An Improved Line-Point Low-Degree Test

Prahladh Harsha[*]   Mrinal Kumar[*]   Ramprasad Saptharishi[*]   Madhu Sudan[†]

Abstract

We prove that the most natural low-degree test for polynomials over finite fields is "robust" in the high-error regime for linear-sized fields. Specifically we consider the "local" agreement of a function $f \colon \mathbb{F}_q^m \to \mathbb{F}_q$ from the space of degree-d polynomials, i.e., the expected agreement of the function from univariate degree-d polynomials over a randomly chosen line in $\mathbb{F}_q^m$, and prove that if this local agreement is $\varepsilon \geq \Omega((d/q)^\tau))$ for some fixed $\tau > 0$, then there is a global degree-d polynomial $Q \colon \mathbb{F}_q^m \to \mathbb{F}_q$ with agreement nearly ε with f. This settles a long-standing open question in the area of low-degree testing, yielding an $O(d)$-query robust test in the "high-error" regime (i.e., when $\varepsilon < 1/2$). The previous results in this space either required $\varepsilon > 1/2$ (Polishchuk & Spielman, STOC 1994), or $q = \Omega(d^4)$ (Arora & Sudan, Combinatorica 2003), or needed to measure local distance on 2-dimensional "planes" rather than one-dimensional lines leading to $\Omega(d^2)$-query complexity (Raz & Safra, STOC 1997).

Our analysis follows the spirit of most previous analyses in first analyzing the low-variable case ($m = O(1)$) and then "bootstrapping" to general multivariate settings. Our main technical novelty is a new analysis in the bivariate setting that exploits a previously known connection between multivariate factorization and finding (or testing) low-degree polynomials, in a non "black-box" manner. This connection was used roughly in a black-box manner in the work of Arora & Sudan — and we show that opening up this black box and making some delicate choices in the analysis leads to our essentially optimal analysis. A second contribution is a bootstrapping analysis which manages to lift analyses for $m = 2$ directly to analyses for general m, where previous works needed to work with $m = 3$ or $m = 4$ — arguably this bootstrapping is significantly simpler than those in prior works.

Contents

1 Introduction

In this paper we consider the classical "line-point" "low-degree test" for multivariate polynomials over finite fields and give a "near-optimal" analysis in the "high-error" regime. We expand on these terms below.

The "low-degree testing" problem is one of the oldest and classical problems in property testing. The key underlying question is whether the natural "global measure" of distance of a function from the space of low-degree polynomials can be inferred from "local measures". The global measure of distance of functions is simply the normalized Hamming distance. Specifically, given a finite domain D, a set R and a family of function $\mathscr{F} \subseteq \{P \colon D \to R\}$ and a function $f \colon D \to R$, the global measure of interest is the quantity $\delta(f, \mathscr{F}) := \min_{P \in \mathscr{F}}\{\delta(f, P)\}$ where $\delta(f, P)$ denotes the normalized Hamming distance between f and P. For integer t, t-local measures are given by a distributions supported on "small" sets $S \subseteq D$ with $|S| \leq t$, and given by the quantity $\mathbb{E}_S[\delta(f|_S, \mathscr{F}|_S)]$ where $f|_S$ denotes the restriction of f to the domain S and $\mathscr{F}|_S$ denotes the family of restrictions $\{P|_S \colon P \in \mathscr{F}\}$. The broad goal of low-degree testing is to give tight bounds on the relationship between the global and local measures when $\mathscr{F} = \mathbb{F}_q^d[X_1, \ldots, X_m]$, namely the space of m-variate polynomials of degree at most d over $\mathbb{F}_q$, viewed as functions from $\mathbb{F}_q^m$ to $\mathbb{F}_q$.

The relationship between the local and global measures is usually studied under the label of "robustness": A t-query $\alpha(\cdot)$-robust test for $\mathscr{F}$ is a distribution supported on sets $S \subseteq D$ with $|S| \leq t$ such that for every function $f \colon D \to R$ we have $\delta(f, \mathscr{F}) \leq \alpha(\mathbb{E}_S[\delta(f|_S, \mathscr{F}_S)])$. The *error (tolerance)* of a test is roughly the highest value of δ such that $\alpha(\delta)$ is non-trivial, i.e., noticeably bounded away from 1. Once it is shown that $\alpha(\delta)$ is non-trivial, it is usually possible to clean up and show $\alpha(\delta) \approx \delta$. So the critical parameter describing α is just the error-tolerance and this is what we will use to describe the history (and importance) of low-degree testing.

The low-degree testing problem is a classical problem with enormous impact in the early as well as state-of-the-art constructions of probabilistically checkable proofs (PCPs). It was introduced by Rubinfeld and Sudan [RS96] who, in our language, gave an $O(d)$-query test that had an error-tolerance of $O(1/d)$ provided $q = \Omega(d^2)$. The test they introduced is now called the "lines-point" test and is given by the uniform distribution on lines in $\mathbb{F}_q^m$. Arora, Lund, Motwani, Sudan and Szegedy [ALMSS98], building on the work of Arora and Safra [AS98], improved the error-tolerance of the lines-point test to $\Omega(1)$ (provided $q = \Omega(d^3)$). This analysis of the test with $\Omega(1)$ error-tolerance was a critical ingredient in getting PCPs with $O(1)$-query complexity. Subsequent improvements to PCP parameters were also closely related to improvements to (analyses of) low-degree tests. For instance the first constructions of nearly-linear sized PCPs by Polishchuk and Spielman [PS94] rely on getting error $= \Omega(1)$ for $q = O(d)$ and $m = O(1)$ in the lines-point test. This was extended to general m by Friedl and Sudan [FS95] which till this work remains the optimal analysis for the lines-point test when $q = O(d)$.

To get better PCP constructions and to improve parameters in inapproximability results, at-

tention in the late 90s shifted to the "high-error" regime where $\delta \to 1$. Higher error low-degree tests and analyses, with error tending to 1, were obtained by Raz and Safra [RS97] and Arora and Sudan [AS03]. The former introduced the "planes-point test", where the underlying distribution is uniform on "planes" (i.e., 2-dimensional affine subspaces of $\mathbb{F}_q^m$) and showed that it had error $1 - (d/q)^\tau$ for some constant τ. Thus, this obtains essentially optimal error, but at the cost of q^2-queries. In a parallel (and independent) line of work to that of Raz and Safra, Arora and Sudan [AS03] improved the analysis of the line-point test, but only in the regime where q was super-quartic in d and showed roughly that it had error $1 - (d^4/q)^{\Omega(1)}$. So this reduced the query complexity to $O(q)$, but only when $q = \Omega(d^4)$. Thus, the three results above, namely [PS94, FS95], [RS97] and [AS03], are essentially incomparable and represent the the three state-of-the-art low-degree tests today.

We remark there is also a vast body of related questions starting with the work of Alon, Kaufman, Krivelevich, Litsyn and Ron [AKKLR05] that might be termed "moderate degree testing" where one considers the setting $q < d \ll m$. These results and their motivations are quite distinct from those in this work and we do not cover those results here. A direction of study more related to the setting of this paper (i.e., when $d \ll q$) involves derandomizing the low-degree test [BSVW03, MR08]. This direction turns out to be crucial in getting PCPs of small (near-linear) size [BGHSV06, BGHSV05, MR10a, MR10b]. We do not pursue this direction in this paper, though it can be a subject of further study.

The main result of this paper is a single analysis of low-degree testing that qualitatively subsumes all previous low-degree tests. We analyze the lines-point low-degree test and give an analysis showing it has error tolerance $1 - (d/q)^{\Omega(1)}$ when $q \geq C \cdot d$ for a sufficiently large constant C, thereby giving the first q-query high-error analysis of a low-degree test when $d/q = \Omega(1)$ (see Theorem 1.2). This work thus settles a 25-year old open question in the field of PCPs and low-degree testing.

1.1 Technical Contributions

Most analyses of low-degree tests follow the following paradigm: One first analyzes the low-degree test in the setting of $m = O(1)$, i.e., with a constant number of variables. And then a second step of analysis "bootstraps" the result from $O(1)$-variables to general m variables. Our improvement follows the same paradigm and contributes to both steps, though the improvement to the first step is the more significant step. We explain our contribution to the two steps below by contrasting with the previous works.

$\mathbf{m = O(1)}$. Previous techniques in the $O(1)$-variable setting come in two distinct flavors: The Raz-Safra analysis [RS97] is very coding theoretic. For instance when testing 3-variate functions by picking planes uniformly, the analysis relies on the fact that two typical planes intersect in a line, and on this line the nearest polynomial is a codeword of a code with $1 - o(1)$ distance (if $d = o(q)$).

This very high distance of the underlying code is critical to their analysis. The Raz-Safra analysis is thus very clean, but it simply can not work with the lines-point test where lines intersect in at most one point.

The $O(1)$-variable analyses for the line-point tests in [PS94] and in [AS03] are both very algebraic. Both rely on ideas used to decode Reed-Solomon codes, but need some heavy machinery to work with these. [PS94] in particular use properties of the derivatives of resultants to effect their analysis; and [AS03] uses effective Hilbert irreducibility to effect their analysis. While such use of heavy machinery is inevitable given the approach, the weakness in previous results comes from the black box use of the tools that they use. For instance [PS94] effectively only uses the fact that when $m = 2$, the low-degree test effectively gives two directions in which the function looks like a low-degree polynomial in that direction, but is unable to use the fact that $\Omega(1)$-fraction of directions actually have this property. [AS03] suffers from a black box use of Hilbert irreducibility. Our key contribution is to open up this black box and adapt it to our setting as elaborated below.

For $m = 2$, given a function $f : \mathbb{F}_q^2 \to \mathbb{F}_q$ we construct a trivariate polynomial $A(X, Y, Z)$ such that $A(a, b, f(a, b)) = 0$ for a large fraction of $(a, b) \in \mathbb{F}_q^2$. Using the ideas from Reed-Solomon decoding [?] we note that if f has non-trivial agreement with a degree d polynomial P on some line $\ell := \{(t, at+b) : t \in \mathbb{F}_q\}$ the bivariate polynomial $A(T, aT+b, Z)$ has a factor of the form $Z - P_\ell(T)$. Upto this point, except for the choices of the degrees of monomials in A our approach is the same as in [AS03]. At this stage, [AS03] use "effective Hilbert Irreducibility" due to Kaltofen [Kal89], which asserts that a multivariate irreducible polynomital cannot have non-trivial factors on two many 2-dimensional planes, essentially as a black box[1] to conclude that the polynomial $A(X, Y, Z)$ must itself have a factor of the form $Z - P(X, Y)$. And one of the multiple potential factors of this form also can be shown to have non-trivial agreement with f itself. (While this is the overall approach in both works one has to be careful in constructing the A polynomial to ensure its degree is low-enough and a function of d alone and not a function of q. We will skip these details for now.) Using Hilbert Irreducibility as a black box suffers from several defects. In our case we are only interested in "roots" of the polynomial A, i.e., factors of the form $Z - P(\cdots)$, and usually proofs of existence of roots can be simpler (for instance the proof of correctness of the Reed-Solomon decoder of [?] is a special case of Bezout's theorem but much simpler). More importantly they can be optimized by weighing degrees of variables appropriately and this often can lead to better dependence between field size and degrees where the results hold.

The proof of Hilbert Irreducibility uses the machinery of Hensel lifting/Newton's iterations and these machineries were in the past not viewed as amenable to adaptations or variations. In recent years however this view has changed and a number of recent results (see e.g., [?, ?, ?, ?, ?, ?]) in polynomial factorization as well in algebraic "hardness vs. randomness" have successfully opened this black box to derive improved results in these fields. While none of these works directly affect

---

[1]They do open the box to confirm that the special class of the irreducibility result also works for the special class of bivariate planes arising in their substitutions.

our work, emboldened by their success we also open the Hensel lifiting black box. We start with polynomials whose Z degree is weighted heavily as opposed to degrees in variables X, Y, or T. We then apply the lifting machinery over X, Y and T but manage to find a much simpler condition for the lifting to work than in Kaltofen's proof, due to the fact that we only aim to lift a root (and not an arbitrary factor). Our final termination condition is also thus simpler which further supplies a necessary step for our final result. When put together with the ideas in previous works this ends up being sufficient to get a more effective version of irreducibility in this setting – and one that eventually enables us to get the linear dependence between degree and field size. (We note that at the end, our proofs are completely self-contained, and arguably simpler than those in [AS03].)

Bootstrapping. Turning to the "bootstrapping" — here there are roughly three previous works to compare against. The Raz-Safra bootstrapping [RS97] turns out to be the weakest and shows that an error tolerance of $1 - \varepsilon$ in the $O(1)$-variable setting can be converted to an error tolerance of $1 - O(m\varepsilon)$ in the m-variate setting. While this may be adequate in some settings, this is certainly not the right answer. In the low-error setting, Friedl and Sudan [FS95], essentially building on Rubinfeld and Sudan [RS96], show that an error upper bound of ε in the 2-variable setting can be converted to an error bound of ε/C (for some universal constant C) in the m-variate setting; but their proof is very highly tailored to the unique-decoding setting, i.e., the proofs use the fact that in such settings there is at most one polynomial at ε distance from the given function. Arora and Sudan [AS03] extend this analysis to the list-decoding setting but their proof is complex and finally is only able to reduce to the 3-variable case. Our analysis shows that an error-tolerance of $1 - \varepsilon$ in the 2-variable setting implies a $1 - \varepsilon^{\Omega(1)}$ error tolerance in the general m. While our analysis is in the same spirit as the previous works, is substantially cleaner and manages to cleanly reduce to the 2-variable case. We remark that our analysis is inspired by the clean local-list-decoder for multivariate polynomial codes due to Sudan, Trevisan and Vadhan [STV01] who in turn clean up a similar result from [AS03], an idea that does not seem to have been explored much in the interim period.

We remark that both the [FS95] analysis and the [AS03] analysis can be interpreted as exploiting some expansion properties of underlying high-dimensional expanders (HDXs). The [FS95] HDX consists of three layers, the points in $\mathbb{F}_q^m$, the lines in $\mathbb{F}_q^m$ and some complicated 2-dimensional surfaces in $\mathbb{F}_q^m$. The [AS03] HDX is simply the Grasmannian (points, lines, planes and cubes), but now it has four layers. The cleanliness of our analysis is highlighted by the fact that we also work with the (affine) Grassmannian in $\mathbb{F}_q^m$, but now with just three layers (points, lines and planes).

Finally we remark that we do not optimize the dependence between q and the ε in the error-tolerance we obtain here. Recent works due to Bhangale, Dinur and Navon [BDN17] and Minzer and Zheng [MZ23] explore this connection and obtain near-optimal dependence between ε and q, albeit at the cost of an even larger query test – a "cubes-point" test. However, these results do not optimize the dependence on d. These results are proved by using deeper expansion properties

of the 4-layered Grassmannian (points, lines, planes and cubes) in $\mathbb{F}_q^m$. If one obtains an optimal ε-vs-q tradeoff for the lines-point test, with $q = O(d)$ as in our proof, this might yield hardness of labelcover of the following type: NP-hard to distinguish between labelcover instances with alphabet size $O(q^2)$ with perfect completeness from those that are at most $O((\log q)/q)$-satisfible. The current best result, in this context, is due Siu on Chan [Cha16], albeit with imperfect completeness. We believe these questions merit further study and view our work as a step towards understanding them.

1.2 Our results

To state our results which work in the high-error regime of the lines-point test, it will be best to first recall similar results in the low-error regime. We begin with the following version of the low-error lines-point low-degree test (LDT) as used in the original proof of the PCP Theorem and subsequently refined by the works of Polischuk & Spielman [PS94] and Friedl & Sudan [FS95]. Below, we use the notation $P_{m,d}$ to refer to the set of m-variate polynomials in $\mathbb{F}_q[x_1, \ldots, x_m]$ of total degree at most m.

Theorem 1.1 (Multivariate low-error LDT [RS96, AS98, ALMSS98, PS94, FS95]). *There is a large enough constant C such that for every finite field $\mathbb{F}_q$, $m > 0$, degree d satisfying $q > C \cdot d$, the following holds for every $f \colon \mathbb{F}_q^m \to \mathbb{F}_q$ and $\delta \leq 0.01$,*

$$\mathbb{E}_\ell[\delta(f|_\ell, P_{1,d})] \leq \delta \ \text{implies} \ \delta(f, P_{m,d}) \leq 4\delta .$$

Since then, there have been several high-error extensions of the above theorem, notably the works of Arora & Sudan [AS03] and Raz & Safra [RS97]. As mentioned earlier, these works either suffer from a large field-size requirment ($q = O(d^4)$ in [AS03]) or a large query complexity (the use of the q^2-query planes-point test in [RS97]). Our main theorem is the following high-error version of the lines-point test for the natural q-query lines-point test with linear-sized fields, thus obtaining a common strengthening of the results of [AS03, RS97]. While presenting high-error results, it will be convenient to state them in terms of agreement instead of distance. The agreement function is naturally defined as: $\text{agree}(f,g) := 1 - \delta(f,g)$ and $\text{agree}(f, \mathscr{F}) := 1 - \delta(f, \mathscr{F}) = \max_{g \in \mathscr{F}} \text{agree}(f,g)$.

Theorem 1.2 (Multivariate low-degree test). *There exists a constant $\tau \in (0, 1]$ such that for every finite field $\mathbb{F}_q$, $\varepsilon_0 \in (0, 1)$, m, and degree d satisfying $\varepsilon_0 > (d/q)^\tau$, the following holds for every function $f \colon \mathbb{F}_q^m \to \mathbb{F}_q$ and $\varepsilon \in (\varepsilon_0, 1)$:*

$$\mathbb{E}_\ell[\text{agree}(f|_\ell, P_{1,d})] \geq \varepsilon \ \text{implies} \ \text{agree}(f, P_{m,d}) \geq \varepsilon - \varepsilon_0 .$$

1.2.1 Technical Lemma

We now explain the technical ingredients that leads to our main theorem. To this end, let us recall the following lemma (essentially) due to Polishchuk and Spielman [PS94][2], which is the backbone of the proof of Theorem 1.1 that led to linear-sized fields (i.e., $q = O(d)$).

Lemma 1.3 (Polishchuk-Spielman Lemma [PS94, Lemma 8], see also [?, Lemma 4.4]). *Let $\mathbb{F}$ be any field and let $A(x, y, z) \in \mathbb{F}[x, y, z]$ be a non-zero trivariate polynomial with z-degree at most 1 and $(1, 0, d)$-degree and $(0, 1, d)$-degree at most D. Furthermore, suppose there exist two sets $U, V \subseteq \mathbb{F}$ such that (1) for every $u \in U$, there exists a degree-d univariate polynomial $C_u(y)$ such that $A(u, y, C_u(y)) \equiv 0$ and similarly (2) for every $v \in V$, there exists a a degree-d univariate polynomial $R_v(x)$ such that $A(x, v, R_v(x)) \equiv 0$.*

If the sets U, V are of size at least $2D$ each then there exists a polynomial $Q(x, y)$ of individual degree at most d in each variable such that

$$A(x, y, Q(x, y)) \equiv 0 \ .$$

Given this, a natural approach to yield a high-error version of Theorem 1.1 would be to extend the Polishchuk-Spielman Lemma 1.3 to higher degrees in z. However, the natural generalization happens to be false. Consider the trivariate polynomial:

$$A(x, y, z) := (xz - C(y)) \cdot (yz - R(x)),$$

where $C \in \mathbb{F}[y], R \in \mathbb{F}[x]$ are two polynomials of degree at most d. Clearly for each $u \neq 0$, $A(u, y, {}^{C(y)}/u) \equiv 0$ and for each $v \neq 0$, $A(x, v, {}^{R(x)}/v) \equiv 0$. Yet, there exists no $Q(x, y)$ such that $A(x, y, Q(x, y)) \equiv 0$. This counterexample exists as the z-degree of A is at least the number of parallel directions[3]. Our main lemma shows that if the number of parallel directions is considerably more than the z-degree, then such counterexamples do not exist. We find it more convenient to state our lemma for sets of lines passing through a point rather than sets of parallel lines[4].

Lemma 1.4 (Main technical lemma (Informal)). *Let $\mathbb{F}$ be any finite field and $A(x, y, z) \in \mathbb{F}[x, y, z]$ be a non-zero trivariate polynomial with z-degree at most d_z and $(1, 1, d)$-weighted degree at most D such that the characteristic of $\mathbb{F}$ is greater than d_z.*

If there is a set $B \subseteq \mathbb{F}^2$ that satisfies (1) $|B| > 2d_z D|\mathbb{F}|$, and (2) for every $(\alpha, \beta) \in B$, there is a set $S_{(\alpha,\beta)} \subseteq \mathbb{F}^2$ of size greater than $(d_z D|\mathbb{F}|)$ such that for every $(u, v) \in S_{(\alpha,\beta)}$, there exists a

---

[2]This lemma is stated somewhat differently in [PS94] in terms of bivariate functions rather than in terms of the triviariate relation A as we state here, but our version can be easily seen to be impled by theirs.

[3]There exist similar counterexamples for every $r \geq 2$ with z-degree being r and r sets of parallel directions. The above example with $r = 2$ has 2 sets of parallel directions (lines parallel to the x-axis and those parallel to the y-axis).

[4]A set of parallel lines in any direction can be viewed as a set of lines in different directions through a single point at infinity. In this sense, working with lines through a fixed point (a configuration that we refer to as a pencil), and sets of parallel lines are essentially equivalent. For our arguments, the former happens to be a bit more natural.

univariate polynomial $P_{((\alpha,\beta),(u,v))}(t) \in \mathbb{F}[t]$ *of degree at most d for which*

$$A\left(\alpha + tu, \beta + tv, P_{((\alpha,\beta),(u,v))}(t)\right) \equiv 0\,,$$

then, there exists a polynomial $P(x,y) \in \mathbb{F}[x,y]$ *of total degree at most d such that*

$$A(x,y,P(x,y)) \equiv 0\,.$$

Remark. *Even though Lemma 1.4 is stated for trivariate polynomials here, the statement is true for multivariate polynomials as it is. Moreover, even though the underlying field $\mathbb{F}$ is assumed to be finite here and with large enough characteristic, a similar statement that suffices for our applications to low-degree testing is true over all fields. We refer to Section 3 for details.* ◇

Using this lemma, we can prove the following bivariate version of our main Theorem 1.2.

Theorem 1.5 (Bivariate low-degree test)**.** *There exists a constant $\tau \in (0,1]$ such that for every finite field $\mathbb{F}_q$, $\varepsilon_0 \in (0,1)$ and degree d satisfying $\varepsilon_0 > (d/q)^\tau$, the following holds for every function $f\colon \mathbb{F}_q^2 \to \mathbb{F}_q$ and $\varepsilon \in (\varepsilon_0, 1)$:*

$$\mathbb{E}_\ell[\mathrm{agree}(f|_\ell, P_{1,d})] \geq \varepsilon \ \text{implies} \ \mathrm{agree}(f, P_{2,d}) \geq \varepsilon - \varepsilon_0\,.$$

An added advantage of the above theorem is that it also works in the low-error regime. This gives a single proof (in the bivariate setting) that works both for both the low-error and high-error regimes. We can then bootstrap the bivariate low-error theorem using an argument similar to (but simpler than) [FS95] to give an alternate proof of multivariate low-error Theorem 1.1. Finally, we bootstrap the high-error version of the above theorem to yield the multivariate high-error result Theorem 1.2. Thus, our presentation yields a completely self-contained treatment of the lines-point test in both error regimes.

