}^{k} \alpha_i x_i + z)$. It is not hard to see that this coefficient is a polynomial of degree e_{k+1} in z. Thus, there are at least $q - e_{k+1}$ values of z for which the coefficient of $\prod_{i=1}^{k} x_i^{e_i}$ in $f|_{L_{\bar{\alpha},z}=0}$ is nonzero. Thus, there are at least $q - e_{k+1}$ values of z for which $\deg(f|_{L_{\bar{\alpha},z}=0}) \geq e_1 + \ldots + e_k \geq k(q - q/p) \geq d + 1$. Thus the probability that $L_{\bar{\alpha},z}$ is 'good' is at least $\frac{q-1}{q} \cdot \frac{q-e_{k+1}}{q}$, where the first multiplicand comes from choosing a non-zero coefficient for x_{k+1} and the second comes from picking z. We consider two cases. If $e_{k+1} < q - 1$ then the probability is at least $\frac{q-1}{q} \cdot \frac{q-e_{k+1}}{q} \geq 2(q-1)/q^2 \geq 1/q$. On the other hand, if $e_{k+1} = q - 1$ then we also have $e_1 = \ldots = e_{k+1} = q - 1$ and thus $\deg(f) = (k+1)(q-1)$. In this case however, it is not hard to show, using similar arguments, that for any non-zero linear function $L = \sum_{i=1}^{k+1} \alpha_i x_i + z$ there is a choice of z such that $\deg(f|_{L=0}) = \deg(f) - (q-1) = k(q-1) \geq d+1$. Thus, in this case as well we get that with probability at least $1/q$ the function L is such that $\deg(F|_{L=0}) > d$.

This completes the proof of the lemma. □

We now use this lemma iteratively to 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 $\rho_d(f, k) \geq \rho_d(f, k') \cdot q^{-(k'-k)}$.*

Proof. The proof follows immediately from Lemma 4.6 by induction on n. For $n = k$ the result is trivial as $\deg(f) > d$ and hence $\rho_d(f, n) = 1$. So assume that $n \geq k+1$. Consider the following way for sampling a random k-dimensional flat. First we choose at random a hyperplane A and then we choose a random k-dimensional flat $B \subseteq A$. By Lemma 4.6 the probability that $f|_A$ has degree larger than d is at least $1/q$. Conditioning on $\deg(f|_A) > d$ we get by the induction hypothesis that $\Pr_B[\deg((f|_A)|_B) > d] \geq q^{-((n-1)-k)} = q \cdot q^{-(n-k)}$. Thus, $\rho_d(f, k) \geq \rho_d(f, n-1) \cdot (q \cdot q^{-(n-k)}) \geq q^{-(n-k)}$.

To prove the 'moreover' part we use a similar argument. Let A be a random k'-dimensional flat and let $B \subseteq A$ be a random k-flat. The probability that $f|_A$ has degree larger than d is exactly $\rho_d(f,k')$. Conditioning on this event, we get by the first part of the claim that $\rho_d(f|_A,k) \geq q^{-(k'-k)}$. Combining the two results we obtain $\rho_d(f,k) \geq \rho_d(f,k') \cdot q^{-(k'-k)}$. □

We now give the proof of Proposition 1.2.

Proof of Proposition 1.2. The fact that $t_{q,d} \leq \lceil (d+1)/(q-q/p) \rceil$ follows easily from Lemma 4.6. To see that $t_{q,d} \geq (d+1)/(q-q/p)$ we let t be such that $d+1 = (t-1)(q-q/p)+r$, where $0 < r \leq q-q/p$. Consider the function $f(x_1,\ldots,x_t) = \left(\prod_{i=1}^{t-1} x_i^{q-q/p}\right) \cdot x_t^r$. Observe that f has degree $(t-1)(q-q/p)+r = d+1$ but when we restrict f to any $(t-1)$-dimensional affine subspace its degree drops to at most $(t-1)(q-q/p) = d+1-r \leq d$ (it is not hard to check that the smallest decrease in degree is obtained for some substitution $x_t = \alpha$). Thus, the testing dimension is at least $t = (d+1-r)/(q-q/p)+1 \geq (d+1)/(q-q/p)$. Since t is an integer it follows that $t_{q,d} \geq t \geq \lceil (d+1)/(q-q/p) \rceil$. □

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

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

$$q^{t_{q,d}-1} \leq N_0(q,d) = q^{\left\lfloor \frac{d-q/p}{q-q/p} \right\rfloor + 1} < q^{t_{q,d-1}+1} \leq q^{t_{q,d}+1}.$$

As a first step we will 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 will 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.

Proof. Consider all monomials of degree d in f that involve x_1 and contain at most $t-1$ other variables. By our assumption, there is at least one such monomials. Let $e_{\max}$ be the maximal degree of x_1 in those monomials. W.l.o.g. assume that $M = x_1^{e_{\max}} \cdot \prod_{i=2}^t x_i^{e_i}$ is such a monomial in f. For a linear function $L(x_1\ldots,x_n) = x_1 + \sum_{i=2}^n \alpha_i x_i + \alpha_0$ denote $L_0(x_2\ldots,x_t) = -(\sum_{i=2}^t \alpha_i x_i)$ and $L_1(x_{t+1},\ldots,x_n) = -(\sum_{i=t+1}^n \alpha_i x_i + \alpha_0)$. Clearly, $L = x_1 - (L_0 + L_1)$. We would like to 'fix' L_0 and count how many different L_1 are there so that the degree of f decreases when we set $L=0$.

Consider the polynomial $g(x_1,\ldots,x_n) = f(x_1+L_0,x_2,\ldots,x_n)$. Notice that

$$g|_{x_1-L_1=0} = g(L_1,x_2,\ldots,x_n) = f(L_1+L_0,x_2,\ldots,x_n) = f|_{x_1=L_0+L_1} = f|_{L=0}.$$

Furthermore, observe, that M also appears in g (because it is of maximal degree in x_1 among all monomials with only t variables). We now express g as a polynomial in $x_2, \ldots, x_t$ with coefficients in $\mathbb{F}_q[x_1, x_{t+1}, \ldots, x_n]$. Namely,

$$g(x_1, \ldots, x_n) = \sum_{\bar{r} \in \{0,\ldots,q-1\}^{t-1}} \left(\prod_{i=2}^{t} x_i^{r_i} \right) \cdot g_{\bar{r}}(x_1, x_{t+1}, \ldots, x_n).$$

As L_1 does not involve any variable among $x_2, \ldots, x_t$ it holds that

$$\deg(g|_{x_1 - L_1 = 0}) < d \iff \deg(g|_{x_1 = L_1}) < d \iff \forall \bar{r} \quad \deg(g_{\bar{r}}|_{x_1 = L_1}) < d - \sum_{i=1}^{t} r_i.$$

