Polynomial Testing and Related Questions

Research Thesis

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Elad Haramaty

Submitted to the Senate of Technion - Israel Institute of Technology

Av, 5774 Haifa August, 2014
It is a great pleasure to thank my thesis advisor Prof. Amir Shpilka for his devoted guidance, encouragement, and good advice.

Special thanks are due to Madhu Sudan, Keren Censor-Hillel and Noga Ron-Zewi for their continuous help and support throughout the course of my Ph.D.

I would also like to thank my fellow students, staff members, professors, secretaries and others that helped me academically, administratively and mentally. They made the Technion feel like home and their help is greatly appreciated.

On a personal level, I want to thank my good friends and my loving family that made this period in life as amazing as it was. I cannot imagine this period without them. Last but definitely not least I thank my girlfriend Yael for her love and devotion, and her willingness to support me in the future in any road I may take.

The generous financial help of the Technion is gratefully acknowledged.
List of Publications


Contents

Abstract

Abbreviations and Notations

1 Introduction

1.1 The t-dimensional test ................................... 1
1.2 Structure of functions that pass the test .................... 2
  1.2.1 Acceptance probability 1 − O(q^t) ....................... 3
  1.2.2 Acceptance probability 1 − O(1) ......................... 3
  1.2.3 Non-negligible acceptance probability ................... 3
1.3 Lifted codes .......................................... 5

2 On the Structure of Cubic and Quartic Polynomials

2.1 Introduction .......................................... 6
  2.1.1 Our results ...................................... 8
  2.1.2 Proof Technique ................................... 10
  2.1.3 Organization ..................................... 11
2.2 Preliminaries ....................................... 11
  2.2.1 Subadditive functions ............................... 12
2.3 The structure of cubic polynomials ......................... 14
  2.3.1 Restricting the polynomial to a ‘good’ subspace ........ 14
  2.3.2 The structure of low rank spaces .................... 15
  2.3.3 Completing the proofs ................................ 16
2.4 The structure of biased 4 degree polynomials ............. 18
  2.4.1 Restricting the polynomial to a ‘good’ subspace ....... 18
  2.4.2 Computing \( \tilde{f} \) using a few quadratics .......... 19
  2.4.3 The structure of \( f \) .................................. 22
2.5 Quartic polynomials with high \( U^4 \) norm ............... 24
  2.5.1 The case of the symmetric polynomial ............... 24
  2.5.2 Finding a ‘basis’ for a space of low rank cubic polynomials .... 25
  2.5.3 Completing the proof ................................ 29
2.6 Conclusions ......................................... 31

3 Optimal testing of multivariate polynomials over small prime fields

3.1 Introduction ......................................... 32
  3.1.1 Our main results .................................. 33
  3.1.2 Comparison to [BKS+10] ................................ 35
3.2 Overview of our proof ................................ 36
3.3 Preliminaries ....................................................... 39
  3.3.1 The Distance Between Polynomials ......................... 39
  3.3.2 Density Hales-Jewett Theorem ............................. 40
3.4 Restrictions to Hyperplanes ...................................... 41
  3.4.1 Canonical Monomials ........................................ 41
  3.4.2 Monotonicity ................................................ 43
  3.4.3 Relating Different Dimensions .............................. 43
  3.4.4 The Case of Polynomials of Degree $d + 1$ .................. 45
  3.4.5 Interpolating from Exact Agreement ....................... 47
  3.4.6 Interpolating from Approximate Agreement ................ 54
3.5 Analysis of the low-degree tests ................................ 55
4 Absolutely Sound Testing of Lifted Codes ....................... 57
  4.1 Introduction ................................................... 57
    4.1.1 Our work: motivation and results ....................... 58
  4.2 Overview of Proof ............................................ 61
    4.2.1 Some natural tests ....................................... 61
    4.2.2 Overview of proof of Main Theorem 4.1.1 ................. 61
    4.2.3 Overview of proof of Theorem 4.2.1 ..................... 62
  4.3 Background and preliminary material ......................... 63
    4.3.1 The structure of affine-invariant linear codes ......... 64
    4.3.2 Lifts of affine-invariant linear codes .................. 64
  4.4 Proof of Main Technical Theorem 4.2.1 ...................... 65
    4.4.1 The code $\mathcal{F}^+$ .................................. 65
    4.4.2 Restrictions of functions in $\mathcal{F}^+$ to hyperplanes
    4.4.3 Restrictions of general functions to hyperplanes ........ 72
  4.5 Proof of Main Theorem 4.1.1, $Q = q$ case .................. 79
  4.6 Proof of Main Theorem 4.1.1, general $Q$ .................... 81
  4.7 New testable codes ........................................... 82
5 On $r$-Simple $k$-Path ........................................... 84
  5.1 Introduction .................................................. 84
    5.1.1 Our results .............................................. 84
    5.1.2 Finding a path with many distinct vertices ............. 85
  5.2 Overview of the proof of Theorem 5.1.2 ....................... 85
  5.3 Definitions and Preliminary Results .......................... 86
  5.4 Gap .......................................................... 87
  5.5 From $r$-Simple $k$-Path to Multivariate Polynomial ......... 88
  5.6 Low Degree Tester ............................................ 88
  5.7 Testing if $P_G$ contains an $r$-monomial ..................... 89
  5.8 Proof of Theorem 5.3.1 ...................................... 92
  5.9 Proof of Lemma 5.6.2 ....................................... 93
6 Summary .......................................................... 96
Abstract

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

Previous works showed that this natural test rejected functions that were $\Omega(1)$-far from degree $d$-polynomials with probability at least $\Omega(q^{-t})$. Thus to get a constant probability of detecting functions that are at constant distance from the space of degree $d$ polynomials, the tests made $q^{2t}$ queries. It is also known that when $q$ is prime, then $q^t$ queries are necessary. Thus these tests were off by at least a quadratic factor from known lower bounds. Bhattacharyya et al. (FOCS 10) gave an optimal analysis of this test for the case of the binary field and showed that the natural test actually rejects functions that were $\Omega(1)$-far from degree $d$-polynomials with probability $\Omega(1)$.

In this work we extend this result for all fields showing that the natural test does indeed reject functions that are $\Omega(1)$-far from degree $d$ polynomials with $\Omega(1)$-probability, where the constants depend only on $q$ the field size. Thus our analysis thus shows that this test is optimal (matches known lower bounds) when $q$ is prime.

Moreover, we extend this result to a more general notion called lifted codes. Given a base code over a $t$-dimensional space, its $n$-dimensional lift consists of all words whose restriction to every $t$-dimensional affine subspace is a codeword of the base code. Lifting not only captures the most familiar codes, which can be expressed as lifts of low-degree polynomials, it also yields new codes when lifting medium-degree polynomials whose rate is better than that of corresponding polynomial codes, and all other combinatorial qualities are no worse. We shows that the natural test for those codes also has constant rejection probability for functions that are $\Omega(1)$-far from the code.

We also study functions that pass a similar test, called Gowers Norm, with non-negligible probability. We present structural results for such polynomials with noticeable Gower norm, showing that they can be represented as "nice" function of lower degree polynomials.

We conclude this thesis with an application, showing that by using our low degree tester one can detect efficiently simple paths in a graphs.
Abbreviations and Notations

\[ \mathbb{N} \] — The set of natural numbers
\[ [n] \] — The set of natural numbers \{1,\ldots,n\}
\[ \mathbb{F} \] — An algebraic field
\[ \mathbb{F}^n \] — The \( n \) dimensional vector space over an algebraic field \( \mathbb{F} \)
\[ f|_S \] — The restriction of the function \( f \) to the set \( S \)
\[ \mathcal{P}(n,d,q) \] — The set of all \( n \)-variate polynomial functions over \( \mathbb{F}_q \) of total degree at most \( d \)
\[ \delta(f,g) \] — The normalised Hamming distance between \( f \) and \( g \), i.e. \( \Pr[x \mid f(x) \neq g(x)] \)
\[ \deg(f) \] — The minimal degree of polynomial that compute function \( f \)
\[ \dim(A) \] — The dimension of the affine subspace \( A \)
Chapter 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 thesis we focus on this basic problem in the setting of degree $d$ multivariate polynomials over fields of constant size, as considered before in [AKK+05, KR06, JPRZ09, BKS+10, RZS13]. We will present the state of the art (and optimal over prime order fields) test for this problem. Additionally, we will generalise this result for general lifted code, as we will describe later. Moreover, we will conclude with an application, showing that by using the low degree tester, one can efficiently detect simple paths in graphs (and even solve an interesting generalisation of this problem called $r$-simple path).

In order to describe the problem more precisely we introduce some basic notations. For an integer $t$, let $[t]$ denote the set $\{1, \ldots, t\}$. Let $\mathbb{F}_q$ denote the finite field of cardinality $q$. We consider the task of testing functions mapping $\mathbb{F}_q^n$ to $\mathbb{F}_q$. Let $\mathcal{P}(n,d,q)$ denote the set of all $n$-variate polynomial functions over $\mathbb{F}_q$ of total degree at most $d$. We let $\delta(f,g) = \Pr_x[f(x) \neq g(x)]$ denote the distance between $f$ and $g$, where the probability is over $x$ chosen uniformly at random from $\mathbb{F}_q^n$. Let $\delta_d(f) = \min_{g \in \mathcal{P}(n,d,q)} \{\delta(f,g)\}$ denote the distance of $f$ from the space of degree $d$ polynomials. We say $f$ is $\delta$-far from $g$ if $\delta(f,g) \geq \delta$ and $\delta$-close otherwise. We say $f$ is $\delta$-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 between the case where $\delta_d(f)$ is zero and 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 $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$ and accepts $f \in \mathcal{P}(n,d,q)$ with probability one. It has $\delta$-soundness $\epsilon$ if it rejects every function $f$ with $\delta_d(f) \geq \delta$ with probability at least $\epsilon$. 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 $\delta$-soundness $\epsilon$.

1.1 The $t$-dimensional test

The state of the art test, was introduced by [AKK+05, KR06]. To describe more formally the test and our theorem we require few more notations. For an affine subspace $A$ in $\mathbb{F}_q^n$, let $\dim(A)$ denote its dimension. For a function $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$ and affine subspace $A$, let $f|_A : A \rightarrow \mathbb{F}_q$ denote the restriction of $f$ to $A$. For a function $f$, we let $\deg(f)$ denote its degree as a polynomial. We use the fact that $f|_A$ can be viewed as a $\dim(A)$-variate polynomial with $\deg(f|_A) \leq \deg(f)$. A special subclass of tests for $\mathcal{P}(n,d,q)$ would uniformly pick a random affine subspace $A$ of $\mathbb{F}_q^n$ and verify that $\deg(f|_A) \leq d$.1

1 over general prime order fields we will consider similar test called the Gowers norm.
We introduce the concept of testing dimension which attempts to explore the minimal dimension for which such a test has positive soundness.

**Definition 1.1.1 (Testing dimension).** For prime power \( q \) and non-negative \( d \), the testing dimension of polynomials of degree \( d \) over \( \mathbb{F}_q \) is the smallest integer \( t \) satisfying the following: For every positive integer \( n \) and every function \( f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q \) 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.

**Proposition 1.1.2.** The testing dimension \( t_{q,d} = \left\lceil \frac{d+1}{q-1/q} \right\rceil \), where \( p \) is a prime number such that \( q \) is a power of \( p \).

As it also follows easily from our results we prove it in Section 3.4.3. To ease notations, when \( q \) and \( d \) are clear from the context we will use \( t \) to describe the testing dimension \( t_{q,d} \).

The test proposed by [KR06] is the following:

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

Over \( \mathbb{F}_2 \) this test can be viewed by the notion of discrete partial derivatives, described as follows.

**Definition 1.1.3.** (Discrete partial derivative) For a function \( f : \mathbb{F}^n \rightarrow \mathbb{F} \) and a direction \( y \in \mathbb{F}^n \) we define \( \Delta_y f(x) = f(x+y) - f(x) \) to be the discrete partial derivative of \( f \) in direction \( y \) at point \( x \).

It is not difficult to see that if \( \deg(f) = d \) then for every \( y \) \( \deg(\Delta_y f) \leq d - 1 \). So one can view the \( t \)-dimensional test over \( \mathbb{F}_2 \) as verifying that a random \( d \)th derivative in a random point is zero. The difference between the acceptance and the rejection of the test over \( \mathbb{F}_2 \) can be viewed as

\[
|\mathbb{E}_{x,y_1,...,y_d}[(-1)^{\Delta_{y_d}...\Delta_{y_1}(f)(x)}]|.
\]

More generally, we define the notion of Gowers norm over general prime order field.

**Definition 1.1.4 (Gowers norm over a field of prime order [AKK+05, Gow98, Gow01]).** The \( d+1 \)-th Gowers norm, \( U^{d+1} \), of a function \( f \) over a field of prime order \( \mathbb{F} \) is defined as

\[
\|f\|_{U^{d+1}} = \|\mathbb{E}_{x,y_1,...,y_{d+1}}[\omega^{\Delta_{y_{d+1}}...\Delta_{y_1}(f)(x)}]|^{1/2^{d+1}},
\]

where \( \omega = e^{2\pi i / \mathbb{F}} \).

In addition to the properties of the original test, we will also be interested in properties of Gowers norm, for general prime ordered fields.

### 1.2 Structure of functions that pass the test

This test is complete since any restriction of a degree \( d \) function to a subspace has a degree of at most \( d \). The other direction is also true. It is easy to verify that functions that pass this test with probability one, must be degree \( d \) polynomials.

A more interesting direction is the study of functions that pass this test with high probability. Strong results regarding such functions will result in the soundness and even the absoluteness of this test. This question was considered before for various ranges of acceptance probability: greater than \( 1 - O(q^t) \), greater than \( 1 - O(1) \) and non-negligible probability.
1.2.1 Acceptance probability $1 - O(q^t)$

The first results in polynomial testing, in the setting where the degree of the polynomial is larger than the field size, considered functions that pass the test with probability $1 - O(q^t)$. It was first studied by Alon et al. [AKK+05] which considered the setting of $q = 2$. They considered the $t$-dimensional test that made $O(2^d)$ queries.\(^2\) (Observe that $t_{2,d} = d$). Their analysis showed that functions that pass the test with probability $\delta 2^d$ are $\delta$ close to some degree $d$ polynomial. 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 that no test with $o(2^d)$ queries could test this family, thus obtaining 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. [JPRZ09]. They showed that over a general field, functions that pass the $t$-dimensional test with probability $\delta q^t$ are also $\delta$ close to some degree $d$ polynomials. Similarly, this yield an absolutely sound test with query complexity $O(q^{2t})$. They also show that $q^t$ is a lower bound on the number of queries if $q$ is prime. Again, showing that the analysis of the test is off by a quadratic factor. The proof techniques of [AKK+05] and [KR06, JPRZ09] 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.

1.2.2 Acceptance probability $1 - O(1)$

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 even functions that pass the test with probability $1 - \delta$ are $O(\delta)$ far from being degree $d$ polynomial. This showed 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, JPRZ09]. (Indeed the work of [KS08] confirms that the central ingredient in the proofs in [AKK+05, KR06, JPRZ09] 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 thesis we manage to overcome these obstacles and show in Theorem 3.1.3 that the $O(q^t)$ query tester of [KR06] is also absolutely sound (though as it turns out, the constant grows extremely fast as a function of $q$). 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.

In Theorem 5.1.2 we present an application to this result. Showing a simple algorithm that determines if a graph contains a simple path (or a general $r$-simple path) using a reduction to low degree testing. We will elaborate more about this result in Chapter 5.

1.2.3 Non-negligible acceptance probability

Another interesting range is considering functions that pass this test with non-negligible probability. For simplicity we first focus on functions over $\mathbb{F}_2$. In this case, as described earlier, the test can be described as verifying that a random $(d+1)$-th derivative in a random point is zero. Therefore, one can verify that if the function is "random" then the probability that this derivative be zero is as likely as the probability that it is one. In other words the acceptance probability of a random function is $\approx 50\%$ and $\|f\|_{U^{d+1}} \approx 0$.\(^2\)

\(^2\)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.
Now assume that we are given a function $f$ that behaves differently than a random function, namely, $\|f\|_{U^{d+1}}$ is bounded away from 0. Is there anything that we can deduce about the structure of $f$ just by knowing this fact? Since we saw before that high $U^{d+1}$ norm implies closeness to degree $d$ polynomial, we expect from $f$ to behave like a polynomial in some weaker sense. This question is very interesting over general prime order fields and even for $d = 0$. We may assume further structure on the function (e.g., it has bounded degree) to get more interesting structural results. Several themes of such polynomial behaviours were considered.

**Such functions correlate with polynomials**

Samorodnitsky and Green and Tao conjectured that such functions with noticeable Gowers norm correlate with degree $d$ polynomials [Sam07, GT08]. This conjecture has become known as the inverse conjecture for the Gowers norm. Samorodnitsky [Sam07] proved that if $\|f\|_{U^3} = \delta$ where $f : \mathbb{F}_2^2 \to \mathbb{F}_2$ is an arbitrary function, then $f$ has an exponentially high (in $1/\delta$) correlation with a quadratic polynomial. Namely, there exists a quadratic polynomial $q$ such that $\Pr_{x \in \mathbb{F}_2}[f(x) = q(X)] \geq 1/2 + \exp(-\text{poly}(1/\delta))$. Independently, Green and Tao [GT08] obtained similar results for fields of odd characteristic. These results give an affirmative answer for the case of the $U^3$ norm. More generally, Green and Tao proved that if $d < |\mathbb{F}|$ and $f$ is a degree $d$ polynomial with a high $U^d$ norm then $f$ is indeed correlated with a lower degree polynomial [GT07].

On the other hand, for the $U^4$ norm it was shown, independently, by Lovett, Meshulam and Samorodnitsky [LMS08] and by Green and Tao [GT07] that no such result is possible, when $\mathbb{F} = \mathbb{F}_2$. Namely, [LMS08] proved that the symmetric polynomial $S_4(x_1, \ldots, x_n) = \sum_{T \subseteq [n], |T| = 4} \prod_{i \in T} x_i$, which is of degree four, has a high $U^4$ norm but has an exponentially (in $n$) small correlation with any lower degree polynomial. Similar examples were given for other fields (when $d$ is large enough compared to the size of the field). These examples show that for small fields the inverse conjecture for the Gowers norm is not true in its current form.

Nevertheless, by extending the range of allowed polynomials to also include phase polynomials Tao and Ziegler were able to prove the conjecture for the large characteristic case using ideas from ergodic theory [TZ08]. That is, Tao and Ziegler proved that if $f : \mathbb{F}^n \to \mathcal{D}$ (where $\mathcal{D}$ is the unit disk in $\mathbb{C}$) is a function with high $U^d$ norm and $d \leq |\mathbb{F}|$ then $f$ isn correlated with a degree $d - 1$ phase polynomial. This completely settled the conjecture (allowing phase polynomials) for the case $d \leq |\mathbb{F}|$. In a follow up work, Green, Tao and Ziegler settled the conjecture for the general case (albeit they give correlation to a phase-polynomial rather than to a low degree polynomial).

**Such functions can be expressed as functions of degree $d$ polynomials**

Now consider the case where $f$ is a degree $d + 1$ polynomial such that $\|f\|_{U^{d+1}} > \delta$. Green and Tao [GT07] show that in this case $f = \sum_{i=1}^C p_i q_i$, where $\deg(p_i), \deg(q_i) \leq d$. Thus, if $f$ has a noticeable Gowers norm, unlike a random degree $d$ polynomial, then $f$ is in fact very far from being random. A simple counting argument shows that most polynomials of degree $d$ cannot be represented as functions of a few lower degree polynomials. One drawback of the results of [GT07] is the dependence of the number of lower degree polynomials on the bias of $f$. In particular when $\deg(f) = 3$, [GT07] get

---

4We stress that over general prime order fields the Gowers norm does not necessarily describe the acceptance probability of the $t$-dimensional test, although it is similar and still a natural notion.

4A phase polynomial of degree $d - 1$ is a function of the form $e^{2\pi i \theta^q}$, for some polynomial $g$ of degree $d - 1 - (|\mathbb{F}| - 1)t$, where $\theta \in [0, 1)$ and $\omega = e^{2\pi i/|\mathbb{F}|}$.

5In fact [TZ08] only obtain a qualitative result. No explicit connection is known between the Gowers norm and the correlation.
the bound \( C_3 = \exp(\text{poly}(1/\text{bias}(f))) \) and for \( \text{deg}(f) = 4 \) they bound\(^6\) \( C_4 \) by a tower of height \( C_3 \).

In Theorems 2.1.6,2.1.10 we improve these results significantly, improving \( C_3 \) to a polylogarithmic function and \( C_4 \) to a quasi-polynomial function. Moreover, we strengthen our results for functions that have noticeable bias (the \( U^1 \) norm) in Theorems 2.1.5,2.1.7.

**Such functions are polynomial like on a large subspace**

Another theme is showing that even if those functions do not correlate with polynomials on all of the space, they are at least correlated with it on a large subspace. [GT08] showed that if \( f : \mathbb{F}_q^s \to \mathbb{F}_5 \) satisfies \( \|f\|_{U^3} = \delta \) then there exists a subspace \( V \) of codimension \( \text{poly}(1/\delta) \), such that on an ‘average’ coset of \( V \), \( f \) is correlated with a quadratic polynomial. Wolf [Wol09] proved a similar result for the case of characteristic two, thus extending Samorodnitsky’s argument [Sam07].

In Theorem 2.1.8 we, again, consider a degree 4 polynomial with a high \( U^4 \) Norm and show that such a polynomial becomes a degree 3 polynomial on a subspace with dimension \( \Omega(n) \). The main difference between the previous results and our result is that ours only holds for polynomials of degree four whereas the results of [GT08, Sam07, Wol09] hold for arbitrary functions. On the other hand our result holds for the \( U^4 \) norm compared to the \( U^3 \) norm studied there. We also note that the inverse theorem for the Gowers norm implies a global structure - i.e. correlation with a low degree phase polynomials over the entire space - whereas our theorem only gives a local structure. However, the structure that we show is more refined as it gives a high dimensional space on which the function equals a lower degree polynomial.

### 1.3 Lifted codes

The \( t \)-dimensional tester can be extended to general lifted codes, describes as follows. Given a linear affine-invariant base code \( \mathcal{B} \) on \( t \) variables and integer \( n \geq t \), the \( n \)-dimensional lift of \( \mathcal{B} \), denoted \( \text{Lift}_n(\mathcal{B}) \), is the set \{ \( f : \mathbb{F}_q^n \to \mathbb{F}_q \mid f|_V \in \mathcal{B} \) for every \( t \)-dimensional affine subspace \( V \subseteq \mathbb{F}_q^n \}\).

Those codes was defined by Ben-Sasson et al. [BSMSS11]. Similarly to polynomials, a natural test for lifted codes is verifying that on a random \( t \) dimensional affine subspace \( V \), \( f|_V \in \mathcal{B} \). This test generalises the \( t \)-dimensional test for polynomials since, by the definition of the testing dimension, \( \mathcal{P}(n, d, q) = \text{Lift}_n(\mathcal{P}(t, d, q)) \).

The soundness of this test follows from the work of [KS08], showing that the rejection probability of functions that are \( \delta \) far from the code is \( O(\delta q^t) \) so one needs to repeat this test \( O(q^t) \) times to get an absolutely sound test. The lower bound for such a test is \( O(q^t) \), so [KS08] were off by quadratic factor. An intriguing question is whether this \( t \) dimensional tester is even absolute in the broader context of general lifted codes. In Theorem 4.1.1 we give an affirmative answer to this question. Thus, we give an optimal test for general lifted code (without extra information on \( \mathcal{B} \)). The importance of lifted codes follow from the work of Guo et al. [GKS08a]. They showed that lifted codes can yield the state of the art codes, in the context of local correction and testability, for some interesting ranges of parameters. To stress this point, in Section 4.7 we demonstrate how our result can be used in the construction of a new family of locally testable codes.

---

\(^6\)These numbers are not explicitly computed there, but this is what the recursive arguments in the papers imply.
Chapter 2

On the Structure of Cubic and Quartic Polynomials

2.1 Introduction

Assume that we are given a degree $d$ polynomial $f$ that, in some sense, ‘behaves’ differently from a random degree $d$ polynomial. Is there anything that we can deduce about the structure of $f$ just by knowing this fact? Recently this question received a lot of attention, where the ‘behavior’ of $f$ was examined with respect to its bias or the more general notion of the Gowers norm.

Definition 2.1.1. Let $f : \mathbb{F}^n \to \mathbb{F}$ be a function. The bias of $f$ is defined as

$$\text{bias}(f) = \mathbb{E}_{a \in \mathbb{F}^n} [\omega^{f(a)}],$$

where $\omega = e^{2\pi i/|\mathbb{F}|}$ is a complex primitive root of unity of order $|\mathbb{F}|$.

Intuitively, the bias of $f$ measures how far is the distribution induced by $f$ from the uniform distribution. We expect a random polynomial to have a vanishing small bias (as a function of the number of variables), so it is interesting to know what can be said when the bias is not too small. Indeed, Green and Tao [GT07] showed that if $f$ is a degree $d$ polynomial over $\mathbb{F}$, such that $d < |\mathbb{F}|$, and $\text{bias}(f) = \delta$ then $f$ can be written as a function of a small number of lower degree polynomial. Formally, $f(x) = F(g_1, \ldots, g_{c_d})$ for some function $F$ and $c_d = c_d(\text{bias}(f), |\mathbb{F}|)$ polynomials $\{g_i\}$ satisfying $\deg(g_i) < d$. Note that $c_d = c_d(\text{bias}(f), |\mathbb{F}|)$ does not depend on the number of variables, i.e. it is some constant. This result was later extended by Kaufman and Lovett [KL08] to arbitrary finite fields (i.e. without the restriction $d < |\mathbb{F}|$). Thus, if $f$ has a noticeable bias, unlike a random degree $d$ polynomial, then $f$ is in fact very far from being random; simple counting arguments show that most degree $d$ polynomials cannot be represented as functions of a few lower degree polynomials. This result is also interesting as it gives an average case - worst case reduction. Namely, if $f$ has correlation $\delta$ with a lower degree polynomial then it is a function of a small number of lower degree polynomials. One drawback of the results of [GT07, KL08] is the dependance of the number of lower degree polynomials on the bias of f. In particular when $\deg(f) = 3$, [GT07, KL08] get the bound $c_3 = \exp(\text{poly}(1/\text{bias}(f)))$ and for $\deg(f) = 4$ they bound $c_4$ by a tower of height $c_3$. On the other hand if $\deg(f) = 2$ and $\text{bias}(f) = \delta$ then it is known that $f$ can be written as a function of at most $2 \log(1/\delta) + 1$ linear functions. This can be immediately deduced from the following well known theorem.

These numbers are not explicitly computed there, but this is what the recursive arguments in the papers imply.
Theorem 2.1.2 (Structure of quadratic polynomials). (Theorems 6.21 and 6.30 in [LN97]). For every quadratic polynomial \( f : \mathbb{F}^n \rightarrow \mathbb{F} \) over a prime field \( \mathbb{F} \) there exists an invertible linear transformation \( T \), a linear polynomial \( \ell \), and field elements \( \alpha_1, \ldots, \alpha_n \) (some of which may be 0) such that:

1. If \( \text{char}(\mathbb{F}) = 2 \) then \( (q \circ T)(x) = \sum_{i=1}^{|n/2|} \alpha_i \cdot x_{2i-1} \cdot x_{2i} + \ell(x) \),

2. If \( \text{char}(\mathbb{F}) \) is odd then \( (q \circ T)(x) = \sum_{i=1}^n \alpha_i \cdot x_i^2 + \ell(x) \).

Moreover, the number of non zero \( \alpha_i \)-s is invariant and depends only on \( f \).

We thus see that there is a sharp contrast between the result for quadratic polynomials and the results for polynomials of degrees as low as three or four. We also note that the results of Kaufman and Lovett only guarantee that \( f \) can be represented as \( f(x) = F(g_1, \ldots, g_c) \) but no nice structure like the one in Theorem 2.1.2 is known. It is thus an intriguing question whether a nice structural theorem exists for biased polynomials and what is the correct dependence of the number of lower degree polynomials on \( \deg(f) \) and \( \text{bias}(f) \).

As mentioned above, a more general measure of randomness that was considered is the so called Gowers norm of \( f \). Intuitively, the \( U^d \) Gowers norm tests whether \( f \) behaves like a degree \( d-1 \) polynomial on \( d \)-dimensional subspaces. To define the Gowers norm we first define the notion of a discrete partial derivative.

Definition 2.1.3. (Discrete partial derivative) For a function \( f : \mathbb{F}^n \rightarrow \mathbb{F} \) and a direction \( y \in \mathbb{F}^n \) we define \( \Delta_y(f)(x) = f(x+y) - f(x) \) to be the discrete partial derivative of \( f \) in direction \( y \) at the point \( x \).

It is not difficult to see that if \( \deg(f) = d \) then for every \( y \), \( \deg(\Delta_y(f)) \leq d - 1 \). We now define the \( d \)-th Gowers norm of a function \( f \).

Definition 2.1.4 (Gowers norm [AKK+05, Gow98, Gow01]). The \( d \)-th Gowers norm, \( U^d \), of \( f \) is defined as

\[
\| f \|_{U^d} = \left[ \mathbb{E}_{x,y_1,\ldots,y_d} [\omega^{\Delta_y \Delta_{y_1} \ldots \Delta_{y_d} (f)(x)} ] \right]^{1/2^d},
\]

where again \( \omega = e^{2\pi i} \).

Note that \( \| f \|_{U^0} = \| f \|_{U^1} = \text{bias}(f) \). It is also clear that if \( \deg(f) = d - 1 \) then \( \| f \|_{U^d} = 1 \). For more properties of the Gowers norm we refer the reader to [Gow98, Gow01, GT08, Sam07, VW07].

In [AKK+05] Alon et al. showed that if \( \| f \|_{U^d} > 1 - \delta \), for some small \( \delta \) (depending on \( d \)), then \( f \) can be well approximated by a degree \( d - 1 \) polynomial. This raises the question whether any function that has a noticeable \( U^d \) norm is somewhat correlated with a lower degree polynomial and indeed in [Sam07, GT08] this was conjectured to be the case. This conjecture has become known as the inverse conjecture for the Gowers norm. Samorodnitsky [Sam07] proved that if \( \| f \|_{U^3} = \delta \) where \( f : F_2^n \rightarrow F_2 \) is an arbitrary function, then \( f \) has an exponentially high (in \( \delta \)) correlation with a quadratic polynomial. Namely, there exists a quadratic polynomial \( q \) such that \( \Pr_{x \in F_2^n} [ f(x) = q(X) ] \geq 1/2 + \exp(-\text{poly}(1/\delta)) \).

Independently, Green and Tao [GT08] obtained similar results for fields of odd characteristic. These results gave an affirmative answer for the case of the \( U^3 \) norm. More generally, Green and Tao proved that if \( d < |\mathbb{F}| \) and \( f \) is a degree \( d \) polynomial with a high \( U^d \) norm then \( f \) is indeed correlated with a lower degree polynomial [GT07]. Recently, the case of large characteristic was solved by Tao and Ziegler [TZ08].

\[ ^2 \]In fact, [TZ08] only get a qualitative result. No explicit connection is known between the Gowers norm and the correlation with polynomials.
unit disk in $\mathbb{C}$) is a function with high $U^d$ norm and $d \leq |F|$ then $f$ is correlated with a degree $d - 1$ phase polynomial.\footnote{A degree $d - 1$ phase polynomial is a function of the form $e^{2\pi i d \theta} \omega^g$, for some degree $d - 1 - (p - 1)t$ polynomial $g$ where $\theta \in [0, 1]$ and $\omega = e^{2\pi i / |F|^t}$.} This completely settled the conjecture for the case $d \leq |F|$. On the other hand, for the $U^4$ norm it was shown, independently, by Lovett, Meshulam and Samorodnitsky [LMS08] and by Green and Tao [GT07] that no such result is possible when $F = F_2$. Namely, [LMS08] proved that the symmetric polynomial $S_d(x_1, \ldots, x_n) = \sum_{T \subseteq [n], |T| = 4} \prod_{i \in T} x_i$, which is of degree four, has a high $U^4$ norm but has an exponentially (in $n$) small correlation with any lower degree polynomial. Similar examples were given for other fields (when $d$ is large enough compared to the size of the field). These examples show that for small fields the inverse conjecture for the Gowers norm is not true in its current form. In their work, Tao and Ziegler [TZ08] proved a variant of the conjecture for the case of low characteristic ($d > |F|$). Namely, that if a function $f$ has high $U^d$ norm then $f$ is correlated with a phase polynomial of a certain constant degree (but not necessarily smaller than $d$). We note however, that if $\deg(f) = d$ then the results of [TZ08] do not give any information on $f$. In fact, even if $\deg(f) = 4$ and $f$ has a high $U^4$ norm then nothing is known on the structure of $f$. It is thus a very interesting question to understand the structure of low degree polynomials having high Gowers norm over small fields.

Besides being natural questions on their own, results on the Gowers norms had many applications in mathematics and computer science. In his seminal work on finding arithmetic progressions in dense sets, Gowers first defined the $U^d$ norm (for functions from $\mathbb{Z}_n$ to $\mathbb{Z}_n$) and proved an inverse theorem for them that was instrumental in his proofs [Gow98, Gow01]. Bogdanov and Viola [BV07] attempt for constructing a pseudo random generator for constant degree polynomials relied on the (erroneous) inverse conjecture for the Gowers norm, yet it paved the way for other papers solving the problem [Lov08, Vio08]. In [ST06] applications of an inverse theorem for the Gowers norm to PCP constructions was given. Samorodnitsky’s proof of the inverse theorem for the $U^3$ norm [Sam07] implies a low degree test distinguishing quadratic functions from those that do not have a non trivial correlation with a quadratic function. This result also gives a test for checking the distance of a given word from the 2nd order Reed-Muller code, beyond the list decoding radius. For a more elaborate discussion of the connection between additive combinatorics and computer science see [Tre09].

### 2.1.1 Our results

In this work we are able to show analogs of Theorem 2.1.2 for polynomials of degree three and four. We also prove a structural result for the case that such a polynomial has a high Gowers norm. Our first main result is the following.

**Theorem 2.1.5. (biased cubic polynomials)** Let $F$ be a finite field and $f \in F[x_1, \ldots, x_n]$ a cubic polynomial ($\deg(f) = 3$) such that $\bias(f) = \delta$. Then there exist $c_1 = O(\log(1/\delta))$ quadratic polynomials $q_1, \ldots, q_{c_1} \in F[x_1, \ldots, x_n]$ and linear functions $\ell_1, \ldots, \ell_{c_1} \in F[x_1, \ldots, x_n]$ and another $c_2 = O(\log^4(1/\delta))$ linear functions $\ell'_1, \ldots, \ell'_{c_2} \in F[x_1, \ldots, x_n]$ such that $f = \sum_{j=1}^{c_1} \ell_j \cdot q_j + g(\ell'_1, \ldots, \ell'_{c_2})$, where $g$ is cubic.

We note that if it weren’t for the $g(\ell'_1, \ldots, \ell'_{c_2})$ part then this result would be quantitatively the same as Theorem 2.1.2 (and tight of course). It is an interesting open question to decide whether we can do only with the $\sum_{j=1}^{O(\log(|F|))} 1/\delta} \ell_i \cdot q_i$ part. Using the same techniques we show a similar result for the case that $\|f\|_{U^3} > \delta$.

**Theorem 2.1.6. (cubic polynomials with high $U^3$ norm)** Let $F$ be a finite field and $f \in F[x_1, \ldots, x_n]$ a cubic polynomial such that $\|f\|_{U^3} = \delta$. Then there exist $c + 1 = O(\log^2(1/\delta))$ quadratic polynomials $q_0, \ldots, q_{c_0} \in F[x_1, \ldots, x_n]$ and $c$ linear functions $\ell_1, \ldots, \ell_c \in F[x_1, \ldots, x_n]$ such that $f = \sum_{j=1}^{c} \ell_j \cdot q_j + q_0$. 
Note that the difference between the structure of $f$ in Theorems 2.1.5 and 2.1.6 is the number of quadratic function required. Recall that in [Sam07] Samorodnitsky proved that if an $F_2$ function $f$ has a high $U^3$ norm then it has an exponentially (in $\|f\|_{U^3}$) high correlation with a quadratic polynomial. Thus, our theorem shows that when $f$ is a cubic polynomial then a much stronger statement holds. Namely, $f$ has correlation $\exp(\log^2(1/\delta))$ with a quadratic polynomial, and further, has a nice structure.

Our second main result is an analog of Theorem 2.1.5 for the case of quartic polynomials (i.e. $\deg(f) = 4$).

**Theorem 2.1.7.** (biased quartic polynomials) Let $F$ be a finite field and $f \in F[x_1, \ldots, x_n]$ a quartic polynomial ($\deg(f) = 4$) such that $\bias(f) = \delta$. Then there exist $4c = \text{poly}(\|F\|/\delta)$ polynomials $\{\ell_i, q_i, q_i', g_i\}_{i=1}^c$, where the $\ell_i$-s are linear, the $q_i$-s and $q_i'$-s are quadratic and the $g_i$-s are cubic such that $f = \sum_{i=1}^c \ell_i \cdot g_i + \sum_{i=1}^c q_i \cdot q_i'$.

As mentioned above, prior to this result it was known that there exist cubic polynomials $g_1, \ldots, g_C$ and a function $F$ such that $f = F(g_1, \ldots, g_C)$, where $C$ is a tower of height $\exp(\text{poly}(1/\delta))$ [GT07, KL08]. Thus, our result greatly improves the dependance on $\delta$ and gives a nice structure for the polynomial. We note that in their work Green and Tao do show that such a nice structure exists when $d < |F|$ [GT07], but no such result was known for smaller fields (in addition $C$ needs to be even larger for such a nice representation to hold).

Our third main result is for the case where $\deg(f) = 4$ and $\|f\|_{U^4} = \delta$. In such a case it is known [LMS08, GT07] that we cannot hope to get a nice structure as in Theorem 2.1.6 as it may be the case that $f$ has an exponentially small (in $n$) correlation with all lower degree polynomials. However, we do manage to show that there is some subspace $U \subseteq F^n$ such that when restricted to $V$, $f|_U$ is equal to some degree three polynomial. Thus, $f$ does not have a correlation with a cubic polynomial in the entire space but instead there is a large subspace on which it is of degree three. In fact we show a more general result. Namely, that there is a large subspace $V$, of dimension $n - O(\log(1/\delta))$, that can be partitioned to subspaces of dimension $n/\exp(\log^2(1/\delta))$ such that the restriction of $f$ to any of the subspaces in the partition is of degree three.

**Theorem 2.1.8.** (quartic polynomials with high $U^4$ norm) Let $F$ be a finite field and $f \in F[x_1, \ldots, x_n]$ a degree four polynomial such that $\|f\|_{U^4} = \delta$. Then there exists a partition of a subspace $V \subseteq F^n$, of dimension $\dim(V) \geq n - O(\log(1/\delta))$, to subspaces $\{V_\alpha\}_{\alpha \in I}$, satisfying $\dim(V_\alpha) = \Omega(n/|F|^{\log^2(1/\delta)})$, such that for every $\alpha \in I$, $f|_{V_\alpha}$ is a cubic polynomial.

**Remark 2.1.9.** Note that the structure guaranteed in Theorem 2.1.8 is shared by very few polynomials. Specifically, a random polynomial of degree four is unlikely to be equal to any degree three polynomial on any subspace of dimension larger than, say, $n^{0.9}$. To see this note that if $|F| = p$ and $\dim(V) = d$ then there are roughly $p^{d^3}$ cubic polynomials and $p^{d^4}$ quartic polynomials over $V$. Furthermore, the map taking a quartic polynomial over $F^n$ to its restriction is a linear map and so the fraction of quartic polynomials that equal a degree three polynomial on $V$ is (roughly) $p^{d^4 - d^3}$. As the total number of subspaces can be bounded by $p^{n^2}$ we get that the fraction of quartic polynomials that are equal to a degree three polynomial on some subspace of dimension greater than $n^{0.9}$ is at most $p^{n^2 - n^{0.9} + n^2} = \exp(-n^{3.6})$.

This result has the same flavor as the inverse $U^3$ norm theorem of [GT08]. There it was shown that if $f : F_5^n \rightarrow F_5$ satisfies $\|f\|_{U^3} = \delta$ then there exists a subspace $V$ of codimension $\text{poly}(1/\delta)$, such that on an ‘average’ coset of $V$, $f$ is correlated with a quadratic polynomial. Recently, Wolf [Wol09] proved a similar result for the case of characteristic two, thus extending Samorodnitsky’s argument [Sam07]. The main difference between these results and our result is that ours only holds for polynomials of degree four whereas the results of [GT08, Sam07, Wol09] hold for arbitrary functions. On the other
hand our result holds for the $U^4$ norm compared to the $U^3$ norm studied there. Moreover, when $\text{char}(\mathbb{F}) > 4$, using the same techniques we can actually show that $f$ must have a structure similar to the one guaranteed by Theorem 2.1.7.

**Theorem 2.1.10.** Let $\mathbb{F}$ be a finite field with $\text{char}(\mathbb{F}) > 4$ and $f \in \mathbb{F}[x_1, \ldots, x_n]$ a degree four polynomial such that $\|f\|_{U^4} = \delta$. Then

$$f = \sum_{i=1}^{R} \ell_i \cdot g_i + \sum_{i=1}^{r} q_i \cdot q'_i,$$

for $r = O(\log^2(1/\delta))$ and $R = \exp(\log^2(1/\delta))$ where $\ell_i$ is linear, $q_i, q'_i$ are quadratic and $g_i$ cubic.

### 2.1.2 Proof Technique

The main approach in all the proofs is to consider the space of discrete partial derivatives of $f$ and look for some structure there. We will explain the idea for the case of degree three polynomials and then its extension to degree four polynomials.

Let $f$ be a degree three polynomials. Assume that $f$ has high bias (respectively high $U^3$ norm). By a standard argument it follows that a constant fraction of its derivatives, which are degree 2 polynomials, have high bias (high $U^2$ norm). By Theorem 2.1.2 it follows that for a constant fraction of the directions, the partial derivatives depends on a small number of linear functions (same for the $U^3$ norm). Hence, in the space of partial derivatives, a constant fraction of the elements depend on a few linear functions. We now show that there must be a small number of linear functions that ‘explain’ this. More accurately, we show that there exist a subspace $V \subset \mathbb{F}^n$, of dimension $\dim(V) = n - O(1)$, and $O(1)$ linear functions $\ell_1, \ldots, \ell_c$, such that for every $y \in V$ it holds that $\Delta_y(f) = \sum_{i=1}^{c} \ell_i \cdot \ell_i^{(y)}$, where the $\ell_i^{(y)}$-s are linear functions determined by $y$.

We are now basically done. Consider the subspace $U = \{x : \ell_1(x) = \ldots = \ell_c(x) = 0\}$. Then, for every $y \in V$ it holds that $\Delta_y(f)|_U = \ell_0^{(y)}|_U$. This implies that $f|_V = \sum_{i=1}^{c} \ell_i \cdot q_i + q_0$, where the $q_i$-s are quadratic polynomials. As $\dim(V) = n - O(1)$ we obtain the same structure (with a different constant $c$) for $f$.

To prove the result for biased degree four polynomials we follow the footsteps of [KL08] with two notable differences. Let $f$ be such a polynomial. First, we pass to a subspace on which all the partial derivatives of $f$ have low rank as degree three polynomials. This steps relies on our results for biased degree three polynomials. Then, as in [KL08], we show that $f$ can be approximated by a function of a few of its derivatives. Because of the properties of the derivatives, this means that $f$ can be approximated well by a function of a few quadratics and linear functions. We then show, again following [KL08], that in such a case $f$ is actually a function of a few quadratics and linear functions. Here we heavily rely on properties of quadratic functions to avoid the blow up in the number of polynomials approximating $f$ that occurs in [KL08, GT07]. Finally, we show that if a degree four polynomial is a function of several quadratic and linear functions then it actually have a nice structure.

The proof for the case of degree four polynomials with high $U^4$ norm is more delicate. Assume that $f$ is such a polynomial. As before, a constant fraction of the partial derivatives of $f$ are degree three polynomials with high $U^3$ norm. By the result for degree three polynomials we get that each of those partial derivatives is of the form $\Delta_y(f) = \sum_{i=1}^{c} \ell_i^{(y)} \cdot q_i^{(y)} + q_0^{(y)}$. Again we find a subspace $V$, of constant co-dimension, such that for every $y \in V$, $\Delta_y(f)$ has a nice structure. We now show that there exist a small number of linear and quadratic functions $\{\ell_i, q_i\}_{i=1}^{c}$ such that for every $y \in V$ it holds that $\Delta_y(f) = \sum_{i=1}^{c} \ell_i \cdot q_i^{(y)} + \sum_{i=1}^{c} \ell_i^{(y)} \cdot q_i + q_0^{(y)}$, where the polynomials $\{\ell_i^{(y)}, q_i^{(y)}\}$ depend on $y$. This is the technical heart of the proof. It now follows quite easily that there is a subspace $U \subseteq V$ of dimension
n/\exp(c) \) such that when restricted to \( U \) all the functions \( \{\ell_i, q_i\} \) are fixed to constants. Thus, for every \( y \in U \) it holds that \( \deg(\Delta_y(f)) = 2 \). So we get that \( \deg(f|_U) = 3 \). In fact, by closely examining the argument above we show an even stronger result. Namely, that we can partition a large subspace of \( \mathbb{F}^n \) to (affine) subspaces of dimension \( n/\exp(c) \) such that on each of the subspaces \( f \) is equal to some cubic polynomial (that may depend on the subspace).

### 2.1.3 Organization

In Section 4.3 we give some basic definitions and discuss properties of subadditive functions. In Section 2.3 we prove the theorems concerning degree three polynomials. In Section 2.4 we prove Theorem 2.1.7 and in Section 2.5 we prove Theorems 2.1.8 and 2.1.10.

### 2.2 Preliminaries

In this chapter \( \mathbb{F} \) will always be a prime field. We denote with \( \mathbb{F}_p \) the field with \( p \) elements. As we will be considering functions over \( \mathbb{F}_p \) we will work modulo the polynomials \( x_i^p - x_i \). In particular, when we write \( f = g \), for two polynomials, we mean that they are equal as functions and not just as formal expressions. This will be mainly relevant when we consider quadratic polynomials (or higher degree polynomials) over \( \mathbb{F}_2 \). More generally, we shall say that a function \( f \) has degree \( d \) if there is a degree \( d \) polynomial \( g \) such that \( f = g \). Note that this does not have an affect on the bias and the Gowers norm. Namely, the bias and \( U^d \) norm of \( f \) do not change when adding multiplies of \( x_i^p - x_i \). Finally we note that if all the partial derivative of \( f \) have degree at most \( d - 1 \) then there is a polynomial \( g \) of degree at most \( d \) such that \( f = g \) (this is easily proved by observing that a degree \( k \) polynomial, all of whose individual degrees are smaller than \( |\mathbb{F}| \), always has a partial derivative whose degree is \( k - 1 \)). From this point on we shall use the notion of a function and a polynomial arbitrarily without any real distinction.

The Fourier transform of a function \( f : \mathbb{F}^n \to \mathbb{F} \) is defined as

\[
\hat{f}(\alpha) = \mathbb{E}_{x \in \mathbb{F}^n} [f(x)\overline{\chi_\alpha(x)}],
\]

where for \( \alpha = (\alpha_1, \ldots, \alpha_n) \), \( \chi_\alpha(x) = \omega^{\sum_{i=1}^n \alpha_i x_i} \) where \( \omega = e^{2\pi i/p} \) is a complex primitive root of unity of order \( |\mathbb{F}| \). For more on Fourier transform see [Ste03].

We say that a function \( h \) \( \epsilon \)-approximates a function \( f \) if \( \Pr_x[f(x) \neq h(x)] \leq \epsilon \).

**Definition 2.2.1.** Following [KL08] we say that the distribution induced by a set of functions \( \{h_i\}_{i=1}^m \) (all from \( \mathbb{F}^n \) to \( \mathbb{F} \)) is \( \gamma \) close to the uniform distribution if for every \( \alpha_1, \ldots, \alpha_m \in \mathbb{F} \) it holds that

\[
\left| \Pr_{x \in \mathbb{F}^n} [\forall 1 \leq i \leq m, h_i(x) = \alpha_i] - |\mathbb{F}|^{-m} \right| \leq \gamma |\mathbb{F}|^{-m}.
\]

The following well known lemma bounds the distance between distributions using the Fourier transform.