1.3 Proof overview

We give a high-level overview of the main ideas in the proofs of our results. The overall structure of our argument proceeds as follows — (1) analyse the bivariate low-degree test, and (2) *bootstrap* the argument to m-variate low-degree tests.

All of our key algebraic ideas already make an appearance in the proof of the bivariate low-degree test, which is then *lifted* to the multivariate setting using combinatorial techniques, in particular the expansion properties of the points-lines-planes affine Grassmannian.

1.3.1 Bivariate low-degree tests

Suppose are given a 'points table' $f\colon \mathbb{F}_q^2 \to \mathbb{F}_q$ that passes the low-degree test with probability ε. The argument for Theorem 1.5 proceeds as follows:

1. **Find an 'explainer':** We will show that there is a trivariate polynomial $A(x, y, z)$ of $(1, 1, d)$-weighted degree at most $D = O_\varepsilon(d)$ that satisfies $A(a, b, f(a, b)) = 0$ for a structured set S of $\text{poly}(\varepsilon)$ fraction of the points $(a, b) \in \mathbb{F}_q^2$.

2. **Show that A has 'low-degree roots' on many lines:** In this step, we define an appropriate notion of a *good* point $(\alpha, \beta) \in F^2$, and conclude that there is a *good* point such that on *many* lines $\ell_{(\alpha,\beta),(u,v)} := \{(\alpha + tu, \beta + tv) : t \in \mathbb{F}_q\}$ through this point, we have a degree d polynomial $P_{((\alpha,\beta),(u,v))}(t)$ such that $A(\alpha + tu, \beta + tv, P_{((\alpha,\beta),(u,v))}(t)) \equiv 0$.

3. **Show that A must have a global 'low-degree root':** With the hypothesis from the previous step, we use Lemma 1.4 to show that there must be a degree d polynomial $Q(x, y)$ such that $A(x, y, Q(x, y)) \equiv 0$.

4. **Show that f must have non-trivial agreement with $Q(x, y)$:** From the polynomial obtained in the previous step, and the structure of S, we show that f and Q must agree on $\Omega(\varepsilon^4 \cdot q^2)$ points.

Once we have non-trivial agreement as in Item 4, standard reductions in the low-degree testing literature yields the stronger form as stated in Theorem 1.5.

Items 1 and 2 proceed along similar lines as in the results of Arora and Sudan [AS03], with some additional care to ensure that the z-degree of A is function of just ε (and not d). With this additional care, the argument is able to interpolate such an 'explainer' when the density of S is $O(D/|\mathbb{F}|)$ as opposed to $\text{poly}(D)/|\mathbb{F}|$ in [AS03]. The key technical step is Item 3 which uses Lemma 1.4. This lemma, can be thought of as a generalization of the celebrated lemma of Polischuk & Spielman [PS94] in the high error setting and might be of independent interest. Item 4 is almost an immediate consequence of the previous steps.

We now sketch the main ideas in the proof of Lemma 1.4.

Sketch of the proof of Lemma 1.4: The hypothesis of the lemma assumes that there is a *large* set $B \in \mathbb{F}^2$ such that for each $(\alpha, \beta) \in B$, there are *many* lines $\ell_{(\alpha,\beta),(u,v)} = \{(\alpha + tu, \beta + tv) : t \in \mathbb{F}_q\}$ through (α, β) for which there exists a degree d polynomial $P_{((\alpha,\beta),(u,v))}(t)$ satisfying

$$A(\alpha + tu, \beta + tv, P_{((\alpha,\beta),(u,v))}(t)) \equiv 0.$$

As a first step of the proof, we zero in on a specific point (α, β) in this set B. The property desired from this point is that the discriminant of the polynomial $A(x, y, z)$, when viewing A as a univariate in z with coefficients from the field $\mathbb{F}(x, y)$ is non-zero at (α, β). To show that such an $(\alpha, \beta) \in B$ exists, we first show that by being careful in the interpolation step, we can conclude (essentially) without loss of generality, that the discriminant of A with respect to z must be non-zero as a formal polynomial in $\mathbb{F}[x, y]$. Once we have this, we apply the Schwartz-Zippel lemma on this

discriminant to bound its number of zeroes in $\mathbb{F}^2$. At this point, the lower bound on the size of the set B turns out to be large enough for us to conclude that there must be an $(\alpha, \beta) \in B$ on which this discriminant evaluates to non-zero.

The rest of the proof of the lemma exclusively focuses on the behaviour of A on lines through this good point (α, β) and discards the information about other points of B. From the hypothesis, we know that for *many* lines $\ell_{(\alpha,\beta),(u,v)}$ through (α, β), there is a degree d polynomial $P_{((\alpha,\beta),(u,v))}(t)$ such that $A(\alpha + tu, \beta + tv, P_{((\alpha,\beta),(u,v))}(t)) \equiv 0$. Our goal is to somehow stitch these *local roots* of degree d of A into a *global* root $P(x, y)$. We refine the hypothesis further and focus on a large subset of lines through $\ell_{(\alpha,\beta),(u,v)}$ through (α, β) such that the corresponding roots $P_{((\alpha,\beta),(u,v))}(t)$ have the same constant term γ. Our final argument shows that all of these roots must in fact come from a common degree d root $P(x, y)$ of $A(x, y, z)$, which also has the same constant term γ.

The proof of this part is based on general ideas from the literature on polynomial factorization, and in particular the notions of Newton Iteration or Hensel lifting and their properties. When applied to the polynomial $A(\alpha + tu, \beta + tv, z) \in \mathbb{F}[t][z]$, Newton Iteration takes as input a root γ of $A(\alpha, \beta, z)$ that is not a root of the derivative $\frac{\partial A(\alpha,\beta,z)}{\partial z}$ and outputs a power series $\psi_{(u,v)}$ in t that satisfies two properties, the first being that for every $k \in \mathbb{N}$, $A(\alpha + tu, \beta + tv, \psi_{(u,v)}(t)) \equiv 0$ modulo t^k, and the second that the constant term of $\psi_{(u,v)}$ equals γ. Alternatively, this means that for every $k \in \mathbb{N}$, the polynomial $\psi_{(u,v),k}(t) \equiv \psi_{(u,v)}(t) \mod t^k$ is an approximate root of order k of $A(\alpha + tu, \beta + tv, z)$. The uniqueness of this lifting process crucially relies on the fact that the starting point γ is not a root of $\frac{\partial A(\alpha,\beta,z)}{\partial z}$, which in-turn follows from the fact that (α, β) is a non-zero of the discriminant of $A(x, y, z)$ with respect to z; thereby explaining our care in the choice of (α, β). At this point, we invoke this uniqueness of the lifting process to conclude that this power series $\psi_{(u,v)}$ must really be the polynomial $P_{((\alpha,\beta),(u,v))}(t)$ since $P_{((\alpha,\beta),(u,v))}(t)$ is also a lift of the root γ of $A(\alpha + tu, \beta + tv, z)$.

We now repeat this whole exercise, but with the trivariate polynomial $A(\alpha + x, \beta + y, z) \in \mathbb{F}[x, y][z]$. The starting point is that γ as chosen above also satisfies $A(\alpha + x, \beta + y, \gamma) \equiv 0$ modulo the bivariate ideal $\langle x, y \rangle$. And, since γ is not a root of $\frac{\partial A(\alpha,\beta,z)}{\partial z}$, we can again use Newton Iteration, but this time in the ring $\mathbb{F}[x, y]$ to obtain a power series $\Phi(x, y)$ that satisfies that for every $k \in \mathbb{N}$, $A(\alpha + x, \beta + y, \Phi(x, y)) \equiv 0$ modulo $\langle x, y \rangle^k$, and that the constant term of $\Phi(x, y)$ equals γ. Finally, to complete the proof of the lemma, we show that $\Phi(x, y)$ is not just an infinite power series, but an honest-to-god polynomial of degree at most d in $\mathbb{F}[x, y]$ and satisfies $A(\alpha + x, \beta + y, \Phi(x, y)) \equiv 0$. This part relies on relating the two instances of Newton iteration that we described here, the first being for the bivariate polynomial $A(\alpha + tu, \beta + tv, z)$ and the second being for the tri-variate polynomial $A(\alpha + x, \beta + y, z)$ and crucially uses the uniqueness guarantees of Newton Iteration. Using the fact that these relations hold for many lines through (α, β), we complete the proof of the lemma. This final step is again an application of the Schwartz-Zippel lemma for an appropriate polynomial identity instance, and crucially, this instance happens to be simpler and of lower degree than analogous

instances generated by a blackbox application of the polynomial factorization machinery, as was originally done in [AS03].

Finally, we note that the local-root-to-global-root lifting that Lemma 1.4 shows is technically subtle in the sense that our proof does not seem to extend immediately to the case when the lines on which the polynomial has a local root do not form a pencil (i.e. these are lines through the same point (α, β)), even though we expect a qualitatively similar statement to be true. Furthermore, we suspect that many of the quantitative bounds in the lemma, for instance the size of the set B are not optimal and understanding these quantitative issues and variants of the lemma are questions of great interest.

In a broad sense, the above proof opens up the use of Hilbert's Irreducibility Theorem in [AS03], optimizes certain parts and bypasses some bottlenecks there to ensure a better dependence in parameters. A more detailed discussion on the differences is provided below.

Technical differences with the analysis of Arora-Sudan: While our proof in the bivariate case is conceptually similar to that of Arora-Sudan [AS03], there are some differences that lead to better quantitative bounds. The first technical difference is that in the interpolation step of the analysis, we interpolate a trivariate polynomial $A(x, y, z)$ of $(1, 1, d)$-degree D, whereas Arora-Sudan work with polynomials of total degree D. As a consequence, the z-degree of A can be bounded by at most D/d and not just D. In particular, when $D = O(d)$, the z-degree of A happens to be a constant for us, whereas this is not the case in [AS03].

The second technical difference is in the statement of a lemma analogous to Lemma 1.4 that Arora & Sudan prove. In their lemma (Lemma 39 in [AS03]), they construct a polynomial of degree $O(D^3)$ that essentially detects whether for $u, v \in \mathbb{F}$, the restriction $A(ut, vt, z)$ of $A(x, y, z)$ has a factor that is linear in z. More precisely, if $A(x, y, z)$ does not have a factor that is linear in z, then they construct a non-zero polynomial $Q(x, y)$ of degree at most $O(D^3)$ such that if $Q(u, v)$ is non-zero, then $A(ut, vt, z)$ *does not* have a factor that is linear in z. Two points to note here are that the degree of Q is polynomially larger than D, and that this statement talks about factors that are linear in z, but is not sensitive to the degree of the such a factor in x, y (which could be as large as D). As it turns out, the analysis of the bivariate low-degree test only ever cares about factors that are linear in z and have degree at most d in x, y. Being aware of (x, y)-degree of these factors of interest, and additionally considering polynomials of $(1, 1, d)$-degree at most D lets us construct an analog of the polynomial Q above whose degree is just D. This reduction in degree of Q then naturally translates to a reduction in the field size requirement in the overall analysis.

Dealing with fields of small positive characteristic: In general, proofs obtained via derivative based techniques like Taylor expansion or Newton iteration suffer from technical issues when applied over fields of small characteristic. Among examples of such results are the results on polynomial factorization [Kal89, KSS15], results on list decoding of multiplicity codes [GW13, Kop14, Kop15, ?]

and hardness-randomness tradeoffs in algebraic complexity [GKSS22, And20]. Intuitively, the issue stems from the fact over a field of small characteristic, a polynomial can depend on a variable, but its partial derivative with respect to this variable can still be identically zero, e.g. if the polynomial is a function of x^p over a finite field of characteristic p. Our outline also suffers from these issues. However, we observe that these issues can be resolved and that the results extend to all finite fields via one simple additional idea — while interpolating an explainer, ensure that $A(x, y, z)$ has a non-zero partial derivative with respect to z by just ignoring all monomials whose exponent in z is divisible by p. As it turns out, this modification still allows us to proceed with the rest of the argument, up to some small changes in the parameters involved.

1.3.2 Bootstrapping to higher dimensions

Having proved the low-degree test theorem for dimension $m = 2$, we now need to bootstrap it to higher dimensions. There are several possible ways to do this. One potential route is the following. Given the lines-point bivariate LDT (i.e., $m = 2$), we obtain a degree-d planes oracle that passes the planes-point test with non-trivial probability and we can then use the Raz-Safra bootstrapping analysis [RS97] to bootstrap to arbitrary dimensions. This however causes ε_0 (in Theorem 1.2) to be at least $\Omega(m \cdot (d/q)^\tau)$. An alternate route to get around this linear dependence in m is the following route. Bootstrap using the Raz-Safra analysis to dimension $m = 3$ and then use the Bhangale-Dinur-Navon bootstrapping argument [BDN17] to bootstrap to arbitrary dimensions. The [BDN17] analysis (as written in their paper) requires ε_0 to be at least $\Omega((d^8/q)^\tau)$. While this avoids the m-dependence, it requires the field size q to be at least d^8. The [BDN17] analysis can be tightened to yield a $q = O(d)$ dependence by using the Friedl-Sudan analysis [FS95] instead of the Rubinfeld-Sudan analysis [RS96] which they use. To avoid these issues and give a self-contained proof, we give a direct bootstrapping argument from $m = 2$ to arbitrary dimensions (inspired of course by bootstrapping analyses of [FS95, AS03, RS97, BDN17]). However, since we intend to lift from the bivariate LDT (that is, the line-point test in a plane), some effort is required to make the broad ideas of bootstrapping work in our regime. As mentioned earlier in the introduction, this direct bootstrapping is inspired by the clean local-list-decoder for multivariate polynomial codes due to Sudan, Trevisan & Vadhan [STV01].

As a warmup, let us first explain the bootstrapping for the low-error regime, a la Friedl-Sudan [FS95]. In this setting, we are given 'points table' $f : \mathbb{F}_q^m \to \mathbb{F}_q$ and the best degree-d lines oracle $P^{(f,d)}$ that fails the LDT with probability at most δ, we wish to show that there is a degree-d polynomial $Q(x_1, \ldots, x_m)$ that is $O(\delta)$-close to f. Following Friedl-Sudan, we define a corrected function f_{corr} as follows: for any point $\mathbf{y} \in \mathbb{F}_q^m$, $f_{\text{corr}}(\mathbf{y})$ is the most popular value of $P_\ell^{(f,d)}(\mathbf{y})$ among all lines ℓ passing through $\mathbf{y}$ where $P_\ell^{(f,d)}$ is the best-fit degree-d polynomial $P_\ell^{(f,d)}$ agreeing with f on the line ℓ (breaking ties arbitrarily). Let δ_f be the rejection probability of low-degree test when run on the points table f and the best-fit lines oracle $P_\ell^{(f,d)}$ for f. We will show that as

long as $\delta_f \leq \delta_0$ for some fixed constant (dependent on d/q), the rejection probability of the corrected function f_{corr}, namely $\delta_{f_{\text{corr}}}$ is significantly smaller than that of f. More precisely, $\delta_{f_{\text{corr}}} \leq \delta_f/2$. Once we have shown this, we can repeat the self-correction procedure several times to eventually arrive at a corrected function f_* such that $\delta_{f_*} = 0$, in which case it is a low-degree polynomial. So, it suffices to show that the corrected function f_{corr} passes the low-degree test with significantly better probability than the original function f. To prove this, it suffices for us to show that for a random point $\mathbf{x}$, the probability that two random lines ℓ, ℓ' passing through $\mathbf{x}$ satisfy $P_\ell^{(f,d)}(\mathbf{x}) = P_{\ell'}^{(f,d)}(\mathbf{x})$. For any such triple $(\mathbf{x}, \ell, \ell')$, consider the plane π containing ℓ and ℓ'. If the function f restricted to this plane π passes the low-degree with high probability, then the bivariate low-error LDT theorem states that restricted to this plane, the function f, mostly behaves like a degree-d polynomial Q_π and both $P_\ell^{(f,d)}$ and $P_{\ell'}^{(f,d)}$ are in fact restrictions of Q_π and hence equal to each other on the point $\mathbf{x}$. The the above arguments work only on average and not for every triple $(\mathbf{x}, \ell, \ell')$. Nevertheless, we show that this suffices to bootstrap to higher dimensions m and here we use the expansion properties of the points-lines-planes affine Grassmannian. This proof is adapted from (and arguably simpler than) the corresponding bootstrapping proof due to Friedl and Sudan [FS95], who use a different high-dimensional expander consisting of points-lines-surfaces.

We now turn to bootstrapping in the high-error regime. Given a 'points table' $f : \mathbb{F}_q^m \to \mathbb{F}_q$ that passes the LDT with probability ε, we wish to show that there is a degree d polynomial $Q(x_1, \ldots, x_m)$ that agrees with f on $\text{poly}(\varepsilon)$ places (proceeding from a 'weak-agreement' statement to the statement in Theorem 1.2 again follows from standard reductions). As in the low-error case, we would like to define a corrected function such that the corrected function passes the low-degree test with significantly better probability than the original function. The primary issue here is that there could be *many* candidate Q's that have $\text{poly}(\varepsilon)$ agreement and hence several different choices for the corrected function. We use an additional advice (a random point $\mathbf{x}$ and the value of the function f at the point $\mathbf{x}$, namely $f(\mathbf{x})$) to disambiguate among the several different choices. One such corrected function $g_{\text{corr}}^{(\mathbf{x})} : \mathbb{F}^m \to \mathbb{F}$ is as follows. For any $\mathbf{y} \in \mathbb{F}_q^m$, $g_{\text{corr}}^{(\mathbf{x})}(\mathbf{y})$ is the most popular value of $Q_\pi^{(\mathbf{x})}(\mathbf{y})$ among all planes π containing both $\mathbf{x}$ and $\mathbf{y}$ such that $Q_\pi^{(\mathbf{x})}$ is the unique degree-d polynomial that "explains" the plane π and satisfies $Q_\pi^{(\mathbf{x})}(\mathbf{x}) = f(\mathbf{x})$. Why does such a "explaining" polynomial Q_π exist? If the restriction of the function f to the plane π passes the low-degree test with probability ε (which happens if the plane π is random), then the bivariate LDT theorem states that there exists such a polynomial Q_π. This is precisely the bootstrapping argument of Arora and Sudan [AS03]. However, then to show that this corrected function $g_{\text{corr}}^{(\mathbf{x})}$ passes the low-degree test with significantly better probability than f, one needs to consider a cube and this is why the [AS03] bootstrapping argument required as base case both the $m = 2$ and $m = 3$ cases. To get around this dependence on the $m = 3$ case, we define an alternate corrected function $f_{\text{corr}}^{(\mathbf{x})}$ (also disambiguated using the advice $(\mathbf{x}, f(\mathbf{x}))$), that we describe informally here:

Pick a random $\mathbf{x} \in \mathbb{F}_q^m$ to use for constructing the correction. Define $f_{\text{corr}}^{(\mathbf{x})}(\mathbf{x}) := f(\mathbf{x})$.

For every other point $\mathbf{y} \in \mathbb{F}_q^m \setminus \{\mathbf{x}\}$, consider the line $\ell_{\mathbf{x},\mathbf{y}}$ passing through $\mathbf{x}$ and $\mathbf{y}$. Find the list of all polynomials $P(t)$ that agree with f on this line on at least $\text{poly}(\varepsilon)$-fraction of places. Among those, if there is a *unique* P that takes value $f(\mathbf{x})$ on $\mathbf{x}$, set $f_{\text{corr}}^{(\mathbf{x})}(\mathbf{y}) = P(\mathbf{y})$ and set it to $\perp$ (or a random value in $\mathbb{F}_q$) otherwise.

In other words, the line joining $\mathbf{x}, \mathbf{y}$ is being used to 'correct' the value at $\mathbf{y}$ but we are disambiguating between the potential possibilities on this line by the value those polynomials take at the point $\mathbf{x}$.

Having defined this corrected function $f_{\text{corr}}^{(\mathbf{x})}$, it is not hard to show that with non-trivial probability over the choice of the advice point $\mathbf{x}$, we have that the corrected function $f_{\text{corr}}^{(\mathbf{x})}$ has non-trivial $\text{poly}(\varepsilon)$ agreement with the function f.

We now try to understand the behaviour of the corrected function on a random line ℓ. Consider the plane π containing the advice point $\mathbf{x}$ and the line ℓ. Using the bivariate LDT theorem, we can show that for a random plane π a non-trivial fraction of the points $\mathbf{x}$ in π, there exists a degree-d polynomial $Q_\pi^{(\mathbf{x})}$ that has $\text{poly}(\varepsilon)$-agreement with f on π and $Q_\pi^{(\mathbf{x})}(\mathbf{x}) = f(\mathbf{x})$. We use this coupled with the expansion properties of the lines-points-planes affine Grassmannian to show that for a random $\mathbf{x}$ and a random line ℓ, the corrected function $f_{\text{corr}}^{(\mathbf{x})}$ (using advice $(\mathbf{x}, f(\mathbf{x}))$) agrees with the polynomial $Q_\pi^{(\mathbf{x})}$ overwhelmingly on the line where π is the plane containing $\mathbf{x}$ and ℓ. This step requires a careful analysis using spectral properties of various natural subgraphs of the Affine Grassmanian. We thus, have, $\text{Pr}_{\mathbf{x},\ell}[\delta(f_{\text{corr}}^{(\mathbf{x})}|_\ell, Q_\pi^{(\mathbf{x})}|_\ell) \leq \gamma] \geq 1 - 2\gamma$. This implies that for a random $\mathbf{x}$, the corrected function $f_{\text{corr}}^{(\mathbf{x})}$ passes the low-degree test overwhelmingly (in fact with probability $1 - 3\gamma$). Hence, by Theorem 1.1, there exists a degree-d m-variate polynomial $Q^{(x)}$ such that $\delta(f_{\text{corr}}^{(\mathbf{x})}, Q^{(\mathbf{x})}) \leq O(\gamma)$. Combining this with the fact that $f_{\text{corr}}^{(\mathbf{x})}$ has $\text{poly}(\varepsilon)$ agreement with the function f, we obtain that f has $\text{poly}(\varepsilon) - O(\gamma) = \text{poly}(\varepsilon)$ agreement with the polynomial $Q^{(\mathbf{x})}$ provided $\gamma \ll \text{poly}(\varepsilon)$. This completes the overview of the bootstrapping argument.

1.4 Organisation of the paper

We begin with notation and preliminaries in Section 2 and then proof the main technical lemma (Lemma 1.4) in Section 3. We then proceed to the analysis of the bivariate low-degree tests in Section 4. Finally, we bootstrap the bivariate low-degree tests to m-variate low-degree tests in Section 5.

2 Preliminaries

Notation:

1. For a polynomial $f(x_1, \ldots, x_m)$ and a vector $\mathbf{a} = (a_1, \ldots, a_m) \in \mathbb{N}^m$, we use the term **e**-weighted degree to denote the largest value of $a_1 e_1 + \cdots + a_m e_m$ among monomials $x_1^{e_1} \cdots x_m^{e_m}$ in the support of f.