Let $\bar{e} = (e_2 \ldots, e_t)$. Consider $g_{\bar{e}}$, recalling that the monomial $M = x_1^{e_{\max}} \cdot \prod_{i=2}^{t} x_i^{e_i}$ appears in g. In particular, $\deg(g_{\bar{e}}) = \deg(M) = e_{\max} \leq q - 1$. Thus, if $\deg(g|_{x_1 - L_1 = 0}) < d$ then it must be the case that

$$\deg(g_{\bar{e}}|_{x_1 - L_1 = 0}) < e_{\max} \leq q - 1.$$

Consider the homogeneous part of degree $e_{\max}$ of $g_{\bar{e}}$, denoted $g_{\bar{e}}^{(e_{\max})}$. It clearly contains $x_1^{e_{\max}}$ as a monomial. Observe further that $\deg(g_{\bar{e}}|_{x_1 = L_1}) < e_{\max} \iff \deg(g_{\bar{e}}^{(e_{\max})}|_{x_1 = L_1}) < e_{\max}$. However, since $g_{\bar{e}}^{(e_{\max})}$ is homogeneous of degree strictly smaller than q, this happens *if and only if* $g_{\bar{e}}^{(e_{\max})}|_{x_1 = L_1} = 0$. Indeed, substituting a linear function to a homogeneous polynomial of degree $D < q$ either makes it zero, or does not affect its degree. However, since $\deg(g_{\bar{e}}^{(e_{\max})}) \leq q - 1$, this means that, if we think of it as a polynomial in x_1 with coefficients in $\mathbb{F}_q[x_{t+1}, \ldots, x_n]$, then it has $L_1 \in \mathbb{F}_q[x_{t+1}, \ldots, x_n]$ as a root. In particular, there are at most $q - 1$ different L_1's that are roots of $g_{\bar{e}}^{(e_{\max})}$.

Concluding, we just proved that for every L_0 there are at most $q - 1$ different L_1's such that $\deg(f|_{x_1 - L_0 - L_1 = 0}) < d$. Hence, there are at most $(q-1) \cdot q^{t-1}$ different linear functions L involving x_1 such that $\deg(f|_{L=0}) < d$, as required. □

The following lemma extends the argument 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* $\deg(f|_{L=0}) \leq \deg(f) - 1$.

Proof. The proof is by induction on t. The case $t = 0$ is trivial. Assume that we proved it for $t - 1$ and let f be a degree d polynomial that contains a max-monomial with t variables. Assume w.l.o.g. that the monomial is $M = \prod_{i=1}^{t} x_i^{e_i}$. Lemma 4.8 implies that there are at most $(q-1) \cdot q^{t-1}$ linear functions L, involving x_1, such that $\deg(f|_{L=0}) < d$.

We now bound the number of linear functions that decrease the degree of f and that do not involve x_1. For that end, express f as a polynomial in x_1. $f = \sum_{e=0}^{q-1} x_1^e g_e(x_2, \ldots, x_n)$. As before, we have that $\deg(f|_{L=0}) < d$ if and only if $\forall 0 \leq e \leq q - 1$, $\deg(g_e|_{L=0}) < d - e$. In particular for $g = g_{e_1}$, where e_1 is the degree of x_1 in M, it must be the case that $\deg(g|_{L=0}) < d - e_1$.

At this point we use the fact that g_{e_1} has a max-monomial with only $t - 1$ variables, $M_1 = \prod_{i=2}^{t} x_i^{e_i}$, and conclude from the induction hypothesis that the number of linear functions L such that $\deg(g_{e_1}|_{L=0}) < \deg(g_{e_1})$ is at most q^{t-1}. Hence, overall there are at most $(q-1) \cdot q^{t-1} + q^{t-1} = q^t$ linear functions L such that $\deg(f|_{L=0}) < \deg(f)$. □

We are now ready to prove Theorem 1.5. For sake of readability we repeat it 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 there are at most $N_0(q,d) = q^{\left\lfloor \frac{d-q/p}{q-q/p} \right\rfloor + 1}$ linear functions L such that $\deg(f|_{L=0}) < d$.*

Proof. Notice that it is enough to prove the theorem for the polynomial $f \circ A$ where $A : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is an invertible linear transformation. Let A be the linear transformation guaranteed by Lemma 4.4. Namely, it is such that $f \circ A$ contains a canonical monomial. To simplify notations we assume from now on that f has a canonical monomial. Let $M = \prod_{i=1}^t x_i^{e_i}$ be some canonical monomial in f. Since M is a canonical monomial, it must be the case that $e_{t-1} + e_t \geq q$. Therefore, $d = \sum_{i=1}^t e_i = (e_1 + \ldots + e_{t-2}) + (e_{t-1} + e_t) \geq (t-2)(q - q/p) + q$ and hence, $t \leq \frac{d-q}{q-q/p} + 2 = \frac{d-q/p}{q-q/p} + 1$. Since t is an integer we actually get that $t \leq \left\lfloor \frac{d-q/p}{q-q/p} \right\rfloor + 1$. Invoking Lemma 4.9 we conclude that there are at most $q^t \leq q^{\left\lfloor \frac{d-q/p}{q-q/p} \right\rfloor + 1} = N_0(q,d)$ linear functions L such that $\deg(f|_{L=0}) < d$. □

Corollary 4.10. *Let n, d, q, K be integers such that $K > N_0(q,d)$. 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 $\deg f \leq d'$.*

Proof. Assume for contradiction that $d' < \deg(f) = \tilde{d} \leq d$. Then, by Theorem 1.5 there are *at most* $N_0(q,\tilde{d}) \leq N_0(q,d) < K$ hyperplanes A on which $\deg(f|_A) < \tilde{d}$. This contradicts our assumption that there are *at least* K hyperplanes $\{A_i\}$ on which $\deg(f|_{A_i}) \leq d'$. □

4.5 Interpolating from Exact Agreement

In this section we prove 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.

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*

$$K \geq N_1(q,d) = 2N_0(q, d+q) \cdot q^{\lambda_{q,4}} = 2q^{\left\lfloor \frac{d}{q-q/p} \right\rfloor + 2 + \lambda_{q,4}},$$

where $\lambda_{q,4}$ is the constant $\lambda_{q,c}$ from Theorem 3.4 for $c = 4$, then there exists a polynomial Q, of degree d, such that $Q|_{A_i} = P_i|_{A_i}$ for every $i \in [K]$.