**Lemma 2.2.2.** For \( i = 1 \ldots m \) let \( h_i : \mathbb{F}^n \to \mathbb{F} \) be a function. Then, the distribution induced by the \( h_i \)-s is \( \gamma \) close to uniform if for every nontrivial linear combination \( h_\alpha = \sum_{i=1}^m \alpha_i h_i \), we have that \( \text{bias}(h_\alpha) \leq \gamma/|\mathbb{F}|^{3m/2} \).
Proof. Let \( H : \mathbb{F}^n \to \mathbb{F}^m \) be defined as \( H(x) = (h_1(x), \ldots, h_m(x)) \). For \( y \in \mathbb{F}^m \) let \( f(y) = \Pr_{x \in \mathbb{F}^n}[H(x) = y] \). We have that

\[
|\hat{f}(\alpha)| = \left| \mathbb{E}_{y \in \mathbb{F}^m} f(y) \chi_\alpha(y) \right| = \left| \mathbb{E}_{y \in \mathbb{F}^m} \Pr_{x \in \mathbb{F}^n}[H(x) = y] \chi_\alpha(y) \right| = |\mathbb{F}|^{-m} \sum_{x \in \mathbb{F}^n} \chi_\alpha[H(x)] = |\mathbb{F}|^{-m} \text{bias} \left( \sum_{i=1}^m \alpha_i h_i \right).
\]

Therefore,

\[
\left( \sum_{y \in \mathbb{F}^m} |f(y) - |\mathbb{F}|^{-m}| \right)^2 \leq |\mathbb{F}|^m \sum_{y \in \mathbb{F}^m} |f(y) - |\mathbb{F}|^{-m}|^2 = |\mathbb{F}|^m \sum_{y \in \mathbb{F}^m} f(y)^2 - 2|\mathbb{F}|^{-m} f(y) + |\mathbb{F}|^{-2m} = \left( \sum_{\alpha \in \mathbb{F}^m} |\mathbb{F}|^{2m} \hat{f}(\alpha)^2 \right) - 1 = \left( \sum_{0 \neq \alpha \in \mathbb{F}^m} |\mathbb{F}|^{2m} \hat{f}(\alpha)^2 \right) = \sum_{0 \neq \alpha \in \mathbb{F}^m} \text{bias} \left( \sum_{i=1}^m \alpha_i h_i \right)^2 < |\mathbb{F}|^{-2m} \gamma^2.
\]

Hence, for every \( y \in \mathbb{F}^m \) it holds that \( |f(y) - |\mathbb{F}|^{-m}| < |\mathbb{F}|^{-m} \gamma \), which is what we wanted to prove. ☐

### 2.2.1 Subadditive functions

As described in Section 2.1.2 our proofs are based on finding a structure for the space of partial derivatives of the underlying polynomial \( f \). For this end we need a special case of a lemma of Bogolyubov (see e.g. [Gre]).

For a set \( A \subseteq \mathbb{F}^n \) denote with \( kA - kA \) the set

\[
kA - kA = \{a_1 + \ldots + a_k - a_{k+1} - \ldots - a_{2k} \mid \forall i \ a_i \in A \}.
\]

**Lemma 2.2.3** (Bogolyubov). Let \( A \subseteq U \) be a subset of a linear space \( U \) such that \( |A| = \mu_0 \cdot |U| \). Then, for some \( k \leq \max(1, \lceil \frac{1}{2} \left( \log \frac{|U|}{|A|} (2/\mu_0 + 2) \right) \rceil) \), \( kA - kA \) contains a subspace \( W \) of co-dimension at most \( \log \frac{|U| - 1/2}{|U| - 1} (1/2\mu_0) \).

For completeness we give the proof here.

**Proof.** For \( \mu \in (0,1) \) define \( \rho(\mu) = \frac{|\mathbb{F}| - 1/2}{|\mathbb{F}|} \cdot \mu \). We shall think of \( A \) also as the characteristic function of the set \( A \) and denote with \( \{\hat{A}(\alpha)\} \) its fourier coefficients. Assume that there is some \( \alpha \neq 0 \) such that \( |\hat{A}(\alpha)| \geq \rho(\mu_0) \). This means that there is some (affine) subspace \( W \) of co-dimension at most one such that

\[
|A \cap W|/|W| \geq \rho(\mu_0) \cdot |\mathbb{F}|/(|\mathbb{F}| - 1) = \frac{|\mathbb{F}| - 1/2}{|\mathbb{F}| - 1} \cdot \mu_0 = (1 + \epsilon) \mu_0,
\]

where \( \epsilon = \frac{1}{2|\mathbb{F}| - 2} \). In other words, the density of \( A \) on \( W \) is \( (1 + \epsilon) \) larger than its density over the entire space. We continue restricting \( A \) to co-dim one subspaces (updating \( \mu \) and considering \( \rho(\mu) \) at each step) until after at most \( t = \log \frac{|\mathbb{F}| - 1/2}{|\mathbb{F}| - 1} (1/2\mu_0) \) steps we reach one of two possibilities. Either we
Lemma 2.2.7. The following lemma is immediate.

Theorem where

Definition 2.2.6. A typical example of a subadditive function will be the rank of a quadratic polynomial. 

Proof. Let \( r_k(x) \) be the number of representations of \( x \) as \( a_1 + \ldots + a_k - a'_1 - \ldots - a'_k \) where the \( a_i \)-s and \( a'_i \)-s are from \( B \). Clearly, \( r_k(x) \) is equal to the sum, over all \( (y_1, \ldots, y_k, z_1, \ldots, z_k-1) \in B^{2k-1} \), of \( A(y_1, \ldots, y_k) \cdot A(z_1) \cdot \ldots \cdot A(z_{k-1}) \cdot A(y_1 + \ldots + y_k - z_1 - \ldots - z_{k-1} - x) \). Writing the Fourier expansion \( A \) and using routine calculations we conclude that

\[
    r_k(x) = |F|^{2k-1}n \cdot \sum_{\alpha} |\hat{B}(\alpha)|^{2k} \chi_\alpha(x) > |F|^{2k-1}n \cdot \left( \hat{B}(0)^{2k} - \sum_{\alpha \neq 0} |\hat{B}(\alpha)|^{2k} \right) \geq |F|^{2k-1}n \cdot \left( \hat{B}(0)^{2k} - |\hat{B}(0)|^{2k} \sum_{\alpha} |\hat{B}(\alpha)|^{2k} \right) = |F|^{2k-1}n \cdot \left( \hat{B}(0)^{2k} - \rho(\mu)^{2k-2} \sum_{\alpha} |\hat{B}(\alpha)|^{2k} \right) = |F|^{2k-1}n \cdot \left( \hat{B}(0)^{2k} - \rho(\mu)^{2k-2} \right) > 0 ,
\]

where the last inequality follows from the choice of \( k \) (we also used the fact that \( A \) is a \( 0/1 \) function). In particular, \( V \subseteq kA - kA \) as needed.

We will mainly apply the lemma on sets \( A \subseteq \mathbb{F}^n \) containing all directions where the partial derivatives of our underlying polynomial \( f \) are either very biased or have a high Gowers norm. More generally we define the notion of a subadditive function below.

Definition 2.2.4. Let \( V \subseteq \mathbb{F}^n \) be a linear space. \( F : V \rightarrow \mathbb{R}^+ \) is a subadditive function if for every \( u, v \in V \) and \( \alpha \in \mathbb{F} \) it holds that \( F(\alpha \cdot u + v) \leq F(u) + F(v) \).

Lemma 2.2.5. Let \( F : U \rightarrow \mathbb{R}^+ \) be a subadditive function. Define, \( A_r \triangleq \{ x \in U \mid F(x) \leq r \} \). If \( |A_r| \geq \mu|U| \), then there exists a vector space \( V \) of co-dimension at most \( \log(1/2)(1/2\mu) = O(\log(1/\mu)) \) such that for every \( y \in V \) it holds that \( F(y) \leq 2r \cdot \left[ \frac{1}{2} \left( \log(1/|F|^{1/2} (2/\mu) + 2) \right) \right] + 2r = O(r \log(1/\mu)) \).

Proof. The proof is immediate from Lemma 2.2.3. Let \( V \) be the subspace guaranteed by the lemma when applied on \( A_r \). As \( V \subseteq kA_r - kA_r \), for \( k \leq \max(1, \left[ \frac{1}{2} \left( \log(1/|F|^{1/2} (2/\mu) + 2) \right) \right] ) \), we get that \( F(y) \leq 2kr \) for every \( y \in V \).

A typical example of a subadditive function will be the rank of a quadratic polynomial.

Definition 2.2.6. Let \( q \) be a degree two function over a prime field \( \mathbb{F} \). We define \( \text{rank}_2(q) = r \), where \( r \) is the number of \( \alpha_i \)-s that are non zero when considering the canonical representation of \( q \) in Theorem 2.1.2.

The following lemma is immediate.

Lemma 2.2.7. For two quadratic polynomials \( q, q' \) and a constant \( \alpha \in \mathbb{F} \) we have that \( \text{rank}_2(q + \alpha q') \leq \text{rank}_2(q) + \text{rank}_2(q') \).

A more interesting example is given in the following lemma.

Lemma 2.2.8. Let \( f \) be a cubic polynomial over a prime field \( \mathbb{F} \). For every \( y \in \mathbb{F}^n \) define \( F(y) = \text{rank}_2(D_y(f)) \). Then \( F \) is a subadditive function.
Proof. The proof follows from the following simple observation

\[\Delta_y(f) + \Delta_z(f) = f(x + y) - f(x) + f(x + z) - f(x) = f(x + y + z) - f(x) - (f(x + y) - f(x + y) - f(x + z) - f(x)) = \Delta_y(z)(f)(x) - (\Delta_z(f)(x + y) - \Delta_z(f)(x)) = \Delta_y(z)(f)(x) + \Delta_y\Delta_z(f)(x).\]

Indeed, we now get that \( F(y + z) = \text{rank}_2(\Delta_y(z)(f)) = \text{rank}_2(\Delta_y(f) + \Delta_z(f) + \Delta_y\Delta_z(f)(x)) = \text{rank}_2(\Delta_y(f) + \Delta_z(f)) \leq \text{rank}_2(\Delta_y(f)) + \text{rank}_2(\Delta_z(f)) = F(y) + F(z), \) where we used the fact that adding a linear function to a quadratic polynomial does not change its rank. \(\square\)

2.3 The structure of cubic polynomials

In this section we prove Theorems 2.1.5 and 2.1.6. As described in Section 2.1.2 both proofs are based on finding a structure for the space of partial derivatives of \(f\).

2.3.1 Restricting the polynomial to a ‘good’ subspace

In this section we show that if a cubic \(f\) is biased or have a high \(U^3\) norm then there is a subspace \(V \subseteq \mathbb{F}^n\) such that for every \(y \in V\) the rank of \(\Delta_y(f)\) is relatively small. We start by showing that if \(f\) is biased or has a high Gowers norm then so do many of its partial derivatives. The following lemmas are well known and we prove them here for completeness.

Lemma 2.3.1. Let \(f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p\) be such that \(\text{bias}(f) = \delta\). Then a fraction of at least \(\frac{1}{2}\delta^2\) of the partial derivatives \(\Delta_y(f)\) satisfy \(\text{bias}(\Delta_y(f)) \geq \frac{1}{2}\delta^2\).

Proof. We first compute the expected bias of a partial derivative with respect to a random direction.

\[
\mathbb{E}_{y \in \mathbb{F}^n}[\text{bias}(\Delta_y(f))] = \mathbb{E}_{y \in \mathbb{F}^n}\left[\mathbb{E}_{x \in \mathbb{F}^n}\left[\omega \Delta_y(f)(x)\right]\right] \geq \mathbb{E}_{y \in \mathbb{F}^n}\left[\mathbb{E}_{x \in \mathbb{F}^n}\left[\omega f(x+y) - f(x)\right]\right] = \mathbb{E}_{y \in \mathbb{F}^n}\left[\mathbb{E}_{x \in \mathbb{F}^n}\left[\omega f(z)\right]\right] = \mathbb{E}_{x \in \mathbb{F}^n}\left[\mathbb{E}_{x \in \mathbb{F}^n}\left[\omega f(z)\right]\right] = \delta \cdot \delta = \delta^2.
\]

Therefore, by the fact that \(\text{bias}(f) \leq 1\), it follows that

\[
\text{Pr}_{y \in \mathbb{F}^n}\left[\text{bias}(\Delta_y(f)) > \frac{1}{2}\delta^2\right] > \frac{1}{2}\delta^2.
\]

\(\square\)

A similar result holds when \(f\) has a high \(U^d\) norm.

Lemma 2.3.2. Let \(f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p\) be such that \(\|f\|_{U^d} = \delta\). Then a fraction of at least \(\frac{1}{2}\delta^{2d}\) of the partial derivatives \(\Delta_y(f)\) satisfy \(\|\Delta_y(f)\|_{U^{d-1}} \geq \frac{1}{2}\delta^2\).

Proof. The proof is again immediate from the definition.

\[
\delta^{2d} = \|f\|_{U^d}^{2d} = \mathbb{E}_{x, y_1, \ldots, y_d}[\omega^{\Delta_y f}](x) \leq \mathbb{E}_{y_d}[\mathbb{E}_{x, y_1, \ldots, y_{d-1}}[\omega^{\Delta_y f}]] = \mathbb{E}_{y}[\|\Delta_y f\|_{U^{d-1}}^{2d-1}] .
\]
As before we get that 

$$\Pr_{y \in \mathbb{F}^n} \left[ \| \Delta_y(f) \|_{U^d-1} > \frac{1}{2} \delta^2 \right] > \frac{1}{2} \delta^{2d}.$$ 

We thus see that in both cases a constant fraction of all partial derivatives of $f$ have high bias or high $U^2$ norm. From Theorem 2.1.2 we get that if a partial derivative (which is a quadratic function) has a high bias then it depends on a few linear functions.

**Lemma 2.3.3.** Let $q$ be a quadratic polynomial over a prime field $\mathbb{F}$. Then $q$ is a function of at most $\log_{|\mathbb{F}|}(\text{bias}(q)) + 1$ linear functions. More accurately, in the notations of Theorem 2.1.2 the number of non zero $\alpha_i$-s is at most $\log_{|\mathbb{F}|}(1/\text{bias}(q))$.

**Proof.** See e.g. Lemmas 15-17 of [BV07].

The next lemma of Bogdanov and Viola [BV07] shows that a similar result holds when a partial derivative has a high $U^2$ norm.

**Lemma 2.3.4.** (Lemma 15 of [BV07]) Every quadratic polynomial $q$ over a prime field $\mathbb{F}$ is a function of at most $\log_{|\mathbb{F}|}(1/\|q\|_{U^2}) + 1$ linear functions. Further, in the notations of Theorem 2.1.2 the number of non zero $\alpha_i$-s is at most $\log_{|\mathbb{F}|}(1/\|q\|_{U^2})$.

Concluding, we have proved the following lemma (recall Definition 2.2.6).

**Lemma 2.3.5.** Let $f$ be a cubic polynomial.

1. If $\text{bias}(f) = \delta$, then for (at least) a $\frac{\delta^2}{2}$ fraction of $y \in \mathbb{F}^n$ it holds that $\text{rank}_2(\Delta_y(f)) \leq \log_{|\mathbb{F}|}(\frac{2}{\delta^2})$.

2. If $\|f\|_{U^2} = \delta$, then for (at least) a $\frac{\delta^4}{2}$ fraction of $y \in \mathbb{F}^n$ it holds that $\text{rank}_2(\Delta_y(f)) \leq \log_{|\mathbb{F}|}(\frac{2}{\delta^4})$.

We now combine Lemma 2.2.8 with Lemma 2.3.5 and Lemma 2.2.5 and obtain the following corollary.

**Corollary 2.3.6.** Let $f$ be a cubic polynomial. If $\text{bias}(f) = \delta$ or $\|f\|_{U^2} = \delta$, then there exists a subspace $V \subseteq \mathbb{F}^n$ such that $\dim(V) \geq n - O(\log(\frac{1}{\delta}))$ and such that for every $y \in V$ it holds that $\text{rank}_2(\Delta_y(f)) = O(\log^2(\frac{1}{\delta}))$.

### 2.3.2 The structure of low rank spaces

So far we have established the existence of a subspace $V \subseteq \mathbb{F}^n$ such that for every $y \in V$ it holds that $\text{rank}_2(\Delta_y(f)) = O(\log^2(\frac{1}{\delta}))$. We now show that such spaces of low rank polynomials have a very restricted structure. Namely, there exist $r = O(\log^2(\frac{1}{\delta}))$ linear functions $\ell_1, \ldots, \ell_r$ such that every $\Delta_y(f)$ can be written as $\Delta_y(f) = \sum_{i=1}^r \ell_i \cdot \ell_i^{(y)} + \ell_0^{(y)}$, where the $\ell_i^{(y)}$-s are linear functions determined by $y$. The intuition behind this result is that $\text{rank}_2(q+q')$ can be much smaller than $\text{rank}_2(q)+\text{rank}_2(q')$ only if there is some basis with respect to which $q$ and $q'$ share many linear functions when represented in the form of Theorem 2.1.2. From this observation we deduce that if we consider some function of maximal rank, $q = \sum_{i=1}^r \ell_i \cdot \ell_i'$, and set $\{\ell_i, \ell_i'\}_{i=1}^r$ to zero (namely, consider the subspace on which they all vanish), then on this subspace the rank of the remaining quadratic functions decreases by a factor of two. Repeating this argument we get that after setting at most $4r$ linear functions to zero, all our quadratic functions become linear functions.

**Lemma 2.3.7.** Let $M$ be a linear space of quadratic functions satisfying $\text{rank}_2(p) \leq r$ for all $p \in M$. Then there exists a subspace $V \subseteq \mathbb{F}^n$ of co-dimension $\leq 4r$ such that $p|_V$ is a linear function for all $p \in M$. 

15
We shall give the proof for the case $F = \mathbb{F}_2$. The proof for odd characteristics is very similar (except that in the odd characteristic case we have that the co-dimension of $V$ is $2r$ whereas in the even characteristic case it is $4r$).

**Proof.** Let $g \in M$ be such that $\text{rank}_2(g) = r$. By Theorem 2.1.2, $g$ can be expressed as $g = \sum_{i=1}^r \ell_{2i-1} \cdot \ell_{2i} + \ell_0$. Denote $V \overset{\Delta}{=} \{x \mid \ell_1(x) = \ell_2(x) = \ldots = \ell_{2r}(x) = 0\}$. We now show that for every $h \in M$ it holds that $\text{rank}_2(h|_V) \leq \frac{r}{2}$. Repeating this argument we get that after setting at most $2r + 2(r/2) + 2(r/4) + \ldots \leq 4r$ linear functions to zero, the rank of all the quadratic functions in $M$ became zero.

Pick some $h \in M$ and denote $\text{rank}_2(h|_V) = s$. As before, $h|_V$ can be expressed as $h|_V = \sum_{i=1}^s m_{2i-1} \cdot m_{2i} + m_0$ (where the $m_i$-s are linear functions). Clearly the functions $\{\ell_1, \ldots, \ell_{2r}, m_1, \ldots, m_2s\}$ are linearly independent. We can therefore write

$$h = \sum_{i=1}^s m_{2i-1} \cdot m_{2i} + m_0 + \sum_{i=1}^{2r} \ell_i \cdot L_i,$$

where the $L_i$-s are linear functions. Write $L_i = \tilde{m}_i + \tilde{\ell}_i + \tilde{L}_i$ where $\tilde{m}_i \in \text{span}\{m_0, \ldots, m_2s\}$, $\tilde{\ell}_i \in \text{span}\{\ell_0, \ldots, \ell_{2r}\}$ and $\tilde{L}_i$ is linearly independent of the $m_j$-s and $\ell_j$-s. Rearranging terms we get that

$$h = \sum_{i=1}^s (m_{2i-1} + \ell_{2i-1} \cdot (m_{2i} + \ell_{2i})) + (m_0 + \ell_0) + \tilde{h}(\ell_0, \ldots, \ell_{2r}, \tilde{L}_1, \ldots, \tilde{L}_{2r}),$$

where each $\ell_i$ is in the span of the $\ell_i$-s and $\tilde{h}$ is a quadratic polynomial. Denote $m'_i = m_i + \ell'_i$. It is clear that $\ell_0, \ldots, \ell_{2r}, \tilde{L}_1, \ldots, \tilde{L}_{2r}$ are linearly independent of the $m'_i$-s (and vice versa). Consequently,\(^4\)

$$\text{rank}_2(\sum_{i=1}^s m'_{2i-1} \cdot m'_{2i} + m_0) + \text{rank}_2(\tilde{h}) = \text{rank}_2(h) \leq r.$$ 

Hence, $\text{rank}_2(\tilde{h}) \leq r - s$. We now get that

$$r \geq \text{rank}_2(g + h) = \text{rank}_2 \left( \sum_{i=1}^s m'_{2i-1} \cdot m'_{2i} + m_0 + \tilde{h}(\ell_0, \ldots, \ell_{2r}, \tilde{L}_1, \ldots, \tilde{L}_{2r}) + g \right) = \text{rank}_2 \left( \sum_{i=1}^s m'_{2i-1} \cdot m'_{2i} + m_0 \right) + \text{rank}_2 \left( g + \tilde{h}(\ell_0, \ldots, \ell_{2r}, \tilde{L}_1, \ldots, \tilde{L}_{2r}) \right) \geq s + (r - (r - s)) = 2s,$$

where we used the fact that $\text{rank}_2(g + \tilde{h}) \geq \text{rank}_2(g) - \text{rank}_2(\tilde{h})$. As we showed that $r \geq 2s$ the proof is completed. \[\square\]

### 2.3.3 Completing the proofs

We are no ready to complete the proofs of Theorems 2.1.5 and 2.1.6.

**Proof of Theorem 2.1.6.** By Corollary 2.3.6 we get that if $\|f\|_{U^2} = \delta$, then there exists a subspace $V \subseteq \mathbb{F}^n$ such that $\dim(V) \geq n - O(\log(1/\delta))$ and such that for every $y \in V$ it holds that $\text{rank}_2(\Delta_y(f)) = O(\log^2(\frac{1}{\delta}))$. Lemma 2.3.7 implies that there are at most $r = O(\log^2(\frac{1}{\delta}))$ linear functions $\ell_1, \ldots, \ell_r$ such that for every $y \in V$ we have that $\Delta_y(f) = \sum_{i=1}^r \ell_i \cdot \ell_0^{(y)} + \ell_0^{(y)}$. Let $U = \{x \in V \mid \ell_1(x) = \ldots = \ell_r(x) = 0\}$. Then $U$ is a linear space of dimension $\dim(U) \geq n - O(\log^2(\frac{1}{\delta}))$. For every $y \in U$ we have that $\Delta_y(f)|_U = \ell_0^{(y)}|_U$. Hence, for every $y \in U$, $\deg(\Delta_y(f)) \leq 1$. Therefore, $\deg(f|_U) \leq 2$. Let $\ell'_1, \ldots, \ell'_t$ be linearly independent linear functions such that $x \in U$ iff $\ell'_1(x) = \ldots = \ell'_t(x) = 0$. It follows that we can write $f = \sum_{i=1}^t \ell'_i \cdot q_i + q_0$ for some quadratic polynomials $\{q_i\}$. As $t = n - \dim(U) = O(\log^2(\frac{1}{\delta}))$ the result follows. \[\square\]

\(^4\)From Theorem 2.1.2 it is clear that for quadratic polynomials $q_1, q_2$ it holds that $\text{rank}_2(q_1(\hat{x}) + q_2(\hat{y})) = \text{rank}_2(q_1(\hat{x})) + \text{rank}_2(q_2(\hat{y}))$. 

16

Technion - Computer Science Department - Ph.D. Thesis PHD-2014-13 - 2014
The proof of Theorem 2.1.5 is essentially the same except that we make another small optimization that reduces the required number of quadratic functions.

**Proof of Theorem 2.1.5.** By the same argument as above we get that \( f = \sum_{i=1}^{t} \ell_i \cdot q_i + q_0 \) for some quadratic polynomials \( \{q_i\} \) and linear functions \( \{\ell_i\} \), where \( t = O(\log^2(\frac{1}{\delta})) \). For convenience we shall assume w.l.o.g. that

\[
f = \sum_{i=1}^{t} x_i \cdot q_i + q_0 .
\] (2.1)

The following lemma shows that by adding a few more linear functions we can assume that no nontrivial linear combination of the \( q_i \)'s has a low rank.

**Lemma 2.3.8.** Let \( q_1, \ldots, q_t \) be quadratic polynomials over \( \mathbb{F}^n \). Then, for every \( r \) there exist a subspace \( V \subset \mathbb{F}^n \) of dimension \( \dim(V) \geq n - t(r + 1) \), and \( t' \leq t \) indices \( i_1, \ldots, i_{t'} \) such that for every affine shift \( V' \) of \( V \) the following holds

1. For all \( i \), \( q_i|_{V'} \in \text{span}\{1, q_{i_1}|_{V'}, \ldots, q_{i_{t'}}|_{V'}\} \).

2. For any non trivial linear combination we have that \( \text{rank}_2 \left( \sum_{j=1}^{t'} \alpha_j q_j|_{V'} \right) > r \).

**Proof.** The proof is by induction on \( t \). For \( t = 1 \) the claim is clear: If \( \text{rank}_2(q_1) > r \) then we are done. Otherwise we have \( q_1 = \sum_{i=1}^{r} \ell_{2i-1} \ell_{2i} + \ell_0 \). Letting \( V = \{x \mid \ell_0(x) = \ell_2(x) = \ldots = \ell_{2r}(x) = 0\} \) the claim follows (indeed notice that passing to an affine shift of \( V \) simply means fixing the \( \ell_i \)'s to arbitrary values). Assume now that we have \( q_1, \ldots, q_t \) and that (w.l.o.g.) \( \text{rank}_2 \left( q_t + \sum_{i=1}^{t-1} \alpha_i q_i \right) \leq r \).

Write \( q_t + \sum_{i=1}^{t-1} \alpha_i q_i = \sum_{i=1}^{r} \ell_{2i-1} \ell_{2i} + \ell_0 \). Set \( V = \{x \mid \ell_0(x) = \ell_2(x) = \ldots = \ell_{2r}(x) = 0\} \). Then, \( q_t|_V \in \text{span}\{q_1|_V, \ldots, q_{t-1}|_V\} \). As \( \dim(V) = n - (r + 1) \) the claim follows by applying the induction argument to \( q_1|_V, \ldots, q_{t-1}|_V \) (again the claim about any affine shift follows easily).

We continue with the proof of the theorem. Having Equation (2.1) in mind we set \( U = \{(0, \ldots, 0, x_{t+1}, \ldots, x_n)\} \subset \mathbb{F}^n \). Applying Lemma 2.3.8 on \( q_1|_V, \ldots, q_t|_V \) with \( r = \log_{|\mathbb{F}|}(2/\delta) \) we get that there is a subspace \( W \subset U \) and \( t' \leq t \) such that: \( \dim(W) \geq \dim(U) - (r + 1)t \geq n - (r + 2)t = n - O(\log^3(\frac{1}{\delta})) \); w.l.o.g. for every \( i = 1 \ldots t \), \( q_i|_W \in \text{span}\{q_1|_W, \ldots, q_t|_W\} \); any nontrivial linear combination of \( q_1|_W, \ldots, q_t|_W \) has rank larger than \( r \). By applying an invertible linear transformation\(^5\) we can further assume that \( W = \{x \in \mathbb{F}^m \mid x_1 = \ldots = x_m = 0\} \) for some \( m \leq (r + 2)t \). For \( i = 1 \ldots t' \) let \( q'_i = q_i|_W \). Note that \( q'_i \) does not contain any of the variables \( x_1, \ldots, x_m \). We can rewrite Equation (2.1) as\(^6\)

\[
f = \sum_{i=1}^{t'} \ell'_i q'_i + \sum_{i=1}^{t} m_{i,j} x_i x_j \ell_{i,j} ,
\] (2.2)

where the \( \ell'_i \)'s are linearly independent linear functions in \( x_1, \ldots, x_t \). We now show that \( t' < \log_{|\mathbb{F}|}(2/\delta) \). Assume for contradiction that \( t' \geq \log_{|\mathbb{F}|}(2/\delta) \). As

\[
\text{bias}(f) = \mathbb{E}_{\alpha_1, \ldots, \alpha_{t'}} \left[ \text{bias}(f(x_1, \ldots, x_n) | (\ell'_1, \ldots, \ell'_{t'}) = (\alpha_1, \ldots, \alpha_{t'}) \right] ,
\]

there exists an assignment \( (x_1, \ldots, x_m) = (\beta_1, \ldots, \beta_m) \) satisfying \( (\ell'_1, \ldots, \ell'_{t'}) = (\alpha_1, \ldots, \alpha_{t'}) \neq 0 \) such that

\[
\text{bias} \left( \sum_{i=1}^{t'} \alpha_i q'_i + \sum_{i=1}^{t} m_{i,j} \beta_j x_i x_j \right) \geq \delta - \frac{1}{|\mathbb{F}|^{t'}} \geq \delta/2 .
\]

\(^5\)This step is not really required but we continue using it just to make the proofs easier to read.

\(^6\)We will later explain why \( q_0 \) ‘disappeared’ from this expression.
Therefore, for some constants \(\alpha_1, \ldots, \alpha_{t'}\) (where not all \(\alpha_1, \ldots, \alpha_{t'}\) are zero) we have that

\[
\text{bias} \left( \sum_{i=1}^{t'} \alpha_i q_i' + \ell \right) \geq \delta / 2 ,
\]

for some linear function \(\ell\). By Lemma 2.3.3 we get that

\[
\text{rank}_2 \left( \sum_{i=1}^{t'} \alpha_i q_i' \right) = \text{rank}_2 \left( \sum_{i=1}^{t'} \alpha_i q_i' + \ell \right) \leq \log_{|F|}(1/(\delta/2)) = r,
\]

in contradiction to the choice of \(q_1', \ldots, q_{t'}\).

To complete the proof we explain the reason for dropping \(q_0\). Indeed, consider Equation (2.1). Let \(U = \{x \mid x_1 = \ldots = x_t = 0\}\). Set \(\tilde{q}_i = q_i|_U\). Then we can rewrite (2.1) as \(\sum_{i=1}^{t} x_i \tilde{q}_i + \tilde{q}_0 + \sum_{i=1}^{t} x_i \sum_{j=1}^{t} x_j \ell_{i,j}\), for some linear functions \(\ell_{i,j}\). Now, for some \(\alpha_1, \ldots, \alpha_t\) we get that

\[
\text{bias} \left( \sum_{i=1}^{t} \alpha_i \tilde{q}_i + \tilde{q}_0 \right) \geq \delta.
\]

Lemma 2.3.3 implies that \(\text{rank}_3(\sum_{i=1}^{t} \alpha_i \tilde{q}_i + \tilde{q}_0) \leq \log_{|F|}(1/\delta)\) and so we can replace \(\tilde{q}_0\) by a linear combination of the other \(\tilde{q}_i\)-s and a function depending on a few linear functions. By passing to a (possibly affine) subspace of dimension at least \(n - \log_{|F|}(1/\delta) - 1\) we get a representation for \(f\) without \(q_0\). This operation increases \(t'\) in Equation (2.2) by no more than \(\log_{|F|}(1/\delta) + 1\) and so we are done.

\[
\square
\]

2.4 The structure of biased 4 degree polynomials

In this section we prove Theorem 2.1.7 on the structure of biased degree 4 polynomials. As in the case of cubic polynomials, we shall focus our attention on a subspace on which all of derivatives have a small rank (a cubic polynomial is of low rank if it depends on a small number of linear and quadratic functions). By a lemma of Bogdanov and Viola [BV07] (Lemma 2.4.3) we get that \(f\) can be well approximated by a function of a small number of its derivatives (which in our case, are all of low rank). Thus, \(f\) is well approximated by a function of a few linear and quadratic polynomials. By passing to a subspace we can assume that \(f\) is well approximated by a function of a small number of quadratic polynomials. Lemma 2.3.8 implies that (possibly on a slightly smaller subspace) \(f\) can be well approximated by a function of a small number of quadratics, that every nontrivial linear combination of them has a high rank. We then show that in this case those quadratic functions are in fact strongly regular (a notion that we later explain) and therefore by a theorem of Kaufman and Lovett [KL08], \(f\) in fact equals a function in those quadratic (on the subspace). We then finish the proof by showing that in this case \(f\) also have a nice structure.

2.4.1 Restricting the polynomial to a ‘good’ subspace

In this section we prove an analogous result to Corollary 2.3.6. We first define the rank of a cubic polynomial.

Definition 2.4.1. Let \(g\) be a degree three polynomial. We define \(\text{rank}_3(g)\) to be the minimal integer \(r\) for which there are \(r\) linear functions \(\ell_1, \ldots, \ell_r\) and \(r + 1\) quadratic functions \(q_0, \ldots, q_r\) such that \(g = \sum_{i=1}^{r} \ell_i q_i + q_0\).

Lemma 2.4.2. Let \(f\) be a degree four polynomial satisfying \(\text{bias}(f) = \delta\). Then there exist a linear subspace \(V \subseteq \mathbb{F}^n\) of dimension \(\dim(V) \geq n - O\left(\log_{|F|}(1/\delta)\right)\), such that for every \(y \in V\) \(\text{rank}_3(\Delta_y(f)) = \log^{O(1)}(1/\delta)\).
Proof. As before, define $\mathcal{F}(y) \triangleq \text{rank}_3(\Delta_y(f))$. It is again not difficult to see that $\mathcal{F}$ is a subadditive function. By Lemma 2.3.1 we get that there is a subset $S \subseteq \mathbb{F}^n$ of size $\frac{|S|^2}{2^2} \cdot \mathbb{F}^n$ such that for all $y \in S$, $\text{bias}(\Delta_y(f)) \geq \frac{\delta^2}{2}$. Theorem 2.1.5 implies that for every $y \in S$ it holds that $\text{rank}_3(\Delta_y(f)) = O(\log^4(\frac{1}{\delta}))$. From Lemma 2.2.5 it follows that there is a linear subspace $V \subseteq \mathbb{F}^n$ with $\text{dim}(V) \geq n - O(\log_{|\mathbb{F}|}(1/\delta))$, such that for every $y \in V$ $\text{rank}_3(\Delta_y(f)) = O(\log^5(\frac{1}{\delta}))$. □

By applying an invertible linear transformation we can assume that $V = \{x : x_1 = \ldots = x_m = 0\}$ for some $m = O(\log_{|\mathbb{F}|}(1/\delta))$. We now have

$$f = \sum_{i=1}^{m} x_i g_i + f', \quad (2.3)$$

where $f' = f'(x_{m+1}, \ldots, x_n)$. Moreover, by Lemma 2.4.2 it follows that for every $y = (0, \ldots, 0, y_{m+1}, \ldots, y_n)$, $\text{rank}_3(\Delta_y(f)) = O(\log^5(\frac{1}{\delta}))$. Notice that for every such $y$ it holds that

$$\Delta_y(f) = \sum_{i=1}^{m} x_i \Delta_y(g_i) + \Delta_y(f').$$

Hence, $\text{rank}_3(\Delta_y(f')) \leq \text{rank}_3(\Delta_y(f)) + m$. We now fix some value to $x_1, \ldots, x_m$, such that $\text{bias}(f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n)) \geq \delta$. Let

$$\tilde{f}(x_{m+1}, \ldots, x_n) \triangleq f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n). \quad (2.4)$$

It follows that $\text{bias}(\tilde{f}) \geq \delta$ and that for every $y = (y_{m+1}, \ldots, y_n)$, $\text{rank}_3(\Delta_y(\tilde{f})) = \text{rank}_3(\Delta_y(f')) = O(\log^5(\frac{1}{\delta}))$ (note that $\deg(\Delta_y(\tilde{f}) - \Delta_y(f')) = 2$ so they have the same rank). From now on we will only consider $\tilde{f}$ and not $f$. Observe that if we prove Theorem 2.1.7 for $\tilde{f}$ then by considering Equations (2.3) and (2.4) we get the required result for $f$ itself.

### 2.4.2 Computing $\tilde{f}$ using a few quadratics

We now show that there is a large subspace on which $f$ can be approximated by a function of a few quadratic polynomials. The following lemma of Bogdanov and Viola shows that if $f$ is biased then it can be well approximated by a small set of partial derivatives.

**Lemma 2.4.3.** (Lemma 2.4 from [BV07]) Let $f : \mathbb{F}^n \to \mathbb{F}$ be a function over a finite field $\mathbb{F}$ with bias($f$) = $\delta$. Then there are $t$ directions $a_1, \ldots, a_t$ and a function $H$ such that $H(\Delta_{a_1}(f), \ldots, \Delta_{a_t}(f))$ $\epsilon$-approximates $f$, where $t \leq (1 + \log \frac{1}{\epsilon}) (|\mathbb{F}|/\delta)^{O(1)}$.

By the construction of $\tilde{f}$ we know that each of its partial derivatives is of rank $\log^{O(1)}(1/\delta)$ and that bias($\tilde{f}$) $\geq \delta$. Thus, Lemma 2.4.3 guarantees that $\tilde{f}$ can be well approximated using a few quadratics.

**Corollary 2.4.4.** For every $\epsilon > 0$ there are $c = (1 + \log \frac{1}{\epsilon}) (|\mathbb{F}|/\delta)^{O(1)}$ quadratic polynomials $Q_1, \ldots, Q_c$ and a function $H$ such that $\tilde{f}$ is $\epsilon$-approximated by $H(Q_1, \ldots, Q_c)$.

The next lemma, which is the main lemma of [KL08] shows that if the approximation is good enough (i.e. $\epsilon$ is small), and if the quadratics satisfy the strong regularity property then $\tilde{f}$ can in fact be computed by a small number of quadratics.
Definition 2.4.5. (strongly regular quadratic functions) We say that a family of quadratic functions \( \{Q_i\}_{i=1}^m \) is \( \gamma \)-strongly regular if the following holds for every \( x_0 \in \mathbb{F}^n \): for independent uniform random variables \( Y_1, \ldots, Y_5 \) the joint distribution of
\[
\left\{ Q_j \left( x_0 + \sum_{i \in I} Y_i \right) \mid j \in [m], I \subseteq [5], 1 \leq |I| \leq 2 \right\}
\]
is \( \gamma \) close to the uniform distribution (recall Definition 2.2.1).

This definition is a restricted version of Definition 8 of [KL08] for quadratic polynomials. The interested reader is referred to that paper for the general definition for higher degree polynomials.

Lemma 2.4.6. (Lemma 13 from [KL08]) Let \( f(x) \) be a degree \( d \) polynomials, \( h_1, \ldots, h_m \) polynomials of degree less than \( d \) and \( H : \mathbb{F}^m \to \mathbb{F} \) a function such that
- \( H(h_1, \ldots, h_m) \) \( \epsilon \)-approximates \( f \) where \( \epsilon \leq 2^{-2(d+1)} \).
- \( \{h_i\}_{i=1}^m \) is a \( \gamma \)-strongly regular family where \( \gamma \leq \min\{2^{-d}, 2^{-m}\} \).

Then there exists a function \( F : \mathbb{F}^m \to \mathbb{F} \) such that \( f = F(h_1, \ldots, h_m) \).

In other words, the lemma says that if \( f \) is well approximated by a family of strongly regular functions then it can actually be computed everywhere by the functions in the family. We shall now show that if \( q_1, \ldots, q_c \) are quadratic polynomials such that the rank of every nontrivial linear combination of them is high, then they are strongly regular. This will imply (by Corollary 2.4.4) that \( f \) is a function of a few quadratics and therefore so is \( f \).

Lemma 2.4.7. Let \( \{Q_i\}_{i=1}^m \) be a family of quadratic functions such that for every nontrivial linear combination, \( \text{rank}_2(\sum_{i=1}^m a_i Q_i) \geq R \). Then \( \{Q_i\}_{i=1}^m \) is a \( \gamma \)-strongly regular family for \( \gamma = |\mathbb{F}|^{3m/2 - R/4} \).

Proof. The proof is based on the analogy between quadratic functions and matrices.

Definition 2.4.8. Let \( Q : \mathbb{F}^n \to \mathbb{F} \) be a quadratic polynomial and \( A \in \mathbb{F}^{n \times n} \) an \( n \times n \) matrix. We say that \( A \) represents \( Q \) if there exists a linear function \( \ell \) such that \( Q(x) = x^t A x + \ell(x) \).

Notice that there may be many different matrices representing the same polynomial \( Q \). For example, every antisymmetric matrix represents the zero function. More generally, if \( S \) is antisymmetric then \( A + S \) represent the same polynomial.

Lemma 2.4.9. Let \( q \) be a quadratic polynomial. Then \( \text{rank}_2(q) \) (recall Definition 2.2.6) is equal to the minimal rank of a matrix representing \( q \). Moreover, for every matrix \( A \) representing \( q \) we have that \( \text{rank}(A + A^t)/2 \leq \text{rank}_2(q) \leq \text{rank}(A + A^t) \).

We shall prove the lemma for \( \mathbb{F} = \mathbb{F}_2 \). The proof for other fields is similar.

Proof. Let \( \text{rank}_2(q) = r \). Then \( q \) can be expressed as \( \sum_{i=1}^r \left( \sum_{j=1}^n a_{i,j} x_j \right) \left( \sum_{j=1}^n b_{i,j} x_j \right) + \ell(x) \). Set \( A = (a_{i,j}), B = (b_{i,j}) \in \mathbb{F}^{r \times n} \). It is clear that \( A^t B \) represents \( q \) and that \( \text{rank}(A^t B) \leq r \). On the other hand, if \( q \) can be represented by a rank \( r \) matrix \( A \), then let \( \ell_1, \ldots, \ell_r \) be a basis for the rows of \( A \), when interpreted as linear functions.\(^7\) Let \( A_i \) be the \( i \)-th row of \( A \) and denote \( A_i = \sum_{j=1}^r a_{i,j} \ell_j \). We have that for some linear function \( \ell \),
\[
q - \ell = x^t A x = \sum_{i=1}^n x_i A_i(x) = \sum_{i=1}^n x_i \sum_{j=1}^r a_{i,j} \ell_j = \sum_{j=1}^r \ell_j \left( \sum_{i=1}^n a_{i,j} x_i \right) = \sum_{j=1}^r \ell_j \ell_j^t ,
\]
\(^7\)I.e. \((a_1, \ldots, a_n) \leftrightarrow \sum_{i=1}^n a_i \cdot x_i \).
where $\ell'_1, \ldots, \ell'_r$ are linear functions. This implies that $\text{rank}_2(q) \leq r$. Thus, $\text{rank}_2(q) = \min \{ \text{rank}(A) \mid q(x) = x^tAx + \ell(x) \}$.

To prove the second claim, let $A$ be any matrix representing $q$. We first change the basis of the space so that with respect to the new basis $q$ will have the form of Theorem 2.1.2. Let $T$ be an invertible matrix representing the change of basis. Clearly, $T^tAT$ represents $q \circ T = \sum_{i=1}^r x_{2i-1}x_{2i} + \ell$, where $r = \text{rank}_2(q)$. Thus, the matrix $T^tAT$ can be written as $D + S$ where $D$ is a block diagonal matrix consisting of $r$ nonzero blocks of size $2 \times 2$ and $S$ is a symmetric matrix. We also note that for each $2 \times 2$ diagonal block $C$ of $D$ it holds that $C + C^t \neq 0$. We thus get that

$$\text{rank}(A + A^t) = \text{rank}(T^t(A + A^t)T) = \text{rank}(D + S + D^t + S^t) = \text{rank}(D + D^t).$$

Now, for every $2 \times 2$ diagonal block $C$ of $D$ we have that $1 \leq \text{rank}(C + C^t) \leq 2$ and so

$$\text{rank}_2(q) = r \leq \text{rank}(D + D^t) \leq 2r = 2\text{rank}_2(q).$$

This completes the proof of the Lemma.\footnote{From the proof it actually follows that over $\mathbb{F}_2$, $\text{rank}_2(q) = \text{rank}(A + A^t)/2$ but this is not the case for other prime fields.}

We continue the proof of Lemma 2.4.7. Using the above observation we now prove that any nontrivial linear combination $\sum_{k \in [m], I \subseteq [5], 1 \leq |I| \leq 2} \alpha_{k,I}Q_j(x + \sum_{i \in I} Y_i)$ has high rank (as a quadratic polynomial in the variables $Y_1 \cup \ldots \cup Y_5$).

Fix $x = x_0$ and let $A_k$ be a matrix representing $Q_k$. Notice that the quadratic polynomial $Q_k(x_0 + \sum_{i \in I} Y_i)$ (in the variables $\cup_{i=1}^5 Y_i$) can be represented by a block matrix $B^{k,I} \in \mathbb{F}^{5n \times 5n}$. Indeed, consider a $5 \times 5$ matrix that has 1 in the $(i,j)$-position iff $i, j \in I$, and zeros otherwise. Now, replace any 1 by the matrix $A_k$ and every 0 by the $n \times n$ zero matrix. It is an easy calculation to see that this matrix represents $Q_k(x_0 + \sum_{i \in I} Y_i)$. We shall abuse notations and for $i,j \in I$ say that $(B^{k,I})_{i,j} = A_k$, and that otherwise $(B^{k,I})_{i,j} = 0$.

Clearly, the linear combination

$$Q' \triangleq \sum \{ \alpha_{k,I}Q_k(x + \sum_{i \in I} Y_i) \mid k \in [m], I \subseteq [5], 1 \leq |I| \leq 2 \}$$

is represented by the matrix

$$C \triangleq \sum \{ \alpha_{k,I}B^{k,I} \mid k \in [m], I \subseteq [5], 1 \leq |I| \leq 2 \}.$$ 

Observe that for $i \neq j \in [5]$, $C_{i,j} = \sum_{k \in [m]} \alpha_{k,\{i,j\}} A_k$. We now show that if for some $i \neq j \in [5]$ and $k \in [m]$ it holds that $\alpha_{k,\{i,j\}} \neq 0$ then the rank of $C^t + C$ (and hence of $Q'$) is high.

$$\text{rank}_2(Q') = \text{rank}_2 \left( \sum \{ \alpha_{k,I}Q_k(x + \sum_{i \in I} Y_i) \mid k \in [m], I \subseteq [5], 1 \leq |I| \leq 2 \} \right)$$

$$\geq \frac{1}{2} \text{rank}(C + C^t) \geq \frac{1}{2} \text{rank}(C_{i,j} + C_{j,i}^t)$$

$$= \frac{1}{2} \text{rank} \left( \sum_{k \in [m]} \alpha_{k,\{i,j\}} (A_k + A_k^t) \right)$$

$$\geq \frac{1}{4} \text{rank}_2 \left( \sum_{k \in [m]} \alpha_{k,\{i,j\}} Q_k \right) > \frac{1}{4} R.$$

21
If it is not the case, namely, for all \( i \neq j \in [5], k \in [m] \) \( \alpha_{k,\{i,j\}} = 0 \), then there is some \( i \in [5] \) and \( k \in [m] \) such that \( \alpha_{k,\{i\}} \neq 0 \) and we get that same result by considering \( C_{i,i} \) instead.

To conclude, every nontrivial linear combination of \( \{ Q_j(x + \sum_{i \in I} Y_i) \}_{k \in [m], j \subseteq [5], 1 \leq |I| \leq 2} \) has rank grater than \( \frac{1}{4} R \). Lemma 2.3.3 implies that the bias of every such linear combination is bounded by \( |\mathbb{F}|^{-R/4} \).

It now follows by Lemma 2.2.2 that the distribution is \( |\mathbb{F}|^{3m/2-R/4} \) close to the uniform distribution as needed.

We thus get the following corollary.

**Corollary 2.4.10.** Let \( g(x) \) be a degree \( d \) polynomials, \( q_1, \ldots, q_m \) quadratic polynomials and \( H : \mathbb{F}^m \to \mathbb{F} \) a function such that

- \( H(h_1, \ldots, h_m) \) \( \epsilon \)-approximates \( g \) where \( \epsilon \leq 2^{-2(d+1)} \).
- The bias of every non trivial combination of \( h_1, \ldots, h_m \) is \( |\mathbb{F}|^{-\Omega(m+d)} \).

Then there exists a function \( G : \mathbb{F}^m \to \mathbb{F} \) such that \( g = G(h_1, \ldots, h_m) \).

We now show that \( \tilde{f} \) can be computed by a few quadratics.

**Lemma 2.4.11.** Let \( g : \mathbb{F}^a \to \mathbb{F} \) be a quartic polynomial such that for every \( y \), \( \text{rank}_3(\Delta_y(f)) \leq \text{poly}(1/\delta) \). Then there exist a subspace \( W, c = \text{poly}(|\mathbb{F}|/\delta) \) quadratics \( q_1', \ldots, q_c' \) and a function \( G \) such that, \( \dim(W) = n - \text{poly}(|\mathbb{F}|/\delta) \) and \( g|_W = G(q_1', \ldots, q_c') \).