2. For two functions $f, g : \mathbb{F}^m \to \mathbb{F}$, we define $\mathrm{agree}(f, g) = \Pr_{\mathbf{a} \in \mathbb{F}^m}[f(\mathbf{a}) = g(\mathbf{a})]$, the fraction of points that the tables f and g agree on. Similarly, we will use $\delta(f, g)$ to denote $\Pr_{\mathbf{a} \in \mathbb{F}^m}[f(\mathbf{a}) \neq g(\mathbf{a})]$, the fraction of points that the tables f and g disagree on.

Lemma 2.1 (Averaging argument). *Suppose $x_1, \ldots, x_n \in [0, 1]$ such that $\sum x_i \geq \mu n$. Then,*

1. *if $S = \{i \in [n] : x_i \geq \mu/2\}$, then $|S| \geq \mu n/2$, and*

2. *$\sum_{i \in S} x_i \geq \mu n/2$.*

Proof. For Item 1, note that

$$\mu n \leq \sum_i x_i \leq \mu/2 \cdot (n - |S|) + |S| \cdot 1 \leq \mu n/2 + |S|$$
$$\implies \mu n/2 \leq |S|.$$

For Item 2,

$$\sum_{i \in S} x_i \geq \mu n - \sum_{i \notin S} x_i \geq \mu n - (n - |S|) \cdot \mu/2$$
$$\geq \mu n - (1 - \mu/2) \cdot \mu/2 \cdot n \geq \mu n/2. \qquad \square$$

Polynomial identity lemma

Lemma 2.2 ([Ore22, DL78, Sch80, Zip79]). *Let $P(\mathbf{x})$ be a non-zero polynomial of total degree at most D with coefficients over a field $\mathbb{F}$ and let S be an arbitrary subset of $\mathbb{F}$. Then, the number of zeroes of P on the product set $S \times S \times \cdots \times S$ is at most $D|S|^{m-1}$.*

Hasse Derivatives and Properties

Throughout the paper, we use the notion of Hasse derivatives of polynomials and some of their basic properties. We start with the definition.

Definition 2.3. *Let $\mathbb{F}$ be any field and $A(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be an m variate polynomial with coefficients in $\mathbb{F}$. Then, the Hasse derivative of $A(\mathbf{x})$ with respect to a monomial $\mathbf{x}^{\mathbf{e}}$, denoted by $\partial_{\mathbf{x}^{\mathbf{e}}}(A)$, is defined as the coefficient of the monomial $\mathbf{z}^{\mathbf{e}}$ when viewing $A(\mathbf{s} + \mathbf{z})$ as a polynomial in $\mathbf{z}$ variables with coefficients in the ring $\mathbb{F}[\mathbf{x}]$.* $\diamond$

Hasse derivatives are an extremely useful and recurrent theme in computer science with a variety of applications. For our proofs in the paper, we rely on the following properties.

Proposition 2.4. *Let $\mathbb{F}$ be any field. Then, the following are true.*

- *For every $d \in \mathbb{N}$, $\partial_x(x^d) = dx^{d-1}$.*

- *For every pair of polynomials $A, B \in \mathbb{F}[x]$, $\partial(AB) = \partial(A) \cdot B + B \cdot \partial(A)$.*

- *For any polynomial $A \in \mathbb{F}[x]$ of degree at least one in x, and a finite field $\mathbb{F}$ of characteristic p, $\partial_x(A)$ is identically zero if and only if there exists a polynomial $\tilde{A} \in \mathbb{F}[x]$ such that $\tilde{A}^p = A$.*

Discriminant

Definition 2.5 (Resultant and Discriminant). *Let $\mathbb{F}$ be any field and $A, B \in \mathbb{F}[x]$ be non-zero univariate polynomials in x of degree a, b respectively. Let $A_0, A_1, \ldots, A_a, B_0, B_1, \ldots, B_b$ be elements of $\mathbb{F}$ such that $A(x) = \sum_{i=0}^{a} A_i x^i$ and $B(x) = \sum_{j=0}^{b} B_j x^j$. Then, the Sylvester matrix of A and B defined as follows.*

$$\text{Sylvester}(A, B) = \begin{bmatrix} A_0 & A_1 & \ldots & & A_a & & & \\ & \ddots & \ddots & & & \ddots & \ddots & \\ & & A_0 & A_1 & & & \ldots & A_a \\ B_0 & \ldots & & B_b & & & & \\ & B_0 & \ldots & & B_b & & & \\ & & \ddots & \ddots & & & \ddots & \\ & & & B_0 & \ldots & & & B_b \end{bmatrix}$$

And, the resultant of A and B is defined as the determinant of the Sylvester matrix of A and B. Moreover, if $B = \partial_x(A)$, then the resultant of A and B is referred to as the Discriminant of A and B and denoted by $\mathsf{Disc}_x(A, B)$. ◇

Lemma 2.6. *Let $\mathbb{F}$ be any field and $A, B \in \mathbb{F}[x]$ be non-zero univariate polynomials in x. Then, A, B have a non-trivial GCD (the degree of GCD is at least one in x) if and only if their resultant is zero.*

Definition 2.7. *Let $\mathbb{F}$ be any field. A polynomial $A(x) \in \mathbb{F}[x]$ is said to be square-free if there does not exist a polynomial $B \in \mathbb{F}$ of degree at least one such that B^2 divides A.* ◇

Lemma 2.8. *Let $\mathbb{F}$ be any field and $A \in \mathbb{F}[x]$ be a non-zero univariate polynomial such that $\partial_x(A)$ is non-zero. Then, A is square-free if and only if the discriminant of A is non-zero.*

2.1 Structure of minimal interpolating polynomials

In this section, we prove the following simple lemma that will be crucial to our analysis of the bivariate test. The lemma essentially lets us assume some structural properties on the interpolating polynomials without loss of generality, and these properties turn out to be important for our analysis, especially when we are working over fields of small characteristic.

Lemma 2.9. *Let $\mathbb{F}$ be any field and $m \in \mathbb{N}$ be an integer, and let $S \subseteq \mathbb{F}^m$ be a set of points. Consider the space of polynomial $\mathcal{Q}_{S,x_1} = \{Q(x_1, \ldots, x_m) : Q(\mathbf{a}) = 0 \text{ for all } \mathbf{a} \in S \text{ and } \partial_{x_1}(Q) \neq 0\}$. For any vector $\mathbf{w} \in \mathbb{N}^m$, the polynomial $Q \in \mathcal{Q}_{S,x_1}$ of lowest $\mathbf{w}$-weighted degree satisfies $\mathsf{Disc}_{x_1}(Q) \neq 0$ and is hence square-free.*

Proof. Let $Q \in \mathcal{Q}_{S,x_1}$ be the polynomial of smallest $\mathbf{w}$-weighted degree. On the contrary, let us assume that Q is *not* square-free and say $Q = P^\ell \cdot R$ where P is irreducible with $\ell \geq 2$ and $\gcd(P, R) = 1$.

Consider the polynomial $\tilde{Q} = P \cdot R$. Clearly $\tilde{Q}(\mathbf{a}) = 0$ for all $\mathbf{a} \in S$ as well, and if we can argue that $\partial_{x_0}(\tilde{Q}) \neq 0$, we have $\tilde{Q} \in \mathcal{Q}_{S,x_1}$ thus yielding a contradiction to the minimality of Q.

$$\partial_{x_1}(Q) = P^\ell \cdot \partial_{x_1}(R) + \ell P^{\ell-1} \cdot R \cdot \partial_{x_0}(P).$$
$$\partial_{x_1}(\tilde{Q}) = P \cdot \partial_{x_1}(R) + R \cdot \partial_{x_1}(P).$$

Since we have $\partial_{x_1}(Q) \neq 0$, we have that at least one of $\partial_{x_1}(P), \partial_{x_1}(R)$ is non-zero. If exactly one of $\partial_{x_1}(P), \partial_{x_1}(R)$ is non-zero, we immediately have that $\partial_{x_1}(\tilde{Q}) \neq 0$. Otherwise, since P is irreducible and $\gcd(P, R) = 1$, we have that $P \nmid R \cdot \partial_{x_1}(P)$. Therefore, $P \nmid \partial_{x_1}(\tilde{Q})$ and this in particular forces $\partial_{x_1}(Q) \neq 0$ yielding our required contradiction. Thus Q must have been square-free. $\qquad\square$

2.2   Power series roots

The following lemma is a standard application of Newton iteration in multivariate polynomial rings and is an important ingredient of many of the multivariate factorization algorithms. We refer to [Bür00] for a proof.

Lemma 2.10 (Newton iteration for roots)**.** *Let $\mathbb{F}$ be a field and let $A(\mathbf{x}, z) \in \mathbb{F}[\mathbf{x}, z]$ be an $(m+1)$-variate polynomial and let $\alpha \in \mathbb{F}$ be a zero of multiplicity one of $A(\mathbf{0}, z)$, i.e., $A(\mathbf{0}, \alpha) = 0$ and $\partial_z(A(\mathbf{x}, z))(\mathbf{0}, \alpha) = \partial_z(A(\mathbf{0}, z))(\alpha) \neq 0$. Then, for every $k \geq 1$, there is a polynomial $\Phi_k(\mathbf{x})$ such that the following are true.*

- *$A(\mathbf{x}, \Phi_k(\mathbf{x})) \equiv 0 \mod \langle x_1, x_2, \ldots, x_m \rangle^{k+1}$,*

- *$\alpha = \Phi_k(\mathbf{0}) \equiv \Phi_k(\mathbf{x}) \mod \langle x_1, x_2, \ldots, x_m \rangle$,*

- *$\deg(\Phi_k) \leq k$.*

Moreover, there is a unique polynomial that satisfies all these three properties.

3   Low-degree roots on restrictions to global low-degree roots

We prove the following technical lemma (a more formal statement of Lemma 1.4) that is the main technical statement driving the analysis of our bivariate low-degree test. The lemma is true more generally (in higher dimensions) and we state and prove it in that way.

Lemma 3.1. *Let $\mathbb{F}$ be any field and let $A(\mathbf{x}, z) \in \mathbb{F}[\mathbf{x}, z]$ be a non-zero $(m+1)$-variate polynomial with z-degree at most d_z, $(1, 1, \ldots, 1, d)$ degree at most D and let $\mathbf{b} \in \mathbb{F}^m$ be a point such that the univariate polynomial $A(\mathbf{b}, z) \in \mathbb{F}[z]$ has no repeated roots. Let $S \subseteq \mathbb{F}^m$ be a set of directions such that for every $\mathbf{u} \in S$, there exists a univariate polynomial $P_{\mathbf{b},\mathbf{u}}(t) \in \mathbb{F}[t]$ of degree at most d that satisfies the identity*

$$A(\mathbf{b} + t\mathbf{u}, P_{\mathbf{b},\mathbf{u}}(t)) \equiv 0\,.$$

If $\left(|S| > d_z D |\mathbb{F}|^{m-1}\right)$, then there exists a polynomial $P(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ of total degree at most d such that

$$A(\mathbf{x}, P(\mathbf{x})) \equiv 0\,.$$

Moreover, there is a set $S' \subseteq S$ of size at least $|S|/d_z$ such that for all $\mathbf{u} \in S'$, $P(\mathbf{x})$ when restricted to the line in direction $\mathbf{u}$ through the point $\mathbf{b}$ equals $P_{\mathbf{b},\mathbf{u}}(t)$.

Proof. In its essence, the lemma above states that if for many different lines $L_{\mathbf{b},\mathbf{u}}$ through a point $\mathbf{b}$, the restriction $A(L_{\mathbf{b},\mathbf{u}}, z) \in \mathbb{F}[t][z]$ of the polynomial $A(\mathbf{x}, z)$ has a low-degree polynomial root $P_{\mathbf{b},\mathbf{u}}(t)$, then the original unrestricted polynomial $A(\mathbf{x}, z) \in \mathbb{F}[\mathbf{x}][z]$ has a low-degree root $P(\mathbf{x})$. Moreover, we are given that the point $\mathbf{b}$ is non-degenerate in the sense that the univariate polynomial $A(\mathbf{b}, z)$ is square free. Qualitatively, the lemma is of flavor similar to the Hilbert's irreducibility theorem and the proof will be along similar lines but technically simpler and as it eventually turns out, quantitatively better for our eventually applications.

For ease of notation, we will assume that $\mathbf{b} = \mathbf{0}$ (by translating A if necessary) and use $P_{\mathbf{u}}$ to denote $P_{\mathbf{b},\mathbf{u}}$. Let $\langle \mathbf{x} \rangle$ denote the ideal $\langle x_1, \ldots, x_m \rangle$.

From the hypothesis of the lemma, we know that for every $\mathbf{u} \in S$, there is a polynomial $P_{\mathbf{0},\mathbf{u}}(t)$ of degree at most d such that $A(t\mathbf{u}, P_{\mathbf{u}}(t)) \equiv 0$, which implies that $A(\mathbf{0}, P_{\mathbf{u}}(0)) = 0$. Therefore, if $\{\alpha_1, \ldots, \alpha_\ell\}$ is the multiset of roots over $\mathbb{F}$ for the polynomial $A(0, z)$, we must have that $P_{\mathbf{u}}(0) = \alpha_i$ for some $i \in [\ell]$. As the $\deg(A(0, z)) \leq d_z$, we have that $\ell \leq d_z$. Hence, there must exist some i such that $\alpha_i = P_{\mathbf{u}}(0)$ for at least $D |\mathbb{F}|^{m-1}$ many $\mathbf{u} \in S$. Let us refer to this α_i as just α, and let $R = \{\mathbf{u} \in S : P_{\mathbf{u}}(0) = \alpha\}$. Therefore, we have that $P_{\mathbf{u}}(0) = \alpha$ for all $\mathbf{u} \in R$, and also

$$A(t\mathbf{u}, P_{\mathbf{u}}(t)) = 0 \implies A(t\mathbf{u}, P_{\mathbf{u}}(t)) = 0 \bmod \langle \mathbf{x} \rangle^{k+1} \text{ for all } k \geq 0.$$

On the other hand, since we are given that $A(\mathbf{0}, z)$ is square-free, the roots of this polynomial are distinct and hence $\partial_z(A(\mathbf{0}, z)) \neq 0$. Thus, by Lemma 2.10, there exist polynomials $\Phi_k(\mathbf{x})$ for

every $k \geq 0$ such that

$$\Phi_k(\mathbf{0}) = \alpha,$$
$$A(\mathbf{x}, \Phi_k(\mathbf{x})) = 0 \bmod \langle \mathbf{x} \rangle^{k+1}.$$

By making the substitution $x_i \mapsto tx_i$, the above equation transforms to $A(t\mathbf{x}, \Phi_k(t\mathbf{x})) = 0 \bmod t^k$. In particular, we have

$$A(t\mathbf{u}, \Phi_k(t\mathbf{u})) = 0 \bmod t^k \quad \text{for all } \mathbf{u} \in R.$$

Fix an arbitrary $\mathbf{u} \in R$ and consider the polynomial $A'_{\mathbf{u}}(t, z) := A(t\mathbf{u}, z) \in \mathbb{F}[t, z]$. We know that $A'_{\mathbf{u}}(0, \alpha) = 0$ is square-free and $\partial_z(A'_{\mathbf{u}}(t, z))(0, \alpha) \neq 0$. Thus, by applying Lemma 2.10 on $A'_{\mathbf{u}}(t, z)$ to lift from α, there ought to be a unique polynomial $\Psi_{\mathbf{u}}(t)$ of degree at most d such that $\Psi_{\mathbf{u}}(0) = \alpha$ and

$$A'_{\mathbf{u}}(t, \Psi_{\mathbf{u}}(t)) = A(t\mathbf{u}, \Psi_{\mathbf{u}}(t)) = 0 \bmod t^{d+1}.$$

We already have two such candidates for $\Psi_{\mathbf{u}}(t)$, namely the polynomials $P_{\mathbf{u}}(t)$ and $\Phi_d(t\mathbf{u})$. Thus, by the uniqueness asserted by Lemma 2.10, we have that

$$P_{\mathbf{u}}(t) = \Phi_d(t\mathbf{u}) \text{ for all } \mathbf{u} \in R.$$

To finish the proof, we wish to argue that $\Phi_d(\mathbf{x})$ is a true root of $A(\mathbf{x}, z)$, i.e. it satisfies $A(\mathbf{x}, \Phi_d(\mathbf{x})) = 0$ (even without $\bmod \langle \mathbf{x} \rangle^{d+1}$). To this end, consider the polynomial $B(\mathbf{x}) = A(\mathbf{x}, \Phi_d(\mathbf{x}))$. Since $\deg(\Phi_d(\mathbf{x})) \leq d$ and A has $(1, \ldots, 1, d)$-weighted degree at most D, we have $\deg(B(\mathbf{x})) \leq D$.

For every $\mathbf{u} \in R$, we have $\Phi_d(t\mathbf{u}) = P_{\mathbf{u}}(t)$ and hence $B(t\mathbf{u}) = A(\mathbf{u}, \Phi_d(t\mathbf{u})) = A(t\mathbf{u}, P_{\mathbf{u}}(t)) = 0$ as a polynomial in t, and thus $B(\mathbf{u}) = 0$. This implies that the m-variate polynomial B of degree at most D has more than $D \cdot |\mathbb{F}|^{m-1}$ zeros in $\mathbb{F}^m$. By the Polynomial Identity Lemma (Lemma 2.2), conclude that $B(\mathbf{x}) = A(\mathbf{x}, \Phi_d(\mathbf{x}))$ is the zero polynomial. $\square$

4 The bivariate low-degree test

In this section, prove Theorem 1.5 which we recall for convenience.

Theorem 1.5 (Bivariate low-degree test). *There exists a constant $\tau \in (0, 1]$ such that for every finite field $\mathbb{F}_q$, $\varepsilon_0 \in (0, 1)$ and degree d satisfying $\varepsilon_0 > (d/q)^\tau$, the following holds for every function*

$f : \mathbb{F}_q^2 \to \mathbb{F}_q$ and $\varepsilon \in (\varepsilon_0, 1)$:

$$\mathbb{E}_\ell[\mathrm{agree}(f|_\ell, P_{1,d})] \geq \varepsilon \ \text{ implies } \ \mathrm{agree}(f, P_{2,d}) \geq \varepsilon - \varepsilon_0 \ .$$

Remark 4.1 (Version with a points table and a lines table). *Many results in the low-degree testing literature often consider a variant where we are given a 'points table' $f : \mathbb{F}^m \to \mathbb{F}$, and a 'lines table' $P : \mathcal{L}^{(m)} \to \mathbb{F}[x]^{\leq d}$ (where $\mathcal{L}^{(m)}$ denotes the set of lines in $\mathbb{F}^m$) that assigns a polynomial of degree at most d for every line. The version above is a special case where P is the* canonical lines table *that assigns the best-fit degree d polynomial on each line. The discussion in the section carries over in a straightforward manner to the more general setting of points and lines table, and we deal with the special case purely to avoid some additional notational clutter.* ◇

Before we proceed with the proof of the above theorem, we will fix some notation. Let $f : \mathbb{F}^2 \to \mathbb{F}$ be provided as an oracle. Throughout this section, we will be using LDT_d to refer to the line-point test:

- Pick $x \in_R \mathbb{F}^2$ and a random line $\ell \ni x$.
- Query the oracle on all points on the line ℓ and let P_ℓ be the best-fit degree d polynomial.
- Accept if $P_\ell(x) = f(x)$.

We will use ε_x to refer to $\mathrm{Pr}_{\ell \ni x}[P_\ell(x) = f(x)]$.

To begin with, we will prove the following "weaker" statement for the low-degree test.

Theorem 4.2. *Suppose $\varepsilon > 0$, and d is a positive integer. Let $\mathbb{F}$ is a finite field with q elements $q > C \cdot d/\varepsilon^7$ for a large enough constant C. Suppose $f : \mathbb{F}^2 \to \mathbb{F}$ that passes LDT_d with probability at least ε, that is*

$$\Pr_{a, \ell \ni a}[f(a) = P_\ell(a)] \geq \varepsilon.$$

Then, there is a polynomial $Q(x, y)$ of degree at most d such that

$$\Pr_{a \in \mathbb{F}^2}[f(a) = Q(a)] \geq \Omega(\varepsilon^4).$$

Although the above theorem appears to yield a weaker agreement than claimed by Theorem 1.5, it would turn out that Theorem 4.2 yields Theorem 1.5 via standard reductions that we elaborate on towards the end of this section. Thus, we proceed towards proving Theorem 4.2.

As a first step, we show the existence of a low-degree trivariate polynomial that *explains* that the given function passes the line point test with a constant probability.

Theorem 4.3 (Interpolation for LDT). *There exist constant $c_1, c_2 \in \mathbb{N}$ such that for every $\varepsilon \in (0, 1]$, any $d \in \mathbb{N}$ and finite field $\mathbb{F}$ of size q and characteristic p with $q > c_1 \cdot d/\varepsilon^{c_2}$, the following is true.*

Let $f : \mathbb{F}^2 \to \mathbb{F}$ be an oracle that passes LDT_d with probability ε. Then, there is a non-zero polynomial $A(x, y, z)$ and a subset $S \subseteq \mathbb{F}^2$ such that

- $\deg_{1,1,d}(A) = O(d/\varepsilon^2)$ *and* $|S| = \Omega(\varepsilon^2 q^2)$,
- $\sum_{x \in S} \varepsilon_x = \Omega(\varepsilon^3 \cdot q^2)$,
- *for every* $(a, b) \in S$, *we have* $\varepsilon_{(a,b)} = \Pr_{\ell \ni (a,b)}[P_\ell(a, b) = f(a, b)] \geq \varepsilon/2$,
- *for every* $(a, b) \in S$, *we have* $A(a, b, f(a, b)) = 0$.
- $\partial_z(A)$ *and* $\mathsf{Disc}_z(A)$ *are not identically zero.*

The proof is mostly along the lines of Arora and Sudan [AS03] but with a tighter analysis and some care for fields of small characteristic. We present it in its entirety, in Appendix A, for completeness and to make the changes clearer to follow.

We now use Theorem 4.3 to complete the proof of Theorem 4.2.

Proof of Theorem 4.2. Suppose $f : \mathbb{F}^2 \to \mathbb{F}$ is given and we know that f passes LDT_d with probability at least ε. By Theorem 4.3, we have a set $S \subseteq \mathbb{F}^2$ with $|S| = \gamma \cdot q^2 = \Omega(\varepsilon^2 \cdot q^2)$ and a non-zero polynomial $A(x, y, z)$ with $\deg_{1,1,d}(A) = D = O(d/\varepsilon^2)$ such that for all $(a, b) \in S$ we have

- $A(a, b, f(a, b)) = 0$,
- $\Pr_{\ell \ni (a,b)}[P_\ell(a, b) = f(a, b)] \geq \varepsilon/2$.

We also know that $\partial_z(A)$ and $\mathsf{Disc}_z(A)$ are both non-zero polynomials. We would like to find an $(a, b) \in S$ such that the following guarantees hold:

(i) $\Gamma_A(a, b) \neq 0$ where $\Gamma_A(x, y) = \mathrm{disc}_z(A) = \mathrm{Res}_z(A, \partial_z A)$,

(ii) There are $\Omega(\varepsilon^2 \cdot (q^2 - 1))$ choices for $(\alpha, \beta) \in \mathbb{F}^2 \setminus \{(0, 0)\}$, such that there is a degree d polynomial $P_{\alpha, \beta}(t)$ such that

$$A(a + \alpha t, b + \beta t, P_{\alpha, \beta}(t)) = 0$$

with $P_{\alpha, \beta}(0) = f(a, b)$, and $P_{\alpha, \beta}(a + \alpha t, b + \beta t) = f(a + \alpha t, b + \beta t)$ for at least $\Omega(\varepsilon^2 \cdot q)$ many $t \in \mathbb{F}$.