Proof. In fact, we prove a slightly stronger statement. Specifically, we show that the conclusion holds when

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

This is indeed a stronger statement as the denominator above

$$2\prod_{i=1}^{n-\log_q N_1(q,d)-3}\left(1-\frac{N_1(q,d)}{q^{n-i-1}}\right) \geq 2\left(1-\sum_{i=1}^{n-\log_q N_1(q,d)-3}\frac{N_1(q,d)}{q^{n-i-1}}\right)$$

$$= 2-\frac{2N_1(q,d)}{q^{n-1}}\sum_{i=1}^{n-\log_q N_1(q,d)-3} q^i > 2-\frac{2N_1(q,d)}{q^{n-1}}q^{n-\log_q N_1(q,d)-2} = 2-2q^{-1} \geq 1,$$

namely, $\widetilde{N_1}(q,d,n) < N_1(q,d)$ for all n, and so the requirement on K is weaker.

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 and use Theorem 1.5 to claim that the resulting polynomial has degree 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 $n \leq \log_q N_1(q,d)+3$. For such n it holds that $\widetilde{N_1}(q,d,n) = \frac{1}{2}N_1(q,d)$. The analysis of this case, which is the technical heart of the proof, is given in the next lemma.

Lemma 4.11 (Main Lemma)**.** *Let $n \leq \log_q(N_1(q,d)/2)+3$ and $K \geq \widetilde{N_1}(q,d,n) = N_1(q,d)/2$. 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 defer the proof of Lemma 4.11 and continue with the proof of the theorem. 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

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

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 following lemma shows that we can find a hyperplane such that if we restrict our attention to any coset of that hyperplane, then the induction assumption continues to hold.

Claim 4.12. *There is linear function $L \in \mathrm{Aff}_q^n$ such that for every $\gamma \in \mathbb{F}_q$ the number of the distinct affine subspaces $A_{i,L,\gamma} \subseteq B_{L,\gamma}$, such that $A_{i,L,\gamma} \neq \emptyset$, is at least $\widetilde{N_1}(q,d,n-1)$.*

Note that this claim is not trivially true as different hyperplanes may have the same intersection with $B_{L,\gamma}$.

Proof. It is clearly sufficient to prove the claim for K such that $\widetilde{N_1}(q,d,n) \leq K \leq N_1(q,d)$. Observe that $A_i \cap B_{L,\gamma} = A_j \cap B_{L,\gamma}$, for linearly independent L_i and L_j, only if there are $\alpha,\beta \in \mathbb{F}_q^*$ such that $L = \alpha L_i + \beta L_j + \gamma$. Further, observe that $A_i \cap B_{L,\gamma} = \emptyset$ only if $L = L_i + \gamma'$, for some $\gamma' \in \mathbb{F}_q$.

Using these two observations we perform a simple counting argument that shows that there is some $L \in \mathrm{Aff}_q^n$ such that for every γ, the number of distinct $A_i \cap B_{L,\gamma}$, that are not empty, is as required.

Clearly, there are exactly q^{n+1} affine linear functions over $\mathbb{F}_q^n$. For each affine linear function L consider the number of ways that L can be represented as $L = \alpha L_1 + \beta L_2 + \gamma$ where[3] $\alpha, \beta, \gamma \in \mathbb{F}_q$ and $L_1, L_2 \in \mathcal{L}$. Since there are q^3K^2 such possible representations, there exists $L \in \mathrm{Aff}_q^n$ that can be represented in at most $\frac{q^3K^2}{q^{n+1}} = \frac{K^2}{q^{n-2}}$ different ways.

It follows, that for the L that we found and any $\gamma \in \mathbb{F}_q$, there are at least $K' = K - \frac{K^2}{q^{n-2}}$ different non empty flats of the form $A_i \cap B_{L,\gamma}$. Indeed, for every such representation of L we throw away one of the functions in the representation. As L cannot be represented using the remaining functions, we get the desired bound on K'. Calculating we get

$$\begin{aligned}
K' = K - \frac{K^2}{q^{n-2}} &= K\left(1 - \frac{K}{q^{n-2}}\right) \\
&\geq \widetilde{N_1}(q,d)\left(1 - \frac{N_1(q,d)}{q^{n-2}}\right) \\
&= \left(1 - \frac{N_1(q,d)}{q^{n-2}}\right) \frac{N_1(q,d)}{2\prod_{i=1}^{n-\log_q N_1(q,d)-3}\left(1 - \frac{N_1(q,d)}{q^{n-i-1}}\right)} \\
&= \frac{N_1(q,d)}{2\prod_{i=2}^{n-\log_q N_1(q,d)-3}\left(1 - \frac{N_1(q,d)}{q^{n-i-1}}\right)} \\
&= \frac{N_1(q,d)}{2\prod_{i=1}^{n-\log_q N_1(q,d)-4}\left(1 - \frac{N_1(q,d)}{q^{n-i-2}}\right)} \\
&= \widetilde{N_1}(q,d,n-1).
\end{aligned}$$

□

We proceed with the proof of Theorem 1.7. Let $L \in \mathrm{Aff}_q^n$ be as promised by Claim 4.12. Notice that L cannot be the constant function, as each constant function has at most $K^2 > \frac{K^2}{q^{n-2}}$ different representations. Fix $\gamma \in \mathbb{F}_q$ and let $A'_i = A_{i,L,\gamma} = A_i \cap B_{L,\gamma}$ and $P'_i = P_i|_{A'_i}$, for $i \in [K]$. It follows, by the choice of L, that the A'_i and P'_i satisfy the inductive assumption (as there are at least $\widetilde{N_1}(q,d,n-1)$ distinct A'_i). Hence, the induction hypothesis implies that there is a polynomial of degree d, $P_{L=\gamma}$, such that $P_{L=\gamma}|_{A'_i} = P'_i|_{A'_i}$ for every $i \in [K]$.

We are not done yet, as we may have a different polynomial for every $\gamma \in \mathbb{F}_q$. So now we show that by combining the different $P_{L=\gamma}$ we get a degree d polynomial P that is consistent with $P_1, ..., P_k$. Define

$$P(x) \triangleq \sum_{\gamma \in \mathbb{F}_q} \left(\prod_{\alpha \neq \gamma} \frac{L(x) - \alpha}{\gamma - \alpha} \right) \cdot P_{L=\gamma}(x) \ .$$

By construction, the degree of P is at most $d + q - 1$. It is easy to verify that for any $\gamma \in \mathbb{F}_q$, P agrees with $P_{L=\gamma}$ on $B_{L,\gamma} = \left\{ v \in \mathbb{F}_q^n \mid L(v) = \gamma \right\}$. As the hyperplanes $\{B_{L,\gamma}\}_{\gamma \in \mathbb{F}_q}$ cover all of

[3] We could have taken $\alpha, \beta \in \mathbb{F}_q^*$, but we use this counting to also include the case that L is a shift of some L_i.