**Proof.** Applying Lemma 2.4.3, and using the fact that every partial derivative of \( g \) has a low rank, we conclude that for \( \epsilon = 2^{-20} \) there exist \( c = \text{poly}(|\mathbb{F}|/\delta) \) linear and quadratic functions, and a function \( H \), such that \( H(\ell_1, \ldots, \ell_c, q_1, \ldots, q_c) \) \( \epsilon \)-approximates \( g \). Let \( r = \text{poly}(|\mathbb{F}|/\delta) \) and \( U = \{ x : \ell_1(x) = \alpha_1, \ldots, \ell_c(x) = \alpha_c \} \) be some subspace such that \( H(\alpha_1, \ldots, \alpha_c, q_1|_U, \ldots, q_c|_U) \) \( \epsilon \)-approximates \( g|_U \). Applying Lemma 2.3.8 on \( q_1|_U, \ldots, q_c|_U \) and \( r \) we get that there exists a (possible affine) subspace \( W \subseteq U \) and \( c' \leq c \) such that: \( \dim(W) \geq \dim(U) - (r + 1)c \geq n - (r + 2)c = n - \text{poly}(|\mathbb{F}|/\delta) \); w.l.o.g. for every \( i = 1 \ldots c \), \( q_i|_W \in \text{span}\{q_1|_W, \ldots, q_c|_W\} \); any nontrivial linear combination of \( q_1|_W, \ldots, q_c|_W \) has rank larger than \( r \); \( g|_W \) is \( \epsilon \)-approximated by \( H(\ell_1|_W, \ldots, \ell_c|_W, q_1|_W, \ldots, q_c|_W) \) (this follows by picking an adequate shift of the linear space in the lemma). Hence, \( g|_W \) is \( \epsilon \)-approximated by \( H(\ell_1|_W, \ldots, \ell_c|_W, q_1|_W, \ldots, q_c|_W) = H'(q_1|_W, \ldots, q_c|_W) \) for some \( H' \). The reason for passing to \( W \) is that now any nontrivial linear combination of \( q_1|_W, \ldots, q_c|_W \) has rank larger than \( r \). We thus get by Corollary 2.4.10 that there is some function \( G \) such that \( g|_W = G(q_1|_W, \ldots, q_c|_W) \).

Recall that we assume w.l.o.g. that for every \( y \in \mathbb{F}^{n-m} \), \( \text{rank}_3(\Delta_y(\tilde{f})) \leq \text{poly}(1/\delta) \). Thus, the lemma above implies the following corollary.

**Corollary 2.4.12.** In the notations of the proof, there exist a subspace \( Z \subseteq \mathbb{F}^{n-m} \) of dimension \( \dim(Z) \geq n - \text{poly}(|\mathbb{F}|/\delta) \) such that \( \tilde{f}|_Z = F(q_1, \ldots, q_c) \), for \( c = \text{poly}(|\mathbb{F}|/\delta) \) quadratic polynomials and some function \( F \).

### 2.4.3 The structure of \( \tilde{f} \)

We now show that we can represent \( \tilde{f} \) as \( \tilde{f} = \sum_{i=1}^k \ell_i \cdot g_i + \sum_{i=1}^k q_i' \cdot q_i'' \), where \( k = \text{poly}(|\mathbb{F}|/\delta) \), the \( \ell_i \)-s are linear, the \( q_i' \)-s and \( q_i'' \)-s are quadratic and the \( g_i \)-s are cubic polynomials. For this we will transform the quadratic polynomials to be what we denote as disjoint polynomials.
Definition 2.4.13. **We say that the quadratic polynomials** $\{Q_i\}_{i=1}^m$ **are disjoint if there is a linear transformation** $T$, **2m variables** $\{x_i\}_{i=1}^m \cup \{y_i\}_{i=1}^m$, **where possibly for several i-s** $x_i = y_i$, **and quadratic functions** $\{Q'_i\}_{i=1}^m$ **such that for every** $k \in [m]$, $Q_k \circ T = x_ky_k + Q'_k$ **where no degree two monomial in** $Q'_k$ **contains a variable from** $\{x_i\}_{i=1}^m \cup \{y_i\}_{i=1}^m$.

Lemma 2.4.14. **Let** $q_1, \ldots, q_c$ **be quadratic polynomials from** $\mathbb{F}^n$ **to** $\mathbb{F}$. **Assume that the rank of every nontrivial linear combination of them is at least** $r$. **Then there exist a subspace** $V \subseteq \mathbb{F}^n$ **of dimension** $\geq n - 2c^2$ **and** $c' \leq c$ **quadratic polynomials** $q'_1, \ldots, q'_{c'} : V \to \mathbb{F}$ **satisfying: the** $q'_i$-s **are disjoint; every nontrivial linear combination of the** $q'_i$-s **has rank at least** $r - 2c^2$; **span($q_1, \ldots, q_c$) = span($q'_1, \ldots, q'_{c'}$).

Proof. **We prove the lemma by iteratively changing each** $q_i$ **to a `disjoint' form. We shall give the proof over** $\mathbb{F}_2$ **but almost the same proof holds for odd characteristics as well. We start with** $q_1$. **Assume w.l.o.g. that** $x_1 \cdot x_2$ **appears in** $q_1$. **Now, from every other** $q_i$ **subtract an appropriate multiple of** $q_1$ **such that at the end** $x_1 \cdot x_2$ **only appears in** $q_1$. **For simplicity we call the new polynomial** $q_1$ **as well. Now, for** $2 \leq i$ **and** $j \in \{1, 2\}$ **let** $x_j \cdot \ell_{i,j}$ **be the degree two monomials involving** $x_j$ **in** $q_i$. **For** $q_1$ **let** $x_j \cdot \ell_{1,j}$ **be the degree 2 monomials involving** $x_j$ **in** $q_1 - x_1 \cdot x_2$. **Let** $V_1 = \{x \mid \ell_{1,1}(x) = \ldots = \ell_{2,c}(x) = 0\}$. **Notice that none of the** $\ell_{i,j}$-s **contain** $x_1$ **or** $x_2$. **After restricting the polynomials to** $V_1$ **we have that** $x_1 \cdot x_2$ **appears in** $q_1$ **and every other appearance of either** $x_1$ **or** $x_2$ **is in degree one monomials. We now move to (the `new')** $q_2$ **and continue this process. At the end we obtain a subspace** $V$ **and quadratics** $q'_1, \ldots, q'_{c'}$ **($c'$ may be smaller than** $c$ **if some polynomials vanished in the process). As at each step we set at most** $2c$ **linear functions to zero, for a total of at mots** $2c^2$ **linear functions, the claims about the dimension of** $V$ **and the rank of every linear combination of the** $q_i$-s **follow. It is clear that the** $q_i|_V$-s **span the** $q'_i$-s **and so the lemma is proved.

When dealing with odd characteristics instead of looking for** $x_1 \cdot x_2$ **we search for** $x_1^2$. **By applying an invertible linear transformation such a monomial always exists and we continue with the same argument.**

The usefulness of Definition 2.4.13 is demonstrated in the following lemma.

Lemma 2.4.15. **Let** $q_1, \ldots, q_c$ **be disjoint quadratic polynomials. Assume that** $\text{deg}(f) = 2d$ **and** $f = F(q_1, \ldots, q_c)$ **for some function** $F(z_1, \ldots, z_c)$. **Then as a polynomial over** $\mathbb{F}$, $\text{deg}(F) \leq d$.

Proof. **We shall give the proof over** $\mathbb{F}_2$ **but it is again similar over odd characteristic fields. Let** $z_1^{e_1} \cdots z_c^{e_c}$ **be a monomial of maximal degree in** $F$. **When composing it with** $q_1, \ldots, q_c$ **we get that** $q_1^{e_1} \cdots q_c^{e_c}$ **contains the monomial** $\prod_{i=1}^c (x_i \cdot y_i)^{e_i}$. **As** $z_1^{e_1} \cdots z_c^{e_c}$ **is of maximal degree and each** $x_i$ **and** $y_i$ **appear only as linear terms in all the** $q_j$-s **(except the monomial** $x_i \cdot y_i$ **in** $q_i$) **we see that this monomial cannot be cancelled by any other monomial created in** $F(q_1, \ldots, q_c)$. **Therefore the monomial** $\prod_{i=1}^c (x_i \cdot y_i)^{e_i}$ **belongs to** $f$ **as well. Since** $\text{deg}(f) = 2d$ **it must be the case that** $2e_1 + \ldots + 2e_c \leq 2d$. **Hence,** $\text{deg}(F) = \sum_{i=1}^c e_i \leq d$. **We are now ready to complete the proof of Theorem 2.1.7.**

Proof of Theorem 2.1.7. **Combining Corollary 2.4.12, Lemma 2.4.15 and Lemma 2.4.14 we get that for the subspace** $Z$ **of Corollary 2.4.12, there exist a subspace** $Z' \subseteq Z$, **of dimension** $\text{dim}(Z') \geq \text{dim}(Z) - \text{poly}(|\mathbb{F}|/\delta)$, $b = \text{poly}(|\mathbb{F}|/\delta)$ **quadratic polynomials** $Q_1, \ldots, Q_b$ **and a quadratic polynomial** $H$ **such that** $f|_{Z'} = H(Q_1, \ldots, Q_b)$. **In other words** $f|_{Z'} = \sum_{i \leq j} \alpha_{i,j}Q_iQ_j + Q_0$. 

23
As \( f|_{Z'} = \tilde{f}|_{Z'} \) it follows that \( f|_{Z'} = \sum_{i \leq j} \alpha_{ij} Q_i Q_j + Q_0 \). Assume w.l.o.g.\(^9\) that \( Z' \) is defined as 
\[ Z' = \{ x \mid x_1 = \beta_1, \ldots, x_k = \beta_k \} \] for some \( k = \text{poly}(|F|/\delta) \). Then it is clear that we can write 
\[ f = \sum_{i=1}^{k} x_i \cdot g_i + \sum_{i \leq j} \alpha_{ij} Q_i Q_j + g_0 \] for cubic polynomials \( g_0, \ldots, g_k \).

\[ \square \]

### 2.5 Quartic polynomials with high \( U^4 \) norm

In this section we prove Theorems 2.1.8 and 2.1.10. Intuitively, the notion of \( d + 1 \) Gowers norm indicates how close a given function is to a degree \( d \) polynomial. In fact, it was conjectured that if the \( U^{d+1} \) norm is bounded away from zero then the function has a noticeable correlation with a degree \( d \) polynomial. This conjecture turned to be false even when the function is a degree four polynomial and \( d = 3 \) [LMS08, GT07]. Here we will show that for this special case a weaker conclusion holds. Namely, that for any degree four polynomial \( f \) there exists a subspace of dimension \( n/\exp(\text{poly}(1/\|f\|_{U^4})) \) on which \( f|_V \) is equal to some cubic polynomial. In fact an even stronger conclusion holds - there exists a partition of (a subspace of small co dimension of) \( F^n \) to such subspaces on which \( f \) equals a cubic.

To ease the reading we restate Theorem 2.1.8 here.

**Theorem** (Theorem 2.1.8). Let \( F \) be a finite field and \( f \in F[x_1, \ldots, x_n] \) a degree four polynomial such that \( \|f\|_{U^4} = \delta \). Then there exists a partition of a subspace \( V \subseteq F^n \), of dimension \( \dim(V) = n - \text{poly}(|F|/\delta) \), to subspaces \( \{V_\alpha\}_{\alpha \in I} \), satisfying \( \dim(V_\alpha) = \Omega(n/|F|^{\text{poly}(1/\delta)}) \), such that for every \( \alpha \in I \), \( f|_{V_\alpha} \) is a cubic polynomial.

In other words, the theorem says that for \( r = \text{poly}(1/\delta) \) any such \( f \) (possibly after a change of basis of \( F^n \)) can be written as 
\[ f = \sum_{i=1}^{r} x_{n-r+i} g_i(x_1, \ldots, x_n) + f'(x_1, \ldots, x_{n-r}) + g_0, \] where the \( g_i \)'s are degree three polynomials and \( f' \) is a polynomial for which there exists a partition of \( F^{n-r} \) to subspaces \( \{V_\alpha\}_{\alpha \in I} \), satisfying \( \dim(V_\alpha) = \Omega(n/\exp(\text{poly}(1/\delta))) \), such that for every \( \alpha \in I \), \( f'|_{V_\alpha} \) is a cubic polynomial.

As in the proof of Theorem 2.1.6 we start by passing to a subspace of a constant codimension on which every derivative has low rank, i.e \( \Delta_y(f) = \sum_{i=1}^{r} \ell_i Q_i + Q_0 \). Then we shall deduce that there is some common 'basis' \( \{\ell_i\}_{i=1}^{t_2}, \{Q_i\}_{i=1}^{t_1} \) to all the derivatives. Namely, every derivative \( \Delta_y(F) \) can be expressed as 
\[ \sum_{i=1}^{t_1} \ell_i^y Q_i + \sum_{i=1}^{t_2} \ell_i^y Q_i + Q_0^y \] (where \( y \) in the exponent means that the polynomial may depend on \( y \)). This is the main technical difficulty of the proof and it is based on an extension of Lemma 2.3.7 to the case of low rank cubic polynomials. Then, we conclude that for every setting \( \alpha \) of \( \{\ell_i\}_{i=1}^{t_2}, \{Q_i\}_{i=1}^{t_1} \) we obtain a subspace \( V_\alpha \) on which all the derivative are quadratic polynomials, i.e \( f|_{V_\alpha} \) is cubic.

#### 2.5.1 The case of the symmetric polynomial

Let \( S_k(x_1, \ldots, x_n) = \sum_{1 \leq i_1 < \ldots < i_k \leq n} x_{i_1} \cdot x_{i_2} \cdots x_{i_k} \). In [GT07, LMS08] it was shown that over \( F_2 \), it holds that \( \|S_4\|_{U^4} \geq \delta \), for some absolute constant \( 0 < \delta \), but for every degree three polynomial \( g \), \( \Pr[S_4 = g] \leq 1/2 + \exp(-n) \). To make the claim of Theorem 2.1.8 clearer we shall work out the case of \( S_4 \) as an example.

Consider a partial derivative \( \Delta_y(S_4) \). For simplicity assume that \( n = 4m \). Computing we get that 
\begin{equation}
\Delta_y(S_4) = S_2 \cdot \sum_{i \neq j} x_i y_j + S_1 \cdot \sum_{i \neq j} x_i y_j + \sum_{i \neq j} x_i y_j .
\end{equation} \( \text{(2.5)} \)

\(^9\)This is true up to an invertible linear transformation and an affine shift and has no real effect on the result, but rather simplifies the notations.
In particular, $S_2$ is a ‘basis’ for the set of partial derivatives of $S_4$. Continuing, we have that

$$S_2(x_1, \ldots, x_n) = \sum_{k=1}^{2m} \left( \sum_{i=1}^{2k-1} x_i \right) \cdot \left( x_{2k} + \sum_{i=1}^{2k-2} x_i \right) + \sum_{i=1}^{m} (x_{4i-3} + x_{4i-2}). \quad (2.6)$$

For $k = 1, \ldots, 2m$ let $\ell_k = \sum_{i=1}^{2k-1} x_i$. Notice that fixing $\ell_1, \ldots, \ell_{2m}$ reduces the degree of $S_2$ to one and so every partial derivative of $S_4$ will have degree two. For example, consider the space $V_0 = \{ x \mid \ell_1(x) = \ldots = \ell_{2m}(x) = 0 \}$. Rewriting we get $V_0 = \{ (0, y_1, y_1, y_2, y_2, \ldots, y_{2m-1}, y_{2m-1}, y_{2m}) \}$. Computing we get that

$$S_4|_{V_0} = S_2(y_1, \ldots, y_{2m-1}).$$

A closer inspection shows that no matter how we set $\ell_1, \ldots, \ell_{2m}$ we will get that the degree of $S_4$ becomes two.

### 2.5.2 Finding a ‘basis’ for a space of low rank cubic polynomials

In this section we prove the main technical result showing that a subspace of degree 3 polynomials with low rank has a small ‘basis’.

**Lemma 2.5.1** (Main Lemma). Let $M$ be a vector space of cubic polynomials satisfying $\text{rank}_3(f) \leq r$ for all $f \in M$. Then there exists a set of linear and quadratic functions $\{ Q_i \}_{i=1}^{t_1} \cup \{ \ell_i \}_{i=1}^{t_2}$, for $t_1 \leq r$ and $t_2 = 2^{O(r)}$, such that every $f \in M$ can be represented as $f = \sum_{i=1}^{t_1} \ell_i^f Q_i + \sum_{i=1}^{t_2} \ell_i Q_i^f + Q_0^f$ for some linear and quadratic functions $\{ \ell_i \}_{i=1}^{t_1} \cup \{ Q_i^f \}_{i=0}^{t_2}$.

The rest of this section is devoted to proving this lemma. Similarly to the proof of Lemma 2.3.7 we will work modulo a collection of linear and quadratic polynomials. For this we shall need the following definition.

**Definition 2.5.2.** For a cubic polynomial $f$ we say that $\text{rank}_3^c(f) = r$ if $r$ is the minimal integer such that $f$ can be written as

$$f = \sum_{i=1}^{r} \ell_i Q_i + \sum_{i=1}^{c} \ell_i^{(1)} Q_i^{(2)} + \ell_i^{(3)} Q_i^{(4)} + Q_0^f,$$

where the $\ell$-s are linear functions and the $Q$-s are quadratics.

To see that difference from the previous notion of $\text{rank}_3$ (Definition 2.4.1) we observe that if $f$ is a degree three polynomial with $\text{rank}_3(f) = r$ then $f = \sum_{i=1}^{r} \ell_i Q_i + Q_0$. If we also know that some nontrivial linear combination of $Q_1, \ldots, Q_r$ has rank (as a quadratic polynomial) less than $c$ then $\text{rank}_3^c(f) < r$. I.e. $\text{rank}_3^c(f)$ ignores, in some sense, low rank quadratic functions in the representation of $f$.

**Definition 2.5.3.** Let $A = \{ Q_i \}_{i=1}^{t_1} \cup \{ \ell_i \}_{i=1}^{t_2}$ be a set of linear and quadratic functions and let $f : \mathbb{F}^n \to \mathbb{F}$ be a degree three polynomial. Denote

$$[f]_A \triangleq \left\{ f + \sum_{i=1}^{t_1} \ell_i Q_i + \sum_{i=1}^{t_2} \ell_i Q_i^f + Q_0^f \mid \text{for linear and quadratic functions } \{ \ell_i \}_{i=1}^{t_1}, \{ Q_i^f \}_{i=0}^{t_2} \right\}$$

For a linear space $M$ of degree three functions, we define the subspace $[M]_A$ to be

$$[M]_A \triangleq \{ [f]_A \mid f \in M \}.$$

As before we define $\text{rank}_3^c([f]_A)$ to be the lowest rank of functions in $[f]_A$.

$$\text{rank}_3^c([f]_A) \triangleq \min \{ \text{rank}_3^c(g) \mid g \in [f]_A \}.$$
The definition of \([f]_A\) resembles, in some sense, the notion of working modulo an ideal. However, we note that as opposed to the usual definition, where for every \(f\), \(\{Q^t_i\}_{i=1}^{t} \cup \{\ell^t_i\}_{i=1}^{t}\) can be arbitrary functions, in our definition they are restricted to being quadratic and linear functions, respectively.

We are now ready to prove the main lemma of this section that shows the existence of a small ‘basis’ for any linear space of cubic polynomials of low rank.

**Lemma 2.5.4.** Let \(A = \{Q^1_i\}_{i=1}^{t} \cup \{\ell^1_i\}_{i=1}^{t}\) be a set of linear and quadratic polynomials. Let \(M\) be a linear space of cubic polynomials such that for every \([f]_A \in [M]_A\), \(\text{rank}_3^c([f]_A) \leq r\). Then, there are \(r\) linear functions \(\{\ell^1_i\}_{i=1}^{t}\) and a quadratic polynomial \(Q\) such that for \(A' = A \cup \{\ell^1_i\}_{i=1}^{t} \cup \{Q\}\) it holds that every \([f]_{A'} \in [M]_{A'}\) satisfies \(\text{rank}_3^c([f]_{A'}) \leq r - 1\), for \(c' = 11c + 3r + t_1\).

In other words, the lemma says that we can find a small set of linear functions and one quadratic polynomial such that by adding them to \(A\) and increasing \(c\) by a constant factor, we can decrease the \(\text{rank}_3^c\) of every polynomial in \([M]_{A'}\).

**Proof.** Assume that there is some \([g]_A \in [M]_A\) such that \(\text{rank}_3^c(g) = \text{rank}_3^c([g]_A) = r\). If no such \(g\) exists then for every \([f]_A \in [M]_A\), \(\text{rank}_3^c([f]_A) \leq r - 1\) and there is nothing to prove. As \(c < c'\) it also holds that \(\text{rank}_3^c([g]_A) = r\). Hence, \(g\) can be represented as \(\sum_{i=1}^{r} \ell^i f_i^A + \sum_{i=1}^{c} c_i h_i^A\), as otherwise we could replace \(Q^1_i\) with a function of the form \(\sum_{i=1}^{t_1} \alpha_i Q_i + \sum_{i=1}^{t_2} \ell^i f_i + \sum_{i=1}^{c} c_i m_i m_i\), where the \(m\)-s are linear functions, and get that \(\text{rank}_3^c([g]_A) \leq r - 1\).

Set \(A' = A \cup \{\ell^1_i\}_{i=1}^{t} \cup \{Q\}\). Assume for contradiction that there is some \(h \in M\) satisfying \(\text{rank}_3^c([h]_A) = r\). This implies that \(\text{rank}_3^c([h]_A) = r\) and that \(\text{rank}_3^{c' - c}([h + g]_A') = r\) as well. Indeed, if the latter equation was not true then by expressing \(h + g\) as a low rank \(\text{rank}_3^{c' - c}\) polynomial and moving \(g\) to the other side we would get that \(\text{rank}_3^c([h]_A) < r\) in contradiction (recall that \(\{\ell^1_i\} \subset A'\)). From this we get that \(\text{rank}_3^c([h + g]_A) = r\) as well. Let \(f \in [h + g]_A\) be such that \(\text{rank}_3^c(f) = r\). Express \(h + f\) as \(h = \sum_{i=1}^{r} \ell^i f_i^A + \sum_{i=1}^{c} c_i h_i^A\), \(f = \sum_{i=1}^{r} \ell^i f_i^A + \sum_{i=1}^{c} c_i f_i^A\), and \(h = \sum_{i=1}^{r} \ell^i f_i^A + \sum_{i=1}^{c} c_i h_i^A\). Note that we can assume that w.l.o.g \(\{\ell^1_i\}_{i=1}^{t}\) are linearly independent as otherwise we can just replace them with a linearly independent subset. Similarly, we can also assume that \(\{\ell^1_i\}_{i=1}^{t} \cup \{h_i^A\}_{i=1}^{r}\) are linearly independent as otherwise we can find a representation for a function in \([g]_A\) with a smaller rank. Using the same argument again we conclude that \(\{\ell^1_i\}_{i=1}^{t} \cup \{h_i^A\}_{i=1}^{r} \cup \{\ell^1_i\}_{i=1}^{t}\) are linearly independent as well (by considering \([h]_A'\)).

Since \((h + f) - f \in [0]_A\), we can express this polynomial as \((h + f) - f = \sum_{i=1}^{r} \ell^i f_i^A + \sum_{i=1}^{c} c_i h_i^A + \sum_{i=1}^{t_2} \ell^i Q_i + Q_0\). In other words:

\[
\begin{align*}
\sum_{i=1}^{r} \ell^i f_i^A + \sum_{i=1}^{c} c_i h_i^A + \sum_{i=1}^{t_2} \ell^i Q_i + Q_0 &= 0. \tag{2.8}
\end{align*}
\]

To ease notations, using the fact that \(\{\ell^1_i\}_{i=1}^{t} \cup \{h_i^A\}_{i=1}^{r} \cup \{\ell^1_i\}_{i=1}^{t}\) are linearly independent, let us assume w.l.o.g. that \(\forall i, \ell^1_i = x_i, \ell^1_i = x_{r+i}\) and \(\ell_i = x_{2(r+i)}\). Thus, Equation (2.8) becomes

\[
\begin{align*}
\sum_{i=1}^{r} x_i Q_i^g + \sum_{i=1}^{c} c_i h_i^A + \sum_{i=1}^{t_2} x_{2(r+i)} Q_i^g &= 0. \tag{2.9}
\end{align*}
\]

\footnote{By definition of \([g]_A\) we can add any quadratic polynomial to \(g\) so we can assume that there is no extra \(Q^g_0\) term in the representation of \(g\).}
where we remember that variables from \( \{x_i\}_{i=1}^{2r+t_2} \) may appear in the linear and quadratic functions in the expression. Consider all terms involving \( x_1 \) (recall that \( \ell_1^0 = x_1 \)) in Equation (2.9). Clearly they sum to zero, but they can also be written as

\[
0 = Q_1^g + \sum_{i=1}^{r} x_i m_i^g + \sum_{i=1}^{3c} \alpha_i^g m_i^{g,(1)} + \sum_{i=1}^{r} x_{r+i} m_i^h + \sum_{i=1}^{3c} \alpha_i^h m_i^{h,(1)} + \sum_{i=1}^{3c} \alpha_i^h m_i^{h,(2)} - \left( \sum_{i=1}^{r} \beta_i^f Q_i^f + \sum_{i=1}^{r} \beta_i^f m_i^f + \sum_{i=1}^{3c} \beta_i^f m_i^{f,(1)} + \sum_{i=1}^{3c} \beta_i^f m_i^{f,(2)} + \sum_{i=1}^{r} \beta_i^r Q_i + \sum_{i=1}^{t_1} \ell_i m_i + \sum_{i=1}^{t_2} x_{2r+i} m_i + m_0 \right),
\]

where the \( m \)-s are linear functions and the \( \alpha \)-s and \( \beta \)-s are field elements. Rearranging terms we conclude that

\[
\text{rank}_2 \left( Q_1^g - \sum_{i=1}^{r} \beta_i^f Q_i^f - \sum_{i=1}^{t_1} \beta_i^r Q_i - \sum_{i=1}^{t_2} x_{2r+i} m_i \right) \leq 3r + 9c + t_1 = c' - 2c.
\]

This implies that

\[
\text{rank}_2 \left( \sum_{i=1}^{r} \beta_i^f Q_i^f \right)_{A'} \leq c' - 2c .
\]

We now have two cases to consider. If \((\beta_1^f, \ldots, \beta_r^f)\) are not all zero then, by arguments described above, this implies that \(\text{rank}_{3}^{c'}([f]_{A'}) \leq r - 1\). Recalling that \([h + g]_{A'} = [f]_{A'}\) we get a contradiction.

If, on the other hand, \((\beta_1^f, \ldots, \beta_r^f) = 0\) then Equation (2.11) implies that \(\text{rank}_2([Q_i^f]_{A'}) \leq c' - 2c\) and so \(\text{rank}_{3}^{c'}([g]_{A'}) \leq r - 1\) in contradiction to the choice of \(g\). Concluding, we have that for every \(f \in M\), \(\text{rank}_3([f]_{A'}) \leq r - 1\) as required.

By applying Lemma 2.5.4 \(r\) times we obtain the following corollary.

**Corollary 2.5.5.** Let \(M\) be a vector space of cubic polynomials satisfying \(\text{rank}_3(f) \leq r\) for every \(f \in M\). Then there exists a set of quadratic and linear functions \(A = \{Q_i\}_{i=1}^r \cup \{\ell_i\}_{i=1}^{r(r-1)/2}\), such that for \(c = \exp(r)\), \(\text{rank}_3([f]_{A}) = 0\) for every \(f \in M\).

We now have that every function in \(M\), modulo some set \(A\) of linear and quadratic functions, can be expressed as \(\sum_{i=1}^{c} \ell_i^{(1)} + \ell_i^{(2)} + \ell_i^{(3)}\), for some \(c\). Next we show that we can add \(3c\) additional linear functions to \(A\) such that modulo the new set every function becomes zero. We again give an iterative procedure for finding those linear functions.

Before proving this result we define the notion of \(\text{dim}_3([f]_{A})\) that will serve as a potential function in our argument (in a similar way to the role played by \(\text{rank}_3^c\)).

**Definition 2.5.6.** Let \(A\) be a set of quadratic and linear functions and \([f]_A\) a class of cubic functions such that \(\text{rank}_3([f]_{A}) = 0\). We define the dimension of the class as follows:

\[
\text{dim}_3([f]_{A}) = \min \left\{ \dim \left( \text{span} \left\{ \ell_i^{(1)}, \ell_i^{(2)}, \ell_i^{(3)} \right\}_{i=1}^c \right) \mid \sum_{i=1}^{c} \ell_i^{(1)} \ell_i^{(2)} \ell_i^{(3)} \in [f]_{A} \right\} .
\]

To better understand the reason for the definition we note that if \(\text{rank}_3([f]_{A}) = 0\) then \(\sum_{i=1}^{c} \ell_i^{(1)} \ell_i^{(2)} \ell_i^{(3)} + Q \in [f]_{A}\) for some linear functions and quadratic \(Q\). Thus, our goal will be to find a small set of linear functions that, simultaneously, form a basis to all those linear functions for all \(f \in M\). The next lemma shows that by joining \(\left\{ \ell_i^{(1)}, \ell_i^{(2)}, \ell_i^{(3)} \right\}_{i=1}^c\) from some polynomial \(f\), of maximal dimension in \([M]_A\), to \(A\), the dimension of every other element in \([M]_A\) decreases.
Lemma 2.5.7. Let \( A = \{ Q_i \}_{i=1}^{t_1} \cup \{ \ell_i \}_{i=1}^{t_2} \) be a set of linear and quadratic functions. Assume that the rank of any nontrivial linear combination of \( \{ Q_i \}_{i=1}^{t_1} \) is greater than \( 9c + t_1 + t_2 \). Let \( M \) be a linear space of cubic polynomials such that for every \( [f]_A \in [M]_A \), \( \text{rank}^2([f]_A) = 0 \) and \( \text{dim}^2([f]_A) \leq d \). Then, there are \( d \) linear functions \( \{ \ell_i^{(d)} \}_{i=1}^{d} \) such that for \( A' = A \cup \{ \ell_i^{(d)} \}_{i=1}^{d} \), \( \text{dim}^2([f]_{A'}) \leq d - 1 \) for all \( [f]_{A'} \in [M]_{A'} \).

The proof is very similar in nature to the proof of Lemma 2.5.4.

Proof. We start by passing to the subspace \( V = \{ x \mid \ell_1(x) = \ldots = \ell_{t_2}(x) = 0 \} \). When restricting the \( Q_i \)'s to \( V \) the rank of every linear combination can drop by at most \( t_2 \) so it is still at least \( 9c + t_1 \). From now on we shall work over \( V \). Note that if we prove the theorem over \( V \) then it clearly holds over \( \mathbb{R}^n \) as well.

Let \( [g]_A \in [M]_A \) be a class satisfying \( \text{dim}^3([g]_A) = d \). By definition we can assume that \( g \) is such that \( g = \sum_{i=1}^{c} \phi_i^{(1)} \phi_i^{(2)} \phi_i^{(3)} \), and that for some \( d \) linearly independent linear functions \( \{ \ell_i^{(d)} \}_{i=1}^{d} \) it holds that \( \{ \phi_i^{(1)} \phi_i^{(2)} \phi_i^{(3)} \}_{i=1}^{c} \subseteq \text{span}(\{ \ell_i^{(d)} \}_{i=1}^{d}) \). Set \( A' = A \cup \{ \ell_i^{(d)} \}_{i=1}^{d} \). We will show that for every \( f \in M \) it holds that \( \text{dim}^3([f]_{A'}) \leq d - 1 \). Assume for contradiction that there is some \( [h]_{A'} \in [M]_{A'} \) such that \( \text{dim}^3([h]_{A'}) = d \). Clearly, \( \text{dim}^3([h]_A) = d \) as well. W.l.o.g. let \( h = \sum_{i=1}^{c} h_i^{(1)} h_i^{(2)} h_i^{(3)} \). We also denote with \( \{ \ell_i^{(d)} \}_{i=1}^{d} \) a basis for \( \{ h_i^{(1)} h_i^{(2)} h_i^{(3)} \}_{i=1}^{c} \). As \( \text{dim}^3([h]_A) \) does not decrease modulo \( \{ \ell_i^{(d)} \}_{i=1}^{d} \), it follows that \( \{ \ell_i^{(d)} \}_{i=1}^{d} \cup \{ \ell_i^{(d)} \}_{i=1}^{d} \) are linearly independent. By definition of \( A' \) we have that \( \text{dim}^3([g + h]_{A'}) = \text{dim}^3([h]_{A'}) = d \). Let \( f \in [g + h]_A \) be such that \( f = \sum_{i=1}^{c} \ell_i^{(1)} \ell_i^{(2)} \ell_i^{(3)} \) and \( \text{dim}(\text{span}(\{ \ell_i^{(j)} \}_{i=1}^{d})) = d \). Since \( g + h - f \in [0]_A \) we have that \( g + h - f = \sum_{i=1}^{t_1} Q_i \ell_i^{(1)} + Q' \). We now show that all \( \ell_i^{(1)} \)'s are zero. Assume for contradiction that this is not the case. Namely, \( \{ \ell_i^{(1)} \}_{i=1}^{t_1} \) are not all zero. In particular, some \( \ell_i^{(1)} \) depends on some variable \( x \). Write \( g + h - f = \ell_i^{(1)} Q_i + \ell_i^{(2)} m_i + \ell_i^{(3)} m_0 \) for some coefficients \( \{ \alpha_i \}_{i=1}^{t_1} \) (not all of them are zero) and some linear functions \( \{ m_i \}_{i=0}^{t_1} \). Hence, \( \text{rank}^2(F) \) is larger than \( 9c \) (remember that \( \text{rank}^2(\sum_{i=1}^{t_1} \alpha_i Q_i) > 9c + t_1 \) on \( V \)). On the other hand, \( g + h - f = Q \) for some quadratic \( Q \). Thus,

\[
\sum_{i=1}^{c} \phi_i^{(1)} \phi_i^{(2)} \phi_i^{(3)} + \sum_{i=1}^{c} h_i^{(1)} h_i^{(2)} h_i^{(3)} - \sum_{i=1}^{c} \ell_i^{(1)} \ell_i^{(2)} \ell_i^{(3)},
\]

so \( F \) can be expressed as \( \sum_{i=1}^{9c} \hat{m}_i \hat{n}_i + \ell \), i.e. it’s rank is at most \( 9c \), in contradiction. It follows that \( g + h - f = Q \). For simplicity, assume w.l.o.g. that for \( i = 1 \ldots d \), \( \ell_i^{(1)} = y_i, \ell_i^{(2)} = z_i \). We would like to show that if Equation (2.12) holds then \( \text{deg}(h) = 2 \) in contradiction to the choice of \( h \). To further simplify notations we assume w.l.o.g. that the \( \ell_i^{(1)} \)-s are linear functions in the variables \( y_1, \ldots, y_d, z_1, \ldots, z_d \) (as we can set all other variables to zero and still obtain a similar equality). In particular, every \( \ell_i^{(1)} \) can be expressed as \( \ell_i^{(1)} = \ell_i^{(g)}(y) + \ell_i^{(h)}(z) \). Hence, Equation (2.12) can be rewritten as \( Q(y, z) + f(y, z) = g(y) + h(z) \). Therefore, it holds that \( g(y) = f(y, 0) + Q(y, 0) \) and \( h(y) = f(0, z) + Q(0, z) \).11 In particular, there is some representation of \( g \) and \( h \) as sums of products of linear functions such that \( \{ \ell_i^{(g)}(y) \} \) and

---

11We can assume w.l.o.g. that \( \ell_i^{(g)} \) and \( \ell_i^{(h)} \) do not have a constant term.
\{l_i^{f,h}(z)\}$ are their basis, respectively. By applying an invertible linear transformation we can further assume that $l_i^{f,q}(y) = y_i$ and $l_i^{f,h}(z) = z_i$. Thus, the basis for $\{l_i^{f,j}\}$ is $l_i^f = y_1 + z_1, \ldots, l_i^d = y_d + z_d$.

As a consequence we have that $f = \sum_{i=1}^c e_i^{f,(1)}(y + z)l_i^{f,(2)}(y + z)l_i^{f,(3)}(y + z)$.

Define $F : \mathbb{F}^d \to \mathbb{F}$ as $F(u) = \sum_{i=1}^c e_i^{f,(1)}(u)l_i^{f,(2)}(u)l_i^{f,(3)}(u)$. Hence, $f = F(y + z), \ g = f(y,0) + Q(y,0) = F(y) + Q(y)$ and $h = F(z) + Q'(z)$. Thus, for every $\alpha, \beta \in \mathbb{F}$, $F(\alpha + \beta) = F(\alpha) + F(\beta) + Q(\alpha, \beta)$. It is not difficult to check that if $F$ is a polynomial such that $\deg(F(\alpha + \beta) - F(\alpha) - F(\beta)) \leq 2$ then $\deg(F) \leq 2$. Therefore, $[h]_A = [F(z)]_A = [0]_A$ (because $F$ is quadratic), in contrary to the fact that $\dim_{\mathbb{F}}([h]_A) = d$. We thus deduce that for every $[h]_A \in [M]_{\mathbb{F}}^n$, $\dim_{\mathbb{F}}([h]_A) < d$ as required.

Combining Lemma 2.5.4 and Lemma 2.5.7 we are now able to prove Lemma 2.5.1.

**Proof of Lemma 2.5.1.** Corollary 2.5.5 implies that there exists a set of quadratic and linear functions $A = \{Q_i\}_{i=1}^r \cup \{\ell_i\}_{i=1}^{(r-1)/2}$, such that for $c = \exp(r)$, $\text{rank}_{\mathbb{F}}([f]_A) = 0$ for every $f \in M$. By Lemma 2.3.8 we can assume w.l.o.g. that every nontrivial linear combination of the $Q_i$’s have rank larger than $10c$ (possibly after passing to a subspace $V$ of dimension at least $n - \text{poly}(c) = n - \exp(r)$ and throwing some of the $Q_i$’s (without changing the property of $[M]_A$)). By applying Lemma 2.5.7 $d = 3c$ times we get a set $A' = \{Q_i\}_{i=1}^{t_1} \cup \{\ell_i\}_{i=1}^{t_2}$, for $t_1 \leq r$ and $t_2 = \exp(r)$, such that $\dim_{\mathbb{F}}([f]_A') = 0$ for every $[f]_A' \in [M]_{\mathbb{F}}$. In particular, every $f \in M$ can be represented as $f = \sum_{i=1}^{t_1} \ell_i^f \cdot Q_i^f + \sum_{i=1}^{t_2} \ell_i^f \cdot Q_i^f + Q_0^f$ for some linear and quadratic functions $\{\ell_i^f\}_{i=1}^{t_1} \cup \{Q_i^f\}_{i=0}^{t_2}$ depending on $f$.

2.5.3 Completing the proof

We can now complete the proof of Theorem 2.1.8. We first give a lemma summarizing what we have achieved so far.

**Lemma 2.5.8.** Let $f$ be a degree four polynomial with $\|f\|_{\ell^4} = \delta$. Then for $r = O(\log^2(1/\delta))$ there exist a subspace $V$, satisfying $\dim(V) \geq n - O(\log(1/\delta))$, $r$ quadratic polynomials $Q_1, \ldots, Q_r$ and $R = \exp(r)$ linear functions $\ell_1, \ldots, \ell_R$ such that for every $y \in V$ we have that $\Delta_y(f|_V) = \sum_{i=1}^r Q_i \cdot \ell_i^y + \sum_{i=1}^R \ell_i \cdot Q_i^y + Q_0^y$.

**Proof.** Let $f$ be a quartic function such that $\|f\|_{\ell^4} > \delta$. By Lemma 2.3.2, Theorem 2.1.6 and Lemma 2.2.5 there is a subspace $V$, satisfying $\dim(V) \geq n - O(\log(1/\delta))$, such that every partial derivative of $f|_V$ is a cubic polynomial of rank at most $r = O(\log^2(1/\delta))$. Let $f' = f|_V$. Lemma 2.5.1 gives a set $A = \{Q_i\}_{i=1}^r \cup \{\ell_i\}_{i=1}^{\exp(r)}$ such that every $\Delta_y(f')$ can be written as $\Delta_y(f') = \sum_{i=1}^r Q_i \cdot \ell_i^y + \sum_{i=1}^{\exp(r)} \ell_i \cdot Q_i^y + Q_0^y$. Notice that the lemma concerns a linear space of cubic polynomials. In our case the linear space will be the span of all the partial derivatives of $f'$. As for every $y, z \in V$ it holds that $\deg(\Delta_y(f') + \Delta_z(f') - \Delta_{y+z}(f')) = 2$, we see that in order to ‘close’ the space we only need to add quadratic polynomials and so the assumption about the rank of the cubic polynomials in the space does not change.

**Proof of Theorem 2.1.8.** By Lemma 2.5.8 for $r = O(\log^2(1/\delta))$ there exist $r$ quadratics $Q_1, \ldots, Q_r$ and $R = \exp(r)$ linear functions $\ell_1, \ldots, \ell_R$ such that for every $y \in V$ we have that $\Delta_y(f|_V) = \sum_{i=1}^r Q_i \cdot \ell_i^y + \sum_{i=1}^R \ell_i \cdot Q_i^y + Q_0^y$.

We now wish to express each $Q_i$ in the form of Theorem 2.1.2. We have two cases. Assume first that $V = F_2$. Then for every $1 \leq i \leq r$ we have that $Q_i = \sum_{j=1}^{n/3} \ell_{i,j} \cdot \ell_{i,j}^t + \ell_{i,0}$. For $\alpha \in \mathbb{R}^n$ let $V_0 = \{x \in V \mid \forall 1 \leq i \leq R, \ell_i(x) = \alpha_i\}$. Clearly, $\dim(V_0) \geq \dim(V) - R$. Let $f_0 = f|_{V_0}$. Then for every $y \in V_0$, $\Delta_y(f_0) = \sum_{i=1}^r Q_i |_V \cdot \ell_i^y + Q_0^y$. We now repeat the following process for each $1 \leq i \leq r$.\]
Assume that we are working over a subspace $V_{\alpha,\beta^1,...,\beta^{t-1}}$, of dimension $d_{t-1} = \dim \left( V_{\alpha,\beta^1,...,\beta^{t-1}} \right)$. Consider $Q_i|_{V_{\alpha,\beta^1,...,\beta^{t-1}}}$. By Theorem 2.1.2 we can write $Q_i|_{V_{\alpha,\beta^1,...,\beta^{t-1}}} = \sum_{i=1}^{d_{t-1}/2} \ell_{i,j} \cdot \ell_{i,j} + \ell_{i,0}$. For $\beta^i \in \mathbb{F}^{d_{t-1}/2}$ define $V_{\alpha,\beta^1,...,\beta^t} = \{ x \in V_{\alpha,\beta^1,...,\beta^{t-1}} \mid \forall 1 \leq j \leq d_{t-1}/2, \ell_{i,j} = (\beta^i)_j \}$. Note that $\cup_{\beta^i \in \mathbb{F}^{d_{t-1}/2}} V_{\alpha,\beta^1,...,\beta^t} = V_{\alpha,\beta^1,...,\beta^{t-1}}$. Thus, the set $\{ V_{\alpha,\beta^1,...,\beta^r} \}$ forms a partition of $V$. Moreover, observe that for every $\alpha,\beta^1,...,\beta^r$, all the partial derivatives of $f|_{V_{\alpha,\beta^1,...,\beta^r}}$ are of degree two and so $\deg \left( f|_{V_{\alpha,\beta^1,...,\beta^r}} \right) \leq 3$. To finish the proof we note that $\dim \left( V_{\alpha,\beta^1,...,\beta^r} \right) \geq \dim \left( V_{\alpha,\beta^1,...,\beta^{t-1}} / 2 \right)$. Therefore, $\dim \left( V_{\alpha,\beta^1,...,\beta^r} \right) \geq (n - R) / 2^r = n / \exp(\log^2(1/\delta))$.

When $\text{char}(\mathbb{F}) = p > 2$ we have the representation $Q_i|_{V_{\alpha,\beta^1,...,\beta^{t-1}}} = \sum_{i=1}^{d_{t-1}} \ell_{i,j}^2 + \ell_{i,0}$. Rewriting we obtain

$$Q_i|_{V_{\alpha,\beta^1,...,\beta^{t-1}}} = \sum_{i=1}^{r} \ell_{i,j}^2 + \ell_{0}$$

Observe that after fixing $\forall 1 \leq j \leq p - 1$, $\ell_{i,j}^2 + \ell_{pi} = (\beta^i)_j$, $Q_i|_{V_{\alpha,\beta^1,...,\beta^{t-1}}} \text{ becomes linear.}$ Thus, the same argument as before gives the required result here as well.

Combining the idea of the above proof with the notion of disjoint polynomials we prove Theorem 2.1.10.

**Proof Sketch of Theorem 2.1.10.** As in the proof of Theorem 2.1.8 we obtain linear $\{ \ell_i \}_{i=1,...,r}$ and quadratic $\{ q_i \}_{i=1,...,r}$, where $r = O(\log^2(1/\delta))$ and $R = \exp(r)$, that form a ‘basis’ to the set of partial derivatives. By passing to a subspace of codimension $R$ and using Lemma 2.4.14 we can assume w.l.o.g. that the $q_i$-s are disjoint and that every partial derivative has the form $\Delta_y(f) = \sum_{i=1}^{r} q_i \cdot \ell_i^0 + q_0^0$. As $\text{char}(\mathbb{F}) > 4$ we can assume w.l.o.g. that $q_i = x_i^2 + q_i^0$ and that $x_j$ can appear in $q_0^0$ only as a linear term. We now subtract from $f$ terms of the form $\alpha_i q_j$ such that in the resulting polynomial $f'$ there will be no monomial of the form $x_i^2 x_j^2$ for $i \leq j$. Note that $f'$ also has the property that for every $y$, $\Delta_y(f') = \sum_{i=1}^{r} q_i \cdot \ell_i^0 + q_0^0$. We now show that degree four monomials in $f'$ may only contain $x_i$ or $x_i^3$ but not $x_i^2$, for $i \in [r]$. Indeed, assume for a contrary that $x_i^2$ appears in a degree four monomial. Then, $x_i$ appears in $\Delta_{x_i}(f')$ in a degree three monomial. This monomial comes from some $\ell_{j}^0 q_j$ for $j \neq i$. Therefore, we also have the term $x_i x_j^2$ in $\Delta_{x_i}(f')$ (it is not difficult to see that this term cannot be cancelled by any other $\ell_k^0 q_k$). As $\text{char}(\mathbb{F}) > 4$, integration w.r.t. $x_i$ gives that the term $x_i^2 x_j^2$ appears in $f'$ in contradiction. We can thus write $f' = \sum_{i=1}^{r} x_i^3 \ell_i^0 + f''$, where in $f''$ each $x_i$ has degree at most one. Consider any $y$ ‘orthogonal’ to $\{ x_1, \ldots, x_r, \ell_1, \ldots, \ell_r \}$ (namely, substituting $y$ in any of those linear functions gives zero). Then for each $i$, $\Delta_y(x_i^3 \ell_i^0) = 0$. Hence, $x_i$ is the highest power of $x_i$ appearing in $\Delta_y(f')$. As the $q_i$-s are disjoint and $\Delta_y(f') = \sum_{i=1}^{r} q_i \cdot \ell_i^0 + q_0^0$ we obtain that it must be the case that $\deg(\Delta_y(f')) \leq 2$. Thus, $f'$ can be rewritten as a polynomial in at most $2r$ variables plus a degree three polynomial. Therefore, possibly after a change of basis we can write $f = \sum_{i=1}^{r} q_i \cdot q_j + \sum_{i=1}^{2r+R} y_i \cdot g_i + g_0$ as needed.
2.6 Conclusions

In this chapter we gave strong structural results for degree three and four polynomials that have a high bias. It is a very interesting question whether such a structure exists for higher degree biased polynomials. Green and Tao [GT07] proved such a result when \( \text{deg}(f) < |\mathbb{F}| \) (with much worse parameters for degrees three and four), so this question is mainly open for small fields. Another interesting question is improving the parameters in the results of [GT07, KL08]. There it was shown that when \( \text{deg}(f) = d \) and \( f \) is biased then \( f = F(g_1, \ldots, g_{c_d}) \), where \( \text{deg}(g_i) < \text{deg}(f) \). However, the dependence of \( c_d \) on the degree \( d \) and the bias \( \delta \) is terrible. Basically, \( c_3 = \exp(\text{poly}(1/\delta)) \) and \( c_d \) is a tower of height \( c_d-1 \). In contrast, our results give that \( c_3 = \log^2(1/\delta) \) and \( c_4 = \text{poly}(1/\delta) \). Thus, it is an intriguing question to find the true dependence of \( c_d \) on \( \delta \). In particular, as far as we know, it may be the case that \( c_d \) is polynomial in \( 1/\delta \) (where the exponent may depend on \( d \)), or even \( \text{poly}(\log(1/\delta)) \).