If we can find such an (a, b), then Lemma 3.1 would imply that there is a polynomial $Q(x, y)$ of degree at most d such that $A(x, y, Q(x, y)) = 0$ with $Q|_\ell = P_{\alpha, \beta}(t)$ where ℓ is the line $\{(a + \alpha t, b + \beta t) : t \in \mathbb{F}\}$ for any (α, β) satisfying Item (ii). Thus, for any $(a', b') = (a + \alpha t, b + \beta t)$ that satisfy $P_{\alpha, \beta}(t) = f(a', b')$, we have $Q(a', b') = P_{\alpha, \beta}(t) = f(a', b')$. Thus, Q agrees with f on $\Omega(\varepsilon^4 \cdot q^2)$ locations, as

claimed by the theorem.

We now work towards guaranteeing Items (i) and (ii). Note that we have $q^2 + q$ lines in $\mathbb{F}^2$, and each point $x \in \mathbb{F}^2$ has $q + 1$ lines containing it, and each line has q points on it. Consider the following bipartite graph where the set of left vertices is $\mathbb{F}^2$ and the set of right vertices are all lines in $\mathbb{F}^2$. We connect a point $x \in \mathbb{F}^2$ on the left to a line ℓ on the right if $x \in S$, $x \in \ell$ and $P_\ell(x) = f(x)$. Note that every $x \in S$ has $\varepsilon_x \cdot (q + 1)$ edges in this bipartite graph. If $\mu > 0$ such that $\mu q^2 = \sum_{x \in S} \varepsilon_x = \Omega(\varepsilon^3 \cdot q^2)$ (by Theorem 4.3), the number of edges in this graph is $(q + 1) \cdot \sum_{x \in S} \varepsilon_x = \mu \cdot (q^3 + q^2)$.

For a line ℓ, let $y_\ell = {}^{\deg(\ell)}/q$ where $\deg(\ell)$ refers to the degree in this graph. Since each $\sigma_\ell \in [0, 1]$ and $\sum \sigma_\ell = \mu(q^2 + q)$. By Lemma 2.1, if $L = \{\ell : \sigma_\ell > \mu/2\}$ (the 'high-degree lines'), then at least $\mu/2 \cdot (q^3 + q^2)$ edges of the graph are incident on lines in L — let us refer to these edges as "marked edges". Again by Lemma 2.1, there are at least $\mu/4 \cdot q^2$ vertices in the left that have at least $\mu/4 \cdot (q+1)$ "marked edges" incident on it — let us call this set of vertices $S' \subseteq S$.

Thus we now have a set $S' \subseteq S \subseteq \mathbb{F}^2$ with $|S'| \geq \mu/4 \cdot q^2 = \Omega(\varepsilon^3 \cdot q^2)$ such that each $x \in S'$ such that at least $\mu/4$-fraction of lines ℓ through it satisfy the following two properties:

- $P_\ell(x) = f(x)$,

- the line ℓ contains at least $\mu \cdot q/2$ points of S.

Guaranteeing Item (i): The polynomial $A(x, y, z)$ has $\deg_{1,1,d} \leq D = O({}^d/\varepsilon^2)$ and hence we have $d_z = \deg_z(A) = O({}^1/\varepsilon^2)$. Therefore, the z-discriminant of A, namely $\mathsf{Disc}_z(A)$ is the determinant of a $O(d_z) \times O(d_z)$ matrix, each of whose entries is a polynomial in x, y of degree at most D. We recall from the last item of Theorem 4.3 that $\mathsf{Disc}_z(A)$ is not identically zero.

Thus, $\deg(\mathsf{Disc}_z(A)) = O(D \cdot d_z) = O({}^d/\varepsilon^4)$. By Lemma 2.2, there are at most $O({}^{dq}/\varepsilon^4)$ points in $\mathbb{F}^2$ that the polynomial Γ vanishes on. Since $|S'| = \Omega(\varepsilon^3 \cdot q^2)$ and $q > C \cdot {}^d/\varepsilon^7$ for a large enough constant C, there must be some $(a, b) \in S'$ that does not make Γ zero.

Guaranteeing Item (ii): Let $(\alpha, \beta) \in \mathbb{F}^2 \setminus (0, 0)$ such that $\ell = \{(a + \alpha t, b + \beta t) : t \in \mathbb{F}\}$ is one of the 'high-degree' lines in L containing (a, b). Let P_ℓ be the best-fit degree d polynomial on this line. Since $(a, b) \in S$, we have $P_\ell(0) = f(a, b)$.

Let the polynomial $B(t)$ be defined as $B(t) := A(a + \alpha t, b + \beta t, P_\ell(t))$. Clearly, B has degree at most D. For any $t \in \mathbb{F}$ such that $(a', b') = (a + \alpha t, b + \beta t) \in S$, we have $P_\ell(t) = f(a', b')$ and hence $B(t) = A(a', b', f(a', b')) = 0$. Since there are at least $\mu \cdot q/2 = \Omega(\varepsilon^2 \cdot q)$ such points, the polynomial $B(t)$ has $\Omega(\varepsilon^2 \cdot q)$ roots but has degree at most $D = O(d/\varepsilon^2)$. Since $q > C \cdot {}^d/\varepsilon^7$, we must have that $B(t) = A(a + \alpha t, b + \beta t, P_\ell(t))$ is identically zero.

As (a, b) is adjacent to $\mu/4 \cdot (q + 1)$ lines, there are at least $\mu/4 \cdot (q + 1) \cdot (q - 1) = \mu/4 \cdot (q^2 - 1)$

23

choices for $(\alpha, \beta) \in \mathbb{F}^2 \setminus (0,0)$ such that there is some degree d polynomial $P_{\alpha,\beta}$ such that

$$A(a + \alpha t, b + \beta t, P_{\alpha,\beta}(t)) = 0.$$

Thus, we can now invoke Lemma 3.1 to get the desired conclusion. This completes the proof of Theorem 4.2. $\square$

4.1 Some consequences of Theorem 4.2

Given Theorem 4.2, we conclude this section with a few variants that will be useful for us in the proof of the high dimensional case. These variants, despite being seemingly stronger in appearance follows immediately from Theorem 4.2 via standard techniques in the low-degree testing literature. For the sake of completeness, we include formal statements and proofs in Appendix B.

Theorem 4.4 (List-decoding version of the bivariate LDT). *For every $\varepsilon_0 \in (0,1)$, finite field $\mathbb{F}_q$ and degree d satisfying $\varepsilon_0 > \Omega((d/q)^{1/16})$, the following holds:*

For every $f : \mathbb{F}_q^2 \to \mathbb{F}_q$ and every $\varepsilon > \varepsilon_0$, there is a (possibly empty) set $\{Q_1, \ldots, Q_t\}$ of polynomials of degree at most d such that $t \leq 2/\varepsilon^8$ and $\mathrm{agree}(f, Q_i) \geq \varepsilon^8$ for all $i \in [t]$ and

$$\Pr_{a \in \mathbb{F}_q^2}[a \text{ is } \varepsilon\text{-good and } f(x) \notin \{Q_1(x), \ldots, Q_t(x)\}] \leq \varepsilon_0.$$

The other is the 'high-agreement variant', namely Theorem 1.5:

Theorem 1.5 (Bivariate low-degree test). *There exists a constant $\tau \in (0,1]$ such that for every finite field $\mathbb{F}_q$, $\varepsilon_0 \in (0,1)$ and degree d satisfying $\varepsilon_0 > (d/q)^\tau$, the following holds for every function $f \colon \mathbb{F}_q^2 \to \mathbb{F}_q$ and $\varepsilon \in (\varepsilon_0, 1)$:*

$$\mathbb{E}_\ell[\mathrm{agree}(f|_\ell, P_{1,d})] \geq \varepsilon \text{ implies } \mathrm{agree}(f, P_{2,d}) \geq \varepsilon - \varepsilon_0 .$$

Proofs of Theorems 4.4 and 1.5. Both the above theorems are immediate consequences of a generic reduction from the 'weak form of LDT' to the 'list-decoding form of LDT' (Lemma B.4) and the 'high-agreement form of LDT' (Lemma B.6) used along with Theorem 4.2. $\square$

Theorem 1.5 is also of interest in the low error regime, where ε is close to 1. In this case, Theorem 4.2 only lets us conclude a constant fraction agreement between f and a low-degree polynomial, whereas the Theorem 1.5 asserts that f agrees with a low-degree polynomial on almost all inputs. In the low error regime, in order to lift the analysis from the bivariate case to the multivariate case for our proof, we rely on the following theorem, which, as we show in the proof also follows from Theorem 1.5.

To avoid confusion with the low-error regime, we will use δ to denote the probability that f *fails* the low-degree test.

Theorem 4.5 (Low-error high-agreement bivariate LDT). *There is a large enough constant C such that for any finite field $\mathbb{F}_q$ and degree d satisfying $q > C \cdot d$, and any $\delta < 0.01$ the following holds:*

Suppose $f : \mathbb{F}_q^2 \to \mathbb{F}_q$ passes the LDT_d with probability $1 - \delta$, then there is a bivariate degree d polyomial Q such that

$$\Pr_{x \in \mathbb{F}_q^2}[f(x) = Q(x)] \geq 1 - 2\delta.$$

Proof. Since f passes LDT_d with probabaility at least $(1 - \delta)$, we have from [Theorem 1.5](#) that there exists a degree d polynomial Q such that

$$\Pr_{a \in \mathbb{F}^2}[f(a) = Q(a)] \geq 1 - \delta - \varepsilon_0.$$

where $\varepsilon_0 = \Omega\left((d/q)^{1/20}\right)$. We note that this agreement is weaker than the conclusion of the above theorem when δ is smaller than ε_0. Let C be chosen large enough so that $\delta + \varepsilon_0 \leq 0.01 + \varepsilon_0 < 1/20 \cdot (1 - d/q)$.

An important ingredient of the proof of this theorem is the definition of a *corrected* version of f that we denote by f_{corr}, which is defined as follows.

$$f_{\mathrm{corr}}(a) := \mathrm{plurality}_{\ell \ni a}\left\{P_\ell^{(f,d)}(a)\right\}.$$

where ℓ is a line through a and $P_\ell^{(f,d)}$ denotes a univariate polynomial of degree d that is closest to the restriction of f on ℓ.

The theorem is an immediate consequence of the following two claims

Claim 4.6. $\Pr_{a \in \mathbb{F}^2}[f(a) = f_{\mathrm{corr}}(a)] \geq 1 - 2\delta.$

Claim 4.7. *For all $a \in \mathbb{F}^2$, we have $f_{\mathrm{corr}}(a) = Q(a)$.*

Proof of [Claim 4.6](#). Let B denote the set of *bad* points in $\mathbb{F}^2$ defined as follows

$$B := \left\{a \in \mathbb{F}^2 : \Pr_{\ell \ni a}[P_\ell^{(f,d)}(a) \neq f(a)] \geq 1/2\right\}.$$

Clearly, if $a \notin B$, then we have $f_{\mathrm{corr}}(a) = f(a)$. Hence, we have $\Pr_a[f(a) \neq f_{\mathrm{corr}}(a)] \leq \Pr_a[a \in B]$.

On the other hand,

$$\delta \geq \Pr_{x,\ell}[f \text{ fails LDT}_d \text{ on } (x,\ell)]$$

$$\geq \Pr[a \in B] \cdot \Pr_{\substack{a \in \mathbb{F}^2 \\ \ell \ni a}}[P_\ell^{(f,d)}(a) \neq f(a) \mid a \in B]$$

$$\geq \Pr_{a \in \mathbb{F}^2}[f(a) \neq f_{\text{corr}}(a)] \cdot 1/2$$

which implies that $\Pr_a[f(a) \neq f_{\text{corr}}(a)] \leq 2\delta$ and hence $\Pr_a[f(a) = f_{\text{corr}}(a)] \geq 1 - 2\delta$. $\qquad\square$

Proof of Claim 4.7. As mentioned earlier, we know that f has fractional agreement of at least $(1 - \delta - \varepsilon_0)$ with the degree d polynomial Q. Let B be the set of inputs where f and Q disagree. Clearly, $|B| < q^2(\delta + \varepsilon_0)$. Let $a \in \mathbb{F}^2$ be an arbitrary input. Since lines through a are a uniform cover of the space $\mathbb{F}^2 \setminus \{a\}$, we have that for all large enough q,

$$\mathbb{E}_{\ell \ni a}[|\ell \cap B|] \leq q \cdot \left(\frac{q^2(\delta + \varepsilon_0)}{q^2 - 1} \right) \leq 2q(\delta + \varepsilon_0).$$

Thus, by Markov's inequality, $\Pr_{\ell \ni a}[|\ell \cap B| > 8q(\eta + \varepsilon)] < 1/4$.

Now, let ℓ be a line through a such that $|\ell \cap B| < 8q(\delta + \varepsilon_0)$. For every such line ℓ, the disagreement between Q and f on ℓ is at most $8q(\delta + \varepsilon_0) + 1 \leq 10q(\delta + \varepsilon_0) < 1/2 \cdot (1 - d/q)$. Since this is less than half the minimum distance of a Reed-Solomon code of degree d and block-length q, we have that $P_\ell^{(f,d)}$ must equal $Q|_\ell$.

Hence, we have that on $3/4$-th of the lines ℓ through a, the restriction $Q|_\ell$ equals the polynomial $P_\ell^{(f,d)}$. From the definition of f_{corr}, it now follows that $f_{\text{corr}}(a) = Q(a)$. Since a is an arbitrary point in $\mathbb{F}^2$, we have that the f_{corr} and Q agree everywhere on $\mathbb{F}^2$. $\qquad\square$

That concludes the proof of Theorem 4.5. $\qquad\square$

5 Lifting to m-variate low-degree tests

In this section, we bootstrap the low-degree test from 2 dimensions to m dimensions. This bootstrapping is inspired and adapted from the corresponding bootstrapping arguments due to Friedl-Sudan [FS95], Arora-Sudan [AS03] and Bhangale-Dinur-Navon [BDN17] and the local-list-decoder for multivariate polynomial codes due to Sudan, Trevisan & Vadhan [STV01]. The bootstrapping arguments heavily use the expansion properties of the points-lines-planes affine Grassmannian. So, we first mention some preliminaries that we would need..

5.1 Preliminaries

Theorem 5.1 ((Weak) Johnson Bound). *For any function $f \colon \mathbb{F} \to \mathbb{F}$, degree parameter d and $\varepsilon \in (0,1)$, let $P_1, \ldots, P_r$ be the set of all degree-d polynomials that have agreement at least ε with f (i.e, $\Pr_{x \in \mathbb{F}}[f(x) = P(x)] \geq \varepsilon$). We say that a point $x \in \mathbb{F}$ is non-unique with respect to function f, degree d and agreement ε, if there exist two* distinct *polynomials $P_k \neq P_{k'}$, $1 \leq k, k' \leq r$ such that $P_k(x) = P_{k'}(x)$ or in short, "$x \in \mathrm{nonunique}_\varepsilon^d(f)$".*

If $\varepsilon \geq 2\sqrt{d/q}$, then the following bounds hold.

1. $r \leq 2/\varepsilon$.

2. The number of points in $\mathrm{nonunique}_\varepsilon^d(f)$ is at most $\binom{r}{2} \cdot d \leq 2d/\varepsilon^2$.

Definition 5.2. *Let $G = (A, B, E)$ be a bi-regular bipartite graph, and let $M \in \mathbb{R}^{A \times B}$ be the adjacency matrix normalized such that $\|M\mathbb{1}\| = 1$, denote by $\lambda(G)$ the value*

$$\lambda(G) = \max_{\mathbf{v} \perp \mathbb{1}} \left\{ \frac{\|M\mathbf{v}\|}{\|\mathbf{v}\|} \right\}.$$

We will refer to $\lambda(G)$ as the second eigen-value of G. ◇

This is really the second largest singular value of M , with a different normalization (such that the maximal singular value equals 1).

The following is the classic Expander Mixing Lemma (for bipartite graphs).

Lemma 5.3 (Expander Mixing Lemma). *Let $G = (A, B, E)$ be a biregular bipartite graph with second eigen-value λ. Then for any two functions $g \colon A \to \mathbb{R}$ and $h \colon B \to \mathbb{R}$ with means and variances μ_g, μ_h and σ_g^2 and σ_h^2, we have*

$$\left| \Pr_{(a,b) \in E}[g(a) \cdot h(b)] - \mu_g \cdot \mu_h \right| \leq \lambda \cdot \sigma_g \cdot \sigma_h$$

We will be using the following corollary of the above Expander Mixing Lemma.

Lemma 5.4 (Corollary of Expander Mixing Lemma [BDN17, Lemma 6]). *Let $G = (A, B, E)$ be a biregular bipartite graph with second eigen-value λ. Then for any set $A' \subseteq A$ of measure μ and any $E' \subseteq E$, we have*

$$\left| \Pr_{a \sim A', b \sim N_G(a)}[(a,b) \in E'] - \Pr_{b \in B, a \sim N_G(b) \cap A'}[(a,b) \in E'] \right| \leq \lambda/\sqrt{\mu}.$$

The following are well-known properties of the points-lines-planes affine Grassmannian in $\mathbb{F}_q^m$.

Theorem 5.5. *The second eigen-value of the following biparite inclusion graphs is as follows:*

1. For all $m \geq 2$ and $G = G(\mathbb{F}^m, \mathscr{L}^{(m)})$, $\lambda(G) = 1/\sqrt{q}$.

27

2. *For all $m \geq 2$, $G = G(\mathbb{F}^m, \mathscr{P}^{(m)})$, $\lambda(G) = 1/q$.*

3. *For all $m \geq 2$, $G = G(\mathscr{L}^{(m)}, \mathscr{P}^{(m)})$, $\lambda(G) = \frac{1}{\sqrt{q}} \cdot (1 + o(1/\sqrt{q}))$.*

4. *For all $m \geq 2$ and $G = G(\mathscr{L}_x^{(m)}, \mathscr{P}_x^{(m)})$, $\lambda(G) = 1/\sqrt{q+1}$.*

5.2   The bootstrapping statements

Given a function $f\colon \mathbb{F}^m \to \mathbb{F}$ and a line ℓ in $\mathbb{F}^m$, we let $P_\ell^{(f,d)}$ be the best fit degree-d univariate polynomial (presented as a list of evaluations) that agrees with f on the line ℓ (if there is more than one such polynomial, we break ties arbitrarily). We now, define the following quantities for any function $f\colon \mathbb{F}^m \to \mathbb{F}$, line ℓ and plane π:

$$
\begin{aligned}
\delta_f(\ell) &:= \Pr_{x \in \ell}[P_\ell^{(f,d)}(x) \neq f(x)], \\
\delta_f(\pi) &:= \mathop{\mathbb{E}}_{\ell \in \pi}[\delta_f(\ell)] = \Pr_{\substack{\ell \in \pi \\ x \in \ell}}[P_\ell^{(f,d)}(x) \neq f(x)], \\
\delta_f &:= \mathop{\mathbb{E}}_\pi[\delta_f(\pi)] = \Pr_{\substack{\ell \in \mathbb{F}^m \\ x \in \ell}}[P_\ell^{(f,d)}(x) \neq f(x)].
\end{aligned}
\tag{5.6}
$$

Bootstrapping in the low-error regime

In Section 5.3, we bootstrap the low-error bivarariate LDT Theorem 4.5 to prove Theorem 1.1, a similar theorem for m dimensions. This bootstrapping argument is an adaptation of a similar argument due to Friedl and Sudan [FS95], the main difference being that we use the expansion properties of the points-lines-planes affine Grassmannian rather than that of a more complicated points-lines-surfaces HDX used in [FS95].

Theorem 1.1 (Multivariate low-error LDT [RS96, AS98, ALMSS98, PS94, FS95])**.** *There is a large enough constant C such that for every finite field $\mathbb{F}_q$, $m > 0$, degree d satisfying $q > C \cdot d$, the following holds for every $f\colon \mathbb{F}_q^m \to \mathbb{F}_q$ and $\delta \leq 0.01$,*

$$
\mathop{\mathbb{E}}_\ell[\delta(f|_\ell, P_{1,d})] \leq \delta \ \text{implies} \ \delta(f, P_{m,d}) \leq 4\delta \ .
$$

Bootstrapping in the high-error regime

In the subsequent Section 5.4, we perform a similar bootstrapping for the high-error regime to obtain the following theorem from bivariate low-degree test (Theorem 4.2). This bootstrapping argument is inspired from the corresponding arguments due to Arora-Sudan [AS03] and Bhangale-Dinur-Navon [BDN17]. The argument presented here, while elementary, requires a careful analysis using the spectral properties (repeated applications of Lemmas 5.3 and 5.4) of various subgraphs of the points-lines-planes affine Grassmannian. The key improvement from [AS03] and [BDN17] is

that we bootstrap from the base case of $m = 2$ while previous arguments worked with a base case of at least $m \geq 3$.

Theorem 5.7 (high-error regime m-variate LDT). *For every $\varepsilon_0 \in (0, 1)$, and finite field $\mathbb{F}_q$ and degree d satisfying $\varepsilon_0 > \Omega((d/q)^{1/48})$, the following holds.*

If the points table $f \colon \mathbb{F}^m \to \mathbb{F}$ and degree-d lines oracle $\ell \mapsto P_\ell$ satisfy

$$\Pr_{x, \ell \ni x} [f(x) = P_\ell(x)] \geq 5\varepsilon,$$

then there exists an m-variate degree-d polynomial Q such that $\Pr_x[f(x) = Q(x)] \geq \varepsilon^2$.

As in the bivariate case, the above 'weak form of the low-degree test' (Theorem 5.7) can be reduced to the 'high-agreement form of the low-degree test' (Theorem 1.2) using standard transformations (Lemma B.6).

5.3 Low-error regime: Proof of Theorem 1.1

Given a function $f \colon \mathbb{F}^m \to \mathbb{F}$, we define the self-correction f_{corr} of the function f as follows:

$$f_{\mathrm{corr}}(x) := \mathrm{plurality}_{\ell \ni x} \left\{ P_\ell^{(f,d)}(x) \right\}.$$

Let $\varepsilon_0, \delta_0 \in (0, 1)$, the field $\mathbb{F}_q$ and degree d satisfy $\varepsilon_0 \geq \Omega((d/q)^{1/20})$ as in the hypothesis of . Define $\delta_0 := 1/20(1 - d/q) - \varepsilon_0$. By the hypothesis of Theorem 1.1, we have that $\delta_f \leq \delta_0/2$.

The theorem follows from the two claims.

Claim 5.8. $\delta(f, f_{\mathrm{corr}}) \leq 2\delta_f$.