$\mathbb{F}_q^n$, it follows that for every $i \in [K]$ and $u \in A_i$, $P(u) = P_i(u)$. Indeed, if we let $\gamma = L(u)$ then $P_i(u) =^{(*)} P_{L=\gamma}(u) = P(u)$, where (*) holds since, by the induction hypothesis, P_i and $P_{L=\gamma}$ agree on $A'_i = A_i \cap B_{L,\gamma}$.

We are still not done as we only showed that $\deg(P) \leq d + q - 1$. However, as

$$K \geq \widetilde{N_1}(q,d,n) = N_1(q,d)/2 > N_0(q,d+q),$$

Corollary 4.10 implies that the degree of P is, in fact, at most d. This complete the proof of Theorem 1.7 modulo the proof of Lemma 4.11 that we give next. □

Proof of Lemma 4.11 As before, we let $L_i \in \mathrm{Aff}_q^n$ be an affine linear function such that $A_i = \left\{u \in \mathbb{F}_q^n \mid L_i(u) = 0\right\}$ and denote $\mathcal{L} = \{L_1, \ldots, L_K\}$. Again we abuse notations and denote, for $L \in \mathcal{L}$, $P_L = P_i$ and $A_L = A_i$ when $L = L_i$.

We will first use the assumption that $n \leq \log_q(N_1(q,d)/2) + 3$ and $K \geq \widetilde{N_1}(q,d,n) = N_1(q,d)/2 = N_0(q,d+q) \cdot q^{\lambda_{q,4}}$ to show that the set $\mathcal{L}$ contains at least $\log_q(N_0(q,d+q))$ lines in linearly independent directions. Indeed, we can think of $\mathcal{L}$ as a set of points in Aff_q^n which is an $(n+1)$-dimensional space over $\mathbb{F}_q$. By our setting of parameters it follows that

$$\frac{|\mathcal{L}|}{|\mathrm{Aff}_q^n|} = \frac{K}{q^{n+1}} \geq \frac{K}{q^{\log_q(N_1(q,d)/2)+4}} = \frac{K}{N_1(q,d)/2} \cdot q^{-4} \geq q^{-4}.$$

Thus, in order to apply Corollary 3.5 we just need to bound $\dim(\mathrm{Aff}_q^n)$ from below. As we have K different hyperplanes over $\mathbb{F}_q^n$ it must be the case that $\log_q(K) \leq n+1$. Therefore,

$$\dim(\mathrm{Aff}_q^n) = n+1 \geq \log_q(K) \geq \left\lfloor \frac{d}{q - q/p} \right\rfloor + 2 + \lambda_{q,4} = \left\lfloor \frac{d+q-q/p}{q-q/p} \right\rfloor + 1 + \lambda_{q,4}.$$

Corollary 3.5 now implies that there are at least $\left\lfloor \frac{d+q-q/p}{q-q/p} \right\rfloor + 2$ combinatorial lines inside $\mathcal{L}$ whose directions are linearly independent. In particular, there are

$$t \geq \left\lfloor \frac{d+q-q/p}{q-q/p} \right\rfloor + 1 \tag{2}$$

such lines that their direction is not a constant linear function. By applying an invertible linear transformation, we can assume w.l.o.g. that those direction are the linear functions $x_1, ..., x_t$. I.e we can assume that there exist t linear functions $L_1, ..., L_t$ such that for any $i \in [t]$ and $\alpha \in \mathbb{F}_q$ the linear function $L_i - \alpha x_i$ belongs to $\mathcal{L}$. Intuitively, the line whose direction is x_1 is depicted in Figure 1 on page 7.

We will use these lines to construct a polynomial P, of degree d, that is consistent with $P_1, ..., P_K$.

The construction of P is done in three steps. First we construct, for every $i \in [t]$ and $\gamma \in \mathbb{F}_q^*$, a polynomial $P_{x_i=\gamma}$ which is defined on the hyperplane $B_{x_i,\gamma} \triangleq \left\{v \in \mathbb{F}_q^n \mid v_i = \gamma\right\}$ and is consistent with all the P_j's. In the second step we construct, for every $i \in [t]$, a polynomial $P_{x_i \neq 0}$, over the set $\cup_{\gamma \neq 0} B_{x_i,\gamma} = \left\{v \in \mathbb{F}_q^n \mid v_i \neq 0\right\}$, by a simple interpolation of $\left\{P_{x_i=\gamma} \mid \gamma \in \mathbb{F}_q^*\right\}$. The last step consists of combining the different $\{P_{x_i \neq 0}\}_{i \in [t]}$ to a single polynomial P.

Step 1 Fix $i \in [t]$ and $\gamma \in \mathbb{F}_q^*$. Denote

$$P_{x_i=\gamma} \triangleq \sum_{\beta \in \mathbb{F}_q} \left(\prod_{\alpha \neq \beta} \frac{L_i - \alpha}{\beta - \alpha} \right) \cdot P_{L_i - \gamma^{-1}\beta x_i} \ .$$

Clearly, P is a polynomial of degree at most $d+q-1$. We now show that $P_{x_i=\gamma}$ is a polynomial of degree at most d which is consistent with $\{P_1, \ldots, P_K\}$ on $B_{x_i,\gamma}$. Fix $j \in [K]$ and $u \in A_j \cap B_{x_i,\gamma}$. In particular, $u_i = \gamma$. Let $\beta' = L_i(u)$. We have

$$\begin{aligned} P_{x_i=\gamma}(u) &= \sum_{\beta \in \mathbb{F}_q} \left(\prod_{\alpha \neq \beta} \frac{L_i(u) - \alpha}{\beta - \alpha} \right) \cdot P_{L_i - \gamma^{-1}\beta x_i}(u) = \sum_{\beta \in \mathbb{F}_q} \left(\prod_{\alpha \neq \beta} \frac{\beta' - \alpha}{\beta - \alpha} \right) \cdot P_{L_i - \gamma^{-1}\beta x_i}(u) \\ &= \left(\prod_{\alpha \neq \beta'} \frac{\beta' - \alpha}{\beta' - \alpha} \right) \cdot P_{L_i - \gamma^{-1}\beta' x_i}(u) = P_{L_i - \gamma^{-1}\beta' x_i}(u) =^{(*)} P_j(u) \ , \end{aligned}$$

where $(*)$ follows from the fact that

$$L_i(u) - \gamma^{-1}\beta' u_i = L_i(u) - \gamma^{-1}\beta'\gamma = L_i(u) - \beta' = 0.$$

Indeed, this implies that $u \in A_{L_i(u) - \gamma^{-1}\beta' x_i}$ and now $(*)$ follows as $P_{L_i - \gamma^{-1}\beta' x_i}$ and P_j agree on $u \in A_j \cap A_{L_i - \gamma^{-1}\beta' x_i}$ (recall that for any $i \in [t]$ and $\alpha \in \mathbb{F}_q$ the linear function $L_i - \alpha x_i$ belongs to $\mathcal{L}$). To conclude, $P_{x_i=\gamma}$ is a degree $d+q-1$ polynomial that agrees with degree d polynomials on at least $K > N_0(q, d+q)$ flats. Corollary 4.10 now implies that $\deg(P_{x_i=\gamma}) \leq d$ on $B_{x_i,\gamma}$. The same argument also shows that $\{P_{x_i=\gamma}\}_{i \in [t], \gamma \in \mathbb{F}_q^*}$ are consistent with each other.