For the case of degree four polynomials with high \( U^d \) norm we proved an inverse theorem showing that on many subspaces, of dimension \( \Omega(n) \), \( f \) equals to a degree three polynomial (a different polynomial for each subspace). Such a result seems unlikely to be true for higher degrees. However, it may be the case that if \( \text{deg}(f) = d \) and \( f \) has a high \( U^d \) norm then \( f \) is correlated with a lower degree polynomial on a high dimensional subspace.
Chapter 3

Optimal testing of multivariate polynomials over small prime fields

3.1 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 chapter 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 prime size. This setting has been considered before in [AKK+05, KR06, JPRZ09, 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 $\mathbb{F}_n^q$ to $\mathbb{F}_q$. Let $\mathcal{P}(n,d,q)$ denote the set of all $n$-variate polynomial functions over $\mathbb{F}_q$ of total degree at most $d$. We let $\delta(f,g) = \Pr_x[f(x) \neq g(x)]$ denote the distance between $f$ and $g$, where the probability is over $x$ chosen uniformly at random from $\mathbb{F}_n^q$. 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 $\delta$-far from $g$ if $\delta(f,g) \geq \delta$ and $\delta$-close otherwise. We say $f$ is $\delta$-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 $f : \mathbb{F}_n^q \rightarrow \mathbb{F}_q$ and accepts $f \in \mathcal{P}(n,d,q)$ with probability one. It has $\delta$-soundness $\epsilon$ if it rejects every function $f$ with $\delta_d(f) \geq \delta$ with probability at least $\epsilon$. 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 $\delta$-soundness $\epsilon$.

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 studied 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 $\delta$-soundness $\Omega(\delta 2^{-d})$. Thus to get an absolutely sound test, they iterated this test $O(2^d)$ times,

---

1Throughout this chapter 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.
Proposition 3.1.2. \( \text{results} \) we give the proof in Section 3.4.3. This notion was studied in \([\text{AKK}^+05]\) who proved the following fact. As it also follows easily from our attempts to explore the minimal dimension for which such a test has positive soundness. 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. [JPRZ09]. 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-degree 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 \( \delta \)-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 \([\text{AKK}^+05]\) and [KR06, JPRZ09] 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.

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 \([\text{AKK}^+05]\) is absolutely sound. This also gave the first example of a linear-invariant property with tight bounds on query complexity.

The proof of \([\text{BKS}+10]\) was significantly more algebraic than those of \([\text{AKK}^+05, \text{KR06, JPRZ09}]\). (Indeed the work of [KS08] confirms that the central ingredient in the proofs in \([\text{AKK}^+05, \text{KR06, JPRZ09}]\) are all the same and relies on very little algebra.) However, the proof of \([\text{BKS}+10]\) 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 constant grows extremely fast as a function of \( q \)). 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.

3.1.1 Our main results

To state the test of \([\text{AKK}^+05, \text{KR06}]\) 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 \( f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q \) and affine subspace \( A \), let \( f|_A : A \rightarrow \mathbb{F}_q \) denote the restriction of \( f \) to \( A \). For a function \( f \), we let \( \deg(f) \) denote its degree as a polynomial. We use the fact that \( f|_A \) can be viewed as a \( \dim(A) \)-variate polynomial with \( \deg(f|_A) \leq \deg(f) \). A special subclass of tests for \( \mathcal{P}(n,d,q) \) would simply pick an affine subspace \( A \) of \( \mathbb{F}_q^n \) and verify that \( \deg(f|_A) \leq d \). We introduce the concept below of the testing dimension which attempts to explore the minimal dimension for which such a test has positive soundness.

**Definition 3.1.1** (Testing dimension). For prime power \( q \) and non-negative \( d \), the testing dimension of polynomials of degree \( d \) over \( \mathbb{F}_q \) is the smallest integer \( t \) satisfying the following: For every positive integer \( n \) and every function \( f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q \) 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 3.4.3.

**Proposition 3.1.2.** 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 \( f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q \), pick a random affine subspace \( A \) with \( \dim(A) = t \) and accept if \( \deg(f|_A) \leq d \).
Global inconsistency
Pairwise consistency
polynomials
integer such that the following holds: There exists
Definition 3.1.6.
the restrictions of functions to hyperplanes and ask “when does pairwise consistency imply global
Theorem 3.1.5.
number of hyperplanes where a degree
Theorem 3.1.3.
low-degree test, we need a similar theorem that talks about general functions. Extracting the correct
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 $H(q,n)$ 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 3.1.4. For prime power $q$ and non-negative integer $d$, let $N = N_0(q,d)$ be the maximum over all $n$, and all functions $f \in \mathcal{P}(n,d,q)$ of the number of hyperplanes $A_1, \ldots, A_N$ such that $\deg(f|_{A_i}) < \deg(f)$. I.e.,
$$N_0(q,d) = \max_{n,f \in \mathcal{P}(n,d,q)} |\{A \in H(n,q) | \deg(f|_{A}) < \deg(f)\}|.$$ 
A priori it may not be clear that $N_0(d,q)$ is even bounded (i.e., is independent of $n$), but an easy argument from [BKS+10] shows this quantity is at most $q^d$. For our purposes we need a much tighter bound of roughly $q^{q,q,d}$ and our first main technical theorem (of two) shows that this is indeed the case.
Theorem 3.1.5. For every $q,d$, $N_0(d,q) \leq q^{q,q,d+1}$. In other words, if $f \in \mathcal{P}(n,d,q)$, then $|\{A \in H(q,n) | \deg(f|_{A}) < \deg(f)\}| \leq N_0(d,q) \leq q^{q,q,d+1}$. 
We note that $N_0(d,q) > q^{q,q,d-1}$. Indeed, let $d = t(q - \frac{q}{p}) + b$, where $b < q - \frac{q}{p}$, and define $f = \prod_{i=1}^b x_i^{q-2} + \prod_{i=b+1}^{t} x_i^{q-2}$. One can verify that $f$ is a degree $d$ polynomial and on any hyperplane supported only on the first $t \geq q,q,d - 1$ variables, the degree of $f$ decreases when restricted to that hyperplane.
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 3.1.6. 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 $A_1, \ldots, A_N \in H(n,q)$ 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 3.1.7.** For every $q$, there exists a constant $\lambda_q$ such that for every $d$, $N_1(q,d) \leq q^{q,d+\lambda_q}$. In other words if $A_1, \ldots, A_K \in \mathcal{H}(n,q)$ 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 $K > q^{q,d+\lambda_q}$, then there exists $Q \in \mathcal{P}(n,d,q)$ such that $Q|_{A_i} = P_i|_{A_i}$ for every $i \in [K]$.

In Section 3.2, we show how the technical theorems above (Theorems 3.1.5 and 3.1.7) lead to an analysis of the low-degree test.

### 3.1.2 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 there is no novelty in ours in this part.) Their proof roughly shows that for $t = \log_q N_1(q,d)$ the $t$-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 Lemmas 4.5.3 and 3.4.7) that a $t-1$ dimensional test (as long as $t-1 \geq t_q,d$) has $\delta$-soundness at least $1/q$ times the $\delta$-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 3.4.3 and 3.4.8). 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 $\epsilon q^c$ hyperplanes in $c$ dimensions will contain $q$ “near-parallel” ones. (Unfortunately this leads to a horrendous bound on $c_{q,\epsilon}$, but fortunately $\epsilon$ is independent of $n$ and $d$ and so this suffices for Theorem 3.1.3).
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 3.1.2 captures this step. Its proof relies on Lemma 4.5.3 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 \( x_2,\ldots,x_m \) whose coefficient is a non-zero function of \( y_2,\ldots,y_m, y_0 \). The key now is to notice that this coefficient is a polynomial in \( y_2,\ldots,y_m, y_0 \) of degree at most \( q - 1 \) and so is non-zero with probability at least \( 1/q \) when \( y_2,\ldots,y_m, y_0 \) are assigned randomly.

This step is performed in Section 3.4.3. The heart of the proof is given by Lemma 4.5.3, 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 3.4.1). 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 3.4.3. 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 3.4.1.

**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 \( O(q^{\theta_0,d}) \). So if the number of variables \( m \) is really large compared to \( q,d \) then the fraction of hyperplanes where \( f \) drops in degree is tiny.

This bound again views the restriction of \( f \) to hyperplanes of the form \( x_1 = \sum_{i=2}^{m} y_i x_i + y_0 \) as a polynomial in \( x_2, \ldots, x_m \) and \( y_2, \ldots, y_m, y_0 \). We then perform an elementary, though somewhat non-obvious, algebraic analysis of this polynomial to show that there are few hyperplanes where \( f \) loses degree. Roughly, we show that when working with an appropriate basis for the space (i.e. when applying the linear transformation that guarantees the existence of a canonical monomial, found in the previous step) it is the case that for every fixing of \( y_2, \ldots, y_t \), where \( t = \log_q N_0(q,d) \approx t_{q,d} \), there is at most one setting of \( y_{t+1}, \ldots, y_m \) such that the degree of \( f \) decreases on the corresponding hyperplane. Canonical monomials again play a crucial role in the proof. This step is captured by Theorem 3.1.5 that is proved in Section 3.4.4. Lemma 3.4.8 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 \( O(q^{\theta_0,d}) \), 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 that 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 \( f|_{A_0 \cap (\cup_j A_i)} \) 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 3.1.7, that is given in Section 3.4.5.

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 3.4.12) makes sure we do not lose too much in this as long as the number of variables is sufficiently large (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., \( A_i \cap A_j = A \cap A_k \) 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 3.3.4). 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 \( \lambda_{q,6} \)), then this allows to conclude that \( q \) such “near-parallel” hyperplanes do exist. Now, with a linear change of basis, we can assume that the \( A_i \cap A_j \) is contained in the hyperplane \( x_1 = 0 \), and that none of the hyperplanes \( A_1, \ldots, A_q \) is equal to the hyperplane \( x_1 = 0 \). The crux of the idea is that now, on all the \( q - 1 \) hyperplanes, \( x_1 = \alpha, \alpha \in \mathbb{F}_q - \{0\} \), the hyperplanes \( A_1 \cap \{x_1 = \alpha\}, \ldots, A_q \cap \{x_1 = \alpha\} \) are parallel. The situation is perhaps better explained by Figure 3.1 (for the case \( q = 5 \)). This allows us to prove (using arguments similar to the inductive step) that \( f \)

![Figure 3.1: Near parallel hyperplanes](image)

on these hyperplanes is a degree \( d \) polynomial, and roughly tells us what \( Q \mod (x_1^{q-1} - 1) \) is (where \( Q \) is the desired polynomial of degree \( d \) that agrees with \( f \) on the union \( \cup_{i \in [K]} A_i \)). Pushing our luck further, we note that if \( \log N_1(q, d) = t + \lambda_{q,6} \) then we can find \( t \) independent variables \( x_1, \ldots, x_t \) such that we know the polynomial \( Q \mod \prod_{i=1}^t (x_i^{q-1} - 1) \). If \( t > d/(q-1) \) this should tell us exactly what \( Q \) is, and with some careful examination we confirm this intuition, and show that this polynomial \( Q \) agrees with \( f \) on every one of the given hyperplanes, thus concluding the analysis in the base case. The base case is given in Lemma 3.4.11.

**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 chapter.** Section 4.3 contains some notations and basic facts regarding polynomial. We discuss the density Hales-Jewett theorem in Section 3.3.2. The main body of the chapter is Section 3.4. The section is organized as follows. In Section 3.4.1 we give the definition of canonical monomials and shows how to “rotate” the space in order to find one (Lemma 3.4.3). Section 3.4.2 shows the basic and simple fact that the rejection probability of the \(\ell\)-dimensional test is monotone in \(\ell\) and in Section 3.4.3 we prove that although the rejection probability is monotone, it does not decrease too fast when we go from \(\ell\) to \(\ell - 1\) (Lemma 4.5.3). We then give the proofs of our two main technical contributions. Theorem 3.1.5, in which we bound \(N_0(q,d)\), is proved in Section 3.4.4 and Theorem 3.1.7 is proved in Section 3.4.5. Section 3.4.6 contains a strengthening of Theorem 3.1.7 (given as Theorem 3.4.16), that is proved in a relatively direct manner from Theorem 3.1.7. Finally, we analyze the \(t_{q,d}\)-dimensional test in Section 3.5, giving a proof of Theorem 3.1.3 – our main theorem.

### 3.3 Preliminaries

Throughout the chapter \(q = p^k\) 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 \(\alpha^{q-1} = 1\). 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 \(\text{Aff}(q,n)\) 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 \(L = \sum_{i=0}^{n} \alpha_i x_i + \alpha_0\) is a linear function, we call \(\alpha_0\) the free term of \(L\).

Let \(d, e \in \mathbb{N}\) be integers and denote by \(d = \sum_i d_ip^i\) and \(e = \sum_i e_ip^i\) their base \(p\) expansion. Namely, \(\forall i \geq 0 \leq d_i, e_i < p\). 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 3.3.1** (Lucas’ theorem). In the notations above, \(\binom{e}{d} \equiv_p \prod_i \binom{e_i}{d_i}\), where \(\binom{e_i}{d_i} = 0\) if \(e_i < d_i\).

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

\[
(y + z)^e \equiv_p \sum_{d \leq e} \binom{e}{d} y^{e-d} z^d.
\] (3.1)

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.

#### 3.3.1 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 3.3.2.** 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 3.3.3 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 \( e_{\text{max}} \) be the degree of \( f \) as a polynomial in \( x_n \). As \( \text{deg}(g_{e_{\text{max}}}) \leq d - e_{\text{max}} \), the induction hypothesis implies that the number of non-zeros of \( g_{e_{\text{max}}} \) is at least \( q^{-(d-e_{\text{max}})/(q-1)} \cdot q^{n-1} \). 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 \( e_{\text{max}} \) and therefore has at least \( q - e_{\text{max}} \) non-zeros. Consequently,

\[
\delta(f,0) \geq (q-e_{\text{max}}) \cdot q^{-(d-e_{\text{max}})/(q-1)} \cdot q^{n-1}/q = (1 - e_{\text{max}}/q) \cdot q^{-(d-e_{\text{max}})/(q-1)} \]

where inequality \((*)\) follows from Claim 3.3.3.

\[ \square \]

Claim 3.3.3. 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 \( 0 \leq x \leq q-1 \).

\[ \square \]

3.3.2 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 proof 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 \( \Sigma \) as being \( \mathbb{F}_q \). A set \( \mathcal{L} = \{v_1,\ldots,v_q\} \subset \Sigma^n \) is a combinatorial line if we can partition the coordinates \( [n] \) to two disjoint sets \( [n] = I \cup J, \ I \cap J = \emptyset \) such that: (1) For all \( i \in I \) and \( k,k' \in [q], \ (v_k)_i = (v_{k'})_i \). Namely, for all \( i \in I \), the \( i \)’th coordinate of all elements in \( \mathcal{L} \) is fixed. (2) For \( j \in J \) and \( k \in [q], \ (v_{k})_j = a_k \). I.e., the \( j \)’th coordinate advances with \( k \).

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

Theorem 3.3.4 ([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 3.3.5. 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.
The proof is by induction on \( t \). For \( t = 1 \), Theorem 3.3.4 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 3.3.4 and get that \( A \cap \mathcal{M} \) contains a line. It is immediate that the direction of this line is linearly independent of \( e_1, \ldots, e_{t-1} \).

\[\square\]

### 3.4 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 \( x_k \).

#### 3.4.1 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 3.4.1.** 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 \( p^\ell \leq p \leq q \) then for every \( j < m \), \( p^\ell + e_j > q - 1 \). (4) \( 0 < e_m < q \).

Note that Properties 3 and 4 imply Property 2, but for clarity we keep all of them.

The following lemma shows that whenever we have a multivariate polynomial over \( \mathbb{F}_q \) there exists an invertible linear transformation \( A : \mathbb{F}_q^n \to \mathbb{F}_q^n \) such that \( f \circ A \) contains a canonical monomial of maximal degree. In fact, we will prove a slightly stronger property. For that end we will need the following definition.

41
Definition 3.4.2 (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, in this order, appearing in $g$ (with a non-zero coefficient of course).

Lemma 3.4.3. Let $f(x_1, \ldots, x_n)$ be a degree $d \leq n(q-1)$ polynomial over $\mathbb{F}_q$. Let

$$A = \arg \max_{\text{invertible } B} \text{max-monomial of } (f \circ B)(x_1, \ldots, x_n).$$

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 $d$.

First, we will prove Lemma 3.4.3 for the bivariate case.

Lemma 3.4.4. Let $f(x_1, x_2)$ be a degree $d \leq n(q-1)$ polynomial over $\mathbb{F}_q$. Let

$$\alpha = \arg \max_{\beta \in \mathbb{F}_q} \text{max-monomial of } (f(x_1, x_2 + \beta x_1)).$$

Then, the max-monomial of $f(x_1, x_2 + \alpha x_1)$ is a canonical monomial of degree $d$.

Proof. Assume w.l.o.g. that $f'(x_1, x_2) \triangleq f(x_1, x_2 + \alpha x_1) = \sum_{e:0 \leq e,d-e < q} \alpha_e x_1^{e} x_2^{d-e}$ (we can ignore monomials of degree smaller than $d$). Let $e_{\max}$ be the maximal degree of $x_1$ in $f'$. Consider the monomial containing $x_1^{e_{\max}}$. Assume toward a contradiction that there is an $i$ such that $p^i \leq d - e_{\max}$ and $e_{\max} + p^i < q$. Consider the polynomial $f'(x_1, x_2 + z x_1)$. By (3.1) it follows that

$$f'(x_1, x_2 + z x_1) \equiv_p \sum_{e:0 \leq e,d-e < q} \alpha_e x_1^{e} \sum_{r \leq p^i} \left(\frac{d-e}{r}\right) (z x_1)^r x_2^{d-e-r}.$$

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

$$\sum_{r \leq e_{\max} + p^i} \alpha_{e_{\max} + p^i - r} \left(\frac{d - \left(e_{\max} + p^i - r\right)}{r}\right) z^r.$$

$$\alpha_{e_{\max}} \left(\frac{d - e_{\max}}{p^i}\right) z^{p^i} + \sum_{r \leq e_{\max} + p^i \land r \neq p^i} \alpha_{e_{\max} + p^i - r} \left(\frac{d - \left(e_{\max} + p^i - r\right)}{r}\right) z^r,$$

where some of the binomials $\left(\frac{d - \left(e_{\max} + p^i - r\right)}{r}\right)$ 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 $x_1^{e_{\max} + p^i} x_2^{d-\left(e_{\max} + p^i\right)}$ is a non-zero polynomial in $z$. It follows that there is some $\beta \neq 0$ such that the coefficient of $x_1^{e_{\max} + p^i} x_2^{d-\left(e_{\max} + p^i\right)}$ in $f'(x_1, x_2 + \beta x_1) = f(x_1, x_2 + (\alpha + \beta) x_1)$ is non-zero. This contradicts our choice of $\alpha$.

We are now ready to prove Lemma 3.4.3.
Lemma 3.4.6. 

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 \(x_1^{q-p} \cdots x_k^{q-p}\) decreases by \(q - q/p\) on any subspace. Proposition 3.1.2 is an (almost) immediate consequence of this lemma.

**Lemma 3.4.6.** Let \(k \geq (d+1)/(q - q/p)\) and let \(f : \mathbb{F}_q^{k+1} \rightarrow \mathbb{F}_q\) have degree larger than \(d\). Then \(\rho_d(f,k) \geq 1/q\).
Proof. Let $A$ be the invertible linear transformation guaranteed by Lemma 3.4.3. 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 $L_{\bar{\alpha},z}(x_1, \ldots, x_{k+1}) = x_i - (\sum_{j=i+1}^{k+1} \alpha_j x_j + z x_{k+1} + \alpha_0)$. Notice 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, $L_{\bar{\alpha},z}(x_1, \ldots, x_{k+1}) = x_1 - (\sum_{j=2}^{k} \alpha_j x_j + z x_{k+1} + \alpha_0)$.

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

$$\tilde{f} = x_1^{e_1 \cdot e_{\max}} \cdot h_{e_{\max}}(x_2, \ldots, x_m) + \sum_{\nu < e_{\max}} x_1^{\nu} \cdot h_{\nu}(x_2, \ldots, x_m),$$

where $h_{e_{\max}} \neq 0$. Let $\hat{f}$ be such that $\tilde{f} = \hat{f} + \tilde{f}$. 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 $\nu$ 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,

$$\deg(x_{k+1}^{e_{\max}} h_{e_{\max}}(x_2, \ldots, x_m)) = \deg(x_1^{e_1 \cdot e_{\max}} h_{e_{\max}}(x_2, \ldots, x_m)) = \deg(\tilde{f}) = \deg(f).$$

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 $\deg(f|_{L_{\bar{\alpha},z}=0}) \geq e_1 + \ldots + e_k \geq q(k-q/p) \geq d+1$. 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 3.4.7.** Let \(n \geq k \geq (d+1)/(q - q/p)\) and let \(f: \mathbb{F}_q^n \rightarrow \mathbb{F}_q\) 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 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 4.5.3 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 \(\Pr_{B}(\deg(\langle f\rangle|_B)) > d) \geq q^{-(n-(n-1)-k)} = q^{-(n-k)}\). Thus, \(\rho_d(f, k) \geq \rho_d(f, n-1) \cdot q^{-(n-k)} \geq q^{-(n-k)}\).

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 3.1.2.

**Proof of Proposition 3.1.2.** The fact that \(t_{q,d} \leq [(d+1)/(q - q/p)]\) follows easily from Lemma 4.5.3.

To see that \(t_{q,d} \geq [(d+1)/(q - q/p)]\) let \(t = [(d+1)/(q - q/p)]\). Consider the function \(f(x_1, \ldots, x_t) = \prod_{i=1}^t x_i^{q-q/p}\). Observe that \(f\) has degree \(t(q - q/p) \geq d + 1\) but when we restrict \(f\) to any \((t-1)\)-dimensional affine subspace its degree drops to at most \((t-1)(q - q/p) \leq d\) (it is not hard to check that the smallest decrease in degree is obtained for some substitution of the form \(x_i = \alpha\)). Thus, the testing dimension is at least \(t\).

### 3.4.4 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 \(N_0(q, d) \leq \tilde{N}_0(q, d) \triangleq q^{\frac{|d-q/p|}{q-q/p} + 1}\). Observe that

\[
q^{t_{q,d}} - 1 < N_0(q, d) \leq \tilde{N}_0(q, d) = q^{\frac{|d-q/p|}{q-q/p} + 1} < q^{t_{q,d} + 1}
\]

As a first step we will bound the number of such hyperplanes that ‘depend’ on \(x_1\).

**Lemma 3.4.8.** 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 4.5.3. We basically show that after fixing some coefficients in a linear function, the number of completions to linear functions \(L\) that have those fixed coefficients and such that \(\deg(f|_{L=0}) < \deg(f)\) is small.
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 monomial. Let $e_{\text{max}}$ be the maximal degree of $x_1$ in those monomials. W.l.o.g. assume that $M = x_{i_1}^{e_{\text{max}}} \cdot \prod_{i=2}^{t} x_i^{e_i}$ is such a monomial in $f$. For a linear function $L(x_1, \ldots, x_n) = x_1 + \sum_{i=2}^{n} a_i x_i + a_0$ denote $L_0(x_2, \ldots, x_n) = - (\sum_{i=2}^{n} a_i x_i)$ and $L_1(x_{t+1}, \ldots, x_n) = - (\sum_{i=t+1}^{n} a_i x_i + a_0)$. 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,

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

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

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

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

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

Consider the homogeneous part of degree $e_{\text{max}}$ of $g_{\bar{e}}$, denoted $g_{\bar{e}}^{(e_{\text{max}})}$. It clearly contains $x_1^{e_{\text{max}}}$ as a monomial. Observe further that $\deg(g_{\bar{e}}|_{x_1=L_1}) < e_{\text{max}} \iff \deg(g_{\bar{e}}^{(e_{\text{max}})}|_{x_1=L_1}) < e_{\text{max}}$. However, since $g_{\bar{e}}^{(e_{\text{max}})}$ is homogeneous of degree strictly smaller than $q$, this happens if and only if $g_{\bar{e}}^{(e_{\text{max}})}|_{x_1=L_1} = 0$. 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 $\deg(g_{\bar{e}}^{(e_{\text{max}})}) \leq q - 1$, 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 $g_{\bar{e}}^{(e_{\text{max}})}$.

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. \hfill $\square$

The following lemma extends the argument to functions that do not necessarily depend on $x_1$.

**Lemma 3.4.9.** 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 3.4.8 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 ∀0 ≤ e ≤ 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_{e_1}|_{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_1}, 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) · 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 3.1.5. For sake of readability we repeat it here (in a slightly different form).

**Theorem.** Theorem 3.1.5 restated Let f : \mathbb{F}_q^n → \mathbb{F}_q be a polynomial of degree d. Then there are at most \(\tilde{N}_0(q,d) = \left\lceil \frac{d−q/p}{q−q/p} \right\rceil+1\) linear functions L such that deg(f|_{L=0}) < d. In particular, \(N_0(q,d) \leq \tilde{N}_0(q,d)\).

**Proof.** Notice that it is enough to prove the theorem for the polynomial f ∘ A where A : \mathbb{F}_q^n → \mathbb{F}_q^n is an invertible linear transformation. Let A be the linear transformation guaranteed by Lemma 3.4.3. Namely, it is such that f ∘ 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} ≥ q so therefore, \(d = \sum_{i=1}^t e_i = (e_1 + \ldots + e_{t−2}) + (e_{t−1} + e_t) ≥ (t−2)(q−q/p) + q\) and hence, \(t ≤ \frac{d−q}{q−q/p} + 2 = \frac{d−q}{q−q/p} + 1\). Since t is an integer we actually get that \(t ≤ \left\lceil \frac{d−q/p}{q−q/p} \right\rceil+1\). Invoking Lemma 3.4.9 we conclude that there are at most \(q^{t} ≤ \left\lceil \frac{d−q/p}{q−q/p} \right\rceil+1 = \tilde{N}_0(q,d)\) linear functions L such that deg(f|_{L=0}) < d.

**Corollary 3.4.10.** 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 \(\mathbb{F}_q\). If there exist K hyperplanes \(A_1, \ldots, A_K\), such that for all \(i \in [K]\) deg f|_{A_i} ≤ d' < d, then deg f ≤ d'.

**Proof.** Assume for contradiction that d' < deg(f) = d. Then, by Theorem 3.1.5 there are at most \(N_0(q,d') ≤ N_0(q,d) < K\) hyperplanes A on which deg(f|_{A}) < d. This contradicts our assumption that there are at least K hyperplanes \(\{A_i\}\) on which deg(f|_{A_i}) ≤ d'.

### 3.4.5 Interpolating from Exact Agreement

In this section we prove Theorem 3.1.7 that shows that if we have enough ‘pairwise consistent’ polynomials then it is possible to obtain ‘global’ consistency. We first restate the theorem.

**Theorem.** Theorem 3.1.7 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 \(P_i|_{A_i ∩ A_j} = P_j|_{A_i ∩ A_j}\) for every pair \(i, j \in [K]\). If

\[K ≥ \tilde{N}_1(q,d) \overset{Δ}{=} 2\tilde{N}_0(q,d + q) · q^{\lambda_{q,6}} = 2q\left\lceil \frac{d}{q−q/p} \right\rceil+2+\lambda_{q,6},\]

where \(\lambda_{q,6}\) is the constant \(\lambda_{q,c}\) from Theorem 3.3.4 for c = 6, then there exists a polynomial \(Q\), of degree d, such that \(Q|_{A_i} = P_i|_{A_i}\) for every \(i \in [K]\).

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

\[K ≥ \tilde{N}_1(q,d,n) \overset{Δ}{=} \frac{\tilde{N}_1(q,d)}{2\prod_{i=1}^{n−\log_q N_1(q,d)−3} \left(1 − \frac{\tilde{N}_1(q,d)}{q^{d−1}}\right)},\]
This is indeed a stronger statement as the denominator above

\[ 2 \prod_{i=1}^{n-\log_q \tilde{N}_1(q,d)-3} \left( 1 - \frac{\tilde{N}_1(q,d)}{q^{n-i-1}} \right) \geq 2 \left( 1 - \sum_{i=1}^{n-\log_q \tilde{N}_1(q,d)-3} \frac{\tilde{N}_1(q,d)}{q^{n-i-1}} \right) \]

\[ = 2 - \frac{2\tilde{N}_1(q,d)}{q^{n-1}} \sum_{i=1}^{n-\log_q \tilde{N}_1(q,d)-3} q^i > 2 - \frac{2\tilde{N}_1(q,d)}{q^{n-1}} q^{n-\log_q \tilde{N}_1(q,d)-2} = 2 - 2q^{-1} \geq 1, \]

namely, \( \tilde{N}_1(q,d,n) < \tilde{N}_1(q,d) \) for all \( n \), and so the requirement on \( K \) is weaker.

We first set some notation. Let \( L_i \in \text{Aff}_q^n \) 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 \( L = \{ L_1, \ldots, L_K \} \). We will abuse notations and denote, for \( L \in L \), \( P_L = P_i \) and \( A_L = A_i \) when \( L = L_i \). Another important notation is the following. For \( L \in \text{Aff}_q^n \) and \( \gamma \in \mathbb{F}_q \) we denote

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

Note that for \( \gamma_1, \gamma_2 \), the hyperplanes \( B_{L,\gamma_1} \) and \( B_{L,\gamma_2} \) are shifts of each other (they can also be empty sets if \( L \) is a constant function).

The proof is by induction on the number of variables \( n \). The idea of the proof is to find a linear function \( L \) and restrict our attention to the different hyperplanes \( B_{L,\gamma} \). We show that we can find an \( L \) such that the induction assumption holds for every \( B_{L,\gamma} \). By the induction hypothesis, for each \( B_{L,\gamma} \) there is a polynomial \( P_{\gamma} \), of degree \( d \), that is defined over \( B_{L,\gamma} \) and is consistent there with the \( P_i \)'s. Then we ‘glue’ the \( P_i \)'s together and use Theorem 3.1.5 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 \( n < \log_q \tilde{N}_1(q,d) + 4 \). For such \( n \) it holds that \( \frac{\tilde{N}_1(q,d,n)}{2} = \frac{1}{2} \tilde{N}_1(q,d) \). The analysis of this case, which is the technical heart of the proof, is given in the next lemma.

**Lemma 3.4.11** (Main Lemma). Let \( n < \log_q \tilde{N}_1(q,d) + 4 \) and \( K \geq \tilde{N}_1(q,d,n) = \frac{\tilde{N}_1(q,d)}{2} \). Let \( A_1, \ldots, A_K \) be distinct hyperplanes in \( \mathbb{F}_q^n \) and let \( P_1, \ldots, P_K \) be polynomials of degree \( d \) satisfying \( P_i|_{A_i \cap A_j} = P_j|_{A_i \cap A_j} \) for every \( i, j \in [K] \). 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 3.4.11 and continue with the proof of the theorem. As Lemma 3.4.11 takes care of the case \( n < \log_q \tilde{N}_1(q,d) + 4 \) we assume for the rest of the proof that \( n \geq \log_q \tilde{N}_1(q,d) + 4 \). 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 3.4.12.** There is a linear function \( L \in \text{Aff}_q^n \) such that for every \( \gamma \in \mathbb{F}_q \) the number of distinct affine subspaces \( A_{i,L,\gamma} \subseteq B_{L,\gamma} \), such that \( A_{i,L,\gamma} \neq \emptyset \), is at least \( \tilde{N}_1(q,d,n-1) \).

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 \( \tilde{N}_1(q,d,n) \leq K \leq \tilde{N}_1(q,d) \). 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 \( L = \alpha L_i + \gamma' \), for some \( \alpha \in \mathbb{F}_q^* \) and \( \gamma' \in \mathbb{F}_q \). Using these two observations we perform a simple counting argument that shows that there
is some $L \in \text{Aff}_q^n$ such that for every $\gamma$, 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 $\alpha, \beta, \gamma \in \mathbb{F}_q$ and $L_1, L_2 \in \mathcal{L}$. Since there are $q^3 K^2$ such possible representations, there exists $L \in \text{Aff}_q^n$ that can be represented in at most $\frac{q^3 K^2}{q^ {n+1}} = \frac{K^2}{q^{n-2}}$ 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

$$K' = K - \frac{K^2}{q^{n-2}} = K \left(1 - \frac{K}{q^{n-2}}\right)$$

$$\geq \tilde{N}_1(q, d) \left(1 - \frac{\tilde{N}_1(q, d)}{q^{n-2}}\right)$$

$$= \left(1 - \frac{\tilde{N}_1(q, d)}{q^{n-2}}\right) 2 \prod_{i=1}^{n-\log_q \tilde{N}_1(q, d)-3} \frac{\tilde{N}_1(q, d)}{q^{n-1}}$$

$$= \tilde{N}_1(q, d)$$

$$\geq \tilde{N}_1(q, d)$$

$$\geq \tilde{N}_1(q, d)$$

$$= \tilde{N}_1(q, d)$$

$$= \tilde{N}_1(q, d, n-1).$$

We proceed with the proof of Theorem 3.1.7. Let $L \in \text{Aff}_q^n$ be as promised by Claim 3.4.12. 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 $A'_i = A_i \cap B_{L,\gamma}$ 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 $\tilde{N}_1(q, d, n-1)$ 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 $P_1, \ldots, P_k$. 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 $B_{L,\gamma} = \{v \in \mathbb{F}_q^n \mid L(v) = \gamma\}$. As the hyperplanes $\{B_{L,\gamma}\}_{\gamma \in \mathbb{F}_q}$ cover all of $\mathbb{F}_q^n$, 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

\footnote{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 $L_i$.}
$P_i(u) = P_{L_i = \gamma}(u) = P(u)$, where the first equality holds since, by the induction hypothesis, $P_i$ and $P_{L_i = \gamma}$ agree on $A_i' = A_i \cap B_{L_i, \gamma}$.

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

$$K \geq \tilde{N}_1(q, d, n) = \tilde{N}_1(q, d) / 2 > \tilde{N}_0(q, d + q) \geq N_0(q, d + q),$$

Corollary 3.4.10 implies that the degree of $P$ is, in fact, at most $d$. This completes the proof of Theorem 3.1.7 modulo the proof of Lemma 3.4.11 that we give next.

**Proof of Lemma 3.4.11** As before, we let $L_i \in \text{Aff}_q^n$ be an affine linear function such that $A_i = \{ u \in F_q^n \mid L_i(u) = 0 \}$ and denote $\mathcal{L} = \{ L_1, \ldots, L_K \}$. Again we abuse notations and denote, for $L \in \mathcal{L}$, $P_L = P_L$ and $A_L = A_L$ when $L = L_i$.

We will first use the assumption that $n \leq \log_q \tilde{N}_1(q, d) + 4$ and $K \geq \tilde{N}_1(q, d, n) = \tilde{N}_1(q, d)/2 = \tilde{N}_0(q, d + q) \cdot q^{5/3}$ to show that the set $\mathcal{L}$ contains at least $\log_q (\tilde{N}_0(q, d + q))$ lines in linearly independent directions. Indeed, we can think of $\mathcal{L}$ as a set of points in $\text{Aff}_q^n$ which is an $(n + 1)$-dimensional space over $\mathbb{F}_q$. By our setting of parameters it follows that

$$\frac{|\mathcal{L}|}{|\text{Aff}_q^n|} = \frac{K}{q^{n+1}} \geq \frac{K}{q^{\log_q N_1(q, d)+5}} = \frac{K}{N_1(q, d)} \cdot q^{-5} \geq q^{-6}.$$ 

Thus, in order to apply Corollary 3.3.5 we just need to bound $\dim(\text{Aff}_q^n)$ 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,

$$\dim(\text{Aff}_q^n) = n + 1 \geq \log_q(K) \geq \left\lceil \frac{d}{q - q/p} \right\rceil + 2 + \lambda_{q,6} = \left\lceil \frac{d + q - q/p}{q - q/p} \right\rceil + 1 + \lambda_{q,6}.$$ 

Corollary 3.3.5 now implies that there are at least $\left\lceil \frac{d + q - q/p}{q - q/p} \right\rceil + 2$ combinatorial lines inside $\mathcal{L}$ whose directions are linearly independent. In particular, there are

$$t \geq \left\lceil \frac{d + q - q/p}{q - q/p} \right\rceil + 1 \quad (3.2)$$

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 $x_1, \ldots, x_t$. I.e we can assume that there exist $t$ linear functions $L_1, \ldots, L_t$ 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 3.1 on page 38.

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

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 $B_{x_i, \gamma} \overset{\Delta}{=} \{ v \in \mathbb{F}_q^n \mid v_i = \gamma \}$ 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 $\cup_{\gamma \neq 0} B_{x_i, \gamma} = \{ v \in \mathbb{F}_q^n \mid v_i \neq 0 \}$, by a simple interpolation of $\{ P_{x_i = \gamma} \mid \gamma \in \mathbb{F}_q^* \}$. 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

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

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

where $(\ast)$ follows from the fact that $L_i(u) - \gamma - 1, \beta' u_i = L_i(u) - \gamma - 1, \beta' = L_i(u) - \beta' = 0$.

Indeed, this implies that $u \in A_{L_i(u) - \gamma - 1, \beta' x_i}$ and now $(\ast)$ 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 3.4.10 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 $P_1, \ldots, P_k$ on the set $\{v \in \mathbb{F}_q^n | v_i \neq 0\}$. Moreover, $\{P_{x_i \neq 0} \}_{i = 1}^t$ are consistent with each other. I.e., for every $v = (v_1, \ldots, v_n) \in \mathbb{F}_q^n$ 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 $\bar{e} = (e_1, \ldots, e_n)$, $0 \leq e_1, \ldots, e_n \leq q - 1$, define the $\bar{e}$-monomial$^\dagger$ $M_{\bar{e}}(x)$ to be

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

$\dagger$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.
Clearly, $\deg(M_\vec{e}) = \sum_{i=1}^n e_i$. We say that $M_\vec{e}$ 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 $f(x) = \sum c_\vec{e} M_\vec{e}(x)$, where $c_\vec{e} \in \mathbb{F}_q$. 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 3.4.13.** Let $I \subseteq [n]$ be a set of indices. Denote $S_I = \{ v \in \mathbb{F}_q^n \mid \forall i \in I : v_i \neq 0 \}$. Let $g, h : \mathbb{F}_q^n \to \mathbb{F}_q$ be two polynomials that agree on $S_I$, namely, $\forall v \in S_I$, $g(v) = h(v)$. Then, the coefficient of any $\vec{e}$-monomial $M$ that is not of full degree in any $x_i$ for $i \in I$, is the same in both $g$ and $h$.

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

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

$$q^{\lvert \mathbb{F}_q^n \setminus S_I \rvert} = q^{\{v \in \mathbb{F}_q^n \mid \exists \vec{e} \in I, e_i = q - 1\}}.$$ 

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

$$\sum_{\vec{e} : \exists i \in I, e_i = q - 1} c_\vec{e} M_\vec{e}(x).$$

This number is equal to $q^{\{\vec{e} \mid \exists i \in I, e_i = q - 1\}}$. Clearly,

$$\# \{ v \in \mathbb{F}_q^n \mid \exists i \in I, v_i = 0 \} = \# \{ \vec{e} \in \{0, \ldots, q - 1\}^n \mid \exists i \in I, e_i = q - 1 \} .$$

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

We continue with the proof of Lemma 3.4.11. By uniqueness of representation, for any $m \in [t]$, $P_{x_m \neq 0}$ can be expressed as

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

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 3.4.14.** 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 $\{ v \in \mathbb{F}_q^n \mid v_k, v_m \neq 0 \}$. Lemma 3.4.13 implies that they have the same coefficient for any $\vec{e}$-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. $\square$
We continue with the proof of the main lemma. For every $J \subseteq [t]$ define $Q_J = Q_J^n$, where $m \in [t] \setminus J$ is arbitrary. By Claim 3.4.14, $Q_J$ is well defined. Now we can define a polynomial $P$ that is consistent with $\{P_1, \ldots, P_K\}$ on all of $\mathbb{F}_q^n$.

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

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

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

As this holds for every $J \subseteq [t]$ 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 3.4.15.** 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,

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

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. 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,

$$\deg(P - P_k)(x, v) = (q - 1)t \geq (q - 1) \cdot \left[\left\lfloor \frac{(d + q) - q/p}{q - q/p} \right\rfloor + 1 \right]$$

$$\geq (q - 1) \cdot \left( \frac{d}{q - q/p} + 1 \right) \geq d + q - 1,$$

where the first inequality follows from Equation (3.2). 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 3.4.11 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 3.4.10, $\deg(P) \leq d$ as required.
3.4.6 Interpolating from Approximate Agreement

We use Theorem 3.1.7 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 3.4.16.** Let $\delta_1 < \frac{1}{d}q^{-(1+(d/(q-1))}$ and $K \geq N_1(q,d)$. If the function $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$ 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 3.3.2, 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 3.3.2).

Next, we use Theorem 3.1.7, 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 3.1.7, 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 chosen 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 $Z_i = Y_i - \text{Exp}[Y_i] = Y_i - 1/q$. Clearly, $\text{Exp}[Z_i^2] = \text{Exp}[Y_i^2] - \text{Exp}[Y_i]^2 = 1/q - 1/q^2$. 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 $\text{Exp}[Z_i \cdot Z_j] \leq 0$, and so $\text{Exp}[(\sum_{i \in [K]} Z_i)^2] \leq \sum_{i \in [K]} \text{Exp}[Z_i^2] = K(q-1)/q^2$. We thus conclude that

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

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 $\delta_1/q$, 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 $\left(\delta(f,Q) - \tau\right)/(2q)$, 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 $\delta_1 < \frac{1}{2q}q^{-(1+d/(q-1))}$, then $\delta_d(f) \leq 2\delta_1 + 4(q-1)/K$. \qed
3.5 Analysis of the low-degree tests

Lemma 3.5.1. Let \( t \geq d/(q-1) \) be an integer. Then, if \( \delta_d(f) \leq \frac{1}{2}q^{-d/(q-1)} \) then \( \rho_d(f,t) \geq \min\left\{ \frac{1}{4q}, \frac{1}{2} \cdot q^t \cdot \delta_d(f) \right\} \).

Proof. We will use the monotonicity of the rejection probability \( \rho_d(f,\cdot) \) (Lemma 3.4.5) 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 \( \ell, d/(q-1) \leq \ell \leq t \), we claim that the probability that on a randomly chosen \( \ell \)-dimensional affine subspace \( A, f|_A \) and \( g|_A \) disagree on exactly one point is at least \( q^\ell \cdot \delta \cdot (1-(q^\ell-1) \cdot \delta) \).

Indeed, the argument is quite routine so we only sketch it. Let \( x \) be a point on the \( \ell \)-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 \( \Pr[f(x) \neq g(x)] - \sum_{y \in A, y \neq x} \Pr[f(y) \neq g(y)] \) and \( f(x) \neq g(x) \) and \( f(x) \neq g(x) \) is \( \delta - (q^\ell-1)\delta^2 \), 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 exactly in one point is at least \( q^\ell \cdot \delta \cdot (1-(q^\ell-1) \cdot \delta) \).

Note that, in this case, \( f|_A \) is not a degree \( d \) polynomial, since, by Lemma 3.3.2, two polynomials cannot differ on only one point of \( A \).

The fact above allows us to analyze \( \rho_d(f,\ell) \) as follows: Since the \( \ell \)-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 \( \ell \) be the largest integer such that \( \delta \leq \frac{1}{2}q^{-\ell} \) (and so \( \delta > \frac{1}{2}q^{-\ell} \)). 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 the second inequality follows by the previous argument.

Lemma 3.5.2. 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 \( \epsilon = \frac{1}{2} \cdot \frac{1}{2q} \) and \( c = \log_q \hat{N}_1(q,d) - t_{q,d} + \log_q 128 \). Recall that

\[
\hat{N}_1(q,d) = 2\hat{N}_0(q,d + q) \cdot q^{\lambda_{q,6}} = 2q \left[ \frac{d}{q^{d/q}} \right]^{2+\lambda_{q,6}}
\]

and observe that

\[
2q^{\lambda_{q,d+1+\lambda_{q,6}}} \leq 2q \left[ \frac{d}{q^{d/q}} \right]^{2+\lambda_{q,6}} \leq 2q^{\lambda_{q,d+2+\lambda_{q,6}}},
\]

where \( \lambda_{q,6} \) is defined in Theorem 3.3.4, 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 \( \frac{1}{8} \cdot q^t \cdot \sum_{i=n+1}^{\infty} q^{-i} < \frac{1}{2} \) 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 \( \rho_d(f,t) = \text{Exp}_A[\rho_d(f|_A,t)] \) and the inductive hypothesis says that \( \rho_d(f|_A,t) \) is usually large. When \( K \) is large, we use Theorem 3.4.16 to show that \( \delta_d(f) \) is small, and this allows us to use Lemma 3.5.1 to claim \( \rho_d(f,t) \) is large in this case also. Details below.

Case 1: \( K < \frac{1}{2}q^t \). 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
\[
\rho_d(f, t) = \text{Exp}_A(\rho_d(f|A, t)) \\
\geq \epsilon + \frac{1}{8}q^t \sum_{i=n}^{\infty} q^{-i} - \frac{1}{8}q^t/q^n \\
= \epsilon + \frac{1}{8}q^t \sum_{i=n+1}^{\infty} q^{-i}
\]
as desired.

**Case 2:** \( K \geq \frac{1}{8}q^t \). Note that
\[
K \geq \frac{1}{8}q^t \geq \frac{1}{8}q^{t_q,d+c} = \frac{1}{8}q^{\log_q \hat{N}_1(q,d)+\log_q 128} > \hat{N}_1(q,d) \geq N_1(q,d).
\]
We thus have by Theorem 3.4.16, \( \delta_d(f) \leq 2q^{-t}+4(q-1)/K \). Using \( 2q^{-t} \leq \frac{1}{3}q^{-d/(q-1)} \) and \( 4(q-1)/K \leq 32 \cdot q^{-t+1} \leq \frac{1}{4}q^{-d/(q-1)} \) (by our choice of \( t \geq \log_q \hat{N}_1(q,d) + \log_q 128 \geq d/(q-1) + \log_q 128 + 1 \)) we conclude in this case that \( \delta_d(f) \leq \frac{1}{4}q^{-d/(q-1)} \). This allows us to use Lemma 3.5.1 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 3.1.3** We analyze two cases depending on \( \delta_d(f) \).

1. \( \delta_d(f) \leq \frac{1}{2}q^{-d/(q-1)} \): Since \( t_{q,d} = [(d+1)/(q - q/p)] \geq d/(q-1) \) we get from Lemma 3.5.1 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 3.5.2 and conclude that there exists constants \( c, \epsilon > 0 \) such that for \( t = t_{q,d} + c \) it holds that \( \rho_d(f, t) \geq \epsilon + \frac{1}{8}q^t \cdot \sum_{i=n+1}^{\infty} q^{-i} \). Applying Lemma 3.4.7 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 3.5.2, \( \epsilon_2 \) 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. \( \square \)
Chapter 4

Absolutely Sound Testing of Lifted Codes

4.1 Introduction

In this work we present results on the testability of “affine-invariant linear codes”. We start with some basic terminology before describing our work in greater detail. Let $q$ be a prime power, $Q$ be a power of $q$ and let $\mathbb{F}_q(\mathbb{F}_Q)$ denote the finite fields of $q$ ($Q$) elements. Finally, let $\{\mathbb{F}_Q^n \rightarrow \mathbb{F}_q\}$ denote the set of functions mapping $\mathbb{F}_Q^n$ to $\mathbb{F}_q$. In this work a code (or a family) will be a subset of functions $\mathcal{F} \subseteq \{\mathbb{F}_Q^n \rightarrow \mathbb{F}_q\}$. We use $\delta(f,g)$ to denote the normalized Hamming distance between $f$ and $g$, i.e., the fraction of inputs $x \in \mathbb{F}_Q^n$ for which $f(x) \neq g(x)$. We use $\delta(\mathcal{F})$ to denote $\min_{f \neq g, f, g \in \mathcal{F}} \{\delta(f,g)\}$ and $\delta(\mathcal{F})(f)$ to denote $\min_{g \in \mathcal{F}} \{\delta(f,g)\}$. A code $\mathcal{F}$ is said to be a linear code if it is an $\mathbb{F}_q$-subspace, i.e., for every $\alpha \in \mathbb{F}_q$ and $f, g \in \mathcal{F}$, we have $\alpha f + g \in \mathcal{F}$. A function $T : \mathbb{F}_Q^n \rightarrow \mathbb{F}_Q^n$ is said to be an affine transformation if there exists a matrix $B \in \mathbb{F}_Q^{n \times n}$ and vector $c \in \mathbb{F}_Q^n$ such that $T(x) = Bx + c$. The code $\mathcal{F} \subseteq \{\mathbb{F}_Q^n \rightarrow \mathbb{F}_q\}$ is said to be affine-invariant if for every affine transformation $T$ and every $f \in \mathcal{F}$ we have $f \circ T \in \mathcal{F}$ (where $(f \circ T)(x) = f(T(x))$).

When $Q = q$, affine-invariant linear codes form a very natural abstraction of the class of low-degree polynomials: The set of $n$-variable polynomials of degree at most $d$ over $\mathbb{F}_q$ is a linear subspace and is closed under affine transformations. Furthermore, as shown by Kaufman and Sudan [KS08] affine-invariant linear codes retain some of the “locality” properties of multivariate polynomial codes (or Reed-Muller codes), such as local testability and local decodability, that have found many applications in computational complexity. This has led to a sequence of works exploring these codes, but most of the works led to codes of smaller rate than known ones, or gave alternate understanding of known codes [GKS08b, GKS09, BSS11, BSMSS11, BGM+11]. A recent work by Guo et al. [GKS13] however changes the picture significantly. They study a “lifting” operator on codes and show that it leads to codes with, in some cases dramatic, improvement in parameters compared to Reed-Muller codes. Our work complements theirs by showing that one family of “best-known” tests manages to work abstractly for codes developed by lifting.