Claim 5.9. *If $q > 100/\delta_0^2$ and $\delta_f \leq \delta_0/2$, then $\delta_{f_{\mathrm{corr}}} \leq \delta_f/2$.*

Proof of Theorem 1.1. Define a sequence of functions $f^{(0)}, f^{(1)}, \ldots \colon \mathbb{F}^m \to \mathbb{F}$ as follows: $f^{(0)} := f$ and for $i \geq 1$, $f^{(i)} := \left(f^{(i-1)}\right)_{\mathrm{corr}}$. We then have, from Claim 5.9, that $\delta_i := \delta_{f^{(i)}} \leq \delta_f/2^i$. For any fixed q and m, since δ_f can only take a set of finitely many values, we have that δ_i eventually becomes 0, i.e., there exists i_* such that $\delta_{i_*} = \delta_{f^{(i_*)}} = 0$. Then, $f^{(i_*)}$ is a degree-d m-variate polynomial, say Q. Now, by Claim 5.8, we have

$$\delta(f, Q) = \delta\left(f^{(0)}, f^{(i_*)}\right) \leq \sum_{i=1}^{i_*} \delta\left(f^{(i-1)}, f^{(i)}\right) \leq \sum_{i=1}^{i_*} 2\delta_{f^{(i-1)}} \leq 2\sum_{i=1}^{i_*} \delta_f/2^{i-1} \leq 4\delta_f.$$

This completes the proof of Theorem 1.1 (assuming the two claims). $\square$

We now proceed to prove the two claims.

Proof of Claim 5.8. Given the points table $f \colon \mathbb{F}^m \to \mathbb{F}$ and lines oracle $\ell \mapsto P_\ell^{(f,d)}$, we define the set of "bad" points as follows: $\mathrm{BAD} := \{x \in \mathbb{F}^m \colon \Pr_{\ell \ni x}[P_\ell^{(f,d)}(x) \neq f(x)] \geq 1/2\}$. Clearly, if $x \notin \mathrm{BAD}$, we have $f_{\mathrm{corr}}(x) = f(x)$. Hence, $\delta(f, f_{\mathrm{corr}}) \leq \Pr[x \in \mathrm{BAD}]$. On the other hand, we have

$$
\begin{aligned}
\delta_f &= \Pr_{\substack{x \in \mathbb{F}^m \\ \ell \ni x}}[P_\ell^{(f,d)}(x) \neq f(x)] \\
&\geq \Pr[x \in \mathrm{BAD}] \cdot \Pr_{\substack{x \in \mathbb{F}^m \\ \ell \ni x}}[P_\ell^{(f,d)}(x) \neq f(x) \mid x \in \mathrm{BAD}] \\
&\geq \delta(f, f_{\mathrm{corr}}) \cdot 1/2.
\end{aligned}
$$

Hence, $\delta(f, f_{\mathrm{corr}}) \leq 2\delta_f$. $\qquad\square$

5.3.1 Self-correction passes LDT with better probability (Proof of Claim 5.9)

This is the heart of the bootstrapping argument, where we use the (high-dimensional) expansion of the points-lines-planes Affine Grassmannian complex to show that the self-corrected function f_{corr} passes the LDT with even better probability than the original function f.

We begin by showing the following bound on $\delta_{f_{\mathrm{corr}}}$.

$$
\begin{aligned}
\delta_{f_{\mathrm{corr}}} &= \Pr_{x, \ell \ni x}[P_\ell^{(f_{\mathrm{corr}},d)}(x) \neq f_{\mathrm{corr}}(x)] \\
&\leq \Pr_{x, \ell \ni x}[P_\ell^{(f,d)}(x) \neq f_{\mathrm{corr}}(x)] \\
&\leq \Pr_{\substack{x \\ \ell, \ell' \ni x}}[P_\ell^{(f,d)}(x) \neq P_{\ell'}^{(f,d)}(x)].
\end{aligned}
$$

The first inequality follows since $P_\ell^{(f,d)}$ cannot perform any better than the best-fit degree-d polynomial $P_\ell^{(f_{\mathrm{corr}},d)}$. The second inequality follows since for each $x \in \mathbb{F}^m$, $f_{\mathrm{corr}}(x)$ is the most popular value among $P_\ell^{(f,d)}(x)$ as ℓ varies over all lines ℓ through x and hence the probability (over ℓ) that $f_{\mathrm{corr}}(x) = P_\ell^{(f,d)}(x)$ is at least the collision probability that for two independently chosen lines ℓ, ℓ' through x, we have $P_\ell^{(f,d)}(x) = P_{\ell'}^{(f,d)}(x)$.

It thus suffices to bound the probability that $P_\ell^{(f,d)}(x) = P_{\ell'}^{(f,d)}(x)$ where x, ℓ, ℓ' are chosen as follows: x is picked uniformly from $\mathbb{F}^m$, ℓ, ℓ' are independently chosen to be two lines through x in $\mathbb{F}^m$. An equivalent way of picking this triple is first picking a random plane π in $\mathbb{F}^m$, a point x in the plane π and two independent lines ℓ, ℓ' in the plane π that contain x. We now define three (bad) events E1, E2 and E3 (based on the choice of π, ℓ, ℓ') such that (1) if none of the three events happen, then $P_\ell^{(f,d)}(x) = P_{\ell'}^{(f,d)}(x)$ and (2) the probability of each event is at most $\delta_f/6$. This will complete the proof of the claim.

Event E1(π): $\delta_f(\pi) \geq \delta_0$.

Consider the bipartite lines-planes incidence graph $G(\mathscr{L}^{(m)}, \mathscr{P}^m)$ in $\mathbb{F}^m$ which has second eigen-value at most $2/\sqrt{q}$ (by Theorem 5.5-3). We know that $\delta_f(\pi) = \mathbb{E}_{\ell \in \pi}[\delta_f(\ell)]$ and $\delta_f = \mathbb{E}_\ell[\delta_f(\ell)]$. Consider the functions $g: \mathscr{L}^{(m)} \to \mathbb{R}$ and $h: \mathscr{P}^{(m)} \to \mathbb{R}$ defined as follows: $g(\ell) := \delta_f(\ell)$ and $h(\pi) := \mathbb{1}[\delta_f(\pi) \geq \delta_0]$. These functions satisfy $\mu_g = \delta_f$, $\mu_h = \Pr[E1] =: \mu$, $\sigma_g^2 = \mathrm{Var}_\ell[\delta_f(\ell)] \leq \delta_f$ and $\sigma_h^2 = \mathrm{Var}_\pi[\mathbb{1}[\delta_f(\pi) \geq \delta_0]] \leq \mu$. Applying Lemma 5.3 to the graph G with functions g and h as defined above, we have

$$\mu \cdot \delta_0 - \mu \cdot \delta_f \leq \lambda \cdot \sqrt{\mu \cdot \delta_f}.$$

Equivalently, $\mu \leq (\lambda/(\delta_0-\delta_f))^2 \cdot \delta_f$. Choosing $q \geq 100/\delta_0^2$, we have $\mu \leq \delta_f/6$ (since $\lambda \approx 1/\sqrt{q}$ and $\delta_f \leq \delta_0/2$).

Furthermore, if event E1 does not happen then by Theorem 4.5 (the bivariate LDT in the low-error regime), we have that there exists a bivariate degree-d polynomial Q_π on the plane π such that $\delta(f|_\pi, Q_\pi) \leq 2\delta_f(\pi) \leq 2\delta_0$.

Event E2(π, ℓ): $\neg$E1 and $\delta(f|_\ell, Q_\pi|_\ell) \geq 4\delta_0$. (Here Q_π is the bivariate degree-d polynomial that is guaranteed to exist since E1 does not occur.)

To begin with let us fix a plane π such that E1 does not occur. We will later randomize over the choice of the plane. Since E1 does not occur, we have there exists a bivariate degree-d polynomial such that $\delta(f|_\pi, Q_\pi) \leq 2\delta_f(\pi) \leq 2\delta_0$.

Consider the bipartite points-line incidence graph $G(\pi, \mathscr{L}^{(\pi)})$ in the plane π which has second eigen-value at most $1/\sqrt{q}$ (by Theorem 5.5-1). Let BAD_π be the set of lines ℓ in π such that $\delta(f|_\ell, Q_\pi|_\ell) \geq 4\delta_0$. Consider the functions $g: \pi \to \mathbb{R}$ and $h: \mathscr{L}^{(\pi)} \to \mathbb{R}$ defined as follows: $g(x) := \mathbb{1}[f(x) = Q_\pi(x)]$ and $h(\ell) := \mathbb{1}[\ell \in \mathrm{BAD}_\pi]$. These functions satisfy $\mu_g = \delta(f|_\pi, Q_\pi)$, $\mu_h = \mu(\mathrm{BAD}_\pi) =: \mu_\pi$, $\sigma_g^2 = \mathrm{Var}_x[[f(x) \neq Q_\pi(x)]] \leq \mu_g = \delta(f|_\pi, Q_\pi)$ and $\sigma_h^2 = \mathrm{Var}_\ell[\mathbb{1}[\ell \in \mathrm{BAD}_\pi]] \leq \mu_\pi$. Applying Lemma 5.3 to the graph G with functions g and h as defined above, we have

$$\mu_\pi \cdot 4\delta_0 - \mu \cdot \delta(f|_\pi, Q_\pi) \leq \lambda \cdot \sqrt{\mu_\pi \cdot \delta(f|_\pi, Q_\pi)}.$$

Equivalently, $\mu_\pi \leq (\lambda/(4\delta_0 - \delta(f|_\pi, Q_\pi)))^2 \cdot \delta(f|_\pi, Q_\pi) \leq (\lambda^2/2\delta_0^2) \cdot \delta_f(\pi)$ (since $\delta(f|_\pi, Q_\pi) \leq 2\delta_f(\pi) \leq 2\delta_0$). Choosing $q \geq 3/\delta_0^2$, we have that for this choice of π, $\mu_\pi \leq \delta_f(\pi)/6$ (since $\lambda \approx 1/\sqrt{q}$ and $\delta_f \leq \delta_0/2$).

We now average over π as follows:

$$\Pr_{\pi, \ell}[E2] = \mathbb{E}_\pi\left[\neg E1(\pi) \cdot \mu_\pi\right] \leq \mathbb{E}_\pi\left[\neg E1(\pi) \cdot (\delta_f(\pi)/6)\right] \leq \mathbb{E}_\pi[\delta_f(\pi)/6] = \delta_f/6.$$

Observe that if for a particular choice of random π and ℓ, events E1 and E2 do not occur,

then there exists a bivariate degree-d polynomial Q_π such that $\delta(f|_\pi, Q_\pi) \leq 2\delta_f(\pi) \leq 2\delta_0$ and furthermore $\delta(f|_\ell, Q_\pi|_\ell) \leq 4\delta_0$. If $1 - 8\delta_0 \geq d/q$, then $Q_\pi|_\ell$ is the (unique) best-fit degree-d polynomial to $f|_\ell$, i.e., $P_\ell^{(f,d)} = Q_\pi|_\ell$.

Event E3(π, ℓ'): $\neg$E1 and $\delta(f|_{\ell'}, Q_\pi|_{\ell'}) \geq 4\delta_0$.

This event is identical to E2 and hence $\Pr_{\pi, \ell'}[\text{E3}] \leq \delta_f/6$.

Clearly, if events E1, E2 and E3 do not occur (for a particular choice of π, ℓ, ℓ' and x), we have that there exists a degree-d bivariate polynomial Q_π such that $P_\ell^{(f,d)} = Q_\pi|_\ell$ and $P_{\ell'}^{(f,d)} = Q_\pi|_{\ell'}$. Hence, $P_\ell^{(f,d)}(x) = Q(x) = P_{\ell'}^{(f,d)}(x)$. This completes the proof of the claim. $\qquad\square$

5.4 High-error regime: Proof of Theorem 5.7

A degree-d lines oracle assigns to each line in $\mathbb{F}^m$, a a degree-d polynomials (presented as a list of evaluations, i.e., a Reed-Solomom codeword) or $\bot$. We say that the degree-d lines oracle is ε-*well-behaved with respect to the function* $f \colon \mathbb{F}^m \to \mathbb{F}$ if for every line ℓ, P_ℓ is well-defined (i.e, $\neq \bot$) if and only if $\Pr_{x \in \ell}[f(x) = P_\ell(x)] \geq \varepsilon$.

We say a point $x \in \mathbb{F}^m$ is ε-*good* if it agrees with at least an ε-fraction of lines that pass through it, i.e.,

$$\Pr_{\ell \ni x}[f(x) = P_\ell(x)] \geq \varepsilon.$$

We will refer to the set of ε-good points in $\mathbb{F}^m$ as ε-GOOD. We will need the above notation, both when the ambient dimension is $m = 2$ (i.e., a plane) or general m (i.e., $\mathbb{F}^m$). To distinguish these two cases, in the former we say "x is ε-good with respect to plane π" , while in the latter we just say "x is ε-good". Sometimes, we also say "x is ε-locally-good" vs. "x is ε-globally-good".

We say that "x is ε-explained with respect to a plane π" if there exists an bivariate degree-d polynomial Q (defined on the plane π) such that (1) $\Pr_{z \in \pi}[f(z) = Q(z)] \geq \varepsilon$ and (2) $f(x) = Q(x)$.

We will assume the the bivariate LDT, given by Theorem 4.4, which states the following (rewritten in the language of "ε-explained"-ness). Let $\varepsilon_0 := \Omega((d/q)^{1/16})$. For every $\varepsilon \geq \varepsilon_0$, the following holds. For any function $f \colon \mathbb{F}^2 \to \mathbb{F}$,

$$\Pr_x[x \text{ is } \varepsilon\text{-good but not } \varepsilon^8\text{-explained }] \leq \varepsilon_0. \tag{5.10}$$

We prove Theorem 5.7 by defining a self-corrected function as in the low-error regime and showing that the corrected function passes the LDT with significantly higher probability. However, unlike the low-error regime, there are several candidate corrected functions and we disambiguate among them using an additional advice $(x, \sigma) \in \mathbb{F}^m \times \mathbb{F}$.

For $x \in \mathbb{F}^m$ and $\sigma \in \mathbb{F}$, define $f_{\text{corr}_\delta}^{x,\sigma} \colon \mathbb{F}^m \to (\mathbb{F} \cup \bot)$ as follows: For any $y \in \mathbb{F}^m \setminus \{x\}$, let P be the unique degree-d univariate polynomial on the unique line $\ell = \ell_{x,y}$ through x and y such that

(1) $\Pr_{z \in \ell}[f(z) = P(z)] \geq \delta$ and (2) $P(x) = \sigma$. If there is no such polynomial P or there is more than one such polynomial, we set $P := \bot$. Finally, we set $f^{x,\sigma}_{\mathrm{corr}_\delta}(y) := P(y)$. Also, set $f^{x,\sigma}_{\mathrm{corr}_\delta}(x) := \sigma$. Finally we define $f^x_{\mathrm{corr}_\delta} := f^{x,f(x)}_{\mathrm{corr}_\delta}$. We will refer to the function $f^x_{\mathrm{corr}_\delta}$ as the x-corrected function.

The following lemma states for a random ε-good x, the corrected function $f^x_{\mathrm{corr}_\delta}$ passes the LDT with very high probability $1 - \gamma$.

Lemma 5.11. *Let the field $\mathbb{F}$, degree d and $\mu, \gamma, \varepsilon \in (0,1)$ satisfy $q \geq 800d \cdot \max\left\{ 1/(\mu\gamma)^{16}, 1/\mu\gamma^2\varepsilon^{16} \right\}$. There is a integer C and $\tau \in (0,1)$ such Given a function f and ε-well-behaved degree-d lines oracle, let S be a subset of the 2ε-good points in $\mathbb{F}^m$ of density μ. Then, the distribution (x, ℓ) obtained by picking a random point in S and a random line ℓ in $\mathbb{F}^m$ satisfies*

$$\Pr_{x \in S, \ell}\left[\exists \ \text{degree-}d \ \text{polynomial} \ P_\ell \ \text{such that} \ \delta(f^x_{\mathrm{corr}_{\varepsilon^8/2}}|_\ell, P_\ell) \leq \gamma \right] \geq 1 - 2\gamma.$$

We now complete the proof of Theorem 5.7 assuming this lemma.

Proof of Theorem 5.7. We are given a points table f and a degree-d lines oracle $\ell \mapsto P_\ell$ for m dimensions such that $\Pr_{x, \ell \ni x}[f(x) = P_\ell(x)] \geq 5\varepsilon$. We first modify the lines oracle by setting P_ℓ to $\bot$ if $\Pr_{x \in \ell}[f(x) = P_\ell(x)] < \varepsilon$. This ensures, that the lines oracle is ε-well-behaved wrt f. This reduces the acceptance probability of the LDT by at most ε and the modified lines oracle satisfies

$$\Pr_{x, \ell \ni x}[f(x) = P_\ell(x)] \geq 4\varepsilon. \tag{5.12}$$

We will be setting $\mu := \varepsilon$ and $\gamma := \varepsilon^2/12$ while applying Lemma 5.11. There exists a suitably large constant C, such that $q \geq C \cdot d/\varepsilon^{48}$ implies that the field-size is large enough for Lemma 5.11 for this choice of ε, μ, γ.

For any line ℓ, let $P^{(\ell)}_1, \ldots, P^{(\ell)}_r$ be all the univariate degree-d polynomials P that satisfy $\Pr_{z \in \ell}[P(z) = f(z)] \geq \varepsilon^8/2$. Let $\mathrm{nonunique}^d_{\varepsilon^8/2}(f|_\ell)$ be the set of all points x on the line ℓ such that there exist two distinct polynomials $P^{(\ell)}_k \neq P^{(\ell)}_{k'}$ such that $P^{(\ell)}_k(x) = P^{(\ell)}_{k'}(x)$. By Theorem 5.1-2, we have that the number of points in $\mathrm{nonunique}^d_{\varepsilon^8/2}(f|_\ell)$ is at most $8d/\varepsilon^{16}$.

The LDT hypothesis (5.12) implies that

$$\Pr_{\ell, x \in \ell}[f(x) = P_\ell(x) \ \text{and} \ x \notin \mathrm{nonunique}^d_{\varepsilon^8/2}(f_\ell)] \geq 4\varepsilon - 8d/q \cdot (\varepsilon^8)^2 \geq 3\varepsilon,$$

provided $q \geq 8d/\varepsilon \cdot \varepsilon^{16}$. Or equivalently,

$$\mathbb{E}_x\left[\Pr_{\ell \ni x}\left[f(x) = P_\ell(x) \ \text{and} \ x \notin \mathrm{nonunique}^d_{\varepsilon^8/2}(f|_\ell) \right] \right] \geq 3\varepsilon.$$

Define $S \subseteq \mathbb{F}^m$ to be the set of points as follows:

$$S := \left\{ x \in \mathbb{F}^m : \Pr_{\ell \ni x} \left[f(x) = P_\ell(x) \text{ and } x \notin \text{nonunique}^d_{\varepsilon^8/2}(f|_\ell) \right] \geq 2\varepsilon \right\} .$$

We thus have $\Pr_{x \in \mathbb{F}^m}[x \in S] \geq \varepsilon$. Also observe that any $x \in S$ is 2ε-good. We now apply Lemma 5.11 on the set S with $\mu := \varepsilon$ to obtain that $f^x_{\text{corr}_{\varepsilon^8/2}}$ passes the LDT with probability at least 3γ for a random x. Fix any such $x \in S$. We have $\delta_{f^x_{\text{corr}_{\varepsilon^8/2}}} \leq 3\gamma$ where δ_f is as defined in (5.6). Applying Theorem 1.1 (m-variate LDT in the low-error regime), we obtain that there exists a m-variate degree-d polynomial Q^x such that $\delta(Q^x, f^x_{\text{corr}_{\varepsilon^8/2}}) \leq 12\gamma$.

We will now argue that f and $f^x_{\text{corr}_{\varepsilon^8/2}}$ agree on at least $2\varepsilon^2$ fraction of points since $x \in S$. Recall the definition of S and the definition of the x-corrected function $f^x_{\text{corr}_{\varepsilon^8/2}}$. For every line ℓ through x such that $f(x) = P_\ell(x)$ and $x \notin \text{nonunique}^d_{\varepsilon^8/2}$, we have that $f^x_{\text{corr}_{\varepsilon^8/2}}(z) = P_\ell(z)$ for every $z \in \ell$ such that $P_\ell(z) = f(z)$. This is because (1) P_ℓ agrees with f on at least $\varepsilon \geq \varepsilon^8/2$-fraction of the points (since the lines oracle is ε-well-behaved), (2) $P_\ell(x) = f(x)$ and (3) P_ℓ is the unique polynomial satisfying (1) and (2). Hence, for any such line ℓ, we have $\Pr_{z \in \ell}[f^x_{\text{corr}_{\varepsilon^8/2}}(z) = f(z)] \geq \varepsilon$. Furthermore, there are at least 2ε-fraction of such lines through x since $x \in S$. Hence, $\Pr_z[f^x_{\text{corr}_{\varepsilon^8/2}}(z) = f(z)] \geq 2\varepsilon \cdot \varepsilon = 2\varepsilon^2$.

Combining this, with $\delta(Q^x, f^x_{\text{corr}_{\varepsilon^8/2}}) \leq 12\gamma$, we have that $\Pr_z[Q^x(z) = f(z)] \geq 2\varepsilon^2 - 12\gamma \geq \varepsilon^2$ provided $\gamma \leq \varepsilon^2/12$. This completes the proof of the theorem. □

5.4.1 Corrected function passes LDT with high probability (proof of Lemma 5.11)

Our plan is to show that if the subset S of 2ε-globally good points is of density at least μ, then a random point x in S satisfies that the corresponding x-corrected function passes the low-degree test with probability $1 - O(\gamma)$ (for arbitrarily small $\gamma > 0$ provided d/q is small enough). To do so we consider a random point $x \in S$ and a random line ℓ and consider the plane π containing x and ℓ and prove that with probability $1 - \gamma$ the following hold: There is a bivariate polynomial Q^x_π on the plane π such that (1) Q^x_π has agreement at least ε^8 with $f|_\pi$, (2) $Q^x_\pi(x) = f(x)$, and (3) For $1 - \gamma$ fraction of the points $y \in \pi$, if we let ℓ' be the line through x and y, then $Q^x_\pi|_{\ell'}$ is the unique polynomial with $\varepsilon^8/2$ agreement with $f|_{\ell'}$.

Note that the existence of even one point x with properties (1)-(3) above suffice, but the proof essentially forces us to prove that most points in S satisfy properties (1)-(3).

We start with a preliminary statement. Note that the distribution on the triples (x, ℓ, π) of interest to us is the following: D_1 is the distribution obtained by picking x to be a uniform point in S, ℓ a uniformly random line in $\mathbb{F}^m$ and π be a uniform plane containing ℓ and x (w.h.p. π is unique given ℓ and x, but if $x \in \ell$ then π is a uniform plane containing ℓ) and output (x, ℓ, π). Now consider the related distribution D_2 obtained by sampling a plane π uniformly in $\mathbb{F}^m$, then sampling

a line ℓ uniformly in π and a point $x \in S$ on π and outputting (x, ℓ, π). If there is no point of S in π (i.e., $S \cap \pi = \emptyset$), return $\bot$. The following claim establishes that these two distributions are $O(1/(\varepsilon q))$-close in statistical distance.