Step 2 Fix $i \in [t]$. Denote

$$P_{x_i \neq 0} \triangleq \sum_{\gamma \in \mathbb{F}_q^*} \left(\prod_{\alpha \in \mathbb{F}_q^* \setminus \{\gamma\}} \frac{x_i - \alpha}{\gamma - \alpha} \right) \cdot P_{x_i=\gamma} \ .$$

By construction, $P_{x_i \neq 0}$, is a polynomial of degree at most $d+q-2$ (recall, $\alpha \in \mathbb{F}_q^* \setminus \{\gamma\}$). It is not hard to verify that $P_{x_i \neq 0}$ is consistent with $P_1, ..., P_k$ on the set $\left\{v \in \mathbb{F}_q^n \mid v_i \neq 0\right\}$. Moreover, $\{P_{x_i \neq 0}\}_{i=1}^t$ are consistent with each other. I.e, for every $v = (v_1, ..., v_n) \in \mathbb{F}_q^n$ such that $v_i, v_j \neq 0$, the polynomials $P_{x_i \neq 0}$ and $P_{x_j \neq 0}$ satisfy $P_{x_i \neq 0}(v) = P_{x_j \neq 0}(v)$. Indeed, this follows immediately from the consistency of $\{P_{x_i=\gamma}\}_{i \in [t], \gamma \in \mathbb{F}_q^*}$ among themselves.

Step 3 This step is slightly more involved than the first two steps. Intuitively, we will show that if a monomial M appears in both $P_{x_i \neq 0}$ and $P_{x_j \neq 0}$ then it has the same coefficient in both. Hence, we can construct a unique polynomial P as the sum of all monomials, with the appropriate coefficients, that appear in any of the $P_{x_i \neq 0}$. While this is indeed the argument, for the proof we will need to work with slightly less natural basis for the space of polynomials.

For a degree $0 \leq e \leq q-1$ define

$$M_e(x_i) \triangleq \begin{cases} 1 & e = 0 \\ x_i^e & e \neq 0, q-1 \\ x_i^{q-1} - 1 & e = q-1 \end{cases} .$$

Notice that $M_0(x_i), \ldots, M_{q-1}(x_i)$ form a basis to the space of polynomials in x_i. For $\bar{e} = (e_1, \ldots, e_n)$, $0 \leq e_1, ..., e_n \leq q-1$, define the *$\bar{e}$-monomial*[4] $M_{\bar{e}}(x)$ to be

$$M_{\bar{e}}(x) \triangleq \prod_{i=1}^{n} M_{e_i}(x_i) .$$

Clearly, $\deg(M_{\bar{e}}) = \sum_{i=1}^{n} e_i$. We say that $M_{\bar{e}}$ is of full degree in x_i if $e_i = q-1$. As with the standard basis, it is not hard to see that every $f : \mathbb{F}_q^n \to \mathbb{F}_q$ has a unique representation as $f(x) = \sum c_{\bar{e}} M_{\bar{e}}(x)$, where $c_{\bar{e}} \in \mathbb{F}_q$. We will heavily rely on this simple fact in the rest of the proof. The next lemma gives some motivation for working with this less ordinary basis.

Lemma 4.13. *Let X be a set of variables. Denote $S_X = \left\{ v \in \mathbb{F}_q^n \mid \forall x_i \in X : v_i \neq 0 \right\}$. Let $g, h : \mathbb{F}_q^n \to \mathbb{F}_q$ be two polynomials that agree on S_X, namely, $\forall v \in S_X,\ g(v) = h(v)$. Then, the coefficient of any $\bar{e}$-monomial M that is not of full degree in any $x_i \in X$, is the same in both g and h.*

Proof. Consider $f = g - h$. Clearly, $f(v) = 0$ for all $v \in S_X$. We will show that when we represent f in our basis we have that $f = \sum_{x_i \in X} (x_i^{q-1} - 1) f_i$. The lemma will immediately follow by uniqueness of representation, as any monomial in $f = g - h$ has full degree in some $x_i \in X$.

The lemma follows from a standard counting argument. First, the number of functions that vanish on S_X is equal to the number of functions over $\mathbb{F}_q^n \setminus S_X$ which is

$$q^{|\mathbb{F}_q^n \setminus S_X|} = q^{\left|\left\{ v \in \mathbb{F}_q^n \mid \exists x_i \in X | v_i = 0 \right\}\right|}.$$

Secondly, let us count the number of polynomials of the form

$$\sum_{\bar{e}:\ \exists x_i \in X \text{ s.t. } e_i = q-1} c_{\bar{e}} M_{\bar{e}}(x).$$

This number is equal to $q^{|\{\bar{e} | \exists x_i \in X,\ e_i = q-1\}|}$. Clearly,

$$\# \left\{ v \in \mathbb{F}_q^n \mid \exists x_i \in X,\ v_i = 0 \right\} = \# \left\{ \bar{e} \in \{0, ..., q-1\}^n \mid \exists x_i \in X,\ e_i = q-1 \right\} .$$

Hence, the number of functions that vanish on S_X is exactly as the number of polynomials of the form $\sum_{\bar{e}: \exists x_i \in X, e_i = q-1} c_{\bar{e}} M_{\bar{e}}(x)$. Furthermore, any such polynomial $\sum_{\bar{e}: \exists x_i \in X, e_i = q-1} c_{\bar{e}} M_{\bar{e}}(x)$ vanishes on S_X. By uniqueness of representation it follows that any f that vanish on S_X is a polynomial of the form $\sum_{\bar{e}: \exists x_i \in X, e_i = q-1} c_{\bar{e}} M_{\bar{e}}(x)$. □

[4] We use $\bar{e}$-monomials to denote monomial in the new basis. Note, that in the standard basis, an $\bar{e}$-monomial may have more than one monomial.