We start by describing the lifting operation: Roughly a lifting of a base code leads to a code in more variables whose codewords are words of the base code on every affine subspace of the base dimension. We define this formally next. For $f : \mathbb{F}_Q^n \rightarrow \mathbb{F}_q$ and $S \subseteq \mathbb{F}_Q^n$, let $f|_S$ denote the restriction of $f$ to the set $S$. A set $A \subseteq \mathbb{F}_Q^n$ is said to be a $t$-dimensional affine subspace, if there exist $\alpha_0, \ldots, \alpha_t \in \mathbb{F}_Q$ such that $A = \{\alpha_0 + \sum_{i=1}^t \alpha_i x_i | x_1, \ldots, x_t \in \mathbb{F}_Q\}$. We use some arbitrary $\mathbb{F}_Q$-linear isomorphism from $A$ to $\mathbb{F}_Q^t$ to view $f|_A$ as a function from $\{\mathbb{F}_Q^t \rightarrow \mathbb{F}_q\}$. Given an affine-invariant linear base code $\mathcal{B} \subseteq \{\mathbb{F}_Q^n \rightarrow \mathbb{F}_q\}$ and integer $n \geq t$, the $n$-dimensional lift of $\mathcal{B}$, denoted $\text{Lift}_n(\mathcal{B})$, is the set $\{f : \mathbb{F}_Q^n \rightarrow \mathbb{F}_q | f|_A \in \mathcal{B}\}$ for
every \( t \)-dimensional affine subspace \( A \subseteq \mathbb{F}_q^n \).

The lifting operation was introduced by Ben-Sasson et al. [BSMSS11] as a way to build new affine-invariant linear codes that were not locally testable. Their codes were also of much lower rate than known affine-invariant linear codes of similar distance. However in more recent work, Guo et al. [GKS13], showed that lifting could be used positively: They used it to build codes with very good locality properties (especially decodability) with rate much better than known affine-invariant linear ones, and matching qualitatively the performance of the best known codes. Our work attempts to complement their work by showing that these codes, over constant sized alphabets, can be “locally tested” as efficiently as polynomial codes.

Testing and Absolutely Sound Testing A code \( \mathcal{F} \subseteq \{\mathbb{F}_q^n \rightarrow \mathbb{F}_q\} \) is said to be a \((k, \epsilon, \delta)\)-locally testable code (LTC), if \( \delta(\mathcal{F}) \geq \delta \) and there exists a probabilistic oracle algorithm that, on oracle access to \( f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q \), makes at most \( k \) queries to \( f \) and accepts \( f \in \mathcal{F} \) with probability one, while rejecting \( f \not\in \mathcal{F} \) with probability at least \( \epsilon \delta_{\mathcal{F}}(f) \).

For an ensemble of codes \( \{\mathcal{F}_m \subseteq \{\mathbb{F}_q^n \rightarrow \mathbb{F}_q\}\}_m \) for infinitely many \( m \), where the code \( \mathcal{F}_m \) is a \((k(m), \epsilon(m), \delta(m))\)-LTC, we say that the code has an absolutely sound tester if there exists \( \epsilon > 0 \) such that \( \epsilon(m) \geq \epsilon \) for every \( m \).

Any tester can be converted into an absolutely sound one by repeating the test \( 1/\epsilon(m) \) times. However this comes with an increase in the query complexity (the parameter \( k(m) \)) and so it makes sense to ask what is the minimum \( k \) one can get for an absolutely sound test.

Chapter 3 raised this question in the context of multivariate polynomial codes (Reed-Muller codes) and showed that the “natural tester” for multivariate polynomial codes is absolutely sound, without any repetitions! The natural test here is derived as follows for prime fields:

To test if a function \( f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q \) is a polynomial of degree at most \( d \), let \( t \) be the smallest integer such that there exist functions of degree greater than \( d \) in \( t \) variables. Pick a random \( t \)-dimensional affine subspace \( A \) and verify that \( f|_A \) is a degree \( d \) polynomial.

The natural test thus makes roughly \( q^t = q^{(d+1)/(q-1)} \) queries. This number turns out to be optimal for prime fields in that every function looks like a degree \( d \) polynomial if queried at at most \( q^{t-1} \) points. Such optimal analyses of low-degree tests turn out to have some uses in computational complexity: In particular one of the many ingredients in the elegant constructions of Barak et al. [BGH+12] is the absolutely sound analysis of the polynomial codes over \( \mathbb{F}_2 \).

Returning to the natural test above, it ends being a little less natural, and not quite optimal when dealing with non-prime fields. Turns out one needs to use a larger value of \( t \) than the one in the definition above (specifically, \( t = (d+1)/(q-q/p) \) where \( p \) is the characteristic of the field \( \mathbb{F}_q \)). While it is unclear if sampling all the points in the larger dimensional space is really necessary for absolutely sound testing the results so far seem to suggest working with prime fields is a better option.

4.1.1 Our work: motivation and results

The motivation for our work is two-fold: Our first motivation is to understand “low-degree testing” better. Low-degree testing has played a fundamental role in computational complexity and yet its proofs are barely understood. They tend to involve a mix of probabilistic, algebraic, and geometric arguments, and the only setting where the mix of these features seems applicable seems to be the setting of low-degree polynomials. Affine-invariant codes naturally separate the geometry of subspaces in high-dimensional spaces, from the algebra of polynomials of low-degree. Thus extending a proof or analysis method from the setting of low-degree polynomials to the setting of generic geometric...
arguments has the nice feature that it has the potential to separate the geometric arguments from the algebraic ones.

Within the theme of low-degree testing, the previous works have revealed interesting analyses. And several of these variations in the resulting theorems have played a role in construction of efficient PCPs or more recently in other searches for explicit objects. In particular the literature includes tests such as those originally given by Blum, Luby and Rubinfeld [BLR90] for testing linearity and followed by [RS96, AKK+05, KR06, JPRZ09] for testing higher degree polynomials. The aspects of this family of tests are well abstracted in Kaufman and Sudan [KS08]. But the literature contains other very interesting theorems, such as those of Raz and Safra [RS97] and Arora and Sudan [AS03] which tend to work in the “list-decoding” regime. The analysis of the former in particular seems especially amenable to a “generic proof” in the affine-invariant setting and yet such a proof is not yet available. Our work explores a third such paradigm in the analysis of low-degree tests, which was introduced in the above-mentioned “absolutely-sound testers” of Bhattacharyya et al. and generalised in Chapter 3.

Our work starts by noticing that the natural tests above are really “lifting tests”: Namely, the test could be applied to any code that is defined as the lift of a base code with the test checking if a given function is a codeword of the base code when restricted to a random small dimensional affine subspace of the base dimension. Indeed this is the natural way of interpreting almost all the previous results in low-degree testing (with the exception of that of [RZS13]). If so, it is natural to ask if the analysis can be carried out to show the absolute soundness of such tests.

The second, more concrete, motivation for our work is the work of Guo et al. [GKS13]. Over prime fields, it was well-known that lifts of low-degree polynomials lead only to polynomials of the same degree (in more variables). Guo et al. show that lifting over non-prime fields leads to better codes than over prime fields! (Prior to their work, it seemed that working with non-prime fields was worse than working with prime fields.) The improved rate gives motivation to study lifted codes in general, and in particular one class of results that would have been nice to extend was the absolutely-sound tester of chapter 3.

In this chapter we show that the natural test of lifted codes is indeed absolutely sound. The following theorem spells this statement out precisely.

**Theorem 4.1.1 (Main).** For every prime power $q$ and $Q$ a power of $q$, there exists $\epsilon_Q > 0$ such that the following holds: Let $t \leq n$ be positive integers and let $B \subseteq \{ \mathbb{F}_Q^t \rightarrow \mathbb{F}_q \}$ be any affine-invariant linear code. Then $F = \text{Lift}_n(B)$ is $(Q^t, \epsilon_Q, Q^{-1})$-locally testable.

We also note that the literature on property testing includes other general theorems which would imply the testability of the families covered in Theorem 4.1.1 above (see, for instance, [BFH+13]) who use only the affine invariance of the families (and do not need the fact that $F$ is a $\mathbb{F}_q$-vector space). However such testability results are quantitatively much weaker than those of [KS08] and do not satisfy our (somewhat strong) definition of $(k, \epsilon, \delta)$-testability for any positive $\epsilon$.

We stress that the importance of the above is in the absolute soundness, i.e., the fact that $\epsilon_Q$ does not depend on $t$ or $B$. If one is willing to let $\epsilon_Q$ depend on $t$ and $B$ then such a result follows from the main theorem of [KS08].

This chapter also sets into proper light the previous chapter who show that the “natural test” for degree $d$ polynomials over the field $\mathbb{F}_q$ of characteristic $p$ makes $q^{(d+1)/(q-q/p)}$ queries and is absolutely sound. Our result does not mention any dependence on $p$, the characteristic of the field. It turns out that such a dependence comes due to the following proposition.

Let $\text{RM}(n, d, q)$ denote the set of polynomials of degree at most $d$ in $n$ variables mapping $\mathbb{F}_q^n$ to $\mathbb{F}_q$. 

59
Proposition 4.1.2. For positive integers $d$ and $q$ where $q$ is a power of a prime $p$, let $t = t_{d,q} = \left\lceil \frac{d+1}{q^{1/p}} \right\rceil$. Then for every $n \geq t$, the Reed-Muller code $\text{RM}(n, d, q)$ equals the code $\text{Lift}_n(\text{RM}(t, d, q))$.

Applying Theorem 4.1.1 to $\text{RM}(n, d, q)$ we immediately obtain the main results of [BKS+10] and chapter 3. And the somewhat cumbersome dependence on the characteristic of $q$ can be blamed on the proposition above, rather than any weakness of the testing analysis. Furthermore, as is exploited by Guo et al. [GKS13] if one interprets the proposition above correctly, then one should use lifts of Reed-Muller codes over non-prime fields with dimension being smaller than $t_{d,q}$. These will yield codes of higher rate while our main theorem guarantees that testability does not suffer.

One concrete consequence of our result is in the use of Reed-Muller codes in the work of Barak et al. [BGH+12]. They show how to construct small-set expander graphs with many large eigenvalues and one of the ingredients in their result is a tester of Reed-Muller codes over the domain $\mathbb{F}_2^n$ (codes obtained by lifting an appropriate family of base codes over domain $\mathbb{F}_2^t$). Till this work, the binary Reed-Muller code seemed to be the only code with performance good enough to derive their result. Our work shows that using codes over the domain $\mathbb{F}_4^n$ or $\mathbb{F}_8^n$ (or any constant power of two) would serve their purpose at least as well, and even give slight (though really negligible) improvements. We elaborate on these codes and their exact parameters in Section 4.7. (In particular, see Theorem 4.7.3.)

Finally, unlike the works of Bhattacharyya et al., and Haramaty et al., we cannot claim that our testers are “optimal”. This is not because of a weakness in our analysis, rather it is due to the generality of our theorem. For some codes, including the codes considered in the previous works, our theorem is obviously optimal (being the same test and more or less same analysis as previously). Other codes however may possess special properties making them testable much better. In such cases we can not rule out better tests, though we hope our techniques will still be of some use in analyzing tests for such codes.

Future research directions As noted earlier, the field of low-degree testing has seen several different themes in the analyses. Combined with the work of Kaufman and Sudan [KS08] our work points to the possibility that much of that study can be explained in terms of the geometry of affine-invariance, and the role of algebra can be encapsulated away nicely. One family of low-degree tests that would be very nice to include in this general view would be that of Raz and Safra [RS97]. Their work presents a very general proof technique that uses really little algebra; and seems ideally amenable to extend to the affine-invariant setting. We hope that future work will address this. We also hope that future work improve the dependence of $\epsilon_Q$ on $Q$ in Theorem 4.1.1 (which is unfortunately outrageous). Indeed it is not clear why there should be any dependence at all and it would be nice to eliminate it if possible.

Organization We give an overview of the proof of Theorem 4.1.1 in Section 4.2, where we also introduce the main technical theorem of this paper (Theorem 4.2.1). We also describe our technical contributions in this section, contrasting the current proof Theorem 3.1.3, which we modify. The remaining sections are devoted to the formal proof of Theorem 4.1.1. Specifically we introduce some of the background material in Section 4.3. We then prove Theorem 4.2.1 in Section 4.4. In Section 4.5 we show how to prove Theorem 4.1.1 for the special case in which $Q = q$ using Theorem 4.2.1. In Section 4.6 we then prove Theorem 4.1.1 for general $Q$ via a simple reduction to the case in which $Q = q$. Finally, in Section 4.7 we give an example of a family of lifted codes for which our main theorem applies.
4.2 Overview of Proof

From now on we shall deal only with the special case in which $Q = q$. In Section 4.6 we show a simple reduction from the general $Q \neq q$ case to the case in which $Q = q$.

4.2.1 Some natural tests

Our proof of Theorem 4.1.1 follows the paradigm used in Chapter 3. There, we consider a natural family of tests (and not just the “most” natural test), and analyze their performance by studying the behavior of functions when restricted to “hyperplanes”. We introduce the family of tests first.

From now onwards all codes we consider will be linear and affine-invariant unless we explicitly say otherwise. Given a base code $B \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ and $n \geq \ell \geq t$, we let $\mathcal{L}_\ell = \text{Lift}_\ell(B)$, with $\mathcal{F} = \mathcal{L}_n$. The $\ell$-dimensional test for membership in $\mathcal{F}$ works as follows: Pick a random $\ell$-dimensional affine subspace $A$ in $\mathbb{F}_q^n$ and accept $f$ if and only if $f|_A \in \mathcal{L}_\ell$.

Let $\text{Rej}_\ell(f)$ denote the probability with which the $\ell$-dimensional test rejects. Our main theorem aims to show that $\text{Rej}_\ell(f) = \Omega(\delta(f))$ when $\ell = t$. As in previous works, our analysis will first lower bound $\text{Rej}_\ell(f)$ for $\ell = t + O(1)$ and then relate the performance of this test to the performance of the $t$-dimensional test.

4.2.2 Overview of proof of Main Theorem 4.1.1

The analysis of the performance of the $\ell$-dimensional tests is by induction on the number of variables $n$ and based on the behaviour of functions when restricted to “hyperplanes”. A hyperplane in $\mathbb{F}_q^n$ is an affine subspace of dimension $n - 1$.

The inductive strategy to analyzing $\text{Rej}_\ell(f)$ is based on the observation that $\text{Rej}_\ell(f) = \mathbb{E}_H[\text{Rej}_\ell(f|_H)]$ where $H$ is a uniform hyperplane. If we know that on most hyperplanes $\delta = \delta(\ell,f|_H)$ is large, then we can prove the right hand side above is large by induction. Thus the inductive strategy relies crucially on showing that if $f$ is far from $\mathcal{F}$, then $f|_H$ can not be too close to $\mathcal{L}_n$ on too many hyperplanes. We state this technical result in the contrapositive form below.

**Theorem 4.2.1 (Main technical).** For every $q$ there exists $\tau < \infty$ such that the following holds: Let $B \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$ be an affine-invariant linear code and for $\ell \geq t$ let $\mathcal{L}_\ell = \text{Lift}_\ell(B)$. For $n > t$, let $f : \mathbb{F}_q^n \to \mathbb{F}_q$ be a function and $H_1, \ldots, H_k$ be hyperplanes in $\mathbb{F}_q^n$ such that $\delta(f|_{H_i}) \leq \delta$ for every $i \in [k]$ for $\delta < \frac{1}{2}q^{-(t+1)}$. Then, if $k \geq q^{t+\tau}$, we have $\delta = \delta(f) \leq 2\delta + 4(q - 1)/k$.

The theorem thus states that if $f$ is sufficiently close to a lift of $B$ on a sufficiently large number of hyperplanes, yet a very small number (independent of $n$) of hyperplanes, then $f$ is close to a lift of $B$. The dependence of the number of hyperplanes on $q$ and $t$ is actually important to our (and previous) analysis. The fact that it is some fixed multiple of $q^t$, where the multiple depends only on $q$ and not on $t$, is crucial to the resulting performance.

Going from Theorem 3.1.3 above to Theorem 4.1.1 is relatively straightforward. In particular using Theorem 4.2.1 we can get a lower bound on $\text{Rej}_{t+\tau}(f)$ without any changes to the proof of Theorem 3.1.3. However going from such an analysis to a lower bound on $\text{Rej}_t(f)$ involves some extra work, with complications similar to (but simpler than), those in the proof of Theorem 4.2.1 so we omit a discussion here. Section 4.5 contains all the details.

The main contribution of this paper is the proof of Theorem 4.2.1. Here, the proof of Theorem 3.1.3 crucially relied on properties of polynomials and in particular the first step in the proof, when testing degree $d$ polynomials, is to consider the case of $f$ being a degree $d+1$ (or a degree $d+q$) polynomial. In our case there is no obvious candidate for the notion of a degree $d+1$ polynomial and it is abstracting...
such properties that forms the bulk of our work. In what follows we give an overview of some of the issues arising in such steps and how we deal with them.

4.2.3 Overview of proof of Theorem 4.2.1

To understand our proof of Theorem 4.2.1 we need to give some background, specifically to the proofs from Chapter 3. Recall the analogous statement there attempting to show that if \( f \) was far from being a polynomial of degree \( d \), then the number of hyperplanes where \( f \) turns out to be close to being a degree \( d \) polynomial is at most \( O(q^d) \) (where \( t \approx d/q \), the exact number will not be important to us). In Chapter 3 we reasoned about this in a sequence of steps: (1) We first showed that any function of degree greater than \( d \), stays of degree greater than \( d \) on at least \( 1/q \) fraction of all hyperplanes (provided \( n > t \)). (2) Next we reasoned about functions of degree \( d + 1 \) and showed that such a function reduces its degree on at most \( O(q^d) \) hyperplanes. (3) In the third step we consider a general function \( f \) that is far from being of degree \( d \) and show that the number of hyperplanes on which \( f \) becomes a degree \( d \) polynomial exactly is \( O(q^d) \). (This is the step where the big-Oh becomes a really big-Oh.) (4) Finally, we show that for functions of the type considered in the previous step the number of hyperplanes where they even get close to being of degree \( d \) is at most \( O(q^d) \), thus yielding the analog of Theorem 4.2.1.

In implementing the program above (which is what we will end up doing) in our more general/abstract setting, our first bottleneck is that, for instance in Step (2) above, we don’t have a notion of degree \( d + 1 \) or some notion of functions that are “just outside our good set \( \mathcal{F} \)”. Natural notions of things outside our set do exist, but they don’t necessarily satisfy our needs. To understand this issue better, let us see why polynomials of degree \( d + O(1) \) appear in the analysis of a theorem such as Theorem 4.2.1. Consider a simple case where \( H_1, \ldots, H_q \) are parallel hyperplanes completely covering \( \mathbb{F}_q^n \) and \( \delta = 0 \) so \( f \) is known to be a good function (member of \( \mathcal{F} \), or degree \( d \)) when restricted to these hyperplanes. So, in the setting of testing polynomials of degree at most \( d \), the hypothesis asserts that \( f \) restricted to these hyperplanes is a polynomial of degree at most \( d \). For notational simplicity we assume that \( H_i \) is the hyperplane given by \( x_1 = \eta_i \) where \( \mathbb{F}_q = \{\eta_1, \ldots, \eta_q\} \). Then \( f|_{H_i} = P_i(x_2, \ldots, x_n) \) for some polynomial \( P_i \) of degree \( d \). By polynomial interpolation, it follows that \( f \) can be described as a degree \( d + q - 1 \) polynomial in \( x_1, \ldots, x_n \). The bulk of the analysis of Theorem 3.1.3 now attempts to use the remaining \( K - q \) hyperplanes on which \( f \) reduces to degree at most \( d \), in conjunction with the fact that \( f \) is a polynomial of degree at most \( d + q - 1 \) to argue that \( f \) is of degree at most \( d \).

For us, the main challenge is that in the generic setting of the lift of some code \( \mathcal{B} \), we don’t have a ready notion of a degree \( d + q - 1 \) polynomial and so we have to define one. Thus the first step in this work is to define such a code. The formal definition appears in Section 4.4.1: For our current discussion it suffices to say that there is an affine-invariant linear code, which we denote \( \mathcal{F}^+ \), which contains all “interpolating functions” of elements of \( \mathcal{F} \) (so \( \mathcal{F}^+ \) contains every function \( f \) for which there exist some \( q \) parallel hyperplanes \( H_1, \ldots, H_q \) such that \( f|_{H_i} \) is a function in \( L_{n-1} \) for all \( i \)). Of course such a set is not useful if it does not have some nice structure. The key property of our definition of \( \mathcal{F}^+ \) is that it is the lift of a non-trivial code on at most \( t + q - 1 \) dimensions. We prove this in Section 4.4.1. This definition of \( \mathcal{F}^+ \) and its analysis rely centrally on some of the structural understanding of affine-invariant linear codes derived in previous works [KS08, GKS08b, GKS09, BSS11, BSMSS11, BGM+11].

Lemma 4.4.5 allows us to say that \( \mathcal{F}^+ \) is almost as nice as \( \mathcal{F} \), roughly analogous to the way the set of degree \( d + q - 1 \) polynomials is almost as nice as the set of degree \( d \) polynomials.

The notion of \( \mathcal{F}^+ \) turns out to be easy enough to use to be able to carry out the steps (3) and (4) in the program above by directly mimicking the proof of Theorem 3.1.3, assuming Step (2) holds. (See Section 4.4.3.) But Steps (1) and (2) turn out to be more tricky. So we turn to these, and in particular Step (2) next.
Our next barrier in extending the proof Theorem 3.1.3 is the notion of “canonical monomials” which plays a crucial role in Step (2) of Theorem 3.1.3. For a function $f$ of degree $d + 1$, the canonical monomial is a monomial $M$ of degree $d + 1$ supported on very few variables such that $M$ is in the support of $f \circ T$ for some affine transformation $T$. The fact that the number of variables in the support is small, while the monomial remains a “forbidden one” turns out to be central to their analysis and allows them to convert questions of the form: “Does $f$ become a polynomial of smaller degree on the hyperplane $H$?” (which are typically not well-understood) to questions of the form “Does $g$ become the zero polynomial when restricted to $H$?” (which is a very well-studied question).

In our case, we need to work with some function $f$ in $\mathcal{F}^+$ which is not a function of $\mathcal{F}$. The fact that $\mathcal{F}^+$ is a lift of a “few-dimensional” code, in principle oughts to help us find a monomial supported on few variables that is not in $\mathcal{F}$. But isolating the “right one” to work with for $f$ turns out to be a subtle issue and we work hard, and come up with a definition that is very specific to each function $f \in \mathcal{F}^+ \setminus \mathcal{F}$. (In contrast the canonical monomials of the previous theorem were of similar structure for every function $f$.) Armed with this definition and some careful analysis we are able to simulate Step (2) in the program above. Details may be found in Section 4.4.2. Finally, Step (1) is also dealt with similarly, using some of the same style of ideas as in the proof of Step (2). (See Lemma 4.5.3.)

### 4.3 Background and preliminary material

In this section we fix some notation and provide some background material on affine-invariant linear codes, needed later on. We start with some basic notation.

Recall we are working with the field $\mathbb{F}_q$ where $q = p^r$, for prime $p$ and integer $r$. Throughout we will consider $q$ as a constant, and so asymptotic notations such as $O(\cdot), \Omega(\cdot)$ in this work may neglect dependence on $q$. All linear-algebraic terminology as subspaces, dimension, span, etc. will be over the field $\mathbb{F}_q$.

and $\mathbb{N}$ denote the set of non-negative integers. For $n > t$, we think of $\{\mathbb{F}_q^n \to \mathbb{F}_q\}$ as a subset of $\{\mathbb{F}_q^n \to \mathbb{F}_q\}$ by using the standard embedding $E : \{\mathbb{F}_q^n \to \mathbb{F}_q\} \to \{\mathbb{F}_q^n \to \mathbb{F}_q\}$ given by $(E(f))(x_1, \ldots, x_n) = f(x_1, \ldots, x_t)$.

We let $\text{Aff}_n \subseteq \{\mathbb{F}_q^n \to \mathbb{F}_q\}$ represent the set of all the affine functions. i.e.,

$$\text{Aff}_n = \left\{ L : \mathbb{F}_q^n \to \mathbb{F}_q \mid \exists \alpha_0, \ldots, \alpha_n \in \mathbb{F}_q \text{ such that } L(x) = \sum_{i=1}^{n} \alpha_i x_i + \alpha_0 \forall x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n \right\}.$$ 

For $L \in \text{Aff}_n$ define $H_L \subseteq \mathbb{F}_q^n$ to be the hyperplane $\{x \in \mathbb{F}_q^n \mid L(x) = 0\}$. We let $\text{Aff}_{n \times n}$ represent the set of affine transformations from $\mathbb{F}_q^n$ to $\mathbb{F}_q^n$. i.e.,

$$\text{Aff}_{n \times n} := \left\{ T : \mathbb{F}_q^n \to \mathbb{F}_q^n \mid \exists B \in \mathbb{F}_q^{n \times n}, c \in \mathbb{F}_q^n \text{ such that } T(x) = Bx + c \forall x \in \mathbb{F}_q^n \right\}.$$ 

For a function $f \in \{\mathbb{F}_q^n \to \mathbb{F}_q\}$ and $T \in \text{Aff}_{n \times n}$, we denote by $f \circ T$ the composition of $f$ and $T$. i.e.,

$$\forall x \in \mathbb{F}_q^n : f \circ T(x) = f(T(x)).$$

We view monomials defined on variables $x_1, \ldots, x_n$ as functions mapping $\mathbb{F}_q^n$ to $\mathbb{F}_q$, given by the evaluations of the monomials. The set $\mathcal{M}_n \subseteq \{\mathbb{F}_q^n \to \mathbb{F}_q\}$ denotes the set of such monomial functions. For $M = \prod_{i=1}^{n} x_i^{a_i} \in \mathcal{M}_n$ where $\{a_i\}_{i=1}^{n} \subseteq \{0, \ldots, q - 1\}$, $\deg_x(M) = a_i$. As usual, $\deg(M) = \sum_{i=1}^{n} \deg_x(M)$.

Note that for $a \in \mathbb{N}$, the monomials $M = x_i^a$ and $M' = x_i^{a \mod q - 1}$ are equivalent when $q - 1 \nmid a$ or $a = 0$, while when $q - 1 \mid a$ and $a \neq 0$ the monomials $M = x_i^a$ and $M' = x_i^{q - 1}$ are equivalent. Motivated by this, we define the operation $a \mod \ast k$ as follows

$$a \mod \ast k = \begin{cases} a \mod k, & a = 0 \text{ or } k \nmid a \\ k, & \text{otherwise} \end{cases}$$
For every function \( f \in \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \) there is a unique representation as a polynomial \( f = \sum_{M \in \mathcal{M}_n} c^f_M M \) for some coefficients \( \{ c^f_M \mid M \in \mathcal{M}_n \} \subseteq \mathbb{F}_q \). We define the support of such a function \( f \) to be \( \text{supp}(f) := \left\{ M \in \mathcal{M}_n \mid c^f_M \neq 0 \right\} \), and we let \( \deg(f) = \max \{ \deg(M) \mid M \in \text{supp}(f) \} \).

4.3.1 The structure of affine-invariant linear codes

One main feature of affine-invariant linear codes is that they can be characterized by the set of monomials in the support of the functions in these codes. Let \( \mathcal{F} \subseteq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \) be an affine-invariant linear code. The support \( \text{supp}(\mathcal{F}) \) of \( \mathcal{F} \) is simply the union of the supports of the functions in \( \mathcal{F} \), i.e., \( \text{supp}(\mathcal{F}) = \bigcup_{f \in \mathcal{F}} \text{supp}(f) \). The following theorem from \cite{KS08} says that every affine-invariant linear code is uniquely determined by its support.

**Lemma 4.3.1** (Monomial extraction lemma, \cite[Lemma 4.2]{KS08}). Let \( \mathcal{F} \subseteq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \) be an affine-invariant linear code. Then \( \mathcal{F} \) has a monomial basis, that is, \( \mathcal{F} = \text{span}(\text{supp}(\mathcal{F})) \).

For a monomial \( M \in \mathcal{M}_n \), let \( \text{Aff}_{n \times n}(M) \) denote the set of all monomials that can be obtained from \( M \) by applying an affine transformation \( T \in \text{Aff}_{n \times n} \) on \( M \), that is,

\[
\text{Aff}_{n \times n}(M) = \left\{ M' \in \mathcal{M}_n \mid \exists T \in \text{Aff}_{n \times n} \text{ such that } M' \in \text{supp}(M \circ T) \right\}.
\]

We will call \( \text{Aff}_{n \times n}(M) \) the \( n \)-dimensional affine set of \( M \). When the dimension \( n \) is clear from the context we will omit the subscript \( n \times n \). Note that if \( M \in \mathcal{F} \), \( M' \in \text{Aff}_{n \times n}(M) \) and \( \mathcal{F} \) is an affine-invariant linear code then \( M' \in \mathcal{F} \). The following lemma, also from \cite{KS08}, gives a sufficient condition under which a monomial belongs to \( \text{Aff}_{n \times n}(M) \).

**Lemma 4.3.2**. [Monomial spread lemma, \cite[Lemma 4.6]{KS08}] Let \( M' = \prod_{i=1}^n x_i^{a_i}, M = \prod_{i=1}^n x_i^{b_i} \) be a pair of monomials in \( \mathcal{M}_n \), where \( a_i, b_i \in \{0, ..., q-1\} \) for all \( 1 \leq i \leq n \). For \( 1 \leq i \leq n \), let \( a_i = \sum_j a_j^{(i)} p^j \), \( b_i = \sum_j b_j^{(i)} p^j \) be the base-\( p \) representation of \( a_i, b_i \) respectively. Assume that for all \( j \), \( \sum_{i=1}^n a_j^{(i)} \leq \sum_{i=1}^n b_j^{(i)} \). Then \( M' \in \text{Aff}_{n \times n}(M) \).

We shall also use the following theorem from \cite{GKS12} which says that if a linear code is invariant under invertible affine transformations then it is also invariant under general affine transformations.

**Theorem 4.3.3**. \cite[Theorem A.1]{GKS12} If \( \mathcal{F} \subseteq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \) is an \( \mathbb{F}_q \)-linear code invariant under invertible affine transformations, then \( \mathcal{F} \) is invariant under all affine transformations.

4.3.2 Lifts of affine-invariant linear codes

The following claim relates the support of the base code to the support of its lift.

**Claim 4.3.4.** Let \( \mathcal{B} \subseteq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \) be an affine-invariant linear base code and let \( \mathcal{F} = \text{Lift}_n(\mathcal{B}) \) be its \( n \)-dimensional lift. Then the following hold:

1. \( \text{supp}(\mathcal{B}) = \text{supp}(\mathcal{F}) \cap \mathcal{M}_t. \)

2. \( \text{supp}(\mathcal{F}) = \left\{ M \in \mathcal{M}_n \mid \text{Aff}_{n \times n}(M) \cap \mathcal{M}_t \subseteq \text{supp}(\mathcal{B}) \right\}. \)

**Proof.** For the proof of the first part of the claim, suppose first that \( M \in \text{supp}(\mathcal{F}) \cap \mathcal{M}_t \) and let \( A \subseteq \mathbb{F}_q^n \) be the \( t \)-dimensional subspace containing all vectors supported on the first \( t \) coordinates. The fact that \( M \in \mathcal{F} = \text{Lift}_n(\mathcal{B}) \) implies that \( M|_A \in \mathcal{B} \). Since \( M \in \mathcal{M}_t \) we thus have that \( M \in \text{supp}(\mathcal{B}) \). On the other hand, suppose that \( M \in \text{supp}(\mathcal{B}) \). Then in this case we clearly have that \( M \in \mathcal{M}_t \). To see that \( M \) is also contained in \( \text{supp}(\mathcal{F}) \) let \( A \) be an arbitrary \( t \)-dimensional affine subspace of \( \mathbb{F}_q^n \).
Then the fact that $\mathcal{B}$ is an affine-invariant code and $M \in \mathcal{B}$ implies that $M|_A \in \mathcal{B}$. Since $\mathcal{F} = \text{Lift}_n(\mathcal{B})$ this implies in turn that $M \in \mathcal{F}$, so we conclude that $M \in \mathcal{M}_t \cap \text{supp}(\mathcal{F})$.

We proceed to the proof of the second part of the claim. Suppose first that $M \in \text{supp}(\mathcal{F})$ and let $M' \in \text{Aff}_{n \times n}(M) \cap \mathcal{M}_t$. Then there exists an affine transformation $T \in \text{Aff}_{n \times n}$ such that $M' \in \text{supp}(M \circ T |_{x_{t+1}=0, \ldots, x_n=0})$. But if we let $e_1, \ldots, e_n$ denote the standard basis for $\mathbb{F}_q^n$ and we let $A$ denote the $t$-dimensional subspace containing all vectors of the form $T(0) + \sum_{i=1}^t (T(e_i) - T(0))x_i$ for $x_i \in \mathbb{F}_q$ then $M \circ T |_{x_{t+1}=0, \ldots, x_n=0} \in \mathcal{B}$ if and only if $M|_A \in \mathcal{B}$. Since $\mathcal{F} = \text{Lift}_n(\mathcal{B})$ and $M \in \mathcal{F}$ we have that $M|_A \in \mathcal{B}$ and so $M' \in \text{supp}(\mathcal{B})$.

For the other direction, suppose that $M \in \mathcal{M}_n$ is such that $\text{Aff}_{n \times n}(M) \cap \mathcal{M}_t \subseteq \text{supp}(\mathcal{B})$, we will show that $M \in \text{supp}(\mathcal{F})$. For this we need to show that $M|_A \in \mathcal{B}$ for every $t$-dimensional affine subspace $A$. Let $A$ be a $t$-dimensional affine subspace and let $\alpha_0, \alpha_1, \ldots, \alpha_t$ be such that $A = \{ \alpha_0 + \sum_{i=1}^t \alpha_i x_i \mid x_i \in \mathbb{F}_q \}$. Let $T \in \text{Aff}_{n \times n}$ be the affine transformation defined as $T(e_i) = \alpha_i + \alpha_0$ for all $1 \leq i \leq t$ and $T(e_i) = 0$ for all $t < i \leq n$. Then $\text{supp}(M \circ T |_{x_{t+1}=0, \ldots, x_n=0}) \subseteq \text{Aff}_{n \times n}(M) \cap \mathcal{M}_t$ and so we also have that $\text{supp}(M|_A) \subseteq \text{Aff}_{n \times n}(M) \cap \mathcal{M}_t$. Our assumption that $\text{Aff}_{n \times n}(M) \cap \mathcal{M}_t \subseteq \text{supp}(\mathcal{B})$ implies in turn that $\text{supp}(M|_A) \subseteq \text{supp}(\mathcal{B})$ and so $M|_A \in \mathcal{B}$ as required. □

The following proposition bounds the distance of lifts of general affine-invariant linear codes from $\mathbb{F}_q^t$ to $\mathbb{F}_q$.

**Proposition 4.3.5.** [KS08, Theorem 5.1] Let $Q$ be a power of $q$, let $\mathcal{B} \subseteq \{ \mathbb{F}_Q^t \to \mathbb{F}_q \}$ be an affine-invariant linear base code and let $\mathcal{F} = \text{Lift}_n(\mathcal{B})$ be its $n$-dimensional lift. Then $\delta(\mathcal{B}) \geq \delta(\mathcal{F}) \geq \delta(\mathcal{B}) - Q^{-t}$.

From the above proposition one can derive the following corollary.

**Corollary 4.3.6.** Let $\mathcal{B} \subseteq \{ \mathbb{F}_Q^t \to \mathbb{F}_q \}$ be some non-trivial affine-invariant linear code and let $\mathcal{F} = \text{Lift}_n(\mathcal{B})$ be its $n$-dimensional lift. Then $\delta(\mathcal{F}) \geq Q^{-t}$.

**Proof.** From Proposition 4.3.5 it is enough to show that $\delta(\mathcal{B}) \geq 2Q^{-t}$. Assume toward a contradiction that $\delta(\mathcal{B}) < 2Q^{-t}$. From linearity of $\mathcal{B}$ there is a function $f \in \mathcal{B}$ such that there is only one point $v \in \mathbb{F}_Q$ such that $f(v) \neq 0$. To reach a contradiction we show that any function $g : \mathbb{F}_Q^t \to \mathbb{F}_q$ can be written as a linear combination of affine transformations of $f$. Because $\mathcal{B}$ is affine-invariant linear code it will follow that $\mathcal{B} = \{ \mathbb{F}_Q^t \to \mathbb{F}_q \}$. Indeed, we can express any $g : \mathbb{F}_Q^t \to \mathbb{F}_q$ as $g(x) = \sum_{u \in \mathbb{F}_Q} g(u) f(x+u)$ and the result follows. □

### 4.4 Proof of Main Technical Theorem 4.2.1

In this section we prove our Main Technical Theorem 4.2.1. Our goal then will be to show that if $f$ is far from $\mathcal{F}$ then on most hyperplanes it remains far from $\mathcal{F}$. In particular if $\mathcal{F}$ is the lift of a $t$-dimensional code, then $f$ should get close on at most $q^{t+O(1)}$ hyperplanes. We start by studying the special case where $f$ results from an “interpolation” of several functions in $\mathcal{F}$.

#### 4.4.1 The code $\mathcal{F}^+$

We start with the definition of the code $\mathcal{F}^+$ which contains all functions obtained from interpolations of functions in $\mathcal{F}$. The code $\mathcal{F}^+$ is defined below as the $n$-dimensional lift of a non-trivial code $\mathcal{B}^+$ on $t + q - 1$ variables. We will then show that the code $\mathcal{F}^+$ contains all interpolations of functions in $\mathcal{F}$.
We will show in this case that

First observe that for every
e subspace

and

monomial

e set of the monomial

Finally, let

\[ Y = 1 + 2 + 3 \]

If

such that

We need to show that for every

\[ D^+ = \text{Aff}_{t+ \times t^+} \left\{ M \prod_{i=1}^{q-1} x_{t+i}^{q-1} \mid M \in D \right\}. \]

We first show that \( \mathcal{F}^+ \) is non-trivial (i.e., \( \mathcal{F}^+ \neq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \)), provided the base code \( \mathcal{B} \) is non-trivial.

Claim 4.4.2. If \( \mathcal{B} \neq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \), then \( \mathcal{F}^+ \neq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \).

Proof. Since \( \mathcal{B} \neq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \) and since, by Lemma 4.3.2, every monomial on \( t \) variables is in the affine set of the monomial \( \prod_{i=1}^{t} x_i^{q-1} \), it follows that \( \prod_{i=1}^{t} x_i^{q-1} \not\in D = \text{supp(} \mathcal{B} \). Hence we have that for every monomial \( M' \in D \), \( \text{deg}(M') < t(q - 1) \). From the definition of \( D^+ \) it follows that every monomial \( M \in D^+ \) must have degree strictly less than \( t^+ \cdot (q - 1) \). It follows that \( \mathcal{B}^+ \neq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \) and \( \mathcal{F}^+ \neq \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \).

Next we show that \( \mathcal{F}^+ \) contains all functions resulting from interpolation of functions in \( \mathcal{F} \), as per the following definition.

Definition 4.4.3 (Interpolation of functions in \( \mathcal{F} \)). We say that \( f \) is an interpolation of functions in \( \mathcal{F} \) if there exist \( q \) parallel hyperplanes \( H_1, \ldots, H_q \) (so \( H_i \cap H_j = \emptyset \) for \( i \neq j \) and \( \cup_i H_i = \mathbb{F}_q^n \)) and \( q \) functions \( f_1, \ldots, f_q \in \mathcal{F} \) such that \( f|_{H_i} = f_i|_{H_i} \) for every \( i \in [q] \).

Claim 4.4.4. A function \( f \in \{ \mathbb{F}_q^n \rightarrow \mathbb{F}_q \} \) is an interpolation of functions in \( \mathcal{F} \) if and only if there exists an affine function \( L \in \text{Aff}_n \) and functions \( \{ f_i \in \mathcal{F} \mid a \in \{0, \ldots, q-1\} \} \) such that \( f = \sum_{a \in \{0, \ldots, q-1\}} f_a L^a \).

Proof. The proof is straightforward by polynomial interpolation.

Lemma 4.4.5. If \( f \) is an interpolation of functions in \( \mathcal{F} = \text{Lift}_n(\mathcal{B}) \), then \( f \in \mathcal{F}^+ \).

Proof. Fix \( f = \sum_a f_a L^a \) for affine function \( L \) and \( f_a \in \mathcal{F} \). We need to show that for every \( t \)-dimensional affine subspace \( A \) it is the case that \( \text{supp}(f|_A) \subseteq D^+ \). Equivalently, we need to show that for every \( T \in \text{Aff}_{n \times n} \), the restriction of \( \text{supp}(f \circ T) \) to the subspace \( \{ x \in \mathbb{F}_q^n \mid x_{t+1} = \ldots = x_n = 0 \} \) is contained in \( D^+ \).

First observe that for every \( T \in \text{Aff}_{n \times n} \), \( f \circ T = \sum_{a \in \{0, \ldots, q-1\}} L^a f'_a \), where \( L' \) is an affine function and \( f'_a \in \mathcal{F} \) (so it is of the same form as \( f \)). Note that every monomial in \( \text{supp}(f \circ T) \) is of the form \( M \prod_{j=1}^{q} x_{ij} \) where \( a \in \{0, \ldots, q-1\}, i_1, \ldots, i_a \in [n] \) and \( M \in \text{supp}(\mathcal{F}) \). Further, restricting \( f \circ T \) to the subspace \( \{ x \in \mathbb{F}_q^n \mid x_{t+1} = \ldots = x_n = 0 \} \) allows us to focus only on the cases \( i_1, \ldots, i_a \in [t] \) and \( M \in \mathcal{M}_{t^+} \).

We will show in this case that \( M \prod_{j=1}^{q} x_{ij} \in D^+ \). Fix \( M \in \mathcal{M}_{t^+} \cap \text{supp}(\mathcal{F}) \), \( i_1, \ldots, i_a \in [t] \) and let \( I \subseteq [t] \) be such that \( |I| = q - 1 \) and \( \{i_1, \ldots, i_a \} \subseteq I \). Write \( M = \prod_{k=1}^{t} x_k^{a_k} \) and choose \( M' \in \mathcal{M}_t \) to be a monomial of the form \( \prod_{k=1}^{t} x_k^{b_k} \) where \( \{b_k \mid k \in [t]\} = \{a_k \mid k \in [t^+ \setminus I]\} \). Then by Lemma 4.3.2,

\[
M \prod_{j=1}^{a} x_{ij} = \prod_{k \in I} x_k^{a_k} \prod_{k \in I} x_k^{a_k + \#(i_j = k)} \in \text{Aff}_{t+ \times t^+} \left( \prod_{k=1}^{t} x_k^{b_k} \prod_{i=1}^{q-1} x_{t+i}^{q-1} \right) = \text{Aff}_{t+ \times t^+} \left( M' \prod_{i=1}^{q-1} x_{t+i}^{q-1} \right).
\]

Observe, again by Lemma 4.3.2, that \( M' \in \text{Aff}_{n \times n} \), so \( M' \in \text{supp}(\mathcal{F}) \cap \mathcal{M}_t \). By Claim 4.3.4, this implies in turn that \( M' \in \text{supp}(\mathcal{B}) \). Consequently, \( M \prod_{j=1}^{q} x_{ij} \in D^+ \), which concludes the proof of the lemma.

\[ \square \]
4.4.2 Restrictions of functions in $\mathcal{F}^+$ to hyperplanes

Let $\mathcal{F} = \text{Lift}_q(\mathcal{B})$ for a non-trivial affine-invariant linear code $\mathcal{B} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_q\}$. Let $\mathcal{B}^+$ be the code given by Definition 4.4.1 and let $\mathcal{F}^+ = \text{Lift}_q(\mathcal{B}^+)$. Our goal in this section will be to show that for every $f \in \mathcal{F}^+ \setminus \mathcal{F}$, the number of hyperplanes $H$ for which $f|_H \in \mathcal{F}$ is upper bounded by $O(q^+)$. (See Theorem 4.4.10 below for formal statement.) We remark that throughout this section one could replace $\mathcal{F}^+$ by any affine-invariant linear code that is the lift of a non-trivial affine-invariant linear base code contained in $\{\mathbb{F}_q^+ \to \mathbb{F}_q\}$ and that the same holds for $\mathcal{F}$ (so $\mathcal{F}^+$ does not have to be as in Definition 4.4.1 and $\mathcal{F}$ could be a lift of a base code defined over $t^+$ variables and not only $t$ variables).

The overall strategy is as follows. (1) We will first show in Lemma 4.4.6 that for every such $f$ there exists an invertible affine transformation $T$ and a monomial $M \notin \mathcal{F}$ supported on the first $t^+$ variables such that $M$ is in the support of $f \circ T$. We further assume that $T$ is such that the degree of $M$ is maximal. Since we can just prove the theorem about $f \circ T$, we assume that $M$ is in the support of $f$. (2) Next we partition the space of all possible hyperplanes into $q^{t^++1}$ sets (based on their coefficients on the first $t^+$ variables). Our goal will be to show that in each set in the partition there are at most some constant (depending on $q$) number of hyperplanes such that $f$ restricted to these hyperplanes becomes a member of $\mathcal{F}$. To do so we extract from $f$ a non-zero low-degree function $g$ (this function $g$ depends on $M$ and the set in the partition under consideration), such that for a hyperplane $H$ from this set, $f|_H \in \mathcal{F}$ only if $g|_H \equiv 0$. (See Lemma 4.4.7.) (3) The final task, to bound the number of hyperplanes on which a low-degree polynomial becomes zero, turns out to be relatively easy and we give this bound in Lemma 4.4.8.

Below we state the three lemmas mentioned above. We defer their proofs to later in this section. We show how they imply Theorem 4.4.10 immediately after stating them.

The first of our lemmas isolates a “canonical monomial” for every function $f \in \mathcal{F}^+ \setminus \mathcal{F}$. We note that this is similar to such a step in Chapter 3 with the main difference being that the canonical monomials here can be quite different for different functions $f$ (whereas in Chapter 3 all canonical monomials of functions $f \in \mathcal{F}^+ \setminus \mathcal{F}$ were of a similar structure).

**Lemma 4.4.6.** For every $f \in \mathcal{F}^+ \setminus \mathcal{F}$ there exists an invertible affine transformation $T$ and a monomial $M \in \mathcal{M}_{t^+}$ such that $M \notin \mathcal{F}$ and $M$ is in the support of $f \circ T$.

Our next lemma, which is the bulk of this section, reduces the task of counting hyperplanes where $f$ becomes a member of $\mathcal{F}$, to the task of counting hyperplanes where a related function becomes zero.

**Lemma 4.4.7.** Let $M \in \mathcal{M}_{t^+}$ be a monomial in the support of $f \in \mathcal{F}^+ \setminus \mathcal{F}$ with $M \notin \mathcal{F}$. Suppose furthermore that for every invertible affine transformation $T$ all monomials $M' \in (\text{supp}(f \circ T) \cap \mathcal{M}_{t^+}) \setminus \mathcal{F}$ satisfy that $\deg(M') \leq \deg(M)$. Then for every $\alpha_0, \alpha_1, \ldots, \alpha_{t^+} \in \mathbb{F}_q$ there exists a non-zero function $g$ with $\deg(g) \leq q^2(q-1)$ such that the following holds: For every choice of $\alpha_{t^++1}, \ldots, \alpha_n \in \mathbb{F}_q$ the hyperplane $H = \{x \in \mathbb{F}_q^n | \sum_{i=1}^n \alpha_ix_i + \alpha_0 = 0\}$ satisfies $f|_H \in \mathcal{F}$ only if $g|_H \equiv 0$.

Finally, we bound the number of hyperplanes where a non-zero low-degree function can become zero.

**Lemma 4.4.8.** Let $f : \mathbb{F}_q^t \to \mathbb{F}_q$ be a non-zero polynomial of degree $d$. Then there are at most $q^{d+1}$ hyperplanes $H$ such that $f|_H \equiv 0$.

**Remark 4.4.9.** We remark that any bound that is constant for constant $d$ and $q$ would have been good enough to suffice for our purpose. We also note that the bound above is close to the right one. In particular if $d = t(q-1)$ and $f(x_1, \ldots, x_n) = \prod_{i=1}^t (x_i^{q-1} - 1)$ then $f$ is zero on every hyperplane of the form $x_1 = \sum_{i=1}^{t-1} \alpha_ix_i + \beta$, with $\alpha_i$’s being arbitrary and $\beta$ being non-zero, and there are at least $(q-1) \cdot q^{d/(q-1)-1}$ of these.
We now state and prove our main theorem of this section.