Claim 5.13. *Suppose $|\mathbb{F}| \geq 3/\gamma \cdot \sqrt{\mu}$ where $\mu = \Pr_x[x \in S]$. Then $\|D_1 - D_2\|_{\mathrm{TV}} \leq \gamma$*

Proof. We construct two related distributions $\widetilde{D_1}$ and $\widetilde{D_2}$ as follows. $\widetilde{D_1}$ is the distribution obtained by picking x to be a uniform point in S, ℓ a uniformly random line in $\mathbb{F}^m$ not containing x in $\mathbb{F}^m$ and π be a uniform plane containing ℓ and x (note π is unique given ℓ and x) and output (x, ℓ, π). The distribution $\widetilde{D_2}$ is obtained by sampling a plane π uniformly in $\mathbb{F}^m$, a point $x \in S$ on π, a line ℓ uniformly in π not containing x and then outputting (x, ℓ, π). If there is no point of S in π (i.e., $S \cap \pi = \emptyset$), return $\bot$. The only difference between $D1$ and $\widetilde{D_1}$ is that the line ℓ definitely does not pass through x in $\widetilde{D_1}$ while it may do so with probability with probability $1/q^{m-1}$ in D_1 . Thus, $\|D_1 - \widetilde{D_1}\|_{\mathrm{TV}} \leq 1/q^{m-1}$. Similarly, $\|D_2 - \widetilde{D_2}\|_{\mathrm{TV}} \leq 1/q$. Hence, $\|D_1 - D_2\|_{\mathrm{TV}} \leq \|\widetilde{D_1} - \widetilde{D_2}\|_{\mathrm{TV}} + 1/q + 1/q^{m-1} \leq 2/q$.

Consider the distribution $\widetilde{D_1}$. It can alternatively be sampled as follows. Pick a random point $x \in S$, a random plane π containing x, a line ℓ uniformly in π not containing x and then outputting the triple (x, ℓ, π). Since the generative processes for picking ℓ in $\widetilde{D_1}$ and $\widetilde{D_2}$ are identical given the pair (x, π) which are picked differently, the distance between the distributions $\widetilde{D_1}$ and $\widetilde{D_2}$ is exactly the distance between the marginals of $\widetilde{D_1}$ and $\widetilde{D_2}$ on the the the (x, π) coordinates. By Lemma 5.4, this distance is at most $\lambda/\mu(S)$ where λ is the second eigen-value of the points-plane incidence graph $G = G(\mathbb{F}^m, \mathscr{P}^{(m)})$ which is at most $1/q$ (Theorem 5.5-2). Hence, $\|\widetilde{D_1} - \widetilde{D_2}\|_{\mathrm{TV}} \leq 1/q\sqrt{\mu}$.

Hence, $\|D_1 - D_2\|_{\mathrm{TV}} \leq 2/q + q/q\sqrt{\mu} \leq \gamma$ provided $q \geq 3/\gamma \cdot \sqrt{\mu}$. $\qquad\square$

Given this claim, it suffices to prove the following lemma to prove Lemma 5.11.

Lemma 5.14. *Let the field $\mathbb{F}$, degree d and $\mu, \gamma, \varepsilon \in (0, 1)$ satisfy $q \geq 800d \cdot \max\left\{1/(\mu\gamma)^{16}, 1/\mu\gamma^2\varepsilon^{16}\right\}$. Given a function f and ε-well-behaved degree-d lines oracle, let S be a subset of the 2ε-good points in $\mathbb{F}^m$ of density μ. Then, the distribution D_2 on triples (π, x, ℓ) obtained by picking a random plane π in $\mathbb{F}^m$, a random $x \in S$ in the plane π and a random line ℓ in the plane π satisfies*

$$\Pr_{(\pi, x, \ell) \sim D_2}\left[\exists \text{ degree-}d \text{ polynomial } P_\ell^x \text{ such that } \delta(f^x_{\mathrm{corr}_{\varepsilon^8/2}}|_\ell, P_\ell^x) \leq \gamma\right] \geq 1 - \gamma.$$

Proof of Lemma 5.11. Follows from Claim 5.13 and Lemma 5.14 $\qquad\square$

Proof. To prove Lemma 5.14, we list some bad events E1–E4 such that if none of these occur then $f^x_{\mathrm{corr}_{\varepsilon^8/2}}$ passes the low-degree test on (ℓ, y) with probability $1 - 2\gamma$ over the choice of $y \in \ell$. We then argue that each of these events happens with probability at most $\gamma/4$.

Event E1(π): $\Pr_{z \in \pi}[z \in S] < \mu/2$.

35

Event E2(π): $\Pr_{z \in \pi}[z \in S$ but z is not ε-locally-good wrt. plane $\pi] \geq \mu \cdot \gamma / 16$.

Event E3(π, x): $\neg$(E1 $\vee$ E2) and "x is not ε^8-explained on π".

Now if the plane π and point x are such that E1, E2 and E3 do not occur, then there exists a bivariate degree-d polynomial Q_π^x on the plane π such that $\Pr_{z \in \pi}[Q_\pi^x(z) = f(x)] \geq \varepsilon^8$ and $f(x) = Q_\pi^x(x)$.

Event E4(π, x, ℓ): $\neg$(E1 $\vee$ E2 $\vee$ E3) and $\Pr_{y \in \ell}[f_{\text{corr}_{\varepsilon^8/2}}^x(y) \neq Q_\pi^x(y)] > \gamma$.

Claims 5.15, 5.16, 5.18 and 5.19 imply that each of these events occurs with probability at most $\gamma / 4$. Furthermore, clearly if none of the 4 events occur, then the polynomial Q_π^x disagrees with the corrected function $f_{\text{corr}_{\varepsilon^8/2}}^x$ on at most γ-fraction of the points on the line ℓ. This completes the proof of the lemma assuming Claims 5.15, 5.16, 5.18 and 5.19. $\square$

The rest of this section is devoted towards bounding the probability of the events E1–E4.

Claim 5.15 (Bounding E1). *If* $\mu = \Pr_x[x \in S]$ *and* $|\mathbb{F}| \geq 4/\sqrt{\mu \cdot \gamma}$, *then*

$$\Pr_\pi[E1] = \Pr_\pi[|S \cap \pi| < \mu/2 \cdot q^2] \leq \gamma/4.$$

Proof. Consider the bipartite points-planes incidence graph $G(\mathbb{F}^m, \mathscr{P}^{(m)})$ in $\mathbb{F}^m$ which has second eigen-value at most $1/q$ (by Theorem 5.5-2). Consider the functions $g \colon \mathbb{F}^m \to \mathbb{R}$ and $h \colon \mathscr{P}^{(m)} \to \mathbb{R}$ defined as follows: $g(z) := \mathbb{1}[z \in S]$ and $h(\pi) := \mathbb{1}[|S \cap \pi| < \mu/2 \cdot q^2]$. These functions satisfy $\mu_g = \mu$, $\mu_h = \Pr[E1] =: \alpha$, $\sigma_g^2 \leq \mu$ and $\sigma_h^2 \leq \alpha$. Applying Lemma 5.3 to the graph G with functions g and h as defined above, we have

$$\mu \cdot \alpha - \mu/2 \cdot \alpha \leq \lambda \cdot \sqrt{\mu \cdot \alpha}.$$

Equivalently, $\alpha \leq 4\lambda^2/\mu$, which in turn is at most $\gamma/4$ (since $\lambda \approx 1/|\mathbb{F}|$ and $|\mathbb{F}| \geq 4/\sqrt{\mu \cdot \gamma}$). $\square$

Claim 5.16 (Bounding E2). *If* $|\mathbb{F}| + 1 \geq 100/\varepsilon \cdot \gamma^2$ *and* $\mu = \Pr_x[x \in S]$, *then*

$$\Pr_\pi[E2] = \Pr_\pi[\pi \text{ has at least } \mu \cdot \gamma/16 \cdot q^2 \text{ points which are in } S \text{ but not } \varepsilon\text{-locally-good }] \leq \gamma/4.$$

Proof. We begin by showing the following: If for any 2ε-good point x (in particular, if $x \in S$) and $|\mathbb{F}| + 1 \geq 100/\varepsilon \cdot \gamma^2$, then

$$\Pr_{\pi \in \mathscr{P}_x}[x \text{ is not } \varepsilon\text{-good with respect to plane } \pi] \leq \gamma^2/64. \tag{5.17}$$

Let x be an 2ε-good point. Let $L_x \subseteq \mathscr{L}_x^{(m)}$ be the set of lines ℓ containing x that satisfy $P_\ell(x) = f(x)$. We have $\Pr_{\ell \ni x}[\ell \in L_x] \geq 2\varepsilon$. Consider the bipartite lines-plane incidence graph

36

$G(\mathscr{L}_x^{(m)}, \mathscr{P}_x^{(m)})$ in $\mathbb{F}^m$ between lines and planes containing x. This graph has second eigen-value at most $1/\sqrt{q+1}$ (by Theorem 5.5-4). Consider the functions $g\colon \mathscr{L}_x^{(m)} \to \mathbb{R}$ and $h\colon \mathscr{P}_x^{(m)} \to \mathbb{R}$ defined as follows: $g(\ell) := \mathbb{1}[f(x) = P_\ell(x)] = \mathbb{1}[\ell \in L_x]$ and $h(\pi) := \mathbb{1}[\Pr_{\ell \in \pi}[\ell \in L_x] < \varepsilon]$. These functions satisfy $\mu_h =: \alpha$, $\sigma_g^2 \le \mu_g$ and $\sigma_h^2 \le \alpha$. Applying Lemma 5.3 to the graph G with functions g and h as defined above, we have

$$\mu_g \cdot \alpha - \varepsilon \cdot \alpha \le \lambda \cdot \sqrt{\mu_g \cdot \alpha}.$$

Equivalently, $\alpha \le \lambda^2 \cdot \mu_g/(\mu_g - \varepsilon)^2$ which is at most λ^2/ε (since $\mu_g \ge 2\varepsilon$), which in turn is at most $\gamma^2/64$ (since $\lambda \approx 1/\sqrt{|\mathbb{F}|+1}$ and $|\mathbb{F}| + 1 \ge 64/\varepsilon \cdot \gamma^2$), completing the proof of (5.17).

We now return to the proof of Claim 5.16. Consider the bipartite point-plane inclusion graph $G = G(\mathbb{F}^m, \mathscr{P}^{(m)})$. Let $A := S$ and $\mu := \Pr_z[z \in A]$. Let $B \subset \mathscr{P}^{(m)}$ be the set of planes π which have at least $\mu \cdot \gamma/16 \cdot q^2$ points which are in S but not ε-locally-good. We mark an edge (x, π) if $x \in A$ (i.e., x is 2ε-globally-good) but x is not ε-good with respect to the plane π. The fraction of marked edges in G by Eq. (5.17) is at most $\mu \cdot \gamma^2/64$. On the other hand, the fraction of marked edges is at least $\mu(B) \cdot \mu \cdot \gamma/16$. Hence, $\mu(B) \cdot \mu \cdot \gamma/16 \le \mu \cdot \gamma^2/64$ or equivalently, $\mu(B) \le \gamma/4$. $\qquad\square$

Claim 5.18 (Bounding E3). *If the parameter ε_0 in (5.10) satisfies $\varepsilon_0 \le \mu \cdot \gamma/16$ where $\mu = \Pr_z[z \in S]$, then $\Pr_{\pi,x}[\text{E3}] \le \gamma/4$.*

Proof. To begin with fix a plane π such that neither E1 nor E2 occurs. Hence, π has at least $\mu/2 \cdot q^2$ points in S of which at most $\mu \cdot \gamma/16 \cdot q^2$ are not ε-locally-good. Hence, there are at least $\mu \cdot (1 - \gamma/8)/2 \cdot q^2 \ge \mu/4 \cdot q^2$ points which are ε-locally-good in π.

Now, by the bivariate LDT (5.10), we have that the probability that a random $x \in \mathbb{F}^m$ is ε-locally-good but not ε^8-explained (both with respect to π) is at most ε_0. Hence, the probability that a random point in $S \cap \pi$ is not ε^8-explained is at most

$$\frac{\mu \cdot \gamma/16 + \varepsilon_0}{\mu(S \cap \pi)} \le \frac{\mu \cdot \gamma/16 + \varepsilon_0}{\mu/2} = \gamma/8 + 2\varepsilon_0/\mu \le \gamma/4,$$

where the last inequality follows if $\varepsilon_0 \le \mu \cdot \gamma/16$. We now bound E3 as follows:

$$\Pr_{\pi,x}[\text{E3}] \le \mathbb{E}_\pi[\Pr_x[\text{E3}|\neg(\text{E1} \vee \text{E2})]] = \mathbb{E}_\pi[\Pr_{x \in \pi}[x \text{ is not } \varepsilon^8\text{-explained} \mid \neg(\text{E1} \vee \text{E2})]] \le \gamma/4. \qquad\square$$

Claim 5.19 (Bounding E4). *If $|\mathbb{F}| \ge 800d/\mu \cdot \gamma^2 \cdot \varepsilon^{16}$, then $\Pr_{\pi,x,\ell}[\text{E4}] \le \gamma/4.x$*

Proof. The argument for bounding $\Pr[\text{E4}]$ will be far more involved and subtler than the previous cases.

Let us for the rest of the argument fix a plane π. Furthermore, let us assume that this plane π and a random $x \in S$ chosen on it are such that none of E1, E2 and E3 occur. This implies that π has at least $\mu/2$-fraction of points which are in S and there exists a bivariate degree-d polynomial

37

Q_π^x such that (1) $\Pr_{z \in \pi}[Q_\pi^x(z) = f(z)] \geq \varepsilon^8$ and (2) $Q_\pi^x(x) = f(x)$. We now need to bound the probability that when we choose a random line ℓ in the plane π, at least γ-fraction of the points y on ℓ satisfy $Q_\pi^x(y) \neq f_{\mathrm{corr}_{\varepsilon^8/2}}^x(y)$. To this end, let us recall the definition of $f_{\mathrm{corr}_{\varepsilon^8/2}}^x \colon \mathbb{F}^m \to \mathbb{F} \cup \{\bot\}$. For any point $y \in \mathbb{F}^m \setminus \{x\}$, let P be the unique degree-d univariate polynomial P on the line $\ell' = \ell_{x,y}$ through x and y such that (1) $\Pr_{z \in \ell'}[P(z) = f(z)] \geq \varepsilon^8/2$ and (2) $P(x) = f(x)$. If there is no such polynomial P or there is more than one such polynomial, then we set P to $\bot$. Finally, we set $f_{\mathrm{corr}_{\varepsilon^8/2}}^x(y) := P(y)$ and $f_{\mathrm{corr}_{\varepsilon^8/2}}^x(x) := f(x)$. We will now argue that the probability that this polynomial P is $Q_\pi^x|_{\ell'}$ for at least $1 - \gamma$ fraction of the points y in ℓ (this probability will be over the random choice of an ε-good point x in π and a random line ℓ in π).

Let $Q_1, \ldots, Q_t$ be the list of all degree-d bivariate polynomials on π such that $\Pr_{y \in \pi}[Q_j(y) = f(y)] \geq \varepsilon^8$. By Johnson bound Theorem 5.1-1, we have $t \leq 2/\varepsilon^8$ provided $\varepsilon^8 \geq 2\sqrt{d/q}$. Note that Q_π^x is one such polynomial. For $j \in [t]$, let $S_j := \{x \in \pi \colon Q_j(x) = f(x)\}$ be the set of agreement points between f and the polynomial Q_j (by definition, $\mu(S_j) \geq \varepsilon^8$).

We list below some bad events B1, B2 and B3, which if they do not occur would imply that for a random $x \in S$, a random line ℓ and a point $y \in \ell$, the polynomial P is the restricted polyomial $Q_\pi^x|_{\ell'}$ (where $\ell' = \ell_{x,y}$). These bad events would be described over the randomness of the choice of the point $x \in S$ and the random line $\ell' \in \mathscr{L}_x^\pi$ (note $\ell' = \ell_{x,y}$ is a random line through the point x in the plane π)

Event B1(ℓ'): There exists $j \in [t]$ such that $|S_j \cap \ell'| < \varepsilon^8/2 \cdot q$.

Event B2(ℓ'): $|S \cap \ell'| < \mu/4 \cdot q$.

> Let $P_1, \ldots, P_r$ be the list of all univariate degree-d polynomials on the line ℓ' such that $\Pr_{z \in \ell'}[P_k(z) = f(z)] \geq \varepsilon^8/2$. Note if B1 does not occur, then $Q_\pi^x|_{\ell'}$ is one such polynomial. By Johnson bound Theorem 5.1, we have $r \leq 4/\varepsilon^8$ provided $\varepsilon^8/2 \geq 2\sqrt{d/q}$. We say x is a non-unique point on ℓ' if there exist two distinct polynomials $k \neq k' \in [r]$ such that $P_k(x) = P_{k'}(x)$.

Event B3(ℓ', x): $\neg$B2 and x is a non-unique point on ℓ'.

Applying Claim 5.20 with $\delta := \gamma^2/8$, we have $\Pr_{(\ell',x) \sim \mathscr{E}}[\text{B1} \vee \text{B2} \vee \text{B3}] \leq \gamma^2/8$ provided $q \geq 800d/\mu \cdot \gamma^2 \cdot \varepsilon^{16}$. However, the distribution $\mathscr{E}$ over the pairs (x, ℓ') is slightly different from what we need. Our required distribution of (x, ℓ') is as follows: given a plane π, pick a random point $x \in S$ on it and a random line ℓ' in π passing through x. The distribution $\mathscr{E}$, on the other hand, is as follows: given a plane π, pick a random line ℓ' and a random $x \in S$ on ℓ' (if one exists). By Lemma 5.4, these two distributions are $\lambda/\sqrt{\mu(S \cap \pi)}$-close in total variation distance where $\lambda \approx 1/\sqrt{q}$ is the second eigen-value of the points-line incidence graph in the plane π (Lemma 5.4-1) and $\mu(S \cap \pi) \geq \mu/2$ (since event E1 does not hold). Hence, this distance is at most $\sqrt{2/q \cdot \mu}$ which is at most $\gamma^2/8$ provided $q \geq 200/\mu \cdot \gamma^4$. Hence, $\Pr_{x, \ell'}[\text{B1} \vee \text{B2} \vee \text{B3}] \leq \gamma^2/4$.

Now, recall that π and x are such that none of the events E1–E3 occur. Hence, π has at least $\mu/2$-fraction of 2ε-good points and there exists a bivariate degree-d polynomial Q_π^x such that (1) $\Pr_{z \in \pi}[Q_\pi^x(z) = f(z)] \geq \varepsilon^8$ and (2) $Q_\pi^x(x) = f(x)$. Now, suppose furthermore that ℓ and $y \in \ell$ (and the corresponding $\ell' = \ell_{x,y}$) are such that B1–B3 do not occur. It follows from these assumptions that Q_π^x is one of the polynomials Q_j and $Q_\pi^x|_{\ell'}$ one of the polynomials P_k. The uniqueness condition of $\neg$B3 implies that $Q_\pi^x|_{\ell'}$ is the only degree-d polynomial that has agreement at least $\varepsilon^8/2$ with f on ℓ and $Q_\pi^x(x) = f(x)$. Hence, $f^x_{\mathrm{corr}_{\varepsilon^8/2}}(y) = Q_\pi^x(y)$. We have thus shown that

$$\Pr_{x,\ell,y \in \ell}[f^x_{\mathrm{corr}_{\varepsilon^8/2}}(y) \neq Q_\pi^x(y)] \leq \gamma^2/4.$$

We are however interested in the fraction of points in ℓ such that $f^x_{\mathrm{corr}_{\varepsilon^8/2}}(y) \neq Q_\pi^x(y)$. A Markov argument shows that

$$\Pr_{x,\ell}\left[\left|\{y \in \ell : f^x_{\mathrm{corr}_{\varepsilon^8/2}}(y) \neq Q_\pi^x(y)\}\right| > \gamma \cdot q\right] \leq \gamma^2 \cdot q/4 \cdot \gamma \cdot q = \gamma/4.$$

Averaging over π, yields the claim. $\qquad\square$

Claim 5.20. *Let $\rho, \delta, \varepsilon \in (0,1)$, field $\mathbb{F}$ of size q and degree parameter d satisfy $q \geq 100d/\mu \cdot \delta \cdot \varepsilon^{16}$. For any plane π and a set S of 2ε-good points in the plane of density at least $\mu/2$, consider the distribution $\mathcal{E}$ on pairs (ℓ', x) chosen as follows: ℓ' is a random line in the plane π and x is a random point on $S \cap \ell$ (if not such point exists, then the distribution returns $\perp$). Then, $\Pr_{(\ell,x') \sim \mathcal{E}}[\mathrm{B1} \vee \mathrm{B2} \vee \mathrm{B3}] \leq \delta$.*

Proof. We bound the probability of each of the events B1, B2 and B3 by $\delta/3$ as follows:

Event B1(ℓ'): There exists $j \in [t]$ such that $|S_j \cap \ell'| < \varepsilon^8/2 \cdot q$.

For any fixed $j \in [t]$, since $\mu(S_j) \geq \varepsilon^8$ and the set of points in a random line ℓ' are pairwise independent, we have $\Pr[|S_j \cap \ell'| < \varepsilon^8/2 \cdot q] \leq q \cdot \varepsilon^8/(q \cdot \varepsilon^8/2)^2 = 4/q\varepsilon^8$. Hence, $\Pr[\mathrm{B1}] \leq t \cdot 4/q\varepsilon^8 \leq 8/q \cdot \varepsilon^{16}$ (as $t \leq 2/\varepsilon^8$). We, hence, have $\Pr[\mathrm{B1}] \leq \delta/3$ provided $q \geq 24/\delta \cdot \varepsilon^{16}$

Event B2(ℓ'): $|S \cap \ell'| < \mu/4 \cdot q$.

Since $\mu(S \cap \pi) \geq \mu/2$ and the set of points in a random line ℓ' are pairwise independent, we have $\Pr[\mathrm{B2}] = \Pr[|S \cap \ell'| < \mu/4 \cdot q] \leq q \cdot \mu/2/(q \cdot \mu/4)^2 = 8/q\mu$, which is at most $\delta/3$ provided $q \geq 24/\mu \cdot \delta$.

Event B3(ℓ', x): $\neg$B2 and x is a non-unique point on ℓ'.

Since event B2 does not happen, we know that that at least $\mu/4$-fraction of points in ℓ' are 2ε-good. The fraction of points x in ℓ such that there exist two distinct polynomials P_k P'_k which coincide on x (i.e., $P_k(x) = P_{k'}(x)$) is at most $\binom{r}{2} \cdot d/q \leq 8d/q \cdot (\varepsilon^8)^2$ since $r \leq 4/\varepsilon^8$. Hence, the probability of event B3 that a random point in $\ell' \cap S$ happens to be non-unique is at most $4 \cdot 8d/\mu \cdot q \cdot (\varepsilon^8)^2$. Hence, $\Pr[\mathrm{B3}] \leq \delta/3$ provided $q \geq 100d/\mu \cdot \delta \cdot \varepsilon^{16}$. $\qquad\square$

Acknowledgements

References

[AKKLR05] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. *Testing Reed-Muller codes*. IEEE Trans. Inform. Theory, 51(11):4032–4039, 2005. (Preliminary version in *7th RANDOM*, 2003). 4

[ALMSS98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. *Proof verification and the hardness of approximation problems*. J. ACM, 45(3):501–555, May 1998. (Preliminary version in *33rd FOCS*, 1992). `eccc:1998/TR98-008`. 3, 6, 24

[And20] Robert Andrews. *Algebraic hardness versus randomness in low characteristic*. In Shubhangi Saraf, ed., *Proc. 35th Comput. Complexity Conf.*, volume 169 of *LIPIcs*, pages 37:1–37:32. Schloss Dagstuhl, 2020. `arXiv:2005.10885`, `eccc:2020/TR20-081`. 11

[AS98] Sanjeev Arora and Shmuel Safra. *Probabilistic checking of proofs: A new characterization of NP*. J. ACM, 45(1):70–122, January 1998. (Preliminary version in *33rd FOCS*, 1992). 3, 6, 24

[AS03] Sanjeev Arora and Madhu Sudan. *Improved low-degree testing and its applications*. Combinatorica, 23(3):365–426, 2003. (Preliminary version in *29th STOC*, 1997). `eccc:1997/TR97-003`. 3, 4, 5, 6, 9, 10, 11, 12, 19, 22, 24, 37, 43

[BCIKS20] Eli Ben-Sasson, Dan Carmon, Yuval Ishai, Swastik Kopparty, and Shubhangi Saraf. *Proximity gaps for Reed-Solomon codes*. In Sandy Irani, ed., *Proc. 61st IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 900–909. 2020. `eccc:2020/TR20-083`.