We continue with the proof of Lemma 4.11. By uniqueness of representation, for any $m \in [t]$, $P_{x_m \neq 0}$ can be expressed as

$$P_{x_m \neq 0}(x) \triangleq \sum_{J \subseteq [t]} Q_J^m(x) \prod_{i \in J} (x_i^{q-1} - 1) \ ,$$

where, for any $J \subseteq [t]$ and $m \in [t]$, the polynomial Q_J^m contains only x_i's for $i \notin J$ and is not of full degree in any variable x_i, $i \in [t]$. Moreover, we note that $\deg(Q_J^m) \leq \deg(P_{x_m \neq 0}) - (q-1)|J|$. Our next goal is showing $Q_J^k = Q_J^m$ for any $k, m \in [t] \setminus J$.

Claim 4.14. *For every $k, m \notin J$ it holds that $Q_J^k = Q_J^m$*

Proof. Recall that $P_{x_k \neq 0}$ and $P_{x_m \neq 0}$ agree on $\left\{v \in \mathbb{F}_q^n \mid v_k, v_m \neq 0\right\}$. Lemma 4.13 implies that they have the same coefficient for any $\bar{e}$-monomial which is not of full degree in neither x_k nor x_m. I.e

$$\sum_{J \subseteq [t] \setminus \{k,m\}} Q_J^k \prod_{i \in J} (x_i^{q-1} - 1) = \sum_{J \subseteq [t] \setminus \{k,m\}} Q_J^m \prod_{i \in J} (x_i^{q-1} - 1).$$

The result now follows from uniqueness of representation. □

We continue with the proof of the main lemma. For every $J \subsetneq [t]$ define $Q_J = Q_J^m$, where $m \in [t] \setminus J$ is arbitrary. By Claim 4.14, Q_J is well define. Now we can define a polynomial P that is consistent with $\{P_1, \ldots, P_K\}$ on all of $\mathbb{F}_q^n$.

$$P \triangleq \sum_{J \subsetneq [t]} Q_J \prod_{i \in J} (x_i^{q-1} - 1).$$

We first show that $\deg(P) \leq d + q - 2$. Indeed, for $J \subsetneq [t]$ let $m \in [t] \setminus J$. Since $Q_J = Q_J^m$, it follows that

$$\deg\left(Q_J \prod_{i \in J} (x_i^{q-1} - 1)\right) = \deg\left(Q_J^m \prod_{i \in J} (x_i^{q-1} - 1)\right) \leq \deg(P_{x_m \neq 0}) \leq d + q - 2.$$

As this holds for every $J \subsetneq [t]$ we get that $\deg(P) \leq d + q - 2$. Later we will show that $\deg(P) = d$, but first we show that P is consistent with the P_i's.

Claim 4.15. *Every $k \in [K]$ and every $u \in A_k$ satisfy $P(u) = P_k(u)$.*

Proof. We will first prove the claim when for some $m \in [t]$, $u_m \neq 0$. For such u, $u_m^{q-1} - 1 = 0$. Therefore,

$$\begin{aligned} P(u) &= \sum_{J \subsetneq [t]} Q_J \prod_{i \in J} (u_i^{q-1} - 1) = \sum_{J \subseteq [t] \setminus \{m\}} Q_J \prod_{i \in J} (u_i^{q-1} - 1) = \sum_{J \subseteq [t] \setminus \{m\}} Q_J^m \prod_{i \in J} (u_i^{q-1} - 1) \\ &= \sum_{J \subseteq [t]} Q_J^m \prod_{i \in J} (u_i^{q-1} - 1) = P_{x_m \neq 0}(u) = P_k(u) \ , \end{aligned}$$

where in the last equality we used the consistency of $P_{x_m \neq 0}$ and P_k on A_k. It remains to show that $P(u) = P_k(u)$ for u such that $(u_1, ..., u_t) = (0, ..., 0)$. Assume for a contradiction that this is not the

case. I.e that there is $v \in \mathbb{F}_q^{[n]\setminus[t]}$ such that $P(0,v) \neq P_k(0,v)$. Denote $\alpha = P(0,v) - P_k(0,v) \neq 0$. We have that

$$(P - P_k)(x,v) = \begin{cases} 0 & x \neq (0,...,0) \\ \alpha & x = (0,...,0) \end{cases} .$$

Hence, as a polynomial in $x_1, \ldots, x_t$,

$$(P - P_k)(x,v) = \alpha \prod_{i \in t} \left(1 - x_i^{q-1}\right) .$$

Therefore,

$$\begin{aligned} \deg(P - P_k)(x,v) = (q-1)t \quad &\geq^{(*)} \quad (q-1) \cdot \left(\left\lfloor \frac{(d+q) - q/p}{q - q/p} \right\rfloor + 1 \right) \\ &\geq \quad (q-1) \cdot \left(\frac{d}{q - q/p} + 1 \right) \geq d + q - 1, \end{aligned}$$

where $(*)$ follows from Equation (2). On the other hand,

$$\deg(P - P_k)(x,v) \leq \deg(P - P_k) \leq \max\{\deg(P), \deg(P_k)\} \leq d + q - 2$$

which is a contradiction. We thus conclude that for every $k \in [K]$ and $u \in A_k$, $P(u) = P_k(u)$. □

We finish the proof of Lemma 4.11 by the following observation. P is a polynomial of degree at most $d+q-2$ that is equal to degree d polynomials on at least $K > N_0(q, d+q)$ hyperplanes. So, by Corollary 4.10, $\deg(P) \leq d$ as required. □

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 is essentially from [BKS+10]; we merely verify it extends to general q (using our bounds on $N_1(q.d)$). As a result, the description is terse and we skip the proof development.

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 $\delta_d(f) \leq 2\delta_1 + 4(q-1)/K$.*

Proof. We prove the theorem in four steps. Let P_i, defined on A_i, denote the polynomial (which, by Lemma 3.2, is unique on A_i) of degree at most d that satisfies $\delta(f|_{A_i}, P_i) \leq \delta_1$.

First, we claim that for every pair of hyperplanes A_i and A_j, we have $P_i|_{A_i \cap A_j} = P_j|_{A_i \cap A_j}$. If A_i and A_j are parallel, then there is nothing to prove. Else note that $|A_i \cap A_j| = \frac{1}{q}|A_i|$ and so $\delta(f|_{A_i \cap A_j}, P_i|_{A_i \cap A_j}) \leq q\delta_1$. Similarly, $\delta(f|_{A_i \cap A_j}, P_j|_{A_i \cap A_j}) \leq q\delta_1$. We conclude that $\delta(P_i|_{A_i \cap A_j}, P_j|_{A_i \cap A_j}) \leq 2q\delta_1 < q^{-d/(q-1)}$. But since both P_i and P_j are degree d polynomials on $A_i \cap A_j$, they must be identical if their distance is so small (by Lemma 3.2).