**Theorem 4.4.10.** Let \( B \subseteq \{F_q^t \rightarrow F_q \} \) be an affine-invariant linear code and let \( F = \text{Lift}_n(B) \). Let \( B^+ \) be the code given by Definition 4.4.1, let \( F^+ = \text{Lift}_n(B^+) \) and let \( f \in F^+ \setminus F \). Then there are at most \( q^{t+q^2+2} \) hyperplanes \( H \) such that \( f|_H \in F \).

**Proof.** Let \( T \) and \( M \) be the invertible affine transformation and the monomial given by Lemma 4.4.6 above, respectively. Suppose furthermore that \( T \) maximizes the degree of \( M \), in the sense that for every other invertible affine transformation \( T' \) all monomials \( M' \in (\text{supp}(f \circ T') \cap M_{t^+}) \setminus F \) satisfy that \( \deg(M') \leq \deg(M) \).

Applying Lemma 4.4.7 to the function \( f \circ T \) and the monomial \( M \), we get that for every \( \alpha_0, \alpha_1, \ldots, \alpha_{t^+} \) there is a non-zero polynomial \( g \) of degree at most \( (q-1)q^2 \) such that \( g|_H \equiv 0 \) whenever \( (f \circ T)|_H \in F \).

By Lemma 4.4.8 there are at most \( q^{t+q^2+2} \) such hyperplanes \( H \). Summing over all possible choices of \( \alpha_0, \alpha_1, \ldots, \alpha_{t^+} \), we get that there are at most \( q^{t+q^2+2} \) hyperplanes \( H \) such that \( (f \circ T)|_H \in F \). The theorem follows from the fact that there is a one-to-one correspondence between the hyperplanes for which the restriction of \( f \circ T \) is in \( F \) and the hyperplanes for which the restriction of \( f \) is in \( F \). \( \square \)

In the remaining subsections of this section we prove the three lemmas mentioned above.

**Proof of Lemma 4.4.6**

**Lemma 4.4.6 (restated).** For every \( f \in F^+ \setminus F \) there exists an invertible affine transformation \( T \) and a monomial \( M \in M_{t^+} \) such that \( M \not\in F \) and \( M \) is in the support of \( f \circ T \).

**Proof.** Let \( \mathcal{F}_f \subseteq \{F_q^n \rightarrow F_q \} \) be the minimal affine-invariant linear code containing \( f \). Note that \( \mathcal{F}_f = \{\sum_{T \in T} c_T \cdot (f \circ T)(c_T \in F_q)\} \), where \( T \) denotes the set of all invertible affine transformations in \( \text{Aff}_{n \times n} \) (the fact that one can sum only over invertible transformations follows from Theorem 4.3.3).

Let \( B^* \subseteq \{F_q^+ \rightarrow F_q \} \) be the code \( B^* = \{g|_{x_{t+1}=0, \ldots, x_m=0} | g \in \mathcal{F}_f\} \). By definition \( B^* \) is an affine-invariant linear code and \( f \in \text{Lift}_n(B^*) \). Since \( f \not\in \text{Lift}_n(B) = \text{Lift}_n(\text{Lift}_{t^+}(B)) \), it follows that \( B^* \not\subseteq \text{Lift}_{t^+}(B) \). So there must exist a monomial \( M \in B^* \setminus \text{Lift}_{t^+}(B) \) (since \( B^* \) is spanned by the monomials in it, by Lemma 4.3.1). Clearly, we have that \( M \in M_{t^+} \). Furthermore, by Claim 4.3.4 \( \text{supp}(\text{Lift}_{t^+}(B)) = \text{supp}(F) \cap M_{t^+} \) and so \( M \not\in F \). Finally, note that the definition of \( B^* \) implies that \( M \) belongs also to \( \mathcal{F}_f \). By the fact that \( \mathcal{F}_f = \{\sum_{T \in T} c_T \cdot (f \circ T)(c_T \in F_q)\} \) it follows that there exists an invertible affine transformation \( T \) such that \( M \in \text{supp}(f \circ T) \). The lemma follows. \( \square \)

**Proof of Lemma 4.4.7**

**Lemma 4.4.7 (restated).** Let \( M \in M_{t^+} \) be a monomial in the support of \( f \in F^+ \setminus F \) with \( M \not\in F \). Suppose furthermore that for every invertible affine transformation \( T \) all monomials \( M' \in (\text{supp}(f \circ T) \cap M_{t^+}) \setminus F \) satisfy that \( \deg(M') \leq \deg(M) \). Then for every \( \alpha_0, \alpha_1, \ldots, \alpha_{t^+} \in F_q \) there exists a non-zero function \( g \) with \( \deg(g) \leq q^2(q-1) \) such that the following holds: For every choice of \( \alpha_{t+1}, \ldots, \alpha_n \in F_q \) the hyperplane \( H = \{x \in F_q^n | \sum_{i=1}^n \alpha_ix_i + \alpha_0 = 0\} \) satisfies \( f|_H \in F \) only if \( g|_H \equiv 0 \).

**Proof.** As a first step, we perform a change of basis that will allow us to assume, w.l.o.g., that \( \alpha_1 = -1 \) and \( \alpha_0 = \alpha_2 = \cdots = \alpha_{t^+} = 0 \) and so restriction of \( f \) to the hyperplane given by \( \alpha_{t+1}, \ldots, \alpha_n \) is given by the function \( f(\sum_{i=t+1}^n \alpha_ix_i, x_2, \ldots, x_n) \). We will analyze such functions in later steps.

Fix \( \alpha_0, \alpha_1, \ldots, \alpha_{t^+} \in F_q \) and let \( H = H_{\alpha_0, \ldots, \alpha_{t^+}} \) be the set of hyperplanes \( H \) such that there exist \( \alpha_{t+1}, \ldots, \alpha_n \) so that \( H = \{x \in F_q^n | \sum_{i=1}^n \alpha_ix_i + \alpha_0 = 0\} \).
First we dismiss the case $\alpha_1 = \cdots = \alpha_{t^+} = 0$. In this case for every hyperplane $H \in \mathcal{H}$, the function $f|_H$ still has the monomial $M$ in its support and so $f|_H \notin \mathcal{F}$. (So formally, $g = 1$ satisfies the condition of the lemma.) So from here on we assume there exists $c \in [t^+]$ such that $\alpha_c \neq 0$. Without loss of generality we assume $c$ is the minimal such index, and that $\alpha_c = -1$. For notational simplicity we assume below that $c = 1$. Now consider the affine transformation $S \in \text{Aff}_{n \times n}$ such that $S(e_1) = e_1 + \sum_{i=2}^{t^+} \alpha_i e_i + \alpha_0$ and $S(e_i) = e_i$ for all $i \geq 2$. Let $f' = f \circ S$. For hyperplane $H = \{x|\sum_{i=1}^{n} \alpha_i x_i + \alpha_0 = 0\}$, let $H'$ be the hyperplane $H' = \{x|x_1 = \sum_{i=t^++1}^{n} \alpha_i x_i\}$. Notice that $f|_H \in \mathcal{F}$ if and only if $f'|_{H'} \in \mathcal{F}$ and $H'$ corresponds to $\alpha_1' = -1$ and $\alpha_i' = 0$ for $i \in \{0, 2, 3, \ldots, t^+\}$.

Now let $M' \in \text{supp}(f') \cap \mathcal{M}_{t^+}$ be a monomial such that $M \in \text{supp}(M' \circ T)$ for some invertible $T \in \text{Aff}_{t^+ \times t^+}$. Note such a monomial $M'$ must exist since $S$ is an invertible transformation in $\text{Aff}_{t^+ \times t^+}$. Since $M \in \text{supp}(M' \circ T)$ and $M \notin \mathcal{F}$ it follows that $M' \notin \mathcal{F}$. Furthermore, the fact that $M \in \text{supp}(M' \circ T)$ implies that $\deg(M) \leq \deg(M')$ and hence $M'$ is also maximal with respect to degree. That is, for every invertible affine transformation $T'$ it holds that all monomials $M'' \in (\text{supp}(f' \circ T') \cap \mathcal{M}_{t^+}) \setminus \mathcal{F}$ satisfy that $\deg(M'') \leq \deg(M')$.

In what follows we prove that the lemma holds for the polynomial $f'$ with monomial $M'$ and coefficients $\alpha_0, \ldots, \alpha_{t^+}$, i.e., we prove the existence of a non-zero polynomial $g'$ of degree at most $q^2(q-1)$ such that $g'|_{H'} \equiv 0$ whenever $f'|_{H'} \in \mathcal{F}$. The lemma follows for $f$ by setting $g = g' \circ S^{-1}$. For notational simplicity we drop the primes below and simply assume $\alpha_1 = -1$ and $\alpha_i = 0$ for all other $i \leq t^+$ and so $f = f'$, $M = M'$.

Let $M \in \mathbb{F}_q[x_2, \ldots, x_{t^+}]$ and let $a \geq 0$ be an integer such that $M = x_1^a M$. Write $f = g_1 M + r_1$ where $g_1 \in \mathbb{F}_q[x_1, x_{t^++1}, \ldots, x_n]$ is such that $g_1 M$ contains all monomials in $\text{supp}(f)$ whose degrees in variables $x_2, \ldots, x_{t^+}$ equal their degrees in $M$ and $r_1$ is the remaining terms. Further write $g_1 = g + g_2$ where $g$ includes all monomials $M'$ of degree $\deg(M')$ mod $*(q-1) = a$ and $g_2$ includes all monomials $M''$ of degree $\deg(M'')$ mod $*(q-1) \neq a$. Rewriting we have $f = g \cdot M + r$ where $r = g_2 M + r_1$, $g \in \mathbb{F}_q[x_1, x_{t^++1}, \ldots, x_n]$ and $r \in \mathbb{F}_q[x_1, \ldots, x_n]$. We show below, using a series of claims that $g$ satisfies the conditions of the lemma. Specifically, fix $\alpha_{t^++1}, \ldots, \alpha_n$, let $L(x) = \sum_{i=t^++1}^{n} \alpha_i x_i$, and let $H$ be the hyperplane given by $\{x_1 = L(x)\}$. We wish to show that $g|_H \equiv 0$ if $f|_H \in \mathcal{F}$.

Let $\mathcal{F}_f$ be the minimal affine-invariant linear code containing $f$. Let

$$\mathcal{F}_{-\bar{M}} = \{h \in \mathbb{F}_q[x_1, x_{t^++1}, \ldots, x_n]|\bar{M} \cdot h \in \mathcal{F}\},$$

and let

$$\mathcal{F}_{f, -\bar{M}} = \{h \in \mathbb{F}_q[x_1, x_{t^++1}, \ldots, x_n]|\bar{M} \cdot h \in \mathcal{F}_f\}.$$
prove the current claim it suffices to show that there is a monomial of degree at most \(2(q - 1)\) that is not contained in \(\mathcal{F}_{M - \tilde{M}}\). We now show that the monomial \(N = x_{t+1}^a x_{t+1}^{q-1} \not\in \mathcal{F}_{M - M} \). Notice that \(N\) is a monomial of degree at most \(a + (q - 1) \leq 2(q - 1)\), and so with Lemma 4.4.15 this suffices to prove the claim.

Assume for contradiction that \(x_{t+1}^a x_{t+1}^{q-1} \in \mathcal{F}_{M - M}\) and so \(M \cdot x_{t+1}^{q-1} = x_{t+1}^a x_{t+1}^{q-1} \tilde{M} \in \mathcal{F}_f\). Since \(B^+ \neq \{\mathbb{F}_q^+ \to \mathbb{F}_q\}\) and \(M \in \text{supp}(\mathcal{F}^+) \cap \mathcal{M}_t^+ = \text{supp}(\mathcal{B}^+)\), we have \(M \neq \prod_{i=1}^{t+1} x_i^{q-1}\). We conclude there exists \(i \in [t+1] \) such that \(d_i = \deg_{x_i}(M) \neq q - 1\). But if \(x_{t+1}^a x_{t+1}^{q-1} \tilde{M} \in \mathcal{F}_f\) then by exchanging the variables \(x_i\) and \(x_{t+1}\) we also have the monomial \(M x_{t+1}^{q-1-d_i} x_{t+1}^{d_i+1} \in \mathcal{F}_f\) and so \(M x_{t+1}^{q-1-d_i} \in \mathcal{F}_f\). We show below that this contradicts the maximality of \(M\).

Note first that \(M x_{t+1}^{q-1-d_i} \) is a monomial in \(\mathcal{M}_t^+\). Furthermore, since \(\mathcal{F}_f = \{\sum_{T \in \mathcal{T}} c_T \cdot (f \circ T) | c_T \in \mathbb{F}_q\}\) we have that \(M x_{t+1}^{q-1-d_i} \in \text{supp}(f \circ T)\) for some invertible affine transformation \(T\). Finally, by Lemma 4.3.2 we have that \(M \in \text{Aff}_{n \times n}(M x_{t+1}^{q-1-d_i})\) and so the fact that \(M \not\in \mathcal{F}\) implies that \(M x_{t+1}^{q-1-d_i} \not\in \mathcal{F}\).

Concluding, we have just shown that \(M x_{t+1}^{q-1-d_i}\) is a monomial in variables \(x_1, \ldots, x_{t+1}\) contained in \(\text{supp}(f \circ T) \setminus \mathcal{F}\) for some invertible affine transformation \(T\). Given that \(\deg(M x_{t+1}^{q-1-d_i}) > \deg(M)\), this clearly violates the maximality of \(M\).

\[\square\]

Claim 4.4.13. If \(f|_H \in \mathcal{F}\) then \(g|_H \in \mathcal{F}_{M - M}\).

**Proof.** Recall that \(H = \{x \in \mathbb{F}_q^n | x_1 = L(x_{t+1}, \ldots, x_n)\}\). Let \(f'(x_2, \ldots, x_n) = f(L(x_{t+1}, \ldots, x_n), x_2, \ldots, x_n)\) denote the function \(f|_H\). As in the partitioning of \(f\), let \(f' = g'_M + r'_1\) where \(g'_M\) includes all monomials of \(f'\) whose degrees in \(x_2, \ldots, x_{t+1}\) equal their degrees in \(M\). Further let \(g'_1 = g' + g'_2\) where \(g'\) includes all terms of degree \(d\) for \(d \mod *\) \((q - 1) = a\) and \(g'_2\) collects the remaining terms.

The proof of the claim relies crucially on the following property of \(g'\), namely that \(g'(x_{t+1}, \ldots, x_n) = g'(L(x_{t+1}, \ldots, x_n), x_{t+1}, \ldots, x_n)\) is the function \(f|_H\). To see this, note that when substituting \(x_1 = L(x_{t+1}, \ldots, x_n)\) degrees in \(x_2, \ldots, x_{t+1}\) do not change and so we have \(g'_1 = g_1(L(x), x_{t+1}, \ldots, x_n)\).

Next we note that for every monomial of degree \(d\), the reductions modulo \(x_i^q = x_i\) (for every \(i\)) can only change the degree of the monomial to \(d'\) which satisfies \(d' \mod *\) \((q - 1) = d\) and so \(g' = g(L(x), x_{t+1}, \ldots, x_n)\).

The claim now follows easily. From the property of the previous paragraph our claim can be rephrased as asserting that if \(f' \in \mathcal{F}\) then \(g' \in \mathcal{F}_{M - M}\). But if \(f' = g'_M + r' \in \mathcal{F}\), then it follows that \(g'_M\) (with its support being a subset of the support of \(f'\)) is also in \(\mathcal{F}\) and so \(g' \in \mathcal{F}_{M - M}\).

\[\square\]

Claim 4.4.14. If \(g|_H \in \mathcal{F}_{M - M}\) then \(g|_H \equiv 0\).

**Proof.** Assume for contradiction that \(g|_H \in \mathcal{F}_{M - M}\) and \(g|_H \equiv 0\). Let

\[g'(x_{t+1}, \ldots, x_n) = g|_H(x) = g(L(x), x_{t+1}, \ldots, x_n)\].

Every monomial of \(g\) of degree \(d\) where \(d \mod *\) \((q - 1) = a\), and hence the same holds also for \(g'\). For \(\beta = (\beta_{t+1}, \ldots, \beta_n)\), let \(p_{\beta}(x_1) = g'(\beta_{t+1}, \ldots, \beta_n) x_1^q\). Since \(g|_H \equiv 0\), there exists \(\beta = (\beta_{t+1}, \ldots, \beta_n)\) such that \(g'(\beta_{t+1}, \ldots, \beta_n) \equiv 0\) and so \(p_{\beta}(x_1)\) has \(x_1^q\) in its support. Note furthermore that \(p_{\beta}(x_1)\) is obtained by an affine (although non-invertible) transformation of the coordinates of \(g'\) which is given by \(T(e_i) = \beta e_i\) for all \(i \in \{e_{t+1}, \ldots, e_n\}\). Thus the fact that \(g' \in \mathcal{F}_{M - M}\) implies in turn that \(x_1^q \in \mathcal{F}_{M - M}\). But, by the definition of \(\mathcal{F}_{M - M}\) this implies \(M = x_1^q M \in \mathcal{F}\) which contradicts the hypothesis of the lemma.

\[\square\]

This concludes the proof of Lemma 4.4.7.
We now state and prove a lemma which bounds the maximal degree of functions in any affine-invariant linear code given a single monomial not in the code, which was used in the proof above.

**Lemma 4.4.15.** Let $\mathcal{G} \subseteq \{\mathbb{F}_q^n \rightarrow \mathbb{F}_q\}$ be an affine-invariant linear code and let $M$ be a monomial of degree $\ell$ such that $M \notin \mathcal{G}$. Then, for every function $f \in \mathcal{G}$, we have $\deg(f) \leq \frac{1}{2} q^2 \ell$.

*Proof.* We first note that we can assume, without loss of generality, that $\ell q \leq n$. Else (if $n < \ell q$) we can prove the result for the code $\mathcal{G}' = \text{Lift}_{\ell q}(\mathcal{G})$, and then use the identity $\mathcal{G} = \mathcal{G}' \cap \{\mathbb{F}_q^n \rightarrow \mathbb{F}_q\}$ to derive the result for $\mathcal{G}$. So from now we have $n \geq \ell q$.

Let $p$ be a prime number and $t$ be an integer such that $q = p^t$. Let $M' = \prod_{i=1}^{n} x_i^{a_i}$ be a monomial in $f \in \mathcal{G}$, write any degree $a_i$ in base-$p$ as $a_i = \sum_{j=0}^{t-1} a_j^{(i)} p^j$, where $a_j^{(i)} \in \{0, \ldots, p-1\}$ for all $0 \leq j \leq t-1$ and $i \in [n]$.

We will show that $\sum_{i=1}^{n} a_j^{(i)} < \ell p^{t-j}$ for every $0 \leq j \leq t-1$. This will show that

$$
\deg(M') = \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} \sum_{j=0}^{t-1} a_j^{(i)} p^j < \sum_{j=0}^{t-1} \ell p^{t-j} p^j = t q \ell \leq \frac{1}{2} q^2 \ell,
$$

thereby yielding the lemma.

Assume for contradiction that there is some $j$, such that $\sum_{i=1}^{n} a_j^{(i)} \geq \ell p^{t-j}$. Then, by Lemma 4.3.2 the monomial $M_1 = \prod_{i=1}^{t} x_i^{\ell p^{t-j}}$ is in $\mathcal{G}$. By applying the linear transformation $T_1$ given by $T_1(e_i) = e_i \mod \ell$ for every $i$ in the monomial $M_1$, we deduce that $\prod_{i=1}^{t} x_i^{\ell} \in \mathcal{G}$. In turn the resulting monomial is equivalent to the monomial $M_2 = \prod_{i=1}^{t} x_i$ over $\mathbb{F}_q$. Let $\ell_i$ denote the degree of $x_i$ in $M$ so that $\sum_i \ell_i = \ell$. Now consider the transformation $T_2$ defined by $\forall i \in [n], \forall k$ such that $\sum_{j=1}^{t} \ell_j < k \leq \sum_{j=1}^{t} \ell_j$ : $T_2(e_k) = e_i$. We have $M_2 \circ T_2 = M$, yielding $M \in \mathcal{G}$ which contradicts our assumption. The lemma follows. \qed

**Proof of Lemma 4.4.8**

We conclude the section by proving Lemma 4.4.8 which we restate below for convenience.

**Lemma 4.4.8 (restated).** Let $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$ be a non-zero polynomial of degree $d$. Then there are at most $q^{\frac{d}{2}+1}$ hyperplanes $H$ such that $f|_H \equiv 0$.

*Proof.* Let $H_1, \ldots, H_k$ be all the hyperplanes that satisfy $f|_{H_i} \equiv 0$. Consider the set

$$
S \triangleq \{ x \in \mathbb{F}_q^n \mid \forall i \in [k], x \notin H_i \}.
$$

We will find an upper and lower bound on the density of the set $S$ as a function of $k$ and this will yield the claimed bound on $k$.

Consider a point $x \in \mathbb{F}_q^n$, chosen uniformly at random. We first give an upper bound on the probability that $x \in S$. Let $Z_i$ be a random variable such that $Z_i = 1$ if and only if $x \in H_i$. Note that we wish to upper bound the probability that $\sum_{i=1}^{k} Z_i = 0$. We bound this probability using the Chebychev bound.

Clearly, for all $i \in [k]$,

$$
\mathbb{E} [Z_i^2] = \mathbb{E} [Z_i] = \frac{|H_i|}{|\mathbb{F}_q^n|} = \frac{1}{q}.
$$

71
Moreover, for $i \neq j$, $\mathbb{E}[Z_iZ_j] \leq \frac{1}{q}$. ($E[Z_iZ_j] = 1/q^2$ if $H_i$ and $H_j$ are not parallel, and equals zero if they are.) Calculating the variance,

$$\text{Var} \left( \sum_{i=1}^{k} Z_i \right) = \mathbb{E} \left[ \left( \sum_{i=1}^{k} Z_i \right)^2 \right] - \mathbb{E} \left[ \sum_{i=1}^{k} Z_i \right]^2 = 2 \sum_{i<j} \mathbb{E} [Z_iZ_j] + \sum_{i=1}^{k} \mathbb{E} [Z_i^2] - \left( \frac{k}{q} \right)^2$$

$$\leq \frac{k(k-1)}{q^2} + \frac{k}{q} - \frac{k^2}{q^2} = \frac{k(q-1)}{q^2}$$

We bound the density of $S$ by Chebyshev’s inequality,

$$\Pr[x \in S] = \Pr \left[ \sum_{i=1}^{k} Z_i = 0 \right] \leq \Pr \left[ \sum_{i=1}^{k} Z_i - \frac{k}{q} \geq \frac{k}{q} \right] \leq \frac{\text{Var} \left( \sum_{i=1}^{k} Z_i \right)}{\left( \frac{k}{q} \right)^2} \leq \frac{q-1}{k}.$$

On the other hand, by Lemma 3.3.2

$$\Pr \left[ x \in S \right] \geq \Pr \left[ f(x) \neq 0 \right] \geq q^{-\frac{d}{q-1}}$$

Combining the above, we have $q^{-\frac{d}{q-1}} \leq \frac{q-1}{k}$ which yields

$$k \leq q^{\frac{d}{q-1}}(q-1) < q^{\frac{d}{q-1}+1},$$

as claimed.

### 4.4.3 Restrictions of general functions to hyperplanes

We finally turn to the proof of the Main Technical Theorem 4.2.1. The proof of this section is a straightforward adaptation of the proof of Theorem 3.1.7, given Theorem 4.4.10. We give a brief overview of the proof first.

Recall that Theorem 4.2.1 says that if a function $f \in \{\mathbb{F}_q^n \rightarrow \mathbb{F}_q\}$ is $\delta$-close to functions from $\mathcal{F}$ on $k$ hyperplanes (for sufficiently large $k$), then $f$ is itself close to some function from $\mathcal{F}$. It turns out that the central difficulty in proving this theorem already arises when $\delta = 0$, and the theorem for general $\delta$ follows immediately. Theorem 4.4.16 states this special case, which we prove first. The proof of Theorem 4.2.1 follows easily and we prove it later in Section 4.4.3.

The proof of Theorem 4.4.16 is itself by induction on $n$, however now the hardest part is the base case. We prove the base case separately as Lemma 4.4.18 in Section 4.4.3. We then prove Theorem 4.4.16 as a consequence in Section 4.4.3.

**Interpolation from exact agreement**

We start by stating Theorem 4.4.16 which implies the special case of Theorem 4.2.1 for the case of $\delta = 0$.

**Theorem 4.4.16.** For every $q$ there exists $\tau < \infty$ such that the following holds: Let $n > t$, let $\mathcal{B} \subseteq \{\mathbb{F}_q^t \rightarrow \mathbb{F}_q\}$ be an affine-invariant linear code and let $\mathcal{F} = \text{Lift}_n(\mathcal{B})$. Let $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$ be a function and $H_1, \ldots, H_k$ be hyperplanes in $\mathbb{F}_q^n$ such that $f|_{H_i} \in \mathcal{F}$ for every $i \in [k]$. Then, if $k \geq q^{t+\tau}$, there exists a function $h \in \mathcal{F}$ such that $f|_{H_i} = h|_{H_i}$ for all $i \in [k]$.

We will prove Theorem 4.4.16 in Section 4.4.3 by induction on $n$. Our proof will rely on a slightly stronger (smaller) bound on $k$ as $n$ gets smaller. This makes the base case of small values of $n$ more challenging and we deal with this first.
The base case

Here we consider the case where \( n = t + O(1) \). In this case the number of hyperplanes \( k \geq q^{t+r} \) is a “constant” fraction of all the hyperplanes. We view these hyperplanes as points in a \((n+1)\)-dimensional subspace (the hyperplane given by \( \sum_{i=1}^{n} \alpha_i x_i = \alpha_0 \) is associated with the point \((\alpha_0, \ldots, \alpha_n) \in \mathbb{F}_q^{n+1}\)), and then use the well-known Hales-Jewett Theorem from additive combinatorics to infer that there are \( q \) points in a straight line among this set of points. (Indeed by choosing our density to be slightly larger we may conclude that there are many straight lines among the given set of points. We use such a version that was already used in Chapter 3.) In terms of hyperplanes these lines lead to a small set that cover most of \( \mathbb{F}_q^n \). We use this set to derive that there is a function \( h \) from \( \mathcal{F}^+ \) that is consistent with \( f \) on all the given hyperplanes. We then use Theorem 4.4.10 to conclude that \( h \) must actually be an element of \( \mathcal{F} \).

Recall the following definition

**Definition 4.4.17.** Let \( v \in \mathbb{F}_q^n \) and \( u \in \mathbb{F}_q^n \setminus \{0\} \). A line through \( v \) in direction \( u \) is the set \( \{v + \alpha u \mid \alpha \in \mathbb{F}_q\} \). Notice that the direction of a given line is unique up to multiplication by an element of \( \mathbb{F}_q \setminus \{0\} \).

For completeness of this chapter we are restating the following corollary of the Hales-Jewett theorem [FK91, Pol09].

**Theorem** (Corollary 3.3.5 restated). For every prime power \( q \) and every \( c > 0 \) there exists an integer \( \lambda_{q,c} \) such that for every integer \( m \in \mathbb{N} \) the following holds: if \( n \geq \lambda_{q,c} + m \) then every set \( A \subseteq \mathbb{F}_q^n \) of size \( |A| \geq q^{n-c} \) contains \( m \) lines whose directions are linearly independent.

The following lemma now states Theorem 4.4.16 for the special (base) case of \( n \leq t + O(1) \).

**Lemma 4.4.18.** For every \( q \) and constant \( c \), there exists a constant \( \tau_c < \infty \) such that the following holds: Let \( n, t \in \mathbb{N} \) be such that \( t < n \leq t + \tau_c \). Let \( B \subseteq \{\mathbb{F}_q^n \to \mathbb{F}_q\} \) be an affine-invariant linear code and let \( \mathcal{F} = \text{Lift}_n(B) \). Let \( f : \mathbb{F}_q^n \to \mathbb{F}_q \) be a function and \( H_1, \ldots, H_k \) be hyperplanes in \( \mathbb{F}_q^n \) such that \( f|_{H_i} \in \mathcal{F} \) for every \( i \in [k] \). Then, if \( k \geq q^{n-\tau_c-c} \), there exists a function \( h \in \mathcal{F} \) such that \( f|_{H_i} = h|_{H_i} \) for all \( i \in [k] \).

**Proof.** We prove the lemma for \( \tau_c = \lambda + q + c \) where \( \lambda = \max \{\lambda_{q,c+1} + 1, q^2 + 4\} \) and \( \lambda_{q,c+1} \) is as given by Corollary 3.3.5.

**Overview:** We start by giving an overview of the proof. Using a natural correspondence between hyperplanes in \( \mathbb{F}_q^n \) and points in \( \mathbb{F}_q^{n+1} \) (the hyperplane \( \sum_{i=1}^{n} \alpha_i x_i = \alpha_0 \) corresponds to the point \((\alpha_0, \ldots, \alpha_n) \in \mathbb{F}_q^{n+1}\)) and the Hales-Jewett theorem in \( \mathbb{F}_q^n \) we find many hyperplanes of a “somewhat structured” type. We will formally describe these later below, but an example of hyperplanes corresponding to points on a line would be the set of hyperplanes \( x_1 + \lambda x_2 = 0 \) for every \( \lambda \in \mathbb{F}_q \). This set of hyperplanes almost covers the entire region \( \mathbb{F}_q^n \) except the points with \( x_1 \neq 0 \) and \( x_2 = 0 \).

We then proceed in three steps: We first observe that for every hyperplane of the form \( x_2 = \eta \) for \( \eta \neq 0 \), \( f \) restricted to this hyperplane is an element of \( \mathcal{F}^+ \). Observing further that \( f|_{x_2=\eta} \) is a function of \( \mathcal{F} \) when restricted to many hyperplanes in \( \mathbb{F}_q^{n-1} \), we use Theorem 4.4.10 to claim that \( f|_{x_2=\eta} \in \mathcal{F} \).

Now if we only could claim that \( f|_{x_2=0} \) is also an element of \( \mathcal{F} \) we would be done by a similar sequence of observations. However this is not necessarily true. To deal with this, we show in the first step, using the fact that there are many hyperplanes, that for many variables \( x_i \) we have \( f|_{x_i=\eta} \in \mathcal{F} \) for every \( \eta \neq 0 \).

In the second step we apply some algebraic interpolations to show for every \( i \in [m] \) the existence of a function \( h_i : \mathbb{F}_q^n \to \mathbb{F}_q \) such that \( h_i \in \mathcal{F}^+ \) and \( h_i|_{x_i=\eta} = f|_{x_i=\eta} \) for every \( \eta \neq 0 \).
In the third and last step we show how to build a single function \( h \in \mathcal{F}^+ \) that agrees with \( f|_{x_i=\eta} \) for every choice of \( i \) and for every \( \eta \neq 0 \), and then show that this function is in \( \mathcal{F} \) and agrees with \( f \) on every given hyperplane. We note that this step requires some non-trivial extensions of corresponding steps in the proof of Lemma 3.4.11. We now turn to the formal proof.

The formal proof: We start by showing that some very structured set of hyperplanes are included among the given \( k \) hyperplanes. For a point \( \alpha = (\alpha_0, \ldots, \alpha_n) \in \mathbb{F}_{q}^{n+1} \) let \( H_\alpha \) denote the hyperplane \( H_\alpha = \{ x \in \mathbb{F}_{q}^n \mid \sum_{i=1}^{n} \alpha_i x_i = \alpha_0 \} \). For an affine transformation \( T \), let \( H_\alpha \circ T \) denote the hyperplane \( H_{T(\alpha)} \).

Claim 4.4.19. Let \( m = t + q + 1 \). There exists an invertible affine transformation \( T \) and \( m \) invertible affine functions \( L_1, \ldots, L_m : \mathbb{F}_{q}^n \rightarrow \mathbb{F}_{q}^n \) such that for every \( i \in [m] \) and \( \gamma \in \mathbb{F}_{q} \) the hyperplane \( H_{\gamma}(i) \triangleq \{ x \mid L_i(x) + \gamma x_i = 0 \} \) is included in the set \( \{ H_1 \circ T, \ldots, H_k \circ T \} \).

Proof. Let \( P = \{ \alpha^{(1)}, \ldots, \alpha^{(k)} \} \subseteq \mathbb{F}_{q}^{n+1} \) be a set of \( k \) points such that \( H_i = H_{\alpha^{(i)}} \) for every \( i \in [k] \).

We assume without loss of generality that \( \alpha_0^{(i)} \in \{ 0, 1 \} \) for all \( i \in [k] \) since other \( \alpha_j^{(i)} \)’s can be scaled to achieve this. Note furthermore that if \( n < \log_q k \) then there is nothing to prove since in this case one cannot find \( k \) distinct hyperplanes inside \( \mathbb{F}_{q}^n \). Hence we may assume that \( n \geq \log_q k \geq t + \tau_c - c \).

Since the density of \( P \) in \( \mathbb{F}_{q}^{n+1} \) is \( k/q^{n+1} \geq q^{-(c+1)} \) and \( n \geq t + \tau_c - c \geq t + \lambda_q c + 1 + 1 + q \), by Corollary 3.3.5 we have that there are at least \( m = t + q + 1 \) linearly independent lines in \( P \). Since all points in \( P \) have their 0th coordinate in \( \{ 0, 1 \} \) these lines must be constant in the 0th direction.

By applying an invertible linear transformation to the last \( n \) coordinates, we can assume without loss of generality that the lines are parallel to the axes in directions \( x_1, \ldots, x_m \). Let \( T \) be such a transformation and let \( T(P) = \{ T(\alpha) \mid \alpha \in P \} \). Then we get that there are vectors \( \beta^{(1)}, \ldots, \beta^{(m)} \subseteq \mathbb{F}_{q}^{n+1} \) such that for every \( i \in [m] \) and every \( \gamma \in \mathbb{F}_{q} \), the vector \( \beta^{(i)} + \gamma e^{(i)} \in T(P) \), where \( e^{(i)} = (\epsilon_0^{(i)}, \ldots, \epsilon_n^{(i)}) \) has \( \epsilon_i^{(i)} = 1 \) and is 0 on every other coordinate. For \( i \in [m] \), let \( L_i(x) = \sum_{j=1}^{n} \beta_j^{(i)} x_j - \beta_0^{(i)} \). The claim follows for this choice of \( T \) and \( L_i \)’s.

In what follows, we assume that the affine transformation \( T \) above is the identity transform (or else we can simply prove Lemma 4.4.18 for the function \( f \circ T \)).

First step

Claim 4.4.20. For every \( i \in [m] \), \( \eta \in \mathbb{F}_{q}^{*} \), we have \( f|_{x_i=\eta} \in \mathcal{F} \) (i.e., \( f \) is an element of \( \mathcal{F} \) when restricted to the hyperplane given by fixing \( x_i \) to \( \eta \)).

Proof. In order to prove the claim we first prove that \( f|_{x_i=\eta} \) is in \( \mathcal{F}^+ \) and then use Theorem 4.4.10 to deduce that \( f|_{x_i=\eta} \) is actually in \( \mathcal{F} \).

Fix \( x \in \mathbb{F}_{q}^{n} \) such that \( x_i = \eta \) and let \( \beta = L_i(x) \). By definition \( x \in H_{L_i-\eta} \beta x_i \) (note that here we use the fact that \( \eta \neq 0 \)). We thus conclude that \( H_{x_i-\eta} = \bigcup_{\gamma \in \mathbb{F}_{q}} (H_{x_i-\eta} \cap H_{L_i-\gamma x_i}) \). In other words the hyperplane \( H_{x_i-\eta} \) is covered by \( q \) parallel hyperplanes in \( \mathbb{F}_{q}^{n-1} \). Thus, since for every \( \gamma \in \mathbb{F}_{q} \) we have \( f|_{H_{L_i-\gamma x_i}} \in \mathcal{F} \), we get \( f|_{H_{x_i-\eta} \cap H_{L_i-\gamma x_i}} \in \mathcal{F} \) as well. We thus conclude, by Lemma 4.4.5, that \( f|_{x_i=\eta} \in \mathcal{F}^+ \).

Now, consider the set \( S = \{ H_{x_i-\eta} \cap H_j \mid j \in [k] \} \). Allowing for \( q \) of \( H_j \)’s to be parallel to \( H_{x_i-\eta} \) and for \( q \) different \( H_j \)’s to become identical when restricted to \( H_{x_i-\eta} \), we still get \( \frac{k}{q} - 1 > q^{t+q^2+q+2} \) many distinct \((n-2)\)-dimensional affine subspaces of \( H_{x_i-\eta} \) in \( S \). For each such subspace \( H_j \), we have \( f|_{H_{x_i-\eta} \cap H_j} = (f|_{H_j})|_{H_{x_i-\eta}} \in \mathcal{F} \). Therefore, by Theorem 4.4.10 (applied to functions over \( \mathbb{F}_{q}^{n-1} \)), we get \( f|_{x_i=\eta} \in \mathcal{F} \).
Second step

Claim 4.4.21. For every $i \in [m]$ there exists a function $h_i : \mathbb{F}_q^n \to \mathbb{F}_q$ with $h_i \in \mathcal{F}^+$ and $h_i|_{x_i=\eta} = f|_{x_i=\eta}$ for every $\eta \in \mathbb{F}_q^*$.

Proof. Let $h_i$ be defined as $h_i(x) = f(x)$ when $x_i \neq 0$ and $h_i(x) = 0$ otherwise. Clearly we have that $h_i|_{x_i=\eta} = f|_{x_i=\eta}$ for every $\eta \in \mathbb{F}_q^*$, it remains to show that $h_i \in \mathcal{F}^+$. To see this note that Claim 4.4.20 above implies that for every $\eta \neq 0$, $h_i|_{x_i=\eta} = f|_{x_i=\eta} \in \mathcal{F}$. Since $\mathcal{F}$ is linear we also have that the zero function is contained in $\mathcal{F}$ and hence $h_i|_{x_i=0}$ is also contained in $\mathcal{F}$. Thus we have that $h_i$ is an interpolation of functions in $\mathcal{F}$ as per Definition 4.4.3. Lemma 4.4.5 then implies that $h_i \in \mathcal{F}^+$. □

Third step Our final step, which is the major step of this proof, is to collect the $h_i$’s together consistently to form the function $h$. Lemma 4.4.22 below proves that there is a function $h \in \mathcal{F}^+$ such that $h$ agrees with $f$ on all the hyperplanes $H_{x_i=\eta}$ for $\eta \neq 0$ and $i \in [m]$. We now conclude the proof by going back to the $k$ hyperplanes $H_1, \ldots, H_k$ given by the hypothesis. For every $j \in [k]$, we first claim that $h|_{H_j} = f|_{H_j}$. To see this, let $S_j = H_j \cap (\bigcup_{i=1}^{m} \cup_{\eta \neq 0} H_{x_i=\eta})$. On the one hand $h$ and $f$ agree on every point in $S_j$. On the other hand, we have $|S_j| \geq q^{n-1} - (1 - q^{-(m-1)})$. Finally, we also have that $h|_{H_j}, f|_{H_j} \in \mathcal{F} \subseteq \mathcal{F}^+$. Since $\mathcal{F}^+ \geq q^{-(t+q-1)} > q^{-(m-1)}$ (since $m > t + q$) we get $f|_{H_j} = h|_{H_j}$. We now have that $h \in \mathcal{F}^+$ is a function that on $k > q^{t+q^2+q+2}$ hyperplanes $h$ restricted to the hyperplane is a function in $\mathcal{F}$. By Theorem 4.4.10, we have $h \in \mathcal{F}$ as desired. □

Lemma 4.4.22. Let $\mathcal{G}$ be an affine-invariant linear code and let $h_1, \ldots, h_m, f : \mathbb{F}_q^n \to \mathbb{F}_q$ be functions such that for every $\eta \neq 0$ and $i \in [m]$, we have $h_i|_{x_i=\eta} = f|_{x_i=\eta}$, and $h_i \in \mathcal{G}$ for every $i \in [m]$. Then there exists $h \in \mathcal{G}$ such that $h|_{x_i=\eta} = f|_{x_i=\eta}$ for every $i \in [m]$ and $\eta \neq 0$.

For the proof of the above lemma we shall need the following definition of non-standard monomials.

Definition 4.4.23 (Non-standard monomials). For integer $j \in \{0, \ldots, q-1\}$ we define the “non-standard” monomial $N_j(t)$ to be $t^j$ if $j \neq q-1$ and $t^{q-1} - 1$ if $j = q-1$. For a vector $a \in \{0, \ldots, q-1\}^n$ we define the non-standard monomial $N_a(x)$ to be $\prod_{i=1}^{n} N_{a_i}(x_i)$.

It is simple to see that non-standard monomials do form a basis for all functions from $\mathbb{F}_q^n \to \mathbb{F}_q$. We mention this and some other properties we will be using below.

Proposition 4.4.24. 1. For every function $f : \mathbb{F}_q^n \to \mathbb{F}_q$ there exists a unique set of coefficients denoted $\{c_a\}_{a \in \{0, \ldots, q-1\}^n}$ such that $f(x) = \sum_a c_a N_a(x)$.

2. For $I \subseteq [n]$, let $A_I = \{a \in \{0, \ldots, q-1\}^n | a_i \neq q-1 \forall i \in I\}$ and let $S_I = \{x \in \mathbb{F}_q^n | x_i \neq 0 \forall i \in I\}$. Then for every function $f(x) = \sum_a c_a N_a(x)$ the coefficients $\{c_a\}_{a \in A_I}$ are uniquely determined by $f|_{S_I}$.

3. Let $\mathcal{G}$ be an affine-invariant linear code and suppose $f = \sum_a c_a N_a(x)$ is in $\mathcal{G}$. Then for every $a$ such that $c_a \neq 0$, it holds that $N_a(x) \in \mathcal{G}$.

Proof. The first part of the proposition is immediate and the second part follows from Lemma 3.4.13 so it remains to prove the third part. For a vector $a \in \{0, \ldots, q-1\}^n$ denote by $x^a$ the (standard) monomial $\prod_{i=1}^{n} x_i^{a_i}$. Let $a$ be such that $c_a \neq 0$ and let $x^{a'}$ be a monomial of maximal degree such that $c_a' \neq 0$ and $x^{a'} \in \text{supp}(N_a'(x))$. Since $x^{a'}$ is of maximal degree we must have that $x^{a'} \in \text{supp}(f)$ which by Lemma 4.3.1 implies that $x^{a'} \in \mathcal{G}$.

Note that all monomials in $\text{supp}(N_a'(x))$ are of the form $x^b$ where $b_i = a_i'$ if $a_i' \neq q-1$ and $b_i \in \{0, q-1\}$ if $a_i' = q-1$. Since $x^a \in \text{supp}(N_a(x))$, in particular we have that every monomial in $\text{supp}(N_a(x))$ is of this form. By Lemma 4.3.2 this implies in turn that $\text{supp}(N_a(x)) \subseteq \text{Aff}_{n \times n}(x^{a'})$. Since $x^{a'} \in \mathcal{G}$ we conclude that $\text{supp}(N_a(x)) \subseteq \mathcal{G}$ so $N_a(x) \in \mathcal{G}$ as required. □
Proof of Lemma 4.4.22. We now use the non-standard monomials. For $i \in [m]$, let $\{c_a^{(i)}\}_{a \in \{0, \ldots, q-1\}^n}$ be such that $h_i(x) = \sum_a c_a^{(i)} N_a(x)$. Let $D = \{a \in \{0, \ldots, q-1\}^n | \exists i \in [m] \text{ s.t. } a_i \neq q-1 \}$. For $a \in D$, let $i(a) = \min \{ |a_i| \neq q-1 \}$. We define $h(x) = \sum_{a \in D} c_a^{(i(a))} N_a(x)$. We argue below that $h$ is a member of $G$ and $h$ agrees with $f$ on the hyperplanes $H_{x_i - \eta}$ for every $i \in [m]$ and $\eta \neq 0$.

The first part is simple. We first notice that every term in the non-standard expansion of $h = \sum_a c_a N_a(x)$ is in $G$. Suppose $N_a(x)$ has a non-zero coefficient in the expansion of $h$. Then we have that $c_a = c_a^{(i)}$ and so $N_a(x)$ has a non-zero coefficient in the non-standard expansion of $h_{i(a)}$. Since $h_{i(a)} \in G$, it follows, from Part (3) of Proposition 4.4.24, that $N_a(x) \in G$. Thus, every term of $h$ is in $G$ and by linearity of $G$ it follows that $h \in G$.

It remains to argue that $h$ equals $f$ on every hyperplane of the form $x_i = \eta$ for $i \in [m]$ and $\eta \neq 0$. To see this we first claim that for $i \neq j \in [m]$ and $a \in \{0, \ldots, q-1\}^n$ if $a_i, a_j \neq q-1$ then $c_a^{(i)} = c_a^{(j)}$. To see this, note that $h_i |_{x_i \neq 0, x_j \neq 0} = f |_{x_i \neq 0, x_j \neq 0} = h_j |_{x_i \neq 0, x_j \neq 0}$. But now, applying Part (2) of Proposition 4.4.24 to the set $I = \{i, j\}$, we get that $c_a^{(i)} = c_a^{(j)}$ as claimed. Thus, as a consequence, we have that $c_a = c_a^{(i)}$ for every $a$ such that $a_i \neq q-1$. Applying Part (2) of Proposition 4.4.24 again, this time to the set $I = \{i\}$, we have that $h$ and $h_i$ must agree in every $x$ such that $x_i \neq 0$. The lemma follows.

\qed

Proof of Theorem 4.4.16

We are now ready to prove Theorem 4.4.16.

Proof of Theorem 4.4.16. We will prove the theorem for $\tau = \max \{ \tau_4 - 3, q^2 + q + 1 \}$ where $\tau_4$ is the constant given by Lemma 4.4.18 for $c = 4$.

Recall that we wish to prove that if $f$ agrees with a function from $F$ on $k$ hyperplanes (where the agreeing function may be different for each hyperplane), then there is a single function in $F$ with whom $f$ agrees on all the given hyperplanes. We wish to prove this when $k \geq q^{\tau + \tau}$, but we will prove a slightly stronger result for the induction.

Inductive Hypothesis: Let $n' := n - t - \tau$, let $C(t, n) = \frac{q^{t+\tau}}{2^\sum_{i=1}^{n-3} (1-q^{-n'+i+1})}$ and let $k \geq C(t, n)$. Let $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$ be a function such that there exist $k$ hyperplanes $H_1, \ldots, H_k$ in $\mathbb{F}_q^n$ such that $f \mid H_i \in F$ for every $i \in [k]$. Then there exists $h \in F$ such that $f \mid H_i = h \mid H_i$ for every $i \in [k]$.

We first note that the hypothesis does imply the theorem. This is so since the denominator in the above expression is

$$2 \prod_{i=1}^{n-3} \left( 1 - q^{-n'+i+1} \right) \geq 2 \left( 1 - \sum_{i=1}^{n-3} q^{-n'+i+1} \right) = 2 - 2q^{-n'+1} \sum_{i=1}^{n-3} q^i$$

$$> 2 - 2q^{-n'+1}, q^{-n'} = 2 - 2q^{-1} \geq 1,$$

and so $C(t, n) \leq q^{t+\tau}$.

**Base Case** ($n \leq t + \tau_4$): In this case we have $k \geq C(t, n) \geq q^{t+\tau} / 2 \geq q^{t+\tau-1} \geq q^t \tau_4 - 4$. Applying Lemma 4.4.18 with $c = 4$, we find that we have $k \geq q^t \tau_4 - \varepsilon$ and $n \leq t + \tau_4$ and so a function $h$ as desired exists.

**Inductive step:** In Claim 4.4.25 below we prove that there exists a linear function $L$ such that for every $\gamma \in \mathbb{F}_q$ the hyperplane $H_{L(x)-\gamma}$ is “good” in the sense that the set

$$S_\gamma = \{ H_i \cap H_{L(x)-\gamma} | i \in [k], \dim(H_i \cap H_{L(x)-\gamma}) = n - 2 \}$$
is of size at least $C(t, n - 1)$. By induction, we conclude that for every $\gamma$ the functions $f|_{H_{L(x),\gamma}}$ belong to $\mathcal{F}$ (since each agrees with a member of $\mathcal{F}$ on at least $C(t, n - 1)$ hyperplanes). Using interpolation, we conclude the existence of a function $h \in \mathcal{F}^+$ which agrees with $f$ on all hyperplanes $H_i$. Applying Theorem 4.4.10 (using the fact that $\tau \geq q^2 + q + 1$ and hence $k \geq q^{t+q^2+2}$) we now conclude that $h \in \mathcal{F}$. Details follow.