[BDN17] Amey Bhangale, Irit Dinur, and Inbal Livni Navon. *Cube vs. Cube low degree test*. In Christos Papadimitriou, ed., *Proc. 8th Innovations in Theor. Comput. Sci. (ITCS)*, volume 67 of *LIPIcs*. Schloss Dagstuhl, 2017. `arXiv:1612.07491`, `eccc:2016/TR16-205`. 6, 11, 22, 23, 24

[BGHSV05] Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil Vadhan. *Short PCPs verifiable in polylogarithmic time*. In Luca Trevisan, ed., *Proc. 20th IEEE Conf. on Comput. Complexity*, pages 120–134. 2005. Full version available at `http://www.tcs.tifr.res.in/~prahladh/papers/BGHSV2/BGHSV2005.pdf`. 4

[BGHSV06] ———. *Robust PCPs of proximity, shorter PCPs and applications to coding*. SIAM J. Comput., 36(4):889–974, 2006. (Preliminary version in *36th STOC*, 2004). `eccc:2004/TR04-021`. 4

[BHKS23] SIDDHARTH BHANDARI, PRAHLADH HARSHA, MRINAL KUMAR, and MADHU SUDAN. *Decoding multivariate multiplicity codes over product sets*. IEEE Trans. Inform. Theory, 2023. (Early Access, Preliminary version in *53rd STOC*, 2021). arXiv:2012.01530, eccc:2020/TR20-179.

[BSVW03] ELI BEN-SASSON, MADHU SUDAN, SALIL VADHAN, and AVI WIGDERSON. *Randomness-efficient low degree tests and short PCPs via epsilon-biased sets*. In LAWRENCE L. LARMORE and MICHEL X. GOEMANS, eds., *Proc. 35th ACM Symp. on Theory of Computing (STOC)*, pages 612–621. 2003. 4

[Bür00] PETER BÜRGISSER. *Completeness and Reduction in Algebraic Complexity Theory*, volume 7 of *Algorithms and computation in mathematics*. Springer, 2000. 16

[Cha16] SIU ON CHAN. *Approximation resistance from pairwise-independent subgroups*. J. ACM, 63(3):27:1–27:32, 2016. (Preliminary version in *45th STOC*, 2013). eccc:2012/TR12-110. 6

[DL78] RICHARD A. DEMILLO and RICHARD J. LIPTON. *A probabilistic remark on algebraic program testing*. Inform. Process. Lett., 7(4):193–195, 1978. 14

[FS95] KATALIN FRIEDL and MADHU SUDAN. *Some improvements to total degree tests*. In *Proc. 3rd Israel Symp. on Theoretical and Computing Systems*, pages 190–198. 1995. (See arXiv for corrected version). arXiv:1307.3975. 3, 4, 5, 6, 8, 11, 12, 22, 24

[GKSS22] ZEYU GUO, MRINAL KUMAR, RAMPRASAD SAPTHARISHI, and NOAM SOLOMON. *Derandomization from algebraic hardness*. SIAM J. Comput., 51(2):315–335, 2022. (Preliminary version in *60th FOCS*, 2019). arXiv:1905.00091, eccc:2019/TR19-065. 11

[GW13] VENKATESAN GURUSWAMI and CAROL WANG. *Linear-algebraic list decoding for variants of Reed-Solomon codes*. IEEE Trans. Inform. Theory, 59(6):3257–3268, 2013. (Preliminary version in *26th IEEE Conference on Computational Complexity*, 2011 and *15th RANDOM*, 2011). eccc:2012/TR12-073. 11

[Kal89] ERICH L. KALTOFEN. *Factorization of polynomials given by straight-line programs*. In SILVIO MICALI, ed., *Randomness and Computation*, volume 5 of *Advances in Computing Research*, pages 375–412. JAI Press, Greenwich, Connecticut, 1989. 5, 11

[Kop14] SWASTIK KOPPARTY. *Some remarks on multiplicity codes*. In ALEXANDER BARG and OLEG R. MUSIN, eds., *Discrete Geometry and Algebraic Combinatorics*, volume 625 of *Contemporary Mathematics*, pages 155–176. AMS, 2014. arXiv:1505.07547. 11

[Kop15] ———. *List-decoding multiplicity codes*. Theory of Computing, 11:149–182, 2015. eccc:2012/TR12-044. 11

[KSS15] SWASTIK KOPPARTY, SHUBHANGI SARAF, and AMIR SHPILKA. *Equivalence of polynomial identity testing and polynomial factorization*. Comput. Complexity, 24(2):295–331, 2015. (Preliminary version in *29th CCC*, 2014). eccc:2014/TR14-001. 11

[LN96] RUDOLF LIDL and HARALD NIEDERREITER. *Finite Fields*, volume 2 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2nd edition, 1996. 36

[MR08]   Dana Moshkovitz and Ran Raz. *Sub-constant error low degree test of almost-linear size*. SIAM J. Comput., 38(1):140–180, 2008. (Preliminary version in *38th STOC*, 2006). `eccc: 2005/TR05-086`. 4

[MR10a]   ———. *Sub-constant error probabilistically checkable proof of almost-linear size*. Comput. Complexity, 19(3):367–422, 2010. `eccc:2007/TR07-026`. 4

[MR10b]   ———. *Two-query PCP with subconstant error*. J. ACM, 57(5), 2010. (Preliminary version in *49th FOCS*, 2008). `eccc:2008/TR08-071`. 4

[MZ23]   Dor Minzer and Kai Zheng. *Approaching the soundness barrier: A near optimal analysis of the cube versus cube test*. In Nikhil Bansal and Viswanath Nagarajan, eds., *Proc. 34th Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 2761–2776. 2023. `arXiv: 2211.09341`. 6

[Ore22]   Øystein Ore. *Über höhere kongruenzen (German) [About higher congruences]*. Norsk Mat. Forenings Skrifter, 1(7):15, 1922. (see [LN96, Theorem 6.13]). 14

[PS94]   Alexander Polishchuk and Daniel A. Spielman. *Nearly-linear size holographic proofs*. In Frank Thomson Leighton and Michael T. Goodrich, eds., *Proc. 26th ACM Symp. on Theory of Computing (STOC)*, pages 194–203. 1994. 3, 4, 6, 7, 9, 24

[RS96]   Ronitt Rubinfeld and Madhu Sudan. *Robust characterizations of polynomials with applications to program testing*. SIAM J. Comput., 25(2):252–271, April 1996. (Preliminary version in *23rd STOC*, 1991 and *3rd SODA*, 1992). 3, 5, 6, 11, 24

[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 Frank Thomson Leighton and Peter W. Shor, eds., *Proc. 29th ACM Symp. on Theory of Computing (STOC)*, pages 475–484. 1997. 3, 4, 5, 6, 11

[Sch80]   Jacob T. Schwartz. *Fast probabilistic algorithms for verification of polynomial identities*. J. ACM, 27(4):701–717, October 1980. 14

[Ska13]   Matthew Skala. *Hypergeometric tail inequalities: ending the insanity*, 2013. (manuscript). `arXiv:1311.5939`. 38

[STV01]   Madhu Sudan, Luca Trevisan, and Salil P. Vadhan. *Pseudorandom generators without the XOR lemma*. J. Comput. Syst. Sci., 62(2):236–266, 2001. (Preliminary version in *31st STOC*, 1999). 6, 11, 22

[Zip79]   Richard Zippel. *Probabilistic algorithms for sparse polynomials.* In Edward W. Ng, ed., *Proc. International Symp. of Symbolic and Algebraic Computation (EUROSAM)*, volume 72 of *LNCS*, pages 216–226. Springer, 1979. 14

A An Arora-Sudan style interpolation

In this section, we present the full proof of Theorem 4.3. We repeat the statement below for convenience.

Theorem 4.3 (Interpolation for LDT). *There exist constant $c_1, c_2 \in \mathbb{N}$ such that for every $\varepsilon \in (0, 1]$, any $d \in \mathbb{N}$ and finite field $\mathbb{F}$ of size q and characteristic p with $q > c_1 \cdot d/\varepsilon^{c_2}$, the following is true.*

Let $f : \mathbb{F}^2 \to \mathbb{F}$ be an oracle that passes LDT_d with probability ε. Then, there is a non-zero polynomial $A(x, y, z)$ and a subset $S \subseteq \mathbb{F}^2$ such that

- $\deg_{1,1,d}(A) = O(d/\varepsilon^2)$ *and* $|S| = \Omega(\varepsilon^2 q^2)$,
- $\sum_{x \in S} \varepsilon_x = \Omega(\varepsilon^3 \cdot q^2)$,
- *for every* $(a, b) \in S$, *we have* $\varepsilon_{(a,b)} = \Pr_{\ell \ni (a,b)}[P_\ell(a, b) = f(a, b)] \geq \varepsilon/2$,
- *for every* $(a, b) \in S$, *we have* $A(a, b, f(a, b)) = 0$.
- $\partial_z(A)$ *and* $\mathsf{Disc}_z(A)$ *are not identically zero.*

The proof is mostly along the lines of Arora and Sudan [AS03] but a tighter analysis and some care for fields of small characteristic. We present it in its entirety for completeness and to make the changes clearer to follow.

A.1 Finding a structured subset of points

We start with the following claim.

Claim A.1 (Finding good directions). *Let $q > 8/\varepsilon$ and suppose $f : \mathbb{F}^2 \to \mathbb{F}$ passes LDT_d with probability ε. Then, there are two different directions ℓ_1 and ℓ_2, and a set $H \subseteq \mathbb{F}^2$ such that*

- $|H| = \Omega(\varepsilon^2 \cdot q^2)$,
- *for all $x \in H$ we have $\Pr_{\ell \ni x}[P_\ell(x) = f(x)] \geq \varepsilon/2$ and*

$$P_{\ell_1^{(x)}}(x) = P_{\ell_2^{(x)}}(x) = f(x)$$

  where $\ell_i^{(x)}$ is the line through x parallel to ℓ_i.

Proof. For each $x \in \mathbb{F}^2$, let $\varepsilon_x = \Pr_{\ell \ni x}[P_\ell(x) = f(x)]$ and we have $\sum_x \varepsilon_x = \varepsilon \cdot q^2$. Let $H' = \{x : \varepsilon_x \geq \varepsilon/2\}$. For a direction ℓ and a point $x \in \mathbb{F}^2$, let $I(x, \ell)$ be indicator random variable defined as

$$I(x, \ell) = \mathbb{1}\left[P_{\ell^{(x)}}(x) = f(x) \text{ and } \varepsilon_x \geq \varepsilon/2\right]$$

43

where $\ell^{(x)}$ is the line through x parallel to ℓ. We then have $\mathbb{E}_\ell[I(x,\ell)] = \varepsilon_x$ if $x \in H'$ and 0 otherwise. Thus,

$$\mathbb{E}_x \mathbb{E}_\ell[I(x,\ell)] = 1/q^2 \cdot \sum_{x \in H'} \varepsilon_x \geq \varepsilon/2 \quad \text{(by Lemma 2.1)}.$$

Therefore, we have

$$\varepsilon/2 \leq \mathbb{E}_x \mathbb{E}_\ell[I(x,\ell)]$$

$$\varepsilon^2/4 \leq \left(\mathbb{E}_x \mathbb{E}_\ell[I(x,\ell)] \right)^2 \leq \mathbb{E}_x (\mathbb{E}_\ell[I(x,\ell)])^2$$

$$= \mathbb{E}_x \left(\mathbb{E}_{\ell_1,\ell_2}[I(x,\ell_1) \cdot I(x,\ell_2)] \right)$$

$$= \mathbb{E}_x \left(\mathbb{E}_{\ell_1 \neq \ell_2}[I(x,\ell_1) \cdot I(x,\ell_2)] + \frac{1}{q^2+q} \right) \qquad (\because \Pr[\ell = \ell'] = \frac{1}{q^2+q})$$

$$\implies \varepsilon^2/8 \leq \mathbb{E}_x \mathbb{E}_{\ell_1 \neq \ell_2}[I(x,\ell_1) \cdot I(x,\ell_2)] \qquad (\because q > 8/\varepsilon)$$

Therefore, there exist two different directions ℓ_1 and ℓ_2 such that $\mathbb{E}_x[I(x,\ell_1)I(x,\ell_2)] \geq \varepsilon^2/8$. Fixing such directions ℓ_1 and ℓ_2, defining $H = \{x : I(x,\ell_1) = I(x,\ell_2) = 1\}$ satisfies the requirements. $\square$

Without of loss of generality, we assume that ℓ_1 is the x-axis and ℓ_2 is the y-axis. Let γ be chosen such that $H = 2\gamma \cdot q^2 = \Omega(\varepsilon^2 q^2)$.

Lemma A.2 (Structured subsets within H). *Let H be the set specified above with respect to the x and y directions for an oracle f passing LDT_d with probability at least ε. Then, for any r satisfying $\frac{2 \log q}{\gamma^2} \leq r \leq \gamma \cdot q$, there are subsets $S_1, S_2 \subseteq \mathbb{F}$ such that*

1. $|S_1| = r$,

2. $|S_2| \geq \gamma \cdot q$,

3. $|(\mathbb{F} \times S_2) \cap H| \geq |H|/2 = \gamma \cdot q^2$,

4. *For all $b \in S_2$, we have $|\{(a,b) : a \in \mathbb{F}\} \cap H| \geq \gamma \cdot q$,*

5. *For all $b \in S_2$, we have $|\{(a,b) : a \in S_1\} \cap H| \geq \gamma/2 \cdot |S_1|$.*

Proof. Let $\ell_{y=b} = \{(a,b) : a \in \mathbb{F}\}$ and let $\ell_{x=a} = \{(a,b) : b \in \mathbb{F}\}$. Since these lines evenly cover the space, we have that $\mathbb{E}_b[|\ell_{y=b} \cap H|] = 2\gamma \cdot q$. Let $S_2 = \{b \in \mathbb{F} : |\ell_{y=b} \cap H| \geq \gamma q\}$. By Lemma 2.1, we have that $|S_2| \geq \gamma \cdot q$ and $|(\mathbb{F} \times S_2) \cap H| \geq |H|/2 = \gamma \cdot q^2$.

Let S_1 be a set of r distinct elements of $\mathbb{F}$ chosen uniformly at random. Then, for any $b \in S_2$

$$\mathbb{E}_{S_1}[|\{(a,b) : a \in S_1 , (a,b) \in H\}|] = r/q \cdot |\ell_{y=b} \cap H| \geq \gamma \cdot r.$$

44

From standard tail bounds for hypergeometric distributions (cf. [Ska13]), we have

$$\text{For all } b \in S_2, \Pr_{S_1}\left[|\{(a,b) \ : \ a \in S_1 \ , \ (a,b) \in H\}| < \gamma r/2\right] \leq \exp(-\gamma^2 r/2)$$
$$\implies \Pr_{S_1}\left[\exists b \in S_2 \ : \ |\{(a,b) \ : \ a \in S_1 \ , \ (a,b) \in H\}| < \gamma r/2\right] \leq |S_2| \cdot \exp(-\gamma^2 r/2) < 1.$$

Thus, there exists a set S_1 of size r such that for every $b \in S_2$ we have

$$|\{(a,b) \ : \ a \in S_1 \ , \ (a,b) \in H\}| \geq \gamma r/2. \qquad \qquad \square$$

A.2 Properties of the desired interpolating polynomial

As preparation of the interpolating step, we start by describing the desired monomial support of the interpolating polynomials and observing some of their properties.

Definition A.3. *For $n, N, p \in \mathbb{N}$ such that $p \geq 2$, the sets $N_{d,D}$ and $N_{d,D,p}$ are defined as follows.*

- *$N_{d,D}$ is the set of exponent vectors of trivariate monomials with $(1,1,d)$-degree bounded by D, i.e.,*

$$N_{d,D} := \left\{(i,j,k) \in \mathbb{Z}_{\geq 0}^3 \ : \ i + j + dk \leq D\right\} .$$

- *$N_{d,D,p}$ is the set of exponent vectors of trivariate monomials $x^i y^j z^k$ with $(1,1,d)$-weighted degree at most D such that the z degree is either zero, or is not a multiple of p. More formally,*

$$N_{d,D,p} := \left\{(i,j,k) \in \mathbb{Z}_{\geq 0}^3 \ : \ i + j + dk \leq D \text{ and } (k \text{ is either zero or } p \nmid k)\right\}. \qquad \Diamond$$

The following simple claims now give us bounds on the sizes of the sets $N_{d,D}$ and $N_{d,D,p}$. We defer the proofs of the claims to Appendix A.4.

Claim A.4. *For all $d, D \in \mathbb{N}$, we have*

$$(D^3/3d - 5/2D^2 + Dd/6) \leq |N_{d,D}| \leq (D^3/3d + 3/2D^2 + Dd/6).$$

Claim A.5. *For all $d, D \in \mathbb{N}$, with $p \geq 2$, we have $|N_{d,D,p}| \geq |N_{d,D}|/2$.*

Finally, we combine the bounds in Claim A.4 and Claim A.5 to get the following claim.

Claim A.6. *For all $d, D, p \in \mathbb{N}$ such that $D > 20d$ and $p \geq 2$, $|N_{d,D,p}| \geq \frac{D^3}{12d}$.*

Interpolation

We now rely on the above estimates to interpolate a low-degree polynomial that explains the function value in the points table on a constant fraction of points. But first, we recall some

notation: H, S_1, S_2 be as guaranteed by Lemma A.2 and let r be the size of S_1. Define $S = \{(a,b) \in \mathbb{F}^2 : b \in S_2, (a,b) \in H\}$. From Lemma A.2, we have that $|S| \geq \gamma \cdot q^2$. With this notation in place, we have the following lemma.

Lemma A.7. *Let $\mathbb{F}$ be a finite field of characteristic p and let D be a natural number satisfying $|N_{d,D,p}| > r(D+1)$. There exists a non-zero polynomial $A(x,y,z) \in \mathbb{F}[x,y,z]$ with $\deg_{1,1,d}(A) \leq D$ such that*

- *For every $a \in S_1$, we have*

$$A(a, y, P_{\ell_{x=a}}(y)) = 0$$

  as a univariate polynomial in y, where $P_{\ell_{x=a}}$ is the best-fit degree-d polynomial for f the line $\ell_{x=a}$.

- *If p is the characteristic of the field $\mathbb{F}$, then, A is supported on monomials of the form $x^i y^j z^k$ where k is either zero or is not divisible by p.*

Proof. Let $A(x,y,z) = \sum_{i,j,k \in N_{d,D}} A_{i,j,k} x^i y^j z^k$, where $A_{i,j,k}$ are indeterminates. For each $a \in S_1$, the constraint $A(a, y, P_{\ell_{x=a}}(y)) = 0$ can be viewed as $(D+1)$ homogeneous linear constraints on the indeterminates $A_{i,j,k}$ obtained by equating the coefficients of $y^0, y^1, \ldots, y^D$ to zero. Note that since the $(1,1,d)$-weighted degree of every monomial in A is at most D, the degree of $A(a, y, P_{\ell_{x=a}}(y))$ is at most D. Hence, as long as $|N_{d,D,p}| > r \cdot (D+1)$, there must exist a non-zero solution to the system of equations, and therefore a non-zero polynomial $(1,1,d)$-degree D that is entirely supported on monomials from the set $N_{d,D,p}$ and therefore satisfies the second item in the lemma by definition. $\square$

We now infer that any polynomial satisfying the conditions in Lemma A.7 imply that it satisfies many more vanishing conditions.

Lemma A.8. *Suppose $A(x,y,z)$ is a polynomial with $\deg_{1,1,d}(A) \leq D$ such that for all $a \in S_1$ we have $A(a, y, P_{\ell_{x=a}}(y)) = 0$. If $D < \gamma/2 \cdot |S_1|$, then we have that*

$$A(a, b, f(a,b)) = 0 \quad \text{for all } (a,b) \in S = \{(a,b) : b \in S_2, (a,b) \in H\}.$$

Proof. Fix a $b \in S_2$ and let $R_b(x) = P_{\ell_{y=b}}$ and $Q(x) = A(x, b, R_b(x))$. Suppose $a \in S_1$ with $(a,b) \in H$, then $R_b(a) = P_{\ell_{x=a}}(b) = f(a,b)$ and hence

$$A(a, b, R_b(a)) = A(a, b, P_{\ell_{x=a}}(a)) = 0.$$

Since $\deg(Q) \leq \deg_{1,1,d}(A) \leq D$, and $|\{(a, b) \ : \ a \in S_1\} \cap H| \geq \gamma/2 \cdot |S_1| = \gamma r/2$, the condition that $D < \gamma r/2$ implies that $Q(x)$ is identically zero. Since $Q(x) = A(x, b, R_b(x)) = 0$, for any a such that $(a, b) \in H$, we have that $A(a, b, R_b(a)) = A(a, b, f(a, b)) = 0$ as claimed. □

A.3 Proof of the interpolation lemma

We now have the necessary ingredients to prove Theorem 4.3.

Proof of Theorem 4.3. Let H be the set guaranteed by Claim A.1 and $\gamma \in (0, 1]$ such that $H = 2\gamma \cdot q^2 = \Omega(\varepsilon^2 \cdot q^2)$; without loss of generality let the two directions guaranteed by Claim A.1 be the standard axes.

Let $r = 900d/\gamma^2$ and $D = \gamma r/3 = 300d/\gamma$. From these choice of parameters, we have the following inequalities.

- $D > 20d$

- $|N_{d,D,p}| \geq D^3/12d = \frac{(300^3 d^2)}{12\gamma^3}$

- $r(D + 1) \leq 2rD = \frac{6(300)^2 d^2}{\gamma^3}$

- $\gamma r/2 = \frac{450d}{\gamma}$

Thus, we have that $D \geq 20d$, $|N_{d,D,p}| > r(D + 1)$ and $D < \gamma r/2$.

Now, Instantiating Lemma A.2 with this choice of r, we obtain sets S_1, S_2 and let S be defined as

$$S = \left\{ (a, b) \in \mathbb{F}^2 \ : \ b \in S_2 \, , \ (a, b) \in H \right\} .$$

By construction, $|S| \geq \gamma \cdot q^2 = \Omega(\varepsilon^2 \cdot q^2)$. Since $S \subseteq H$ and each $x \in S$ satisfies $\varepsilon_x \geq \varepsilon/2$ (from Claim A.1), we have that $\sum_{x \in S} \varepsilon_x = \Omega(\varepsilon^3 \cdot q^2)$.