Next, we use Theorem 1.7, to claim that there is a degree d polynomial Q that agrees with all the given P_i's. Specifically, we have $Q|_{A_i} = P_i|_{A_i}$ for every $i \in [K]$. Note that to use Theorem 1.7, we need $K \geq N_1(q,d)$, which is true from our hypothesis.

The third claim we make is that there is a large fraction of points that are contained in a noticeable fraction of the K hyperplanes. Specifically, if we say that $x \in \mathbb{F}_q^n$ is *bad* if $|\{i \in [K] \mid x \in A_i\}| \leq K/(2q)$, then the probability that a uniformly chosen $x \in \mathbb{F}_q^n$ is bad is at most $\tau = 4(q-1)/K$. To prove this claim, let $x \in \mathbb{F}_q^n$ be chose uniformly at random and let Y_i be the indicator random variable that is 1 if $x \in A_i$ and 0 otherwise. Note that we need to show that the probability that $\sum_i Y_i \leq K/(2q)$ is at most $4(q-1)/K$. Let $Z_i = Y_i - \mathrm{Exp}[Y_i] = Y_i - 1/q$. Clearly, $\mathrm{Exp}[Z_i^2] = \mathrm{Exp}[Y_i^2] - \mathrm{Exp}[Y_i]^2 = 1/q - 1/q^2$. Furthermore, the expectation of $Y_i \cdot Y_j \leq 1/q^2$ (it is zero if the hyperplanes are parallel and $1/q^2$ otherwise). Thus we have $\mathrm{Exp}[Z_i \cdot Z_j] \leq 0$, and so $\mathrm{Exp}[(\sum_{i\in[K]} Z_i)^2] \leq \sum_i \mathrm{Exp}[Z_i^2] = K(q-1)/q^2$. We thus conclude that

$$\begin{aligned}
\Pr\left[\sum_i Y_i \leq K/(2q)\right] &= \Pr\left[\sum_i Z_i \leq -K/(2q)\right] \\
&\leq \Pr\left[(\sum_i Z_i)^2 \geq K^2/(2q)^2\right] \\
&\leq \frac{4q^2}{K^2} \cdot \frac{K(q-1)}{q^2} \leq \frac{4(q-1)}{K}.
\end{aligned}$$

Finally, we claim that $\delta(f,Q)$ can be bounded by $\tau + 2\delta_1$. To see this, we consider the following experiment: Pick $x \in \mathbb{F}_q^n$ and $i \in [K]$ uniformly and independently and consider the event that "$x \in A_i$ and $f(x) \neq P_i(x)$". On the one hand we have this event happens with probability at most δ_1/q, since probability $x \in A_i$ is exactly $1/q$ and $Pr_{x\in A_i}[f(x) \neq P_i(x)] \leq \delta_1$. On the other hand, this probability can also be seen to be at least $(\delta(f,Q) - \tau)/(2q)$, since the probability that x is not bad and satisfies $f(x) \neq Q(x)$ is at least $\delta(f,Q) - \tau$ and for every x that is not bad, the probability that $A_i \ni x$ for random i is at least $1/(2q)$. The upper bound $\delta(f,Q) \leq 2\delta_1 + \tau$ follows immediately.

Putting the above claims together we get that if $K \geq N_1(q,d)$ and $\delta_1 < \frac{1}{2}q^{-(1+d/(q-1))}$, then $\delta_d(f) \leq 2\delta_1 + 4(q-1)/K$. □

5 Analysis of the low-degree tests

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 $\rho_d(f,t) \geq \min\{\frac{1}{4q}, \frac{1}{2} \cdot q^t \cdot \delta_d(f)\}$.*

Proof. We will use the monotonicity of the rejection probability $\rho_d(f,\cdot)$ (Lemma 4.5) and give a lower bound on the rejection probability $\rho_d(f,\ell)$ for some $\ell \leq t$.

Let $\delta = \delta_d(f)$ and let g be a polynomial of degree at most d satisfying $\delta(f,g) = \delta$.

For every integer ℓ, $d/(q-1) \leq \ell \leq t$, we claim that the probability that on a randomly chosen ℓ-dimensional affine subspace A, $f|_A$ and $g|_A$ disagree on exactly one point is at least $q^\ell \cdot \delta \cdot (1 - (q^\ell - 1) \cdot$

δ). Indeed, the argument is quite routine so we only sketch it. Let x be a point on the ℓ-dimensional flat A. Consider the event that $f(x) \neq g(x)$ but $f(y) = g(y)$ for any other $y \in A$. Clearly its probability is at least $\Pr[f(x) \neq g(x)] - \sum_{y \in A, y \neq x} \Pr[f(y) \neq g(y) \text{ and } f(x) \neq g(x)] = \delta - (q^\ell - 1)\delta^2$, where we have used the fact that the points in A are pairwise independent. Thus, taking the union bound over all $x \in A$ we get that the probability that f and g disagree in exactly one point is at least $q^\ell \cdot \delta \cdot (1 - (q^\ell - 1) \cdot \delta)$.

The fact above allows us to analyze $\rho_d(f, \ell)$ as follows: Since the ℓ-dimensional test rejects whenever it picks an A where $f|_A$ and $g|_A$ disagree on exactly one point, we conclude that $\rho_d(f, \ell) \geq q^\ell \cdot \delta \cdot (1 - (q^\ell - 1) \cdot \delta)$.

Now if $\delta \leq \frac{1}{2} q^{-t}$, then we immediately get $\rho_d(f, t) \geq \frac{1}{2} \cdot q^t \cdot \delta$. Else, let ℓ be the largest integer such that $\delta \leq \frac{1}{2} q^{-\ell}$ (and so $\delta > \frac{1}{2q} q^{-\ell}$). We then get $\rho_d(f, t) \geq \rho_d(f, \ell) \geq^{(*)} \frac{1}{2} \cdot q^\ell \cdot \delta > \frac{1}{4q}$ as desired, where $(*)$ follows by the previous argument. □

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 $\rho_d(f, t) \geq \epsilon + \frac{1}{8} q^t \cdot \sum_{i=n+1}^{\infty} q^{-i}$.*

Proof. We prove the lemma for $\epsilon = \frac{1}{32q}$ and $c = \log_q N_1(q, d) - t_{q,d} + \log_q 128$. Recall by Theorem 1.7 that

$$N_1(q, d) = 2N_0(q, d + q) \cdot q^{\lambda_{q,4}} = 2q^{\left\lfloor \frac{d}{q - q/p} \right\rfloor + 2 + \lambda_{q,4}} \geq 2q^{t_{q,d} + 1 + \lambda_{q,4}},$$

where $\lambda_{q,4}$ is defined in Theorem 3.4, and so c is indeed bounded independent of d.