**Claim 4.4.25.** There exists a linear function $L \in \text{Aff}_n$ such that for every $\gamma \in \mathbb{F}_q$ the set

$$S_\gamma = \{ H_{L(x),\gamma} \cap H_i \mid i \in [k], \dim (H_{L(x),\gamma} \cap H_i) = n - 2 \}$$

has cardinality at least $C(t, n - 1)$

**Proof.** Without loss of generality assume $k = C(t, n)$. Let $L_i \in \text{Aff}_n$ be an affine function such that $H_i = H_{L_i}$. For $L \in \text{Aff}_n$ and $i \neq j \in [k]$ such that $H_i \cap H_j \neq \emptyset$ the sets $H_{L,\gamma} \cap H_i, H_{L,\gamma} \cap H_j$ are the same only if there exist $\alpha, \beta \in \mathbb{F}_q \setminus \{0\}$ such that $L = \alpha L_i + \beta L_j + \gamma$. Moreover, $\dim (H_{L,\gamma} \cap H_i) \neq n - 2$ only if there are $\alpha, \gamma' \in \mathbb{F}_q$ such that $L = \alpha L_i + \gamma'$. There are at most $k^2 q^3$ ways to represent a function in $\text{Aff}_n$ as $L = \alpha L_i + \beta L_j + \gamma$ where $i, j \in [k]$ and $\alpha, \beta, \gamma \in \mathbb{F}_q$. Hence there is some function $L \in \text{Aff}_n$ such that there are at most $\frac{k^2 q^3}{|\text{Aff}_n|} = \frac{k^2}{q^n-2}$ such different ways to represent it (we allow $\alpha, \beta$ and $\gamma \in \mathbb{F}_q$ arbitrary to be zero to deal with the case where $L = \alpha L_i + \gamma$). As we saw, for any hyperplane that we lose in the set $S_\gamma$ there is at least one such representation for $L$. So $|S_\gamma| \geq k - \frac{k^2}{q^n-2}$. Calculating

$$|S_\gamma| \geq k - \frac{k^2}{q^n-2} = k \left( 1 - \frac{k}{q^n-2} \right)
\geq C(t, n) \left( 1 - \frac{q^{t+\tau}}{q^n-2} \right) = C(t, n) \left( 1 - q^{-n'}+2 \right)
= \left( 1 - q^{-n'}+2 \right) 2 \prod_{i=1}^{n'-3} \frac{q^{t+\tau}}{1 - q^{-n'+i+1}}
= \frac{2 \prod_{i=2}^{n'-3} \left( 1 - q^{-n'+i+1} \right)}{2 \prod_{i=1}^{n'-4} \left( 1 - q^{-n'+i+2} \right)} = C(t, n - 1)
$$

We are ready to continue the proof of Theorem 4.4.16. Consider the function $L \in \text{Aff}_n$ as promised by Claim 4.4.25 and fix some $\gamma \in \mathbb{F}_q$. Observe that there are $C(t, n - 1)$ hyperplanes of the space $H_{L,\gamma}$ of the form $H_i \cap H_{L,\gamma}$ where $i \in [k]$. On each one $f|_{H_i \cap H_{L,\gamma}} = (f|_{H_i})|_{H_i \cap H_{L,\gamma}} \in \mathcal{F}$. So, by the induction hypothesis there exists some function $h_\gamma \in \mathcal{F}$ such that $(f|_{H_{L,\gamma}})|_{H_i \cap H_{L,\gamma}} = (h_\gamma)|_{H_i \cap H_{L,\gamma}}$ for all $i \in [k]$ such that $\dim(H_i \cap H_{L,\gamma}) = n - 2$.

Define

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

By Lemma 4.4.5, $h$ is in $\mathcal{F}^+$. Let $i \in [k]$ and $x \in H_i$. Define $\gamma' = L(x)$, so clearly $x \in H_i \cap H_{L,\gamma'}$ and hence

$$h(x) = \sum_{\gamma \in \mathbb{F}_q} \left( \prod_{\alpha \neq \gamma} \frac{\gamma' - \alpha}{\gamma - \alpha} \right) \cdot h_\gamma(x) = h_{\gamma'}(x) = f(x).$$

77
We saw that $h|_{H_i} = f|_{H_i}$ for any $i \in [k]$. We conclude by observing that $h \in \mathcal{F}^+$ is a function such that on $k \geq q^{t^2+q+1}$ hyperplanes $H, h|_H \in \mathcal{F}$. Hence by Theorem 4.4.10 $h \in \mathcal{F}$ and we are done.

The case of general $\delta$

We finally turn to the proof of Theorem 4.2.1. To prove this theorem, we shall also need the following proposition whose proof appears as part of the proof of Theorem 3.4.16

**Proposition 4.4.26.** Let $f, g : \mathbb{F}_q^n \to \mathbb{F}_q$ be a pair of functions such that there are $k$ hyperplanes $H_1, ..., H_k$ which satisfy $\delta(f|_{H_i}, g|_{H_i}) < \delta$ for all $1 \leq i \leq k$. Then $\delta(f, g) \leq 2\delta + \frac{4(q-1)}{k}$.

We include the proof below for completeness.

**Proof.** Let $S \subseteq \mathbb{F}_q^n$ be the set of points given by $S \triangleq \{x \in \mathbb{F}_q^n \mid \Pr_{i \in [k]}[x \in H_i] \leq 1/(2q)\}$. We claim first that $\delta(f, g) \leq 2\delta + |S|/q^n$. To see this consider the following experiment: Pick a random hyperplane $H_i$ by picking $i$ uniformly from $[k]$ and pick a point $x$ uniformly at random from $\mathbb{F}_q^n$ and let $I = I(i, x) = 1$ if $x$ lies on $H_i$ and $f(x) \neq g(x)$. On the one hand we have $\mathbb{E}[I] \leq \delta/q$ since the probability that $x$ lies on $H_i$ is $1/q$ and conditioned on $x \in H_i$ the probability that $f$ and $g$ disagree is at most $\delta$. On the other hand, the probability that $f$ and $g$ disagree on $x$ and $x \notin S$ is at least $\delta(f, g) - |S|/q^n$ and conditioned on $x \notin S$, the probability that $x \in H_i$ is at least $1/(2q)$. We conclude that $\frac{1}{2q}(\delta(f, g) - |S|/q^n) \leq \delta/q$ and so $\delta(f, g) \leq 2\delta + |S|/q^n$. Thus it suffices to show that $|S|/q^n \leq 4(q-1)/k$, which we do next (by an application of Chebychev bound).

Consider picking $x \in \mathbb{F}_q^n$ at random and let $Y_i = Y_i(x) = 1$ if $x \in H_i$. Notice $x \in S$ if and only if $Y(x) \triangleq \sum_{i=1}^k Y_i(x) < k/(2q)$. We have $\mathbb{E}[Y_i] = 1/q$ and $\mathbb{E}[Y_iY_j] \leq 1/q^2$ (we have $\mathbb{E}[Y_iY_j] = 1/q^2$ if the hyperplanes are not parallel and $\mathbb{E}[Y_iY_j] = 0$ if they are). Thus $Y = \sum_{i=1}^k Y_i$ has expectation $k/q$ and variance $\mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \leq k/q + k(k-1)/q^2 = k(1/q)(1-1/q)$. By the Chebychev bound it follows that $\Pr[Y < k/(2q)] \leq \Pr[|Y - \mathbb{E}[Y]| \geq k/(2q)] \leq (2q)^2 k(1/q)(1-1/q)/k^2 = 4(q-1)/k$. The proposition follows.

We can now prove Theorem 4.2.1 as a corollary of Theorem 4.4.16 and Proposition 4.4.26.

**Proof of Theorem 4.2.1.** For all $i \in [k]$ let $g_i$ be a function in $\mathcal{F}$ such that $\delta(g_i|_{H_i}, f|_{H_i}) \leq \delta$. We will show that the functions $g_1, ..., g_k$ are consistent with each other, namely that $g_i|_{H_i \cap H_j} = g_j|_{H_i \cap H_j}$ for all $1 \leq i, j \leq k$.

For any $i, j \in [k]$, if $H_i \cap H_j = \emptyset$ then there is nothing to prove. Else,

$$\begin{align*}
\delta(g_i|_{H_i \cap H_j}, g_j|_{H_i \cap H_j}) &\leq \delta(g_i|_{H_i \cap H_j}, f|_{H_i \cap H_j}) + \delta(f|_{H_i \cap H_j}, g_j|_{H_i \cap H_j}) \\
&\leq q\delta + q\delta < q^{-t}.
\end{align*}$$

But by Corollary 4.3.6, the distance of $\mathcal{F}$ is at least $q^{-t}$, so $g_i$ and $g_j$ must agree on $H_i \cap H_j$. Theorem 4.4.16 then implies the existence of a function $g \in \mathcal{F}$ such that $g|_{H_i} = g_i|_{H_i}$ for every $i \in [k]$. By Proposition 4.4.26, $\delta(g, f) \leq 2\delta + 4(q-1)/k$ and so $\delta_\mathcal{F}(f) \leq 2\delta + 4(q-1)/k$ as required.
4.5 Proof of Main Theorem 4.1.1, $Q = q$ case

In this section we prove our Main Theorem 4.1.1 for the special case in which $Q = q$. In order to prove Theorem 4.1.1 in this case, we first show in Lemma 4.5.1 below, using probabilistic arguments, bounds on the rejection probability of the $\ell$-dimensional test for the case in which $f$ is relatively close to $\mathcal{F}$ and $\ell \geq t + 1$. In Lemma 4.5.2 we then use our Main Technical Theorem 4.2.1 to bound the rejection probability of the $\ell$-dimensional test for the case in which $f$ is relatively far from $\mathcal{F}$ and $\ell \geq t + c$ for some absolute constant $c$. Combining Lemmas 4.5.1 and 4.5.2 one can bound the rejection probability of the $\ell$-dimensional test when $\ell = t + c$. Relating this to the rejection probability of the $\ell$-dimensional test requires some extra work given in Lemma 4.5.3 and Corollary 4.5.4 below.

We start by analyzing the rejection probability of the $\ell$-dimensional test for the case in which $f$ is relatively close to $\mathcal{F}$ and $\ell \geq t + 1$. Recall that $\text{Rej}_\ell(f)$ denotes the probability that the $\ell$-dimensional test rejects the function $f : \mathbb{F}_q^n \to \mathbb{F}_q$.

**Lemma 4.5.1.** Let $\mathcal{F} = \text{Lift}_n(\mathcal{B})$ for an affine-invariant linear base code $\mathcal{B} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_{q^t}\}$. Then for every $n \geq \ell \geq t + 1$, and every $f : \mathbb{F}_q^n \to \mathbb{F}_q$, if $\delta_\mathcal{F}(f) \leq q^{-(t+1)/2}$ then $\text{Rej}_\ell(f) \geq \min\{\frac{1}{4q}, q\delta_\mathcal{F}(f)/2\}$.

**Proof.** The proof is similar to the proof of Lemma 3.5.1. We will use the monotonicity of the rejection probability and prove a bound on $\text{Rej}_\ell(f)$ for some $\ell' \leq \ell$.

Let $\ell'$ be such that $t + 1 \leq \ell' \leq \ell$ and let $A$ be an $\ell'$-dimensional subspace. Let $g \in \mathcal{F}$ be the function closest to $f$, so that $\delta(f, g) = \delta \triangleq \delta_\mathcal{F}(f)$. We now use the fact that $\delta(f|_A, g|_A)$ is the average of $N = q^{\ell'}$ random $\{0, 1\}$-valued variables of expectation $\delta$ that are roughly pairwise independent to derive a lower bound on the probability that $f|_A$ and $g|_A$ disagree on exactly one point. Note that if $f|_A$ and $g|_A$ disagree on exactly one point, then $\delta(f|_A, g|_A) = q^{-\ell'} < q^{-\ell}$ due to our assumption that $\ell' \geq t + 1$. But by Corollary 4.3.6 we have that $\delta(\mathcal{F}) \geq q^{-\ell}$ and consequently $f|_A \notin \mathcal{F}$ in this case. Thus the $\ell'$-dimensional test rejects whenever $f|_A$ and $g|_A$ disagree on exactly one point and this will imply a lower bound on $\text{Rej}_\ell(f)$.

Let $A$ be specified by $\alpha_0, \ldots, \alpha_{\ell'} \in \mathbb{F}_q^n$ such that $A = \{A(\theta) \triangleq \alpha_0 + \sum_{i=1}^{\ell'} \theta_i \alpha_i \mid \theta = (\theta_1, \ldots, \theta_{\ell'}) \in \mathbb{F}_{q^{\ell'}}\}$. Fix $\theta \in \mathbb{F}_{q^{\ell'}}$, and let $X(\theta)$ denote the random variable that is 1 if $f(A(\theta)) \neq g(A(\theta))$ and 0 otherwise, where $A$ is a uniform $\ell'$-dimensional affine subspace. We note that for every $\theta \in \mathbb{F}_{q^{\ell'}}$, we have $\mathbb{E}_A[X(\theta)] = \delta$. Furthermore, for every pair of distinct points $\theta, \eta \in \mathbb{F}_{q^{\ell'}}$ we have $\mathbb{E}_A[X(\theta)X(\eta)] \leq \delta^2$. (If the points $\alpha_1, \ldots, \alpha_{\ell'}$ were not required to be linearly independent, this expectation would be exactly $\delta^2$. But we insist that they are independent we get that $A(\theta)$ and $A(\eta)$ are two distinct random points in $\mathbb{F}_q^n$ and so the bound above is a (strict) inequality.) Furthermore we have $\delta(f|_A, g|_A) = q^{-\ell'} \sum_\theta X(\theta)$.

Thus we have

$$\text{Pr}[\delta(f|_A, g|_A) = q^{-\ell'}] = \text{Pr}_A[q^{-\ell'} \sum_\theta X(\theta) = q^{-\ell'}] \geq q^{-\ell'} \delta(1 - (q^{-\ell'} - 1)\delta) \geq q^{-\ell'} \delta(1 - q^{-\ell'} \delta).$$

When $\delta \leq \frac{1}{2} q^{-\ell'}$ the bound above implies that $\text{Rej}_\ell(f) \geq \frac{1}{2} q^{-\ell'} \delta$. Else, let $\ell'$ be the largest integer such that $\delta \leq \frac{1}{2} q^{-\ell'}$ (and so $\delta > \frac{1}{2q} q^{-\ell'}$). Note that the assumption that $\delta_\mathcal{F}(f) \leq q^{-(t+1)/2}$ implies that $\ell' \geq t + 1$. We then get $\text{Rej}_\ell(f) \geq \text{Rej}_{\ell'}(f) \geq \frac{1}{2q} q^{-\ell'} \delta > \frac{1}{4q}$. The lemma follows.

Next we bound the rejection probability of the $\ell$-dimensional test in the case in which $f$ is relatively far from $\mathcal{F}$ and $\ell \geq t + c$ for some absolute constant $c$.

**Lemma 4.5.2.** Let $\mathcal{F} = \text{Lift}_n(\mathcal{B})$ for an affine-invariant linear base code $\mathcal{B} \subseteq \{\mathbb{F}_q^t \to \mathbb{F}_{q^t}\}$. Then for every $q$ there exist $\epsilon > 0$ and $c < \infty$ such that if $n \geq \ell \geq t + c$ we have the following: For every $f : \mathbb{F}_q^n \to \mathbb{F}_q$ with $\delta_\mathcal{F}(f) \geq q^{-\ell}$ we have $\text{Rej}_\ell(f) \geq \epsilon + \frac{1}{16q} q^\ell \sum_{i=n+1}^{\infty} q^{-i}$. 

79
Proof. The proof is identical to the proof of Lemma 3.5.2. Let $c = \max\{\tau + 3, 9\}$ where $\tau$ is the constant from Theorem 4.2.1.

We prove the lemma by induction on $n$. The base case $n = \ell$ is straightforward since in this case $\text{Rej}_l(f) = 1$ and $1/16 q^\ell \sum_{i=1}^{\infty} q^{-i} \leq 1/16 q^{\ell-1} \leq 1/16$ and so this case holds for every $\epsilon \leq 1/16$.

For the inductive step, let $H_1, \ldots, H_k$ be all hyperplanes which satisfy $\delta_{L_{n-1}}(f[H_i]) \leq q^{-\ell}$. If $k < 1/16 q^\ell$, then we are done by induction since $\text{Rej}_l(f) = \mathbb{E}_H[\text{Rej}_l(f[H])] \geq \epsilon + 1/16 q^\ell \sum_{i=1}^{\infty} q^{-i} - k/q^n \geq \epsilon + 1/16 q^\ell \sum_{i=1}^{\infty} q^{-i}$ as desired.

Finally we are left with the case where $k \geq 1/16 q^\ell$. In this case we use Theorem 4.2.1 to show that $\delta_f(f)$ is small and then use Lemma 4.5.1 to show that $\text{Rej}_l(f)$ is large. Specifically, by Theorem 4.2.1 we have $\delta_f(f) \leq 2q^{-\ell} + 4(q - 1)/k \leq (2 + 64q) \cdot q^{-\ell} \leq (66q) \cdot q^{-\ell}$. Since $\ell \geq t + c$ and $c \geq 9$ we have $\delta_f(f) < q^{-\ell}/2$ and so by Lemma 4.5.1 we have $\text{Rej}_l(f) \geq \min\{1/4q, q^\ell \delta_f(f)/2\} \geq 1/4q \geq \epsilon + 1/4q \geq 1/32q$. So the lemma is true for $\epsilon = 1/32q$.

As noted above, Lemmas 4.5.1 and 4.5.2 suffice to analyze the rejection probability of a sufficiently high dimensional test ($\ell = t + c$), but not the $t$-dimensional test. To relate the two we use a lemma similar to Lemma 3.4.7. We note that the proof again gets new complications since our result is more general.

**Lemma 4.5.3.** Let $F = \text{Lift}_n(B)$ for an affine-invariant linear code $B \subseteq \mathbb{F}_q^n \rightarrow \mathbb{F}_q$. Let $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$ be a function such that $f \notin F$. Then $\text{Pr}_H[f[H] \notin F] \geq 1/32q$ where the probability is over a hyperplane $H$ chosen uniformly in $\mathbb{F}_q^n$.

**Proof.** Since $f \notin F$ there exists an affine transformation $T \in \text{Aff}_{n \times n}$ such that $f \circ T|_{x_1 + \ldots + x_n = 0}$ is not in $B$. To simplify notation, assume $T$ is the identity. We now bound the number of hyperplanes $H$ such that $f[H] \notin F$. Each hyperplane can be written as $H_\alpha = \{x \in \mathbb{F}_q^n \mid x_c = \sum_{i=c+1}^n \alpha_i x_i + \alpha_0\}$ for $\alpha = (\alpha_0, \ldots, \alpha_n) \in \mathbb{F}_q^{n+1}$, where $c \geq 1$ is the first coordinate such that $\alpha_c \neq 0$. For such a hyperplane define $f_\alpha := f(x_1, \ldots, x_{c-1}, x_c, x_{c+1}, \ldots, x_n)$. Observe that $f[H_\alpha] \in F$ if and only if $f_\alpha \in F$.

We divide into cases according to the value of $c$. For any hyperplane $H_\alpha$ for which $c > t$ we will show the existence of a hyperplane $H_{\alpha'}$, where $\alpha'$ may differ from $\alpha$ only by the 0-th coordinate, such that $f[H_{\alpha'}] \notin F$. Similarly, for any hyperplane $H_\alpha$ for which $1 \leq c \leq t$ we will show the existence of a hyperplane $H_{\alpha'}$, where $\alpha'$ may differ from $\alpha$ only by the $n$-th coordinate, such that $f[H_{\alpha'}] \notin F$. This will prove the claim since it will map at most $q$ different hyperplanes to one ‘good’ hyperplane.

First, consider the case where $c > t$. In this case consider $\alpha'$ such that $\forall i > 0 : \alpha'_i = \alpha_i$ and $\alpha'_0 = 0$. Then $f_{\alpha'}|_{x_t = \ldots = x_n = 0} = f|_{x_1 = \ldots = x_n = 0} \notin B$, hence $f_{\alpha'} \notin F$ which implies in turn that $f[H_{\alpha'}] \notin F$.

Next assume that $1 \leq c \leq t$ and for a variable $z$, let $\alpha(z) \in \mathbb{F}_q^{n+1}$ denote the vector which satisfies $\alpha(z)_i = \alpha_i$ for all $i \neq n$ and $\alpha(z)_n = z$. Our goal will be to show that there exists an assignment $\beta \in \mathbb{F}_q$ to $z$ for which $f_{\alpha(z)} \notin F$. In order to do so we shall show that there exists a monomial $M$ in variables $x_1, \ldots, x_n$ in $\text{supp}(f_{\alpha(z)})$ such that $M \notin F$ and the coefficient of $M$ is a non-zero polynomial in the variable $z$. This will imply in turn that there exists an assignment $\beta \in \mathbb{F}_q$ to $z$ such that $M$ has a non-zero coefficient in $f_{\alpha(z)}$ and consequently $f_{\alpha(z)} \notin F$.

Consider the affine transformation $B \in \text{Aff}_{n \times n}$ which satisfies $\forall i \neq c : B(e_i) = e_i$ and $B(e_c) = e_c + \sum_{i=c+1}^{n} \alpha_i e_i + \alpha_0$ and the affine transformation $B' \in \text{Aff}_{t \times t}$ which satisfies $\forall i \neq c : B'(e_i) = e_i$ and $B'(e_c) = e_c + \sum_{i=c+1}^{t} \alpha_i e_i + \alpha_0$. Observe that

$$f \circ B|_{x_t = \ldots = x_n = 0} = (f|_{x_t = \ldots = x_n = 0}) \circ B' \notin B.$$

Therefore, there exists a monomial $M = \prod_{i=1}^{t} x_i^{\alpha_i}$, containing only the variables $x_1, \ldots, x_t$, that is in $\text{supp}(f \circ B)$ but not in $\text{supp}(B)$. By Claim 4.3.4, we also have that $M \notin \text{supp}(F)$. Note next that
Then another lemma yields the following corollary.

Let \( \alpha \) be an \( \mathbb{F}_q \)-invariant linear code, and suppose that \( \alpha \) has the monomial \( x^{z_{\beta} e} \prod_{i \in [l] \setminus \{c\}} x_i^{a_i} \) is a monomial of \( \alpha \) when viewed as a function of the variables \( \{x_i\}_{i \neq c} \) and \( z \).

Now view \( \alpha \) as a function of \( \{x_i\}_{i \neq c} \) with coefficients that are functions of \( z \). Then the coefficient of the monomial \( \alpha' = x^{z_{\beta} e} \prod_{i \in [l] \setminus \{c\}} x_i^{a_i} \) 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 \( \alpha' \) will be non-zero. In particular, \( \alpha(\beta) \) has the monomial \( \alpha' \) in its support. The proof is completed by noting that \( M \in \text{Aff}_{n \times n}(\alpha') \) and hence the fact that \( M \notin \mathcal{F} \) implies that \( M' \notin \mathcal{F} \). Consequently, \( \alpha(\beta) \notin \mathcal{F} \).

By applying the above lemma iteratively we obtain the following corollary.

**Corollary 4.5.4.** Let \( \mathcal{F} = \text{Lift}_n(\mathcal{B}) \) for an affine-invariant linear code \( \mathcal{B} \subseteq \{\mathbb{F}_q^l \to \mathbb{F}_q\} \) and let \( n \geq \ell \geq k \geq t \). Let \( f : \mathbb{F}_q^n \to \mathbb{F}_q \) be a function such that \( f \notin \mathcal{F} \). Then \( \text{Rej}_k(f) \geq \text{Rej}_t(f) \cdot q^{-(k-t)} \).

**Proof.** The proof is by induction on \( k \). The base case, where \( k = \ell \) is trivial. Now assume the corollary holds for \( k = r + 1 \) and we will prove it for \( k = r \). Consider the following way of choosing a random \( r \)-dimensional affine subspace. First choose a random \( (r + 1) \)-dimensional affine subspace \( V' \subseteq \mathbb{F}_q^r \) and then choose a random \( r \)-dimensional affine subspace \( V \subseteq V' \). Then

\[
\text{Rej}_r(f) = \Pr[f|_{V'} \notin \mathcal{F}] \geq \Pr[f|_{V'} \notin \mathcal{F} | f|_{V'} \notin \mathcal{F}] \Pr[f|_{V'} \notin \mathcal{F}]
\geq \frac{1}{q} \cdot \text{Rej}_t(f) \cdot q^{-(k+1-r)} = \text{Rej}_t(f) \cdot q^{-(k-r)},
\]

where the last inequality is obtained by the induction hypothesis and Lemma 4.5.3.

**Proof of Theorem 4.1.1, \( Q = q \) case.** Follows immediately from Lemmas 4.5.1 and 4.5.2 and Corollary 4.5.4.

### 4.6 Proof of Main Theorem 4.1.1, general \( Q \)

In this section we prove Theorem 4.1.1 for the case in which \( Q \neq q \) via a simple reduction to the \( Q = q \) case. For this we shall need several facts concerning affine-invariant linear codes from \( \mathbb{F}_Q^n \) to \( \mathbb{F}_q \). The following lemma is a generalization of the monomial extraction lemma (Lemma 4.3.1) for such codes.

**Lemma 4.6.1.** [GKS12, Proposition 2.2] Let \( \mathcal{F} \subseteq \{\mathbb{F}_Q^n \to \mathbb{F}_q\} \) be an affine-invariant linear code. Then

\[
\mathcal{F} = \left\{ f : \mathbb{F}_Q^n \to \mathbb{F}_q \mid \text{supp}(f) \subseteq \text{supp}(\mathcal{F}) \right\}.
\]

Let \( Q = q^s \) for an integer \( s \), and let \( \text{Trace} \) denote the trace function from \( \mathbb{F}_Q \) to \( \mathbb{F}_q \) given by \( \text{Trace}(x) = x + x^q + \cdots + x^{q^{s-1}} \). The following lemma says that if a monomial belongs to the support of an affine-invariant linear code then its trace belongs to the code.

**Lemma 4.6.2.** [GKS12, Lemma A.7] Let \( \mathcal{F} \subseteq \{\mathbb{F}_Q^n \to \mathbb{F}_q\} \) be an affine-invariant linear code, and suppose that \( M \in \text{supp}(\mathcal{F}) \). Then \( \text{Trace}(\alpha M) \in \mathcal{F} \) for every \( \alpha \in \mathbb{F}_Q \).

The above lemma yields the following corollary.

**Corollary 4.6.3.** Let \( \mathcal{F} \subseteq \{\mathbb{F}_Q^n \to \mathbb{F}_q\} \) be an affine-invariant linear code, and suppose that \( f \in \mathcal{F} \). Then \( \text{Trace}(\alpha f) \in \mathcal{F} \) for every \( \alpha \in \mathbb{F}_Q \).
Proof. Let \( f = \sum_{M \in \mathcal{M}_n} c_M M \) for \( c_M \in \mathbb{F}_Q \). Lemma 4.6.2 implies that \( \text{Trace}(\alpha c_M M) \in \mathcal{F} \) for every \( M \in \mathcal{M}_n \) and \( \alpha \in \mathbb{F}_Q \). By linearity of the trace function, we have that \( \text{Trace}(\alpha f) = \sum_{M \in \mathcal{M}_n} \text{Trace}(\alpha c_M M) \), and hence \( \text{Trace}(\alpha f) \in \mathcal{F} \) by linearity of \( \mathcal{F} \).

Proof of Theorem 4.1.1, \( Q \neq q \) case. Let \( \mathcal{B}' \) be the \( \mathbb{F}_Q \)-span of \( \mathcal{B} \), i.e.,

\[
\mathcal{B}' = \left\{ \sum_i \alpha_i f_i \mid f_i \in \mathcal{B}, \alpha_i \in \mathbb{F}_Q \right\},
\]

and let \( \mathcal{F}' = \text{Lift}_n(\mathcal{B}') \). Note that \( \mathcal{B}' \) is an affine-invariant linear code from \( \mathbb{F}_Q^n \) to \( \mathbb{F}_Q \).

We start by showing that \( \mathcal{B}' \neq \left\{ \mathbb{F}_Q^n \rightarrow \mathbb{F}_Q \right\} \). To see this note that our assumption that \( \mathcal{B} \neq \left\{ \mathbb{F}_Q^n \rightarrow \mathbb{F}_Q \right\} \), together with Lemma 4.6.1, imply that there exists a monomial \( M \in \mathcal{M}_t \) such that \( M \notin \text{supp}(\mathcal{B}) \). Since \( \mathcal{B} \) and \( \mathcal{B}' \) have the same set of monomials in their support, we have that \( M \notin \text{supp}(\mathcal{B}') \) and hence \( \mathcal{B}' \neq \left\{ \mathbb{F}_Q^n \rightarrow \mathbb{F}_Q \right\} \).

We have just shown that \( \mathcal{B}' \subseteq \left\{ \mathbb{F}_Q^n \rightarrow \mathbb{F}_Q \right\} \) is an affine-invariant linear code. Hence Theorem 4.1.1 for the \( Q = q \) case implies a local tester \( T \) for \( \mathcal{F}' \) which makes \( q^t \) queries, accepts every function \( f \in \mathcal{F}' \) and rejects every function \( f \in \mathbb{F}_Q^n \setminus \mathcal{F}' \) with probability at least \( c_Q \delta_{\mathcal{F}'}(f) \).

We claim that \( T \) is also a local tester for \( \mathcal{F} \) with the same soundness parameter \( c_Q \) and consequently \( \mathcal{F} \) is \( (Q^t, c_Q, Q^{-t}) \)-locally testable. For this it suffices to show that \( \delta_{\mathcal{F}}(f) = \delta_{\mathcal{F}'}(f) \) for every function \( f : \mathbb{F}_Q^n \rightarrow \mathbb{F}_Q \).

Fix \( f : \mathbb{F}_Q^n \rightarrow \mathbb{F}_Q \). Since \( \mathcal{B} \subseteq \mathcal{B}' \), we clearly have that \( \mathcal{F} \subseteq \mathcal{F}' \), and consequently \( \delta_{\mathcal{F}'}(f) \leq \delta_{\mathcal{F}}(f) \). It remains to show that \( \delta_{\mathcal{F}'}(f) \geq \delta_{\mathcal{F}}(f) \). To see this, let \( g \in \mathcal{F}' \) be such that \( \delta_{\mathcal{F}'}(f) = \delta(f,g) \). Then clearly \( \delta(f, \text{Trace}(g)) \leq \delta(f,g) \). Next we show that \( \text{Trace}(g) \in \mathcal{F} \) and so \( \delta_{\mathcal{F}}(f) \geq \delta_{\mathcal{F}'}(f) \).

We need to show that \( \text{Trace}(g)|_A \in \mathcal{B} \) for every \( t \)-dimensional affine subspace \( A \). Let \( A \) be a \( t \)-dimensional affine subspace. Since \( g \in \mathcal{F}' \) we have that \( g|_A \in \mathcal{B}' \), so \( g|_A = \sum_i \alpha_i f_i \) for \( f_i \in \mathcal{B} \) and \( \alpha_i \in \mathbb{F}_Q \). Hence \( \text{Trace}(g)|_A = \text{Trace}(g|_A) = \sum_i \text{Trace}(\alpha_i f_i) \) by linearity of the trace function. By Corollary 4.6.3 we have that \( \text{Trace}(\alpha_i f_i) \in \mathcal{B} \) for all \( i \). Finally, by linearity of \( \mathcal{B} \), this implies that \( \text{Trace}(g)|_A \in \mathcal{B} \) which completes the proof of the Theorem.

\[\square\]

4.7 New testable codes

In this section, we give some examples of codes with “nice” parameters that are testable with absolute soundness based on our main theorem (Theorem 4.1.1).

The need for such codes is motivated by the work of Barak et al. [BGH+12]. Their work used appropriate Reed-Muller codes over the domain \( \mathbb{F}_2^n \). Our work gives the second family of codes that is known to satisfy their requirements. We point out that Guo et al. [GKS13] also give codes motivated by the work of [BGH+12], but their codes are not, thus far, known to be testable with absolute soundness and so fail to meet all the requirements of [BGH+12]. Our codes fall within the class of “lifted” codes studied by [GKS13], but were not analyzed there. Here we use analysis similar to their to analyze the rate and distance of our codes, while the testing follows from our main theorem.

The code. Our codes are defined by three parameters: a real number \( \epsilon > 0 \) and two integers \( s \) and \( n \). The code \( \mathcal{F} = \mathcal{F}_{\epsilon,s,n} \) is obtained as follows: Let \( Q = 2^s \), \( q = 2 \) and \( \ell = \lceil \frac{1}{\epsilon} \log 1/\epsilon \rceil \). Let \( \mathcal{B} = \{ f : \mathbb{F}_Q^n \rightarrow \mathbb{F}_q \mid \sum_{x \in \mathbb{F}_Q^{n-\ell}} f(x) = 0 \} \). Let \( \mathcal{F} = \text{Lift}_n(\mathcal{B}) \).
Basic parameters:

**Proposition 4.7.1.** For every $\epsilon, s$ and $n$ the code $\mathcal{F} = \mathcal{F}_{\epsilon,s,n}$ has block length $N = 2^sn$, (absolute, non-normalized) distance at least $1/\epsilon$ and dimension at least $2^sn - \left(\binom{n}{\ell}^s + \sum_{i=0}^{sl-1} \binom{ns}{i}\right)$.

**Proof.** The size of the block length can be easily verified and the distance follows from Proposition 4.3.5 and Lemmas 3.11 and 3.12 in Guo et. al. [GKS13] analyzed the dimension of the code $\mathcal{F}_{\epsilon,s,n}$ for the case in which $s = \log(1/\epsilon)$ (so $\ell = 1$). More specifically, given a degree pattern $a = (a_1, \ldots, a_n)$ with $\{a_i\}_{i=1}^n \subseteq \{0, \ldots, Q-1\}$, let $a_i^{(j)}$ denote the $j$-th bit of the binary expansion of $a_i$. Let $M(a)$ denote the $n \times s$ matrix with entries $M(a)_{i,j} = a_i^{(j)}$. Guo et. al. show that in the special case in which $\ell = 1$ the code $\mathcal{F}_{\epsilon,s,n}$ contains in its support all monomials with degree pattern $a = (a_1, \ldots, a_n)$ such that there exists a column in $M(a)$ with at least two zeroes. This readily implies a bound of $2^sn - (n+1)^\ell$ on the dimension of their code.

A similar analysis shows that our code $\mathcal{F}_{\epsilon,s,n}$ contains all monomials with degree pattern $a = (a_1, \ldots, a_n)$ where the matrix $M(a)$ has at least $sl + 1$ zeroes, or the matrix has $sl$ zeroes and there exists a column in $M(a)$ with at least $\ell + 1$ zeroes. The lower bound on the dimension follows. □

**Testability:** The following is an immediate application of Theorem 4.1.1.

**Proposition 4.7.2.** For every $s$ there exists a constant $\tau > 0$ such that for every $\epsilon$ and $n$ the code $\mathcal{F} = \mathcal{F}_{\epsilon,s,n}$ is testable by a test that makes $\epsilon N$ queries, accepts codewords with probability one, while rejecting all functions $f : \mathbb{F}_Q^n \to \mathbb{F}_q$ with probability at least $\tau \cdot \delta_{\mathcal{F}}(f)$.

We remark that the dimension of our codes, for any choice of $N$ and $\epsilon$ is strictly better than that of the codes used in [BGH+12] which have dimension $2^sn - \sum_{i=0}^{sl} \binom{ns}{i} \approx 2^sn - \frac{1}{\sqrt{2\pi}st} (en/\ell)^{st}$. An important parameter for them is the “co-dimension” of their code (block length minus the dimension, or the dimension of the dual code), which thus turns out to be roughly $\frac{1}{\sqrt{2\pi}st} (en/\ell)^{st}$ from the above expression. (A smaller codimension is better for their application.) Simplifying the dimension of our code from Proposition 4.7.1, we see that the codimension of our code is smaller by a multiplicative factor of roughly $O((s/2^{-1})$, making our codes noticeably better. Unfortunately such changes do not alter the essential relationship between $N = 2^sn$, the parameter $\epsilon$ (which determines the locality of the tester) and the codimension of the code. The following theorem summarizes the performance of our codes.

**Theorem 4.7.3.** For every positive $s$ there exists a constant $\tau$ such that for every sufficiently small $\epsilon$ and sufficiently large $N$ there exists a binary code of block length $N$, codimension $(\log \frac{1}{\epsilon})^{-s} \cdot \left(\frac{e \log N}{\log \frac{1}{\epsilon}}\right)^{\log \frac{1}{\epsilon}}$ that is testable with a tester that makes $\epsilon \cdot N$ queries accepting codewords with probability one, while rejecting words at distance $\delta$ from the code with probability at least $\tau \cdot \delta$.

To contrast, the corresponding result in [BGH+12] would assert the existence of a positive constant $s$ for which the above held.
Chapter 5

On $r$-Simple $k$-Path

5.1 Introduction

Let $G$ be a directed graph on $n$ vertices. A path $\rho$ is called simple if all the vertices in the path are distinct. The SIMPLE $k$-PATH problem, given $G$ as input, asks whether there exists a simple path in $G$ of length $k$. This is a generalization of the well known HAMILTONIAN-PATH problem that asks whether there is a simple path passing through all vertices, i.e., a simple path of length $n$ in $G$. As HAMILTONIAN-PATH is NP-complete, we do not expect to find polynomial time algorithms for SIMPLE $k$-PATH for general $k$. Moreover, we do not even expect to find good approximation algorithms for the corresponding optimization problem: the longest path problem, where we ask what is the length of the longest simple path in $G$. This is because Björklund et al. [BHK04] showed that the longest path problem cannot be approximated in polynomial time to within a multiplicative factor of $n^{1-\epsilon}$, for any constant $\epsilon > 0$, unless P=NP. This motivated finding algorithms for SIMPLE $k$-PATH with running time whose dependence on $k$ is as small as possible. The first result in this venue by Monien [Mon85] achieved a running time of $k! \cdot \text{poly}(n)$. Since then, there has been extensive research on constructing algorithms for SIMPLE $k$-PATH running in time $f(k) \cdot \text{poly}(n)$, for a function $f(k)$ as small as possible [Bod93, AYZ08, KMRR06, CLSZ07, Kou08]. The current state of the art is $2^k \cdot \text{poly}(n)$ by Williams [Wil09] for directed graphs and $O(1.657^k) \cdot \text{poly}(n)$ by Björklund [BHKK10] for undirected graphs.

5.1.1 Our results

In this paper we look at a further generalization of SIMPLE $k$-PATH which we call $r$-SIMPLE $k$-PATH. In this problem instead of insisting on $\rho$ being a simple path, we allow $\rho$ to visit any vertex a fixed number of times. We now formally define the problem $r$-SIMPLE $k$-PATH.

Definition 5.1.1. Fix integers $r \leq k$. Let $G$ be a directed graph.

- We say a path $\rho$ in $G$ is $r$-simple, if each vertex of $G$ appears in $\rho$ at most $r$ times. Obviously, $\rho$ is a simple path if and only if it is a 1-simple path.

- The $r$-SIMPLE $k$-PATH problem, given $G$ as input, asks whether there exists an $r$-simple path in $G$ of length $k$.

At first, one may wonder whether for some fixed $r > 1$, $r$-SIMPLE $k$-PATH always has a polynomial time algorithm. We show this is unlikely by showing that for any $r$, for some $k$ $r$-SIMPLE $k$-PATH is NP-complete. See Theorem 5.3.1 in Section 5.3 for a formal statement and proof of this. Thus, as in
the case of SIMPLE $k$-PATH, one may ask what is the best dependency of the running time on $r$ and $k$ that can be obtained in an algorithm for $r$-SIMPLE $k$-PATH.

Our main result is

**Theorem 5.1.2.** Fix any integers $r, k$ with $2 \leq r \leq k$. There is a randomized algorithm running in time

$$\text{poly}(n) \cdot O\left(\frac{2k}{r} + O(1)\right) = \text{poly}(n) \cdot 2^{O(k \log r / r)}$$

solving $r$-SIMPLE $k$-PATH on a graph with $n$ vertices with one-sided error.

One may ask how far from optimal is the dependency on $k$ and $r$ in Theorem 5.1.2. Theorem 5.3.1 implies that a running time of $\text{poly}(n) \cdot 2^{O(k/r)}$ would give an algorithm with running time $2^{O(n)}$ for HAMILTONIAN-PATH. Moreover, even a running time of $\text{poly}(n) \cdot 2^{c \cdot k/r}$, for a small enough constant $c < 1/2$, would imply a better algorithm for HAMILTONIAN-PATH than those of [Wil09, BHKK10] which are the best currently known. So, in a sense our algorithm is optimal up to an $O(\log r)$ factor. We find closing this $O(\log r)$ gap, e.g. by a better reduction to HAMILTONIAN-PATH, or a better algorithm for $r$-SIMPLE $k$-PATH, to be an interesting open problem.

5.1.2 Finding a path with many distinct vertices

We give more motivation for the $r$-SIMPLE $k$-PATH problem. Suppose we are in a situation where we wish to find a relatively short path passing through many distinct vertices. Note that an $r$-simple path of length $k$ must pass through at least $k / r$ distinct vertices. Thus, in case, for example, a 2-simple path of length $k$ exists, our algorithm for 2-SIMPLE $k$-PATH can be used to find a path of length $k$ with at least $k / 2$ distinct vertices in time $\text{poly}(n) \cdot 2^{k/2}$. One may ask how this would compare to the number of distinct vertices returned by the algorithms for SIMPLE $k$-PATH. We show there can be a large gap. Specifically, for any given $k$, we show there is a graph $G$ where all simple paths are of length less than $4 \cdot \log k$, but $G$ contains a 2-simple path of length $k$. See Theorem 5.4.2 for a precise statement.

5.2 Overview of the proof of Theorem 5.1.2

We give an informal sketch of Theorem 5.1.2. We are given a directed graph $G$ on $n$ vertices, and integers $r \leq k$. We wish to decide if $G$ contains an $r$-simple path of length $k$. There are two main stages in our algorithm. The first is to reduce the task to another one concerning multivariate polynomials. This part, described below, is very similar to [AB13].

**Reduction to a question about polynomials** We want to associate our graph $G$ with a certain multivariate polynomial $p_G$. We associate with the $i$'th vertex a variable $x_i$. The monomials of the polynomial will correspond to the paths of length $k$ in $G$. So we have

$$p_G(x) = \sum_{i_1 \rightarrow i_2 \rightarrow \cdots \rightarrow i_k \in G} x_{i_1} \cdots x_{i_k},$$

where $i_1 \rightarrow i_2 \rightarrow \cdots \rightarrow i_k \in G$ means that $i_1, i_2, \cdots, i_k$ is a directed path in $G$. An important issue is over what field $\mathbb{F}$ is $p_G$ defined? A central part of the algorithm is indeed choosing the appropriate field to work over. Another issue is how efficiently $p_G$ can be evaluated? (Note that it potentially contains $n^k$ different monomials.) Williams shows in [Wil09] that using the adjacency matrix of $G$ it
can be computed in poly(n)-time. See Section 5.5. For now, think of \( p_G \) as defined over \( \mathbb{Q} \), i.e., having integer coefficients. It is easy to see that \( G \) contains an \( r \)-simple path of length \( k \) if and only if \( p_G \) contains a monomial such that the individual degrees of all variables are at most \( r \). Let us call such a monomial an \( r \)-monomial. Thus our task is reduced to checking whether a homogenous polynomial of degree \( k \) contains an \( r \)-monomial.

Checking whether \( p_G \) contains an \( r \)-monomial Let us assume in this overview for simplicity that \( p = r + 1 \) is prime. Let us view \( p_G \) as a polynomial over \( \mathbb{F}_p \). One problem with doing this is that if we have \( p \) directed paths of length \( k \) passing through the same vertices in different order, this translates in \( p_G \) to \( p \) copies of the same monomial summing up to 0. To avoid this we need to look at a variant of \( p_G \) that contains auxiliary variables that prevent this cancelation. For details on this issue see [AB13] and Section 5.5. For this overview let us assume this does not happen. Recall that we have the equality \( a^p = a \) for any \( a \in \mathbb{F}_p \). Let us look at a monomial that is not an \( r \)-monomial, say \( x_1^{r+1} \cdot x_2 = x_1^p \cdot x_2 \). The equality mentioned implies this monomial is equivalent as a function from \( \mathbb{F}_p^n \) to \( \mathbb{F}_p \) to the monomial \( x_1 \cdot x_2 \). By the same argument, any monomial that is not an \( r \)-monomial will be ‘equivalent’ to one of smaller degree. More generally, \( p_G \) that is homogenous of degree \( k \) over \( \mathbb{Q} \) will be equivalent to a polynomial of degree smaller than \( k \) as a function from \( \mathbb{F}_p^n \) to \( \mathbb{F}_p \) if and only if it does not contain an \( r \)-monomial. Thus, we have reduced our task to the problem of low-degree testing. In this context, this problem is as follows: Given black-box access to a function \( f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p \) of degree at most \( k \), determine whether it has degree exactly \( k \) or less than \( k \), using few queries to the function. Here, for a function \( f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p \), by its degree we mean the total degree of the lowest-degree polynomial \( p \in \mathbb{F}_p[x_1, \ldots, x_n] \) representing it. In Chapter 3 we gave an optimal solution (in terms of the number of queries) to this problem for prime fields. An important observation is that for the right choice of parameters, this test can be performed in linear time in the number of queries. See Section 5.6 for details. For details on dealing with the case that \( r + 1 \) is not prime, see Section 5.7.

5.3 Definitions and Preliminary Results

In this section we give some definitions and preliminary results that will be used throughout this paper.
Let \( G(V,E) \) be a directed graph where \( V \) is the set of vertices and \( E \subseteq V \times V \) the set of edges. We denote by \( n = |V| \) the number of vertices in the graph and by \( m = |E| \) the number of edges in the graph. A \( k \)-path or a path of length \( k \) is a sequence \( \rho = v_1, \ldots, v_k \) such that \((v_i,v_{i+1})\) is an edge in \( G \) for all \( i = 1, \ldots, k-1 \). A path is a \( k \)-path for some integer \( k > 0 \). A path \( \rho \) is called simple if all the vertices in the path are distinct. We say that a path \( \rho \) in \( G \) is \( r \)-simple, if each vertex of \( G \) appears in \( \rho \) at most \( r \) times. Obviously, a simple path is a 1-simple path.

Given as input a directed graph \( G \) on \( n \) vertices, the \( r \)-SIMPLE \( k \)-PATH problem asks for a given \( G \) whether it contains an \( r \)-simple path of length \( k \). When \( r = 1 \) then the problem is called SIMPLE \( k \)-PATH. The \( r \)-SIMPLE PATH problem asks for a given \( G \) and integer \( k \) whether \( G \) contains an \( r \)-simple \( k \)-path of length \( k \).

We show that \( r \)-SIMPLE PATH is NP-complete.

**Theorem 5.3.1.** For any \( r \) the decision problem \( r \)-SIMPLE PATH is NP-complete.

The proof of the theorem is deferred to Section 5.8. The above result implies

**Corollary 5.3.2.** If \( r \)-SIMPLE \( k \)-PATH can be solved in \( T(r,k,n,m) \) time then HAMILTONIAN-PATH can be solved in \( T(r,2rn - n + 2,2n,m + 2n) \).
In particular, if there is an algorithm for \textit{r-SIMPLE k-PATH} that runs in time \(\text{poly}(n) \cdot 2^{(c/2)(k/r)}\) then there is an algorithm for \textit{HAMILTONIAN-PATH} that runs in time \(\text{poly}(n) \cdot 2^c\).

## 5.4 Gap

In this section we show that the gap between the longest simple path and the longest \(r\)-simple path can be exponentially large even for \(r = 2\).

We first give the following lower bound for the gap

\textbf{Theorem 5.4.1.} If \(G\) contains an \(r\)-simple path of length \(k\) then \(G\) contains a simple path of length \(\lceil \frac{\log k}{\log r} \rceil\).

\textit{Proof.} Let \(t = \lfloor \frac{\log k-1}{\log r} \rfloor\). Let \(\rho\) be an \(r\)-simple path whose first vertex is \(v_0\). We will use \(\rho\) to construct a simple path \(\bar{\rho}\) of length \(\lfloor \frac{\log k}{\log r} \rfloor\). We denote \(\rho_0 = \rho\). As \(v_0\) appears at most \(r\) times in \(\rho_0\), there must be a subpath \(\rho_1\) of \(\rho_0\) of length at least \((k-r)/r\) where \(v_0\) does not appear. Let \(v_1\) be the first vertex of \(\rho_1\). Similarly, for \(1 < i \leq t\), we define the subpath \(\rho_i\) of \(\rho_{i-1}\) to be a subpath of length at least

\[
(k-r-\ldots-r^i)/r^i \geq (k-r^{i+1})/r^i,
\]

where \(v_1, \ldots, v_{i-1}\) do not appear, and define \(v_i\) to be the first vertex of \(\rho_i\). Note that we can always assume there is an edge from \(v_{i-1}\) to \(v_i\) as we can start \(\rho_{i+1}\) just after an appearance of \(v_{i-1}\) in \(\rho_{i-1}\).

Note that for \(1 \leq i \leq t\), such a \(v_i\) as defined indeed exists as \((k-r^{i+1})/r^i \geq 1\) when

\[
k \geq 2 \cdot r^{i+1} \iff i + 1 \leq (\log k - 1)/\log r
\]

Thus, \(v_0 \cdots v_{t-1}\) is a simple path of the desired length. \(\square\)

Before we give the upper bound we give the following definition. A \textit{full \(r\)-tree} is a tree where each vertex has \(r\) children and all the leaves of the tree are in the same level. The root is on level 1.

\textbf{Theorem 5.4.2.} There is a graph \(G\) that contains an \(r\)-simple path of length \(k\) and no simple path of length greater than \(4\log k/\log r\).