Since $D^3/12d > r(D + 1)$ and $D < \gamma r/2$, we have by Lemma A.7 and Lemma A.8 that there is a non-zero polynomial $\tilde{A}(x, y, z)$ with $\deg_{1,1,d}(\tilde{A}) \leq D$, entirely supported on monomials in the set $N_{d,D,p}$ such that $\tilde{A}(a, b, f(a, b)) = 0$ for all $(a, b) \in S$. In order to proceed further in the proof, we need the following claim whose proof we defer to the end of this section.

Claim A.9. *The polynomial $\tilde{A}$ depends on the variable z, i.e., there exists a monomial of the form $x^i y^j z^k$ with non-zero coefficient such that k is non-zero.*

Since $\tilde{A}$ depends on z and is only supported on monomials where the z degree is either zero or is not a multiple of p, we have from the first item of Proposition 2.4 and the linearity of Hasse derivatives that $\partial_z(\tilde{A})$ must be non-zero. We take $A(x, y, z)$ to be a non-zero polynomial of the minimum $(1, 1, d)$-weighted degree such that it vanishes on $(a, b, f(a, b))$ for all $(a, b) \in S$ and $\partial_z(A)$ is non-zero. We know that the polynomial $\tilde{A}$ interpolated above is one such polynomial. Therefore,

$\deg_{(1,1,d)}(A) \leq \deg_{(1,1,d)}(\tilde{A}) \leq D$. The polynomial A thus obtained satisfies the hypothesis of Lemma 2.9, and therefore, by Lemma 2.9, we get that $\mathsf{Disc}_z(A)$ must be non-zero.

This proves all the properties of A and S claimed by Theorem 4.3. □

To complete the proof of Theorem 4.3, we now prove Claim A.9.

Proof of Claim A.9

Proof of Claim A.9. If $\tilde{A}$ does not depend on z at all, then it is a non-zero bivariate polynomial of $(1, 1)$-weighted degree, i.e., total degree at most D, and it vanishes at all points (a, b) in the set $S \subseteq \mathbb{F} \times \mathbb{F}$. But, from Lemma 2.2, we have that the number of zeroes of $\tilde{A}$ on $\mathbb{F} \times \mathbb{F}$ can be at most Dq, which by the choice of D is at most $300/\gamma \cdot dq \leq O\left(\frac{dq}{\varepsilon^2}\right)$. We also have that $|S| > \Omega(\varepsilon^2 q^2)$. Thus, for any sufficiently large constants c_1, c_2, we have that if $q > c_1 \cdot d/\varepsilon^{c_2}$, then, $|S|$ exceeds $O(dq/\varepsilon^2)$, thereby implying that $\tilde{A}$ must be identically zero, which is a contradiction. Thus, $\tilde{A}$ must depend on z. □

A.4   Proofs of Claim A.4, Claim A.5 and Claim A.6

Proof of Claim A.4. From the definition, we have that

$$
\begin{aligned}
|N_{d,D}| &\geq \sum_{k=0}^{\lfloor D/d \rfloor} \sum_{i=0}^{(D-kd)} \sum_{j=0}^{(D-kd-i)} 1 \,, \\
&\geq \sum_{k=0}^{\lfloor D/d \rfloor} (D - kd)(D - kd + 1)/2 \,, \\
&\geq 1/2 \cdot \sum_{k=0}^{\lfloor D/d \rfloor} (D - kd)^2 \,, \\
&\geq 1/2 \cdot \sum_{k=0}^{\lfloor D/d \rfloor} (D^2 + k^2 d^2 - 2Ddkd) \,, \,, \\
&\geq 1/2 \cdot \left(\left(\sum_{k=0}^{\lfloor D/d \rfloor} D^2 \right) + \left(\sum_{k=0}^{\lfloor D/d \rfloor} k^2 d^2 \right) - \left(\sum_{k=0}^{\lfloor D/d \rfloor} 2Ddk \right) \right)
\end{aligned}
$$

For ease of notation, let $\tilde{D}$ denote $\lfloor D/d \rfloor$. So, we have $(D/d - 1) \leq \tilde{D} \leq (D/d)$. Now, from standard estimates on sums and sums or squares of the first n natural numbers, we have the following.

$$
\sum_{k=0}^{\tilde{D}} D^2 \geq D^2 \cdot (\tilde{D} + 1) \in [D^3/d, D^3/d + D^2] \,.
$$

Similarly,

$$\sum_{k=0}^{\tilde{D}} k^2 d^2 = d^2 \cdot \left(\sum_{k=0}^{\tilde{D}} k^2 \right) \geq d^2 \cdot (\tilde{D})(\tilde{D}+1)(2\tilde{D}+1)/6 \, .$$

Plugging in the bounds for $\tilde{D}$, we get that

$$\sum_{k=0}^{\tilde{D}} k^2 d^2 \in [1/6(2D^3/d - 3D^2 + Dd), 1/6(2D^3/d + 3D^2 + Dd)] \, .$$

Finally,

$$\sum_{k=0}^{\tilde{D}} 2Ddk = 2Dd \cdot \tilde{D}(\tilde{D}+1)/2 \, ,$$

which, again using the upper and lower bounds on $\tilde{D}$ satisfies

$$\sum_{k=0}^{\tilde{D}} 2Ddk \in [D^3/d - D^2, D^3/d + D^2] \, .$$

Putting the estimates together, we get the claim. □

Proof of Claim A.5. For any $d, D \in \mathbb{N}$, and $k \in \mathbb{Z}_{\geq 0}$, let A_k be the set defined as

$$A_k = \left\{ (i, j) \in \mathbb{Z}_{\geq 0}^2 \ : \ i + j + dk \leq D \right\} \, .$$

Clearly, the size of A_k is a non-increasing function of k. Thus, for every integer $\ell \geq 1$ and for every $j \in \{1, 2, \ldots, p-1\}$ we have that $|A_{\ell p}| < |A_{(\ell-1)p+j}|$, thereby implying that

$$|A_{\ell p}| < \frac{1}{p} \left(\sum_{j=1}^{p} |A_{(\ell-1)p+j}| \right) \, .$$

Summing over ℓ, we get that

$$\sum_{\ell=1}^{\lfloor D/d \rfloor} |A_{\ell p}| < \frac{1}{p} \left(\sum_{\ell=1}^{\lfloor D/d \rfloor} \sum_{j=1}^{p} |A_{(\ell-1)p+j}| \right) \, .$$

We now note from the definitions of the sets $N_{d,D}$ and A_k that $|N_{d,D}| = \sum_{k=0}^{\lfloor D/d \rfloor} |A_k|$ and, thus, we

have that

$$\sum_{\ell=1}^{\lfloor D/d \rfloor} |A_{\ell p}| < \frac{1}{p}|N_{d,D}|.$$

Moreover, the size of the set $N_{d,D,p}$ satisfies

$$|N_{d,D,p}| \geq |N_{d,D}| - \sum_{\ell=1}^{\lfloor D/d \rfloor} |A_{\ell p}| \geq \left(1 - \frac{1}{p}\right)|N_{d,D}|,$$

which, for $p \geq 2$ gives $|N_{d,D,p}| \geq \frac{1}{2}|N_{d,D}|$. $\qquad\square$

Proof of Claim A.6. From the lower bound on $|N_{d,D}|$ in Claim A.4, we have that $|N_{d,D}| \geq D^3/3d - 5/2D^2$. Now, if $D/d > 20$, then $|N_{d,D}| \geq D^2(D/3d - 5/2) \geq D^3/6d$. Combining this with the bound in Claim A.5 completes the proof. $\qquad\square$

B   Variations on Low-degree Testing

In this section we show the equivalence of some standard variations of low-degree testing. These results are essentially folklore - our proofs follow closely the proofs from Arora and Sudan [AS03, Section 2.2], confirming along the way that the proofs continue to hold even when the field size is only linear in the degree.

For a function $f\colon \mathbb{F}^m \to \mathbb{F}$ and line ℓ in $\mathbb{F}^m$, let $P_\ell^{(f,d)}$ be the best fit degree d univariate polynomial for f on the line ℓ. We now, define the following quantities for any function $f\colon \mathbb{F}^m \to \mathbb{F}$, line ℓ and plane π

$$\delta_f(\ell) := \Pr_{x \in \ell}[P_\ell^{(f,d)}(x) \neq f(x)],$$

$$\delta_f(\pi) := \mathbb{E}_{\ell \in \pi}[\delta_f(\ell)] = \Pr_{\substack{\ell \in \pi \\ x \in \ell}}[P_\ell^{(f,d)}(x) \neq f(x)],$$

$$\delta_f := \mathbb{E}_\pi[\delta_f(\pi)] = \Pr_{\substack{\ell \in \mathbb{F}^m \\ x \in \ell}}[P_\ell^{(f,d)}(x) \neq f(x)].$$

As mentioned in Remark 4.1, the general low-degree test is provided two oracles $f : \mathbb{F}_q^m \to \mathbb{F}_q$ and $P : \mathscr{L}^{(m)} \to \mathbb{F}_{\bar{q}}^{\leq d}[t]$, the first mapping points in $\mathbb{F}_q^m$ to values in $\mathbb{F}_q$ and the latter mapping lines in $\mathbb{F}_q^m$ to (a table of evaluations of) a degree $\leq d$ univariate polynomial. The probability that the low-degree test accepts is given by $\Pr_{x,\ell \ni x}[f(x) = P[\ell](x)]$.

We say x is β-good for (f, P) if $\Pr_{\ell \ni x}[P[\ell](x) = f(x)] \geq \beta$. Note that if the low-degree test accepts (f, P) with probability β then at least $(\beta/2)$ fraction of the points are $(\beta/2)$-good for (f, P). Conversely, if the low-degree test accepts (f, P) with probability at most $\beta = \beta_1\beta_2$ then at most β_1

fraction of the points are β_2-good for (f, P) (for any choice of β_1, β_2 satisfying $\beta = \beta_1\beta_2$).

Definition B.1 (Weak form of LDT). *Given field $\mathbb{F}_q$, integer parameters $m \geq 1$, $d \geq 0$, real $\beta > 0$ and function $h : \mathbb{R}^{>0} \to \mathbb{R}^{>0}$, we say that* weak low-degree testing *holds for $(\mathbb{F}_q, m, d, \beta_0, h)$ if the following is true:*

> *For all $f : \mathbb{F}_q^m \to \mathbb{F}_q$ and $P : \mathscr{L}^{(m)} \to \mathbb{F}_q[t]^{\leq d}$ and all $\beta > \beta_0$, if (f, P) pass the low-degree test with probability at least β, then there exists some degree d polynomial Q such that $\mathrm{agree}(f, Q) \geq h(\beta)$.*

$\diamond$

Note that Theorem 4.2 asserts that there are universal constants C, α such that weak low-degree testing holds for $(\mathbb{F}_q, 2, d, \beta_0, h)$ for $h : \beta \mapsto \alpha\beta^4$ provided $q \geq Cd/\beta_0^7$.

There are two incomparable ways to strengthen a weak low-degree test and we first define the *list-decoding* variant.

Definition B.2 (List-decoding form of the LDT). *Given a field $\mathbb{F}_q$, integer parameters $m \geq 1$, $d \geq 0$, real $\beta_0 > 0$ and function $h_1 : \mathbb{R}^{>0} \to \mathbb{R}^{>0}$, we say that* list-decoding low-degree testing *holds for $(\mathbb{F}_q, m, d, \beta_0, h_1)$ if the following is true:*

> *For all $\beta > \beta_0$, $f : \mathbb{F}_q^m \to \mathbb{F}_q$ and $P : \mathcal{L}^{(m)} \to \mathbb{F}_d[t]^{\leq d}$, there is a (possibly empty) list of at most $C \leq \frac{2}{h_1(\beta)}$ m-variate degree d polynomials $Q_1, \ldots, Q_C$ such that $\mathrm{agree}(f, Q_i) \geq h_1(\beta)$ for each i, and*
>
> $$\Pr_{x \in \mathbb{F}_q^m} [x \text{ is } \beta\text{-good for } (f, P) \text{ and } f(x) \notin \{Q_1(x), \ldots, Q_C(x)\}] \leq \beta_0.$$

The above must hold regardless of the LDT acceptance probability of (f, P). Indeed, if (f, P) passes the LDT with probability less than $\beta \cdot \beta_0$, then the list of polynomial may be empty as there can be at most β_0 fraction of β-good points for (f, P). $\diamond$

Another strengthening of the weak LDT is in terms of the agreement probability.

Definition B.3 (High-agreement form of the low-degree test). *Given a finite field $\mathbb{F}_q$, integer parameter $m \geq 1$, $d \geq 0$, real $\beta_0 > 0$ and a function $h_2 : \mathbb{R}^{>0} \to \mathbb{R}^{>0}$, we say that $(\mathbb{F}_q, m, d, \beta_0, h_2)$ if the following is true:*

> *For every $\beta > \beta_0$, $f : \mathbb{F}_q^m \to \mathbb{F}_q$ and $P : \mathcal{L}^{(m)} \to \mathbb{F}_q[t]^{\leq d}$ such that (f, L) passes the low-degree test with probability at least β, then there exists an m-variate degree d polynomial Q such that $\mathrm{agree}(f, Q) \geq \beta - h_2(\beta_0)$.*

$\diamond$

In our applications we will assume h, h_1, h_2 are monotone non-decreasing functions. We will see that a weak low-degree test with $h(\beta) \to 0$ as $\beta \to 0$ imply both the list-decoding variant and the

high-agreement variant for some appropriate functions h_1, h_2 that also satisfy $h_1(\beta), h_2(\beta) \to 0$ as $\beta \to 0$.

In our applications we will assume h, h_1, h_2 are monotone non-decreasing functions. Further we assume $h(\beta) \to 0$ as $\beta \to 0$ and show that $h_1(\beta) \to 0$ and $h_2(\beta) \to 0$ as $\beta \to 0$. Note that the two implications above are incomparable and do not directly imply each other. However as we will see in the proof, the implication in (1) is useful to prove (2).

Lemma B.4 (Weak LDT implies list-decoding LDT). *Suppose $\mathbb{F}_q, m, d, \beta_0$ and $h(\cdot)$ are such that weak low-degree testing holds for $(\mathbb{F}_q, m, d, \beta_0, h)$. Then list-decoding low-degree testing holds for $(\mathbb{F}_q, m, d, \beta_0', h_1)$ for*

- *β_0' satisfying $\beta_0' > \sqrt{\beta_0}$ and $h(\beta_0'^2) \geq \max \left(\frac{e}{\sqrt{4q}}, \frac{2(d+1)}{q}, \sqrt{\frac{d}{q}} \right)$,*

- *$h_1 : \beta \mapsto 1/2 \cdot h(\beta \cdot \beta_0)$.*

Proof. Fix any $\beta > \beta_0'$. Let $Q_1, \ldots, Q_C$ be the set of all m-variate degree d polynomials that have agreement at least $\eta = 1/2 \cdot h(\beta \cdot \beta_0') > 1/2 \cdot h(\beta_0'^2)$ with f on $\mathbb{F}_q^m$. By the Johnson bound, we have that $C \leq 2/\eta$ since $\eta \geq 2\sqrt{d/q}$ by the choice of β_0'.

Consider the following randomly chosen function $g : \mathbb{F}_q^m \to \mathbb{F}_q$ given by $g(x) = f(x)$ if $x \notin \cup_i S_i$ and $g(x) \sim \text{Unif}(\mathbb{F}_q)$ if $x \in \cup_i S_i$. Claim B.5 below asserts that with positive probability we have that $\text{agree}(g, Q) < 2\eta$ for every degree d polynomial Q. Fix a g such that this holds. By the weak low-degree test applied to g (in contrapositive form) it follows that (g, P) pass the low-degree test with probability at most $\beta \cdot \beta_0'$ since $\beta \cdot \beta_0' > \beta_0'^2 \geq \beta_0$ and $h(\beta \cdot \beta_0') = 2\eta$. Therefore, $\Pr_{x \in \mathbb{F}_q^m}[x \text{ is } \beta\text{-good for } (g, L)] < \beta_0'$. But now note that $x \notin \cup_i S_i$ satisfies $f(x) = g(x)$ and so such an x is τ-good for (g, P) iff it is β-good for (f, P). We conclude that

$$\Pr_{x \in \mathbb{F}_q^m}[x \text{ is } \beta\text{-good for } (f, L) \text{ and } f(x) \notin \{P_1(x), \ldots, P_C(x)\}]$$

$$= \Pr_{x \in \mathbb{F}_q^m}[x \text{ is } \beta\text{-good for } (g, L) \text{ and } f(x) \notin \{P_1(x), \ldots, P_C(x)\}]$$

$$= \Pr_{x \in \mathbb{F}_q^m}[x \text{ is } \beta\text{-good for } (g, L)] \leq \beta_0'.$$

To complete the proof, we only need to prove the following claim.

Claim B.5. *If $\eta \geq \max\{\frac{e}{\sqrt{q}}, \frac{4(d+1)}{q}\}$ then*

$$\Pr[\exists Q : \mathbb{F}_q^m \to \mathbb{F}_q , \ \deg(Q) \leq d , \ \text{agree}(g, Q) \geq 2\eta] \leq q^{-(\eta/4)q^m} < 1.$$

Proof. Fix a degree d polynomial Q. If $Q \notin \{Q_1, \ldots, Q_C\}$ then $\text{agree}(Q, f) < \eta$ and if $Q \in \{Q_1, \ldots, Q_C\}$ we have for every $x \notin \cup_i S_i$, $Q(x) \neq f(x) = g(x)$. Thus in either

52

case we have $|\{x \notin \cup_i S_i | Q(x) = g(x) = f(x)\}| \le \eta q^m$. Thus to have agreement at least 2η with Q, g must satisfy $|\{x \in \cup_i S_i | P(x) = g(x)\}| \ge \eta q^m$. We show below that the probability that this happens is at most $q^{-(\eta/2)q^m}$.

For a fixed set $S \subseteq \cup_i S_i$ of size ηq^m, the probability that g and Q agree on the set is $q^{-\eta q^m}$. The number of sets S is at most $\binom{q^m}{\eta q^m} \le (e/\eta)^{\eta q^m} \le q^{(\eta/2)q^m}$ (using $\eta \ge e/\sqrt{q}$). We conclude that the probability that there exists a set $S \subseteq \cup_i S_i$ of size at least ηq^m such that g and Q agree on S is at most $q^{-(\eta/2)q^m}$.

Now to conclude the proof we take a union bound over all Q's. The number m variate monomials of degree at most d is clearly at most $(d+1)^m$ and so the number of polynomials is at most $q^{(d+1)^m} \le q^{((\eta/4)q)^m} \le q^{(\eta/4)q^m}$ (where the first inequality uses $\eta \ge 4(d+1)/q$ and the second uses $m \ge 1$). We conclude that the probability that there exists Q of degree at most d such that $|\{x \in \cup_i S_i | Q(x) = g(x)\}| > \eta q^m$ is at most $q^{-(\eta/4)q^m}$. The claim follows. $\qquad\square$

This completes the proof of Lemma B.4. $\qquad\square$

Lemma B.6 (Weak LDT implies high-agreement LDT). *Suppose $\mathbb{F}_q, m, d, \beta_0$ and $h(\cdot)$ are such that weak low-degree testing holds for $(\mathbb{F}_q, m, d, \beta_0, h)$. Then, high-agreement low-degree testing holds for $(\mathbb{F}_q, m, d, \beta_0'', h_2)$ for*

- *β_0'' satisfying $\beta_0'' > \beta_0'$ and $\beta_0''^3 \cdot h_1(\beta_0''^2) > 2/q$ and $\beta_0'' \cdot h_1(\beta_0''^2) \ge 2d/q$,*

- *$h_2 : \beta \mapsto 3\beta$,*

where β_0' and $h_1(\cdot)$ are as implied by Lemma B.4.

In other words, if (f, P) passes the LDT with probability $\beta > \beta_0''$, then there is some m-variate degree d polynomial Q such that $\mathrm{agree}(f, Q) \ge \beta - 3\beta_0''$.

Proof. By applying Lemma B.4 (with $\beta = \beta_0'' \ge \beta_0'$) we have that for any pair (f, P), there exist at most $C = 2/h_1(\beta_0'')$ polynomials $Q_1, \ldots, Q_C$, each with agreement at least $h_1(\beta_0'')$ with f such that

$$\Pr_{x \in \mathbb{F}_q^m} \left[x \text{ is } \beta_0''\text{-good for } (f, P) \text{ and } f(x) \notin \{Q_1(x), \ldots, Q_C(x)\} \right] \le \beta_0'.$$

Let $S_i = \{x : Q_i(x) = f(x)\}$ and let $\eta_i = q^{-m}|S_i|$. Assume w.l.o.g. that $\eta_1 \ge \cdots \ge \eta_C$. Note that we are not guaranteed that $C \ge 1$ and the above list of polynomials be empty. However, we wish to show that if (f, P) passes the LDT with probability $\beta \ge \beta_0''$, then $C \ge 1$ and $\eta_1 \ge \beta - \gamma$, where $\gamma = 3\beta_0''$. We will do so by proving that the acceptance probability of the low-degree test on (f, P) is upper bounded by $\eta_1 + \gamma$.

Define a line ℓ to be *standard* if $P[\ell] \in \{Q_1|_\ell, \ldots, Q_C|_\ell\}$. Say the ℓ is *abnormal* if there exists $i \in [C]$ such that $|\{x \in \ell : Q_i(x) = f(x)\}| \ge (\eta_1 + \beta_0'') \cdot q$. For a non-standard line ℓ say that a

point $x \in \ell$ is *coincidental* for ℓ if there exists $i \in [C]$ s.t. $Q_i(x) = P[\ell](x)$. Finally say that a pair (x, ℓ) with $x \in \ell$ is *unexplained* if $f(x) = P[\ell](x)$ and $f(x) \notin \{Q_1(x), \ldots, Q_C(x)\}$.

We now upper bound the probability that the low-degree test accepts a random pair (x, ℓ) by consider various cases. We first note that for the low-degree test to accept a pair (x, ℓ) at least one of the following must happen:

1. ℓ is standard and normal and the low-degree test accepts, or

2. ℓ is abnormal, or

3. ℓ is non-standard and x is coincidental for ℓ, or

4. (x, ℓ) is unexplained.

The typical case is Item 1 where ℓ is standard and normal. In this case the probability over x that the low-degree test accepts the pair (x, ℓ) is at most $\eta_1 + \beta_0''$.

For Item 2, the probability that a random line ℓ is abnormal is upper bounded by $\gamma_1 := C/(\gamma^2 q)$ by a Chebychev argument (for fixed $i \in C$ the expected fraction of agreement is η_i and a random line contains q pairwise independent random samples of points from $\mathbb{F}_q^m$). Since β_0'' satisfies $\beta_0''^3 \cdot h_1(\beta_0''^2) > 2/q$, we this probability is bounded by β_0''.

For Item 3, the probability that a point x on a non-standard line ℓ is coincidental for the line is at most Cd/q (for every $i \in [C]$ there are at most d points where $Q_i(x) = P[\ell](x)$). Since β_0'' satisfies $\beta_0'' \cdot h_1(\beta_0''^2) > 2d/q$, we this probability is bounded by β_0'' as well.

And finally for Item 4, the probability that a pair (x, ℓ) is unexplained is, by the list-decoding version of the LDT, at most $\beta_0' \leq \beta_0''$.

We thus conclude that (f, P) passes the low-degree test accepts with probability at most $\eta_1 + 3\beta_0''$. $\qquad\square$