The proof is by induction on n. For the base case, we use $n = t$. In this case note that $\rho_d(f, t) = 1$ and $\sum_{i=t+1}^{\infty} q^{-i} = q^{-t}/(q-1)$ and so $\frac{1}{8} \cdot q^t \cdot \sum_{i=t+1}^{\infty} q^{-i} < \frac{1}{2}$ and so the lemma holds for every $\epsilon \leq \frac{1}{2}$.

We now move to the inductive case. Let $A_1, \ldots, A_K$ be all the distinct hyperplanes for which $\delta_d(f|_{A_i}) < q^{-t}$. If K is small, then we are easily done by induction since $\rho_d(f, t) = \mathrm{Exp}_A[\rho_d(f|_A, t)]$ and the inductive hypothesis says that $\rho_d(f|_A, t)$ is usually large. When K is large, we use Theorem 4.16 to show that $\delta_d(f)$ is small, and this allows us to use Lemma 5.1 to claim $\rho_d(f, t)$ is large in this case also. Details below.

Case 1: $K < \frac{1}{8} q^t$. For a hyperplane A such that $\delta_d(f|_A) \geq q^{-t}$ we have, by the induction hypothesis, $\rho_d(f|_A, t) \geq \epsilon + \frac{1}{8} q^t \sum_{i=n}^{\infty} q^{-i}$. Using the fact that the number of hyperplanes in $\mathbb{F}_q^n$ is at least q^n, we get that $\Pr_A[\delta_d(f|_A) < q^{-t}] \leq \frac{1}{8} q^t / q^n$. Combining the two we get

$$\begin{aligned}
\rho_d(f, t) &= \mathrm{Exp}_A(\rho_d(f|_A, t)) \\
&\geq \epsilon + \frac{1}{8} q^t \sum_{i=n}^{\infty} q^{-i} - \frac{1}{8} q^t / q^n \\
&= \epsilon + \frac{1}{8} q^t \sum_{i=n+1}^{\infty} q^{-i}
\end{aligned}$$

as desired.

Case 2: $K \geq \frac{1}{8} q^t$. Note that

$$K \geq \frac{1}{8} q^t \geq \frac{1}{8} q^{t_{q,d} + c} = \frac{1}{8} q^{\log_q N_1(q,d) + \log_q 128} > N_1(q, d).$$

We thus have by Theorem 4.16, $\delta_d(f) \leq 2q^{-t} + 4(q-1)/K$. Using $2q^{-t} \leq \frac{1}{4}q^{-d/(q-1)}$ and $4(q-1)/K \leq 32 \cdot q^{-t+1} \leq \frac{1}{4}q^{-d/(q-1)}$ (by our choice of $t \geq \log_q N_1(q,d) + \log_q 128 \geq d/(q-1) + \log_q 128 + 1$) we conclude in this case that $\delta_d(f) \leq \frac{1}{2}q^{-d/(q-1)}$. This allows us to use Lemma 5.1 in this case and conclude that $\rho_d(f) \geq \min\{\frac{1}{4q}, \frac{1}{2} \cdot q^t \cdot \delta_d(f)\}$. It now follows from the choice of parameters that $\frac{1}{4q} \geq \epsilon + \frac{1}{8}q^t \sum_{i=n+1}^{\infty} q^{-i}$ and $\frac{1}{2} \cdot q^t \cdot \delta_d(f) \geq \frac{1}{2} \geq \epsilon + \frac{1}{8}q^t \sum_{i=n+1}^{\infty} q^{-i}$. □

We now give the proof of our main theorem.

Proof of Theorem 1.3 We analyze two cases depending on $\delta_d(f)$.

1. $\delta_d(f) \leq \frac{1}{2}q^{-d/(q-1)}$: Since $t_{q,d} = \lceil (d+1)/(q - q/p) \rceil \geq d/(q-1)$ we get from Lemma 5.1 that $\rho_d(f, t_{q,d}) \geq \min\{\frac{1}{4q}, \frac{1}{2} \cdot q^{t_{q,d}} \cdot \delta_d(f)\}$.

2. $\delta_d(f) > \frac{1}{2}q^{-d/(q-1)}$: In this case we can apply Lemma 5.2 and conclude that there exists constants $c, \epsilon > 0$ such that for $t = t_{q,d} + c$ it holds that $\rho_d(f,t) \geq \epsilon + \frac{1}{8}q^t \cdot \sum_{i=n+1}^{\infty} q^{-i} > \epsilon$. Applying Lemma 4.7 we obtain that

$$\rho_d(f, t_{q,d}) \geq \rho_d(f,t) \cdot q^{-(t - t_{q,d})} \geq \epsilon \cdot q^{-c}.$$

Set $\epsilon_1 = 1/2$ and $\epsilon_2 = \min\{\frac{1}{4q}, \epsilon \cdot q^{-c}\}$. Note that, by Lemma 5.2, ϵ_2 depends only on q. Combining the two cases we conclude that

$$\rho_d(f, t_{q,d}) \geq \min\{\epsilon_2, \epsilon_1 \cdot q^{t_{q,d}} \cdot \delta_d(f)\}$$

as claimed. □

References

[AKK+05] 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.

[BKS+10] A. Bhattacharyya, S. Kopparty, G. Schoenebeck, M. Sudan, and D. Zuckerman. Optimal testing of reed-muller codes. In *Proceedings of the 51th Annual FOCS*, pages 488–497, 2010.

[DK00] P. Ding and J. D. Key. Minimum-weight codewords as generators of generalized reed-muller codes. *IEEE Transactions on Information Theory*, 46(6):2152–2158, 2000.

[FK91] H. Furstenberg and Y. Katznelson. A density version of the Hales-Jewett theorem. *J. d'Analyse Math.*, 57:64–119, 1991.

[JPRZ04] 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*, pages 423–432, 2004.

[KR06] T. Kaufman and D. Ron. Testing polynomials over general fields. *SIAM J. Comput.*, 36(3):779–802, 2006.

[KS08] T. Kaufman and M. Sudan. Algebraic property testing: the role of invariance. In *Proceedings of the 40th Annual STOC*, pages 403–412, 2008.

[Pol09] D. H. J. Polymath. A new proof of the density Hales-Jewett theorem. *CoRR*, arxiv.org/abs/0910.3926, 2009.

[RS96] R. Rubinfeld and M. Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM J. Comput.*, 25(2):252–271, 1996.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de