\textit{Proof.} We first give the proof for \(r \geq 3\). Consider a full \((r-1)\)-tree of depth \(\lceil \log n/\log(r-1) \rceil\). Remove vertices from the lowest level (leaves) so the number of vertices in the graph is \(n\). Obviously there is an \(r\)-simple path of length \(k \geq n\). Any simple tour in this tree must change level at each step and if it changes from level \(\ell\) to level \(\ell + 1\) it cannot go back in the following step to level \(\ell\). So the longest possible simple path is \(2\lceil \log n/\log(r-1) \rceil - 2 \leq 3.17(\log k/\log r)\).

For \(r = 2\) we take a full binary tree \((2\text{-tree})\) and add an edge between every two children of the same vertex. The 2-simple path starts from the root \(v\), recursively makes a tour in the left tree of \(v\) then moves to the root of the right tree of \(v\) (via the edge that we added) then recursively makes a tour in the right tree of \(v\) and then visit \(v\) again. Obviously this is a 2-simple path of length \(k > n\). A simple tour in this graph can stay in the same level only twice, can move to a higher level or can move to a lower level. Again here if it moves from level \(\ell\) to \(\ell + 1\) it cannot go back in the following step to level \(\ell\). Therefore the longest simple path is of length at most \(4\log n \leq 4\log k\). \(\square\)
5.5 From r-Simple k-Path to Multivariate Polynomial

The purpose of this section is to reduce the question of whether a graph \( G \) contains an \( r \)-simple \( k \)-path, to that of whether a certain multivariate polynomial \( \text{contains an} \ r \text{-monomial} \), as defined below.

**Definition 5.5.1 (r-monomial).** Fix a field \( \mathbb{F} \). Fix a monomial \( M(z) = z_1^{i_1} \cdots z_t^{i_t} \).

- We say \( M \) is an \( r \)-monomial if no variable appears with degree larger than \( r \) in \( M \). That is, for all \( 1 \leq j \leq t, i_j \leq r \).
- Let \( f(z) \) be a multivariate polynomial over \( \mathbb{F} \). We say \( f \) contains an \( r \)-monomial, if there is an \( r \)-monomial \( M(z) \) appearing with a nonzero coefficient \( c \in \mathbb{F} \) in \( f \).

We now describe this reduction. Let \( G(V, E) \) be a directed graph where \( V = \{1, 2, \ldots, n\} \). Let \( A \) be the adjacency matrix and \( B \) be the \( n \times n \) matrix such that \( B_{i,j} = x_i \cdot A_{i,j} \) where \( x_i, i = 1, \ldots, n \) are indeterminates. Let \( 1 \) be the row \( n \)-vector of 1s and \( x = (x_1, \ldots, x_n)^T \). Consider the polynomial \( p_G(x) = 1 \cdot B^{k-1} \cdot x \). It is easy to see

\[
p_G(x) = \sum_{i_1 \rightarrow i_2 \rightarrow \cdots \rightarrow i_k \in G} x_{i_1} \cdots x_{i_k}
\]

where \( i_1 \rightarrow i_2 \rightarrow \cdots \rightarrow i_k \in G \) means that \( i_1, i_2, \ldots, i_k \) is a directed path in \( G \).

Obviously, for field of characteristic zero there is an \( r \)-simple \( k \)-path if and only if \( p_G(x) \) contains an \( r \)-monomial. For other fields the later statement is not true. For example, in undirected graph, \( k = 2 \), and \( r = 1 \) if \( (1,2) \in E \) and the field is of characteristic 2 then the monomial \( x_1x_2 \) occurs twice and will vanish in \( p_G(x) \). We solve the problem as follows.

Let \( B^{(m)} \) be an \( n \times n \) matrices, \( m = 2, \ldots, k \), such that \( B^{(m)}_{i,j} = x_i \cdot y_{m,i} \cdot A_{i,j} \) where \( x_i \) and \( y_{m,i} \) are indeterminates. Let, \( y = (y_1, \ldots, y_k) \) and \( y_m = (y_{m,1}, \ldots, y_{m,n}) \). Let \( x \cdot y = (x_1y_1, \ldots, x_ny_1, \ldots) \).

Consider the polynomial \( P_G(x, y) = 1B^{(k)}B^{(k-1)} \cdots B^{(2)}(x \cdot y) \). It is easy to see that

\[
P_G(x, y) = \sum_{i_1 \rightarrow i_2 \rightarrow \cdots \rightarrow i_k \in G} x_{i_1} \cdots x_{i_k} y_{i_1} \cdots y_{i_k}
\]

Obviously, no two paths have the same monomial in \( P_G \). Note that as \( P_G \) contains only \( \{0,1\} \) coefficients, we can define it over any field \( \mathbb{F} \). It will actually be convenient to view it as a polynomial \( P_G(x) \) whose coefficients are in the field of rational functions \( \mathbb{F}(y) \). Therefore, for any field, there is an \( r \)-simple \( k \)-path if and only if \( P_G(x, y) \) contains an \( r \)-monomial in \( x \). We record this fact in the lemma below.

**Lemma 5.5.2.** Fix any field \( \mathbb{F} \). The graph \( G \) contains an \( r \)-simple \( k \)-path if and only if the polynomial \( P_G \), defined over \( \mathbb{F}(y) \), contains an \( r \)-monomial \( M(x) \).

5.6 Low Degree Tester

In this section we present a tester that determines whether a function \( f : \mathbb{F}_p^n \to \mathbb{F}_p \) of degree at most \( d \) has, in fact, degree less than \( d \). The important point is that the tester will be able to do this using few black-box queries to \( f \). The results of this section essentially follow from Chapter 3. A crucial observation is that for a certain choice of parameters this low-degree test can be performed in linear time in the number of queries.

First, let us say precisely what we mean by the \textit{degree} of a function \( f : \mathbb{F}_p^n \to \mathbb{F}_p \). We define this to be the degree of the lowest degree polynomial \( f' \in \mathbb{F}_p[x] \) that agrees with \( f \) as a function from
$\mathbb{F}_p$ to $\mathbb{F}_p$. It is known from the theory of finite fields that there is a unique such $f'$, and that the individual degrees of all variables in $f'$ are smaller than $p$. Moreover, given any polynomial $g \in \mathbb{F}_p[x]$ agreeing with $f$ as a function from $\mathbb{F}_p^n$ to $\mathbb{F}_p$, $f'$ can be derived from $g$ by replacing, for any $1 \leq i \leq n$, occurrences of $x_i$ with $x_i^p \mod x_i - x_i$ (i.e., $x_i^{(t-1) \mod (p-1)+1} \text{ when } t \neq 0$). We do not prove these basic facts formally here. They essentially follow from the fact that $a^p = a$ for $a \in \mathbb{F}_p$.

This motivates the following definition.

**Definition 5.6.1.** Fix positive integers $n, d$ and a prime $p$. Let $f \in \mathbb{F}_p[x] = \mathbb{F}_p[x_1, \ldots, x_n]$. We define $\deg_p(f)$ to be the degree of the polynomial $f$ when replacing, for $1 \leq i \leq n$, $x_i$ by $(x_i^p \mod x_i - x_i)$. More formally, $\deg_p(f) \triangleq \deg(f')$ where

$$f'(x_1, \ldots, x_n) \triangleq f(x_1, \ldots, x_n) \mod x_1^p - x_1, \ldots, \mod x_n^p - x_n.$$ 

Moreover, for a function $g : V \to \mathbb{F}_p$ where $V \subseteq \mathbb{F}_p^n$ is a subspace of dimension $k$, we define $\deg_p(g) = \min f \deg_p(f)$ where $f \in \mathbb{F}_p[x_1, \ldots, x_n]$ and $f|_V = g$. Here $g$ can be regarded as a function in $\mathbb{F}_p[x_1, \ldots, x_k]$.

We note that this notion of degree is affine invariant, i.e. doesn’t change after affine transformations. In addition it has the property that for any affine subspace $V$, $\deg_p(f|_V) \leq \deg_p(f)$.

We now present the main result of this section.

**Lemma 5.6.2.** There is a randomized algorithm $A$ running in time $\text{poly}(n) \cdot p \left[ \frac{d}{p-1} \right] + 1$ that determines with constant one-sided error whether a function $f$ of degree at most $d$ has degree less than $d$. More precisely, given black-box access to a function $f : \mathbb{F}_p^n \to \mathbb{F}_p$ with $\deg_p(f) \leq d$,

- If $\deg_p(f) = d$, $A$ accepts with probability at least $99/100$.
- If $\deg_p(f) < d$, $A$ rejects with probability one.

The proof of Lemma 5.6.2 is deferred to Section 5.9.

### 5.7 Testing if $P_G$ contains an $r$-monomial

In this section we present a method for testing whether the polynomial $P_G$, described in Section 5.5, contains an $r$-monomial. This is done using the low-degree tester from the previous section.

As stated in Lemma 5.5.2, this is precisely equivalent to whether $G$ contains an $r$-simple $k$-path. Recall we viewed $P_G$ as a polynomial over a field of rational functions $\mathbb{F}_p(y)$. To obtain efficient algorithms, we first reduce the question to checking whether a different polynomial defined over $\mathbb{F}_p$ rather than $\mathbb{F}_p(y)$ contains an $r$-monomial. It is important in the next Lemma that we are able to do this reduction for any $p$, in particular a ‘small’ one.

**Lemma 5.7.1.** Fix any integers $r, k$, with $r \leq k$. Let $p$ be any prime and $t = \lceil \log p 10k \rceil$. Let $G$ be a directed graph on $n$ vertices. Given an adjacency matrix $A_G$ for $G$, we can return in $\text{poly}(n)$-time $\text{poly}(n)$-size circuits computing polynomials $f^1_G, \ldots, f^t_G : \mathbb{F}_p^n \to \mathbb{F}_p$ on inputs in $\mathbb{F}_p^n$ such that

- For $1 \leq i \leq t$, $f^i_G$ is (either the zero polynomial or) homogenous of degree $k$.
- If $G$ contains an $r$-simple $k$-path then with probability at least $9/10$, for some $1 \leq i \leq t$, $f^i_G$ contains an $r$-monomial.
- If $G$ does not contain an $r$-simple $k$-path, for all $1 \leq i \leq t$, $f^i_G$ does not contain an $r$-monomial.
Proof. Note that the discussion in Section 5.5 implies we can compute \( P_G \) in \( \text{poly}(n) \)-time over inputs in \( \mathbb{F}_p^{2^n} \). We choose random \( \mathbf{b} \in \mathbb{F}_p^n \) and let

\[
f_G(x) \triangleq P_G(x, \mathbf{b}).
\]

Suppose \( P_G \), as a polynomial over \( \mathbb{F}(y) \), contains an \( r \)-monomial \( M'(x) \). The coefficient \( c_{M'}(y) \) of \( M' \) in \( P_G \) is a nonzero polynomial of degree \( k \). So, by the Schwartz-Zippel Lemma, \( c_{M'}(\mathbf{b}) = 0 \) with probability at most \( k/p \leq 1/10 \). In the event that \( c_{M'}(\mathbf{b}) \neq 0 \), \( f_G(x) \) is a homogenous polynomial of degree \( k \) in \( \mathbb{F}_p[x] \) containing an \( r \)-monomial. Let us assume from now on, we chose a \( \mathbf{b} \) such that indeed \( a_{M'} \triangleq c_{M'}(\mathbf{b}) \neq 0 \). We now discuss how to end up with polynomials having coefficients in \( \mathbb{F}_p \) rather than \( \mathbb{F}_p^n \).

Let \( T_1, \ldots, T_t : \mathbb{F}_p^n \to \mathbb{F}_p \) be independent \( \mathbb{F}_p \)-linear maps. Suppose \( f_G = \sum_M a_M \cdot M(x) \). For \( 1 \leq i \leq t \), define a polynomial \( f_G^i \in \mathbb{F}_p[x] \) by

\[
f_G^i(x) \triangleq \sum_M T_i(a_M) \cdot M(x).
\]

Note that for all \( 1 \leq i \leq t \), \( f_G^i \) is the zero polynomial or homogenous of degree \( k \). As \( a_{M'} \neq 0 \), for some \( i \), \( T_i(a_{M'}) \neq 0 \). For this \( i \), \( f_G^i \) is homogenous of degree \( k \) and contains an \( r \)-monomial, specifically, the \( r \)-monomial \( a_{M'} \cdot M'(x) \). We claim that for all \( 1 \leq i \leq t \), \( f_G^i \) can be computed by a \( \text{poly}(n) \)-size circuit on inputs \( \mathbf{a} \in \mathbb{F}_p^n \). This is because \( f_G \) and \( T_i \) are efficiently computable, and because for \( \mathbf{a} \in \mathbb{F}_p^n \),

\[
T_i(f_G(\mathbf{a})) = T_i \left( \sum_M a_M \cdot M(\mathbf{a}) \right) = \sum_M T_i(a_M) \cdot M(\mathbf{a}) = f_G^i(\mathbf{a}),
\]

where the second equality is due to the \( \mathbb{F}_p \)-linearity of \( T_i \).

The above lemma implies

**Corollary 5.7.2.** Fix any prime \( p \). Suppose that given black-box access to a polynomial \( g \in \mathbb{F}_p[x] \) that is homogenous of degree \( k \), we can determine in time \( \text{poly}(n) \cdot S \) if it contains an \( r \)-monomial. Then we can also determine in time \( \text{poly}(n) \cdot S \) whether \( P_G \) as a polynomial over \( \mathbb{F}_p(y) \) contains an \( r \)-monomial.

Our reduction to low-degree testing is based on the following simple observation that, for the right \( p \) and for homogenous polynomials, containing an \( r \)-monomial is equivalent to having a certain \( \text{deg}_p \)-degree.

**Lemma 5.7.3.** Suppose \( g \in \mathbb{F}_p[x] \) is a homogenous polynomial of degree \( k \). Suppose \( r = p - 1 \). Then \( \text{deg}_p(g) = k \) if and only if \( g \) contains an \( r \)-monomial.

**Proof.** If \( g \) contains an \( r \)-monomial \( M \) then, as \( r < p \), \( \text{deg}_p(M) = k \), which implies that \( \text{deg}_p(g) = k \). If \( g \) does not contain an \( r \)-monomial, then for every monomial \( M \) in \( g \) there is an \( i \in [n] \) such that the degree of \( x_i \) in \( M \) is at least \( r + 1 = p \). So replacing \( x_i^p \) by \( x_i \) will reduce the degree of \( M \) and therefore \( \text{deg}_p(M) < k \). Since this happens for all monomials of \( g \), \( \text{deg}_p(g) < k \).

We introduce another element on notation that will be convenient in the rest of this section.

**Definition 5.7.4.** Fix integers \( n, d \) and prime \( p \). Let \( f \in \mathbb{F}_p[x] \) be an \( n \)-variate polynomial of degree at most \( d \). We define \( LDT(f, n, d, p) \) to be 1 if \( \text{deg}_p(f) = d \), and 0 otherwise.
Before proceeding, we note that the results of Section 5.6 imply that given \( n, d, p \) and black-box access to \( f \), \( \LDT(f, n, d, p) \) can be computed in time \( \poly(n) \cdot O(p^{d/(p-1)+1}) \). In particular, if given \( a \in \mathbb{F}_p^n \), we can compute \( f(a) \) in \( \poly(n) \)-time, then we can compute \( \LDT(f, n, d, \cdot) \) in time \( \poly(n) \cdot O(p^{d/(p-1)+1}) \).

The following lemma is an easy corollary of Lemma 5.7.3.

**Lemma 5.7.5.** Fix integers \( r, k \) with \( r \leq k \). Suppose \( p = r+1 \) is prime. Let \( g \in \mathbb{F}_p[x] \) be homogenous of degree \( k \) and computable in \( \poly(n) \)-time. There is a randomized algorithm running in time

\[
\poly(n) \cdot O((r + 1)^{\left\lfloor \frac{d}{r} \right\rfloor + 1})
\]

determining whether \( g \) contains an \( r \)-monomial.

**Proof.** The algorithm simply returns \( \LDT(g, n, d = k, p = r + 1) \). The running time follows from the discussion above. The correctness follows from Lemma 5.7.3. \( \square \)

We wish to have a similar result when \( r + 1 \) is not a prime.

**Lemma 5.7.6.** Fix integers \( r, k \) with \( r \leq k \). Let \( p \) be the smallest prime such that \( \frac{p-1}{r} \in \mathbb{Z} \). Let \( g \in \mathbb{F}_p[x] \) be homogenous of degree \( k \) and computable by a \( \poly(n) \)-size circuit. There is a randomized algorithm running in time \( \poly(n) \cdot O(p^{\left\lfloor \frac{d}{r} \right\rfloor + 1}) \) determining whether \( g \) contains an \( r \)-monomial.

**Proof.** Denote \( l \triangleq \frac{p-1}{r} \) and define

\[
h(x_1, x_2, \ldots , x_n) \triangleq g(x_1^l, x_2^l, \ldots , x_n^l).
\]

The algorithm returns \( \LDT(h, n, d = k \cdot l, p) \).

Note that \( h \) is homogenous of degree \( k \cdot l \). Note also that \( h \) contains an \( r \cdot l \)-monomial if and only if \( g \) contains an \( r \)-monomial. As \( r \cdot l + 1 = p \) correctness now follows from Lemma 5.7.3. \( \square \)

The best known bound for the smallest prime number \( p \) that satisfies \( r | p - 1 \) is \( r^{\frac{5}{2}} \) due to Heath-Brown [Rib96]. This gives a randomized algorithm running in time

\[
\poly(n) \cdot O(r^{\frac{5}{2} \cdot k + O(1)})
\]

Schinzel, Sierpinski, and Kanold have conjectured the value to be 2 [Rib96]. In the following Theorem we give a better bound. We first give the following

**Lemma 5.7.7.** Fix integers \( r, k \) with \( r \leq k \). Let \( p \) be the smallest prime such that there is an \( l \in \mathbb{Z} \) for which \( r \cdot l \leq p - 1 \) and \( (r + 1) \cdot l > p - 1 \). Let \( g \in \mathbb{F}_p[x] \) be homogenous of degree \( k \) and computable by a \( \poly(n) \)-size circuit. There is a randomized algorithm running in time

\[
\poly(n) \cdot O(p^{\left\lfloor \frac{d}{r} \right\rfloor + 1})
\]

determining whether \( g \) contains an \( r \)-monomial.

**Proof.** As in the proof of Lemma 5.7.6, we define \( h(x_1, x_2, \ldots , x_n) \triangleq g(x_1^l, x_2^l, \ldots , x_n^l) \). The algorithm returns \( \LDT(h, n, d = k \cdot l, p) \). As in the proof of Lemma 5.7.6, \( h \) is homogenous of degree \( k \cdot l \) and contains an \( (r \cdot l) \)-monomial if and only if \( g \) contains an \( r \)-monomial. Furthermore, as \( r \cdot l \leq p - 1 \) and \( (r + 1) \cdot l \geq p \), \( h \) contains a \( (p - 1) \)-monomial if and only if \( g \) contains an \( r \)-monomial. Correctness now follows from Lemma 5.7.3. \( \square \)
The main result of this section contains two results. The first is unconditional. The second is true if Cramer’s conjecture is true. Cramer’s conjecture states that the gap between two consecutive primes \( p_{n+1} - p_n = O(\log^2 p_n) \), [Cra36].

**Theorem 5.7.8.** *(Unconditional Result)* Fix any integers \( r, k \) with \( 2 \leq r \leq k \). Let \( g \in \mathbb{F}_p[x] \) be homogenous of degree \( k \) and computable by a \( \text{poly}(n) \)-size circuit. There is a randomized algorithm running in time
\[
\text{poly}(n) \cdot O\left(r^{\frac{2k}{r}+O(1)}\right)
\]
determining whether \( g \) contains an \( r \)-monomial.

*(Conditional Result)* If Cramer’s Conjecture is true then the time complexity of the algorithm is
\[
\text{poly}(n) \cdot O\left(r^{\frac{k}{r}+o(\frac{k}{r})}\right).
\]

**Proof.** We will find \( p \) and \( l \) as required in Lemma 5.7.7. Fix a prime \( p \) such that \( r^2 + r + 1 < p < 2r^2 + 2r \leq 3r^2 \). (This can be done as for any positive integer \( t > 3 \), there is always a prime between \( t \) and \( 2t \).)

Define \( l \triangleq \left\lfloor \frac{p-1}{r} \right\rfloor \). We have
\[
 r \cdot l = r \cdot \left\lfloor \frac{p-1}{r} \right\rfloor \leq p - 1
\]
\[
 (r + 1) \cdot l \geq (r + 1) \cdot \left( \frac{p-1}{r} - 1 \right)
 = (p - 1) + \frac{p-1}{r} - r - 1 > (p - 1)
\]

The first claim now follows from Lemma 5.7.7 and Corollary 5.7.2.

If Cramer’s conjecture is true then there is a constant \( c \) such that for every integer \( x \) there is a prime number in \([x, x + c \log^2(x)]\). Then there is a prime number \( p \) in the interval \([2cr \log^2 r, 2cr \log^2 r + c \log^2(2cr \log^2 r)]\) and we can choose \( l = 2c \log^2 r \). Then the time complexity will be
\[
\text{poly}(n) \cdot O\left(r^{\frac{k}{r}+o(\frac{k}{r})}\right).
\]

\[\square\]

### 5.8 Proof of Theorem 5.3.1

**Proof.** We will reduce deciding HAMILTONIAN-PATH on a graph of \( n \) vertices, to deciding \( r \)-SIMPLE \((2rn - n + 2)\)-PATH on a graph of \( 2 \cdot n \) vertices.

Given an input graph \( G = (V, E) \) to HAMILTONIAN-PATH, we define a new graph \( G' = (V', E') \) as follows. We let \( V' = V \cup \bar{V} \), where \( \bar{V} = \{ \bar{v}_1, \bar{v}_2, ..., \bar{v}_n \} \) and \( E' = E \cup \bar{E} \) where

\[ \bar{E} = \{ (\bar{v}_i, v_i), (v_i, \bar{v}_i) | i \in [n] \}. \]

For \( j \in [n] \), it will be convenient to denote by \( \rho_j \), the path of length \( 2r - 1 \) that begins at \( v_j \), goes back and forth from \( v_j \) to \( \bar{v}_j \) and ends in \( v_j \), i.e., \( \rho_j \triangleq (v_j, \bar{v}_j, \ldots, v_j, \bar{v}_j, v_j) \).

We make 2 observations.
1. If a vertex $\tilde{v}_j \in \tilde{V}$ appears $r$ times in an $r$-simple path $\rho$ then it must be the start or end vertex of $\rho$: To see this, note that by construction of $G'$, if $\tilde{v}_j$ is not the start or end vertex of $\rho$, visiting it $r$ times requires visiting $v_j$ $r + 1$ times.

2. Suppose $\rho$ is an $r$-simple path that begins and ends in a vertex of $V$. If $\rho$ visits a vertex $\tilde{v}_j \in \tilde{V}$ $r - 1$ times, then it must contain $\rho_j$ as a subpath: To see this, note that as $\rho$ does not start in $\tilde{v}_j$, any visit to $\tilde{v}_j$ must have a visit to $v_j$ before and after. The only way this would sum up to at most $r$ visits in $v_j$ is if these visits were continuous. In other words, only if $\rho$ contains $\rho_j$.

We want to show that $G$ contains a Hamiltonian path if and only if $G'$ contains an $r$-simple path of length $2rn - n + 2$. Assume first that $G$ contains a Hamiltonian path $v_{i_1} \cdot v_{i_2} \cdot \cdots v_{i_n}$. Define the path $\rho = \tilde{v}_{i_1} \cdot \rho_{i_1} \cdot \rho_{i_2} \cdot \cdots \rho_{i_n} \cdot \tilde{v}_{i_n}$. It is of length

$$n \cdot (2r - 1) + 2 = 2rn - n + 2,$$

and it is $r$-simple.

Now assume that we have an $r$-simple path $\rho$ in $G'$ of length $2nr - n + 2$. We first claim that $\rho$ must start and end with a vertex from $\tilde{V}$: Otherwise, using the first observation above, $\rho$ contains at most $n + 1$ vertices appearing $r$ times, and thus has length at most

$$(n + 1) \cdot r + (n - 1) \cdot (r - 1) = 2rn - n + 1.$$ 

Let $\rho'$ be the path $\rho$ with the first and last vertex deleted. So $\rho'$ has length $2rn - n$ and begins and ends in a vertex of $V$. Note that by the first observation $\rho'$ visits all vertices of $\tilde{V}$ at most $r - 1$ times. We now claim that for every $j \in [n]$, $\rho'$ must contain $\rho_j$ as a subpath. Otherwise, by the second observation, $\rho'$ visits some vertex of $\tilde{V}$ less than $r - 1$ times. In this case it has length less than $n \cdot r + n \cdot (r - 1) = 2nr - n$. A contradiction. Thus $\rho'$ contains every $\rho_j$ as a subpath. It cannot contain anything else ‘between’ the $\rho_j$’s, as then it would visit some vertex of $V$ more than $r$ times. So

$$\rho' = \rho_{i_1} \cdots \rho_{i_n},$$

for some ordering $i_1, \ldots, i_n$ of $[n]$. It follows that $v_{i_1} \cdots v_{i_n}$ is a Hamiltonian path in $G$. 

\[ \square \]

### 5.9 Proof of Lemma 5.6.2

Before proving Lemma 5.6.2, we state some required preliminary lemmas.

**Lemma 5.9.1.** Suppose we have black-box access to a function $f : \mathbb{F}_p^d \rightarrow \mathbb{F}_p$. Then we can determine in deterministic time $O(p^t)$ whether $\deg_p(f) \geq (p - 1) \cdot t$.

**Proof.** Consider the algorithm that yields a positive answer if and only if $\sum_{a \in \mathbb{F}_p^d} f(a) = 0$ . It is clear that the running time is indeed $O(p^t)$. Let us now show correctness. As in Definition 5.6.1, define

$$f'(x_1, \ldots, x_n) \triangleq f(x_1, \ldots, x_n) \mod x_1^p - x_1, \ldots, \mod x_n^p - x_n,$$

so that $\deg_p(f) = \deg(f')$. We show that

1. The only monomial of degree $\geq t(p - 1)$ in $f'$ is $M_{\max} \triangleq \prod_{i=1}^t x_i^{p-1}$ and
2. the coefficient of $M_{\max}$ in $f'$, is $(-1)^t \cdot \sum_{a \in \mathbb{F}_p^d} f(a)$.
From these two items, it is clear that indeed \( \deg_p(f) = \deg(f') \geq t \cdot (p-1) \) if and only if \( \sum_{a \in \mathbb{F}_p} f(a) \neq 0 \). The first item is obvious, as the individual degrees in \( f' \) are at most \( p-1 \).

For the second item, let us calculate the coefficient of \( M_{\text{max}} \) in \( f' \). For every \( a \in \mathbb{F}_p \), consider the function \( g_a : \mathbb{F}_p^t \rightarrow \mathbb{F}_p \) that is one on \( a \) and zero elsewhere. One can verify that \( g_a(x) = \prod_{i=1}^t \frac{\prod_{\alpha \in \mathbb{F}_p \setminus \{a_i\}} (x_i - \alpha)}{\prod_{\beta \in \mathbb{F}_p \setminus \{0\}} \beta} \). Clearly, the coefficient of \( M_{\text{max}} \) in \( g_a \) is \( (\prod_{\beta \in \mathbb{F}_p \setminus \{0\}} \beta)^{-t} = (-1)^t \).

Note that in \( g_a \), all individual degrees are smaller than \( p \). Hence, \( f' = \sum_{a \in \mathbb{F}_p} f(a) \cdot g_a \) and the coefficient of \( M_{\text{max}} \) in \( f' \) is \( (-1)^t \cdot \sum_{a \in \mathbb{F}_p} f(a) \). \( \square \)

The algorithm for Lemma 5.6.2 checks the degree of the function only on a small subspace. The key for its correctness is to show that when you restrict the function to a subspace (even for \( n-1 \) dimensional subspace) the degree does not decrease with high probability. Here we use Theorem 3.1.5. We give a proof sketch here for completeness.

**Theorem** (Theorem 5.6.2). Let \( \mathbb{F}_p \) be a field of prime size \( p \) and \( f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p \) be a function with \( \deg_p(f) = t(p-1) \). The number of hyperplanes \( H \) such that \( \deg_p(f|_H) < t(p-1) \) is at most \( p^{t+1} \).

**Proof sketch.** We will assume w.l.o.g that \( f \) has the monomial \( \prod_{i=1}^t x_i^{p-1} \). One can show that for any degree \( t(p-1) \) polynomial \( f \) there is linear transformation \( A \) such that \( f(Ax) \) has the monomial \( \prod_{i=1}^t x_i^{p-1} \). So it will be enough to prove the lemma for the suitable transformation of \( f \).

We will assume for simplicity that all the hyperplanes are of the form of \( H_\alpha = \{ x \mid x_1 = \sum_{i=2}^n \alpha_i x_i + \alpha_0 \} \) for some \( \alpha_2, \ldots, \alpha_n \). Indeed, there are few more hyperplanes that does not depend on the first coordinate, but they don’t contribute much to the upper bound.

To prove the lemma we will show that for any of the \( p^t \) possible values for \( \alpha_2, \ldots, \alpha_t, \alpha_0 \) there are \( < p \) possibilities for \( \alpha_{t+1}, \ldots, \alpha_n \) such that \( \deg(f|_{H_\alpha}) < t(p-1) \). Fix \( \alpha_2, \ldots, \alpha_t, \alpha_0 \). For simplicity we assume they are all zero, but the same bound goes for any \( \alpha_2, \ldots, \alpha_t, \alpha_0 \) (one can reduce the general case to the zero case by some affine transformation).

Now consider all the monomials \( M \) in \( f \) with the following properties: (1) \( M \) divides \( \prod_{i=2}^t x_i^{p-1} \) and (2) \( \deg(M) = t(p-1) \). We can write the sum of all those monomials as \( \prod_{i=2}^t x_i^{p-1} g(x_1, x_{t+1}, \ldots, x_n) \).

By definition, \( g \) is homogenous polynomial of degree \( p-1 \). Because \( \prod_{i=1}^t x_i^{p-1} \) is a monomial of \( f \), \( x_1^{p-1} \) is a monomial of \( g \).

Because the function \( f \) does not depend on the variables \( x_2, \ldots, x_t \) (recall, we assumed \( \alpha_2 = \cdots = \alpha_t = 0 \)) the degree of \( f \) can decrease on \( H_\alpha \) only if the degree of \( g \) decrees on \( H_\alpha \). Because \( g \) is homogenous of degree \( p-1 \) and we consider only linear hyperplanes of the form \( x_1 = L(x_{t+1}, \ldots, x_n) \), then \( g|_{x_1=L} \) is still homogenous of degree \( p-1 \), so if the degree \( \deg(g|_{x_1=L}) < p-1 \) then \( g|_{x_1=L} \equiv 0 \).

Now consider \( g \) as an univariate polynomial in \( x_1 \) over the field of rational functions in \( x_{t+1}, \ldots, x_n \). In this view our question is: how many field elements \( L \in \mathbb{F}_p(x_{t+1}, \ldots, x_n) \) are there such that \( g(L) = 0 \). From the fundamental theorem of algebra the answer is \( p-1 \) and we are done. \( \square \)

From Lemma 5.9 we get the following corollary.

**Corollary 5.9.2.** Let \( n > t \) and \( f : \mathbb{F}_p^n \rightarrow \mathbb{F}_p \) be a polynomial such that \( \deg_p(f) = t(p-1) \). Then \( \Pr \left[ \deg(f|_V) = t(p-1) \right] \geq \frac{1}{p-1} \prod_{k=1}^{n-t} (1 - p^{-k}) = \Omega\left(\frac{1}{p} \right) \), where \( V \) is a random \( t \)-dimensional affine subspace.

**Proof.** We proceed by induction on \( n \). Consider first the base case, where \( n = t+1 \). In this case the number of \( t \)-dimensional affine subspaces \( V \subseteq \mathbb{F}_p^{t+1} \) is \( \frac{p^{t+2} - 1}{p-1} > p^{t+1} + p^t \). By Lemma 5.9 on at most \( p^{t+1} \) of them \( \deg(f|_V) < t(p-1) \) so the probability that \( \deg(f|_V) = t(p-1) \) is \( \frac{1}{p-1} \) as required.
Now assume the claim is true for \( n - 1 \), and consider the following way of choosing a random \( t \)-dimensional affine subspace \( V \). First choose a random hyperplane \( H \subseteq \mathbb{F}_p^n \) and then choose a random \( t \)-dimensional affine subspace \( V \subseteq H \). There are more than \( p^n \) hyperplanes \( H \subseteq \mathbb{F}_p^n \), so by Lemma 5.9 the probability that \( \deg_p(f|_H) = t(p - 1) \) is at least \( 1 - p^{t+1-n} \). Moreover, in the event that \( \deg_p(f|_H) = t(p - 1) \), we can apply the induction hypothesis to \( f|_H \). Hence,

\[
\Pr[\deg_p(f|_V) = t(p - 1)] = 
\Pr[\deg_p((f|_H)|_V) = t(p - 1) \mid \deg_p(f|_H) = t(p - 1)] \cdot \Pr[\deg_p(f|_H) = t(p - 1)] 
\leq \frac{1}{p+1} \prod_{k=1}^{(n-1)-t-1} (1 - p^{-k}) \cdot (1 - p^{t+1-n}) = \frac{1}{p+1} \prod_{k=1}^{n-t-1} (1 - p^{-k}).
\]

We are now ready to prove Lemma 5.6.2.

**Proof of Lemma 5.6.2.** Let \( t = \left\lceil \frac{d}{p-1} \right\rceil \). We assume without lost of generality that \( d = t(p - 1) \):

Otherwise, let \( a = t(p - 1) - d \) and consider the function \( f'(x_0, x_1, \ldots, x_n) \triangleq x_0^a \cdot f(x_1, \ldots, x_n) \). It is easily checked that \( \deg_p(f') \leq t(p - 1) \). Also \( \deg_p(f') = t(p - 1) \) if and only if \( \deg_p(f) = d \).

We will present an algorithm for the problem with one sided error probability \( 1 - \Omega \left( \frac{1}{p} \right) \) that runs in time \( \text{poly}(n) \cdot O(p^t) \). By repeating it \( O(p) \) times, we can get down to error probability \( 1/100 \) in running time \( \text{poly}(n) \cdot O(p^{t+1}) \) as required.

Consider the following algorithm. Choose a random \( t \)-dimensional affine subspace \( V \). Accept if and only if \( \deg_p(f|_V) < t(p - 1) \). Assume first that \( \deg_p(f) < t(p - 1) \). Then for any affine subspace \( V \), \( \deg_p(f|_V) \leq \deg_p(f) < t(p - 1) \). On the other hand, if \( \deg_p(f|_V) = t(p - 1) \), Corollary 5.9.2 implies we will accept with probability at least \( \Omega \left( \frac{1}{p} \right) \).

We conclude by analyzing the running time. Choosing \( V \) can be done in \( \text{poly}(n) \)-time. For checking whether \( \deg_p(f|_V) = t(p - 1) \), Lemma 5.9.1 gives running \( O(p^t) \) assuming black-box access to \( f|_V \). Given black-box access to \( f \), we can compute \( f|_V(a) \) for \( a \in \mathbb{F}_p^t \) in \( \text{poly}(n) \)-time. The claimed running time of \( \text{poly}(n) \cdot O(p^t) \) follows. \(\square\)
Chapter 6

Summary

In this thesis we studied problems related to polynomial testing. We opened our work with the semi-norm that is closely related to polynomial testing, called Gowers norm. We gave strong structural results for degree three and four polynomials that have a high such norm. It is a very interesting question whether such a structure exists for higher degree biased polynomials. Green and Tao [GT07] proved such a result when \( \deg(f) < |F_p| \) (with much worse parameters for degrees three and four), so this question is mainly open for small fields. Kaufman and Lovett [KL08] proved analogous results for functions that have large bias. Another interesting question is to improve the parameters in the results of [GT07, KL08]. These results showed that when \( \deg(f) = d \) and \( f \) is biased then \( f = F(g_1, \ldots, g_{c_d}) \), where \( \deg(g_i) < \deg(f) \) for every \( 1 \leq i \leq c_d \). However, the dependence of \( c_d \) on the degree \( d \) and the bias \( \delta \) is quite bad. Basically, \( c_3 = \exp(\text{poly}(1/\delta)) \) and \( c_d \) is a tower of height \( c_{d-1} \). In contrast, our results give that \( c_3 = \log^2(1/\delta) \) and \( c_4 = \text{poly}(1/\delta) \). Thus, it is an intriguing question to find the true dependence of \( c_d \) on \( \delta \). In particular, as far as we know, it may be the case that \( c_d \) is polynomial in \( 1/\delta \) (where the exponent may depend on \( d \)), or even \( \text{poly}(\log(1/\delta)) \).

For the case of degree four polynomials with high \( U^4 \) norm we proved an inverse theorem showing that on many subspaces, of dimension \( \Omega(n) \), \( f \) equals to a degree three polynomial (a different polynomial for each subspace). Such a result seems unlikely to be true for higher degrees. However, it may be the case that if \( \deg(f) = d \) and \( f \) has a high \( U^d \) norm then \( f \) is correlated with a lower degree polynomial on a high dimensional subspace.

Our second result showed that the natural test for lifted codes is absolute. Lifted codes are codes that are defined by their restriction to \( t \)-dimensional subspaces, for some dimension \( t \). The natural test checks whether a given word satisfies the code constraints on such a random \( t \)-dimensional subspace. We showed that the rejection probability of such a test depends only on the field size, \( q \), and not on the parameter \( t \). The drawback of the result is the bad dependency of the rejection probability on \( q \). Though we currently know the right dependence of the rejection probability of lifted code testing (and polynomial testing) on the block length and on the base code block length, we are far from understanding its dependence on the field size. So it is a very interesting open problem to improve this dependence.

Another motivation for this problem is the work of Guo et al. [GKS13]. They used lifted codes to construct dense locally testable codes which motivates further investigation of the problem of testing lifted codes. Some of their codes require non-constant field size. Unfortunately our results do not hold in this setting and one needs to improve the dependency of the rejection probability of the test on the field size in order to obtain strong testability results for those dense codes.

The blow up in the parameters comes from Lemma 4.4.18. For simplicity we will describe the lemma in the setting of polynomials. The lemma shows that if on a constant fraction of all hyperplanes, the
function is a degree $d$ polynomial, then there is a single polynomial that agrees with the function on every such hyperplane. In the proof we used the Hales-Jewett theorem which is the reason for the blowup. The statement of the lemma is very natural so it seems likely that one may obtain the lemma without using the Hales-Jewett theorem and without that heavy price.

A natural generalization of the notion of lifted codes are single orbit properties.

**Definition 6.0.3.** The single orbit of a base code $B \subseteq \{F_q^k \to F_q\}$ with respect to the set of points $a_1, \ldots, a_k \in F_q^n$ is the set of all functions $f : F_q^n \to F_q$ that satisfy that $f \circ A|_{\{a_1, \ldots, a_k\}} \in B$ for all affine transformations $A : F_q^n \to F_q^n$.

One can observe that if the set of points $\{a_i\}_{i=1}^k$ forms an $F_q$-subspace, then we get a lifted code. The natural test for single orbit properties checks for a random transformation $A$ that $f \circ A \in B$. Kaufman and Sudan showed in [KS08] that the rejection probability of this test does not depend on $n$. It will be interesting to extend our result and show that this rejection probability does not depend even on $k$, i.e. this test is absolute. Since $k$ queries are clearly necessary (without any extra information on $B$) it will show that this test is asymptotically optimal.

The importance of the above extension follows from the conjecture of Ben Sasson at al. [BHK04] that single orbit properties capture all testable affine invariant linear properties. Assuming this conjecture, the above extension will show that the natural test for every testable affine invariant linear property is absolute. So one may view Theorem 4.1.1 as an intermediate step towards this goal.

We conclude this work with an application to finding $r$-simple paths in graphs. This application uses the polynomial method in the context of testing and algorithms. Roughly speaking, in the polynomial method one associates some carefully chosen polynomial(s) with a given instance of a problem and solves the problem by analyzing the polynomial(s). Surprisingly, in many cases, this method gives short and elegant proofs for combinatorial problems that a priori seem to be difficult and not related to polynomials. Usually the polynomial method is used for showing the existence or non-existence of mathematical objects with certain properties. This is done, usually, by bounding the dimension of a family of polynomials, the number of solutions of a system of polynomial equations or via other algebraic bounds.

In this work we used the polynomial method in the context of algorithms via a reduction to polynomial testing. The connection between polynomial testing and graph algorithms is not fully explored yet. It will be interesting to further explore this connection and find other examples of solutions to algorithmic problems via a reduction to polynomial testing.
Bibliography


[AKK+05] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing Reed-Muller codes. *IEEE Transactions on Information Theory*, 51(11):4032–4039, 2005. 1, 1.1, 1.1.4, 1.2.1, 1.2.2, 2.1.4, 2.1, 3.1, 3.1.1, 4.1.1


[BKS+10] Arnab Bhattacharyya, Swastik Kopparty, Grant Schoenebeck, Madhu Sudan, and David Zuckerman. Optimal testing of reed-muller codes. In *FOCS* [DBL10], pages 488–497. (document), 1, 1.2.2, 3.1, 3.1.1, 3.1.1, 3.1.2, 3.2, 3.2, 3.4.6, 4.1.1


[BV07] Andrej Bogdanov Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. In Proceedings of the 48th Annual FOCS, pages 41–51, 2007. 2.1, 2.3.1, 2.3.4, 2.4, 2.4.3


[Cra36] Harald Cramer. On the order of magnitude of the difference between consecutive prime numbers. In Acta Arithmetica 2, 1936. 5.7


[GKS08b] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. 2-transitivity is insufficient for local testability. In IEEE Conference on Computational Complexity, pages 259–267. IEEE Computer Society, 2008. 4.1, 4.2.3


[GKS12] Alan Guo, Swastik Kopparty, and Madhu Sudan. New affine-invariant codes from lifting. Electronic Colloquium on Computational Complexity (ECCC), 19(149), 2012. 4.3.1, 4.3.3, 4.6.1, 4.6.2


[RS97] Ran Raz and Shmuel Safra. A sub-constant error-probability low-degree test, and a sub-
  constant error-probability pcp characterization of np. In STOC, pages 475–484. ACM, 
  1997. 4.1.1, 4.1.1


[Sam07] Alex Samorodnitsky. Low-degree tests at large distances. In 39th Annual STOC, pages 
  506–515, 2007. 1.2.3, 1.2.3, 2.1, 2.1.1, 2.1.1

[ST06] Alex Samorodnitsky and Luca Trevisan. Gowers uniformity, influence of variables, and 
  pcps. In 38th Annual STOC, pages 11–20, 2006. 2.1

  of Chicago, Department of Computer Science, 2003. 2.2

[Tre09] Luca Trevisan. Additive combinatorics and theoretical computer science. SIGACT News 
  Complexity Column, 63, 2009. 2.1

[TZ08] Terence Tao and T. Ziegler. The inverse conjecture for the gowers norm over finite fields 
  via the correspondence principle. arXiv:0810.5527v1, 2008. 1.2.3, 5, 2.1, 2

[Vio08] Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d. In 
  Proceedings of the 23rd Annual IEEE Conference on Computational Complexity (CCC), 
  pages 124–127, 2008. 2.1

[VW07] Emanuele Viola and Avi Wigderson. Norms, xor lemmas, and lower bounds for gf(2) 
  polynomials and multiparty protocols. In 22nd Annual CCC, pages 141–154, 2007. 2.1

  318, 2009. 5.1, 5.1.1, 5.2

[Wol09] Julia Wolf. An inverse theorem for F_{2}^{n}. In preparation, 2009. 1.2.3, 2.1.1

101
\textbf{Theorem 1.1 (Multipartite Shannon Capacity)}

Let $G = (V, E)$ be a graph with $|V| = n$ vertices and $|E| = m$ edges. The capacity $C(G)$ of the graph $G$, is defined as

$$C(G) = \max_{S \subseteq V} I(S, \overline{S})$$

where $I(S, \overline{S})$ is the mutual information between the sets $S$ and $\overline{S}$.

1. **Lower Bound**

   $$C(G) \geq \log_2\left(1 + \frac{m}{n^2}\right)$$

2. **Upper Bound**

   $$C(G) \leq \frac{m}{n^2}$$

Equality holds if and only if $G$ is a complete bipartite graph.

---

**Proof**

Let $S$ be an optimal set of size $k$. By the union bound,

$$I(S, \overline{S}) \geq k \log_2\left(1 + \frac{m}{k(n-k)}\right)$$

Using the Cauchy-Schwarz inequality, we have

$$\sum_{i=1}^{k} I(S_i, \overline{S}_i) \geq \left(\sum_{i=1}^{k} \frac{1}{\log_2(1 + \frac{m}{k(n-k)})}\right)^{-1} \sum_{i=1}^{k} \log_2(1 + \frac{m}{k(n-k)})$$

Since $\sum_{i=1}^{k} \frac{1}{\log_2(1 + \frac{m}{k(n-k)})} \leq k$, we get

$$I(S, \overline{S}) \geq \sum_{i=1}^{k} I(S_i, \overline{S}_i) \geq k \log_2\left(1 + \frac{m}{k(n-k)}\right)$$

Equality holds if and only if all the sets $S_i$ are equal. Therefore, we can maximize the mutual information by taking $S_i = S$ for all $i$. This gives

$$I(S, \overline{S}) \geq k \log_2\left(1 + \frac{m}{k(n-k)}\right)$$

Equality holds if and only if $G$ is a complete bipartite graph.

---

**Corollary**

Let $G$ be a graph with $n$ vertices and $m$ edges. Then the maximum achievable rate $C(G)$ is given by

$$C(G) = \max_{S \subseteq V} I(S, \overline{S})$$

Equality holds if and only if $G$ is a complete bipartite graph.
תק.IsDBNull

בתיהן מפולחנים יונת האת האבעת המקורית בתוכו לתוך התוכן. יוני ביתואג עלית התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן ועונים בתוכן רוחת אוניברסיטת תוכן למקורות התוכן של אוניברסיטת תוכן

מכאן, בלקט של שים וручית, לוגרחה הכותב הולך ו友情链接 רוחת את התוכן בשתי אוניברסיטאות אוניברסיטאות על תואר אוניברסיטת תוכן. מטרות של שדה מורגל, אנשי התוכן של אוניברסיטת תוכן בן

_contacts רוחת את התוכן למקורות התוכן של אוניברסיטת תוכן

 cerco, צו צמחי מח projeto של אוניברסיטת תוכן התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

בגרה, האת האבעת המקורית בתוכו לתוך התוכן. יוני ביתואג עלית התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן ועונים בתוכן רוחת אוניברסיטת תוכן למקורות התוכן של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן (אוניברסיטת תוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן)

F_q \{1,...,t\} \in \mathbb{P}(n,d,q)

לעמנ ולתא את האיבוד ברורה יונת מתודת קצוף סופר סימול. עומר סופר תחילה עוזר את האת האבעת המקורית בתוכו לתוך התוכן. יוני ביתואג עלית התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן ועונים בתוכן רוחת אוניברסיטת תוכן למקורות התוכן של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניברסיטת תוכן

פומרanian יונת התוכן לשכת התוכן של בתים בתוכנה של אוניבסר
אני שמח ואוגה לשמש ולהודות למנהלה, אהיה שפילקה אמיר שפילקה, על הנחיה המסורה, הסבלנות, העידוד והعونה שהוטבחו שנע כיוון השכין.

תודה מיוחדת מועילה ומלמדת, קרה צנזור יילהonga רוז עבי על העזרות והתמיכות הבלתי הפסקיות aloud כל תקופת הדוקטורט.

בנוסף, אני רוצה להודות למעבדי המשרד, חבייה הסגל האקדמי והנומלחי, הפרופסורים המצריים בחברינו ונסים נספים ושעותר לי акדimité, אדרמインタרישיט וوضحתי. לزهر, אני לאcherche להמלות את כל אלה går yüksel מחלים עם מי מוער זה מואר.

ברמה האישית אני רוצה להודות ללהב הסובים ומשפחתי האוהבתי שישוทอดו והמעודים בראות את היישור זכרים, אני רצה להודות לתוך מדיה הכפייה שלה, אני לא כולם מתווך, אבל אני רוצה להודות לתוך זה במדיה comando כל זמן, כל ריס בבאר.

אני מודע לתכון של התמיכת הכספית ההדיבתי בחת NSMutableית.
בחינת פולינומיים וביעות קשורות נוספת

ألעד הרמתי
בוחן פולינומיים ובשיטות קשורות נספורת

טיעון על מחקר

לשמ מירות חלח
שם הדרישות לכתבת התואר
דוקטור לפילוסופיה

על העמד

הוגש לפנים הטכניון
מוכן טכניologi
לישראל
אב תשע"ד
חיפה
אוגוסט 2014

Technion - Computer Science Department - Ph.D. Thesis PHD-2014-13 - 2014