Improved low-degree testing and its applications

Sanjeev Arora*
Princeton University

Madhu Sudan†
IBM Watson Research Center

Abstract

NP = PCP($\log n, 1$) and related results crucially depend upon the close connection between the probability with which a function passes a *low degree test* and the distance of this function to the nearest degree d polynomial. In this paper we study a test proposed by Rubinfeld and Sudan [RS93]. The strongest previously known connection for this test states that a function passes the test with probability δ for some $\delta > 7/8$ iff the function has agreement $\approx \delta$ with a polynomial of degree d. We present a new, and surprisingly strong, analysis which shows that the preceding statement is true for $\delta \ll 0.5$. The analysis uses a version of *Hilbert irreducibility*, a tool of algebraic geometry.

As a consequence we obtain an alternate construction for the following proof system: A constant prover 1-round proof system for NP languages in which the verifier uses $O(\log n)$ random bits, receives answers of size $O(\log n)$ bits, and has an error probability of at most $2^{-\log^{1-\epsilon} n}$. Such a proof system, which implies the NP-hardness of approximating Set Cover to within $\Omega(\log n)$ factors, has already been obtained by Raz and Safra [RazS96]. Our result was completed after we heard of their claim.

A second consequence of our analysis is a self tester/corrector for any buggy program that (supposedly) computes a polynomial over a finite field. If the program is correct only on δ fraction of inputs where $\delta = 1/|\mathrm{F}|^{\epsilon} \ll 0.5$, then the tester/corrector determines δ and generates $O(\frac{1}{\delta})$ values for every input, such that one of them is the correct output. In fact, our techniques yield something stronger: Given the buggy program, we can construct $O(\frac{1}{\delta})$ randomized programs such that one of them is correct on every input, with high probability.

*Supported by an NSF CAREER award and an Alfred P. Sloan Fellowship.arora@cs.princeton.edu.
†madhu@watson.ibm.com

1 Introduction

The use of algebraic techniques has recently led to new (probabilistic) characterizations of traditional complexity classes. These characterizations involve an interaction between an untrustworthy prover (or many provers) and a polynomial-time verifier. In MIP= NEXPTIME [BFL91], and NP = PCP($\log n, 1$) [AS92, ALMSS92] the verifier has to probabilistically verify the satisfiability of a boolean formula by reading very few bits in a "proof string" presented by a prover. In IP=PSPACE [LFKN92, Sh92] the verifier has to probabilistically verify tautologyhood of a quantified boolean formulae by interacting with a prover. All these results fundamentally rely on the same idea: the verifier first *arithmetizes* (or *algebraizes*) the boolean formula, which involves viewing a boolean assignment not as a sequence of bits but as values of a polynomial [LFKN92]. From then on, verifying satisfiability or tautologyhood involves verifying — using some efficient algebraic procedures — specific properties of a polynomial that has been provided by the prover.

In this paper we present an improved analysis of the *low degree test*, an algebraic procedure used in the result NP=PCP($\log n, 1$). We expect that this result to have many applications; some are already known. For example, the new analysis is known to lead to new characterizations of NP in terms of PCP, which in turn lead to improved results about the hardness of approximation. Recall that NP=PCP($\log n, 1$) implies the hardness of computing approximate solutions to many optimization problems such as CLIQUE [FGLSS91, AS92], CHROMATIC NUMBER and SET COVER [LY93], AND MAX-3SAT [ALMSS92]. For most of these problems it implies NP-hardness, but for some —most notably the problem of approximating SET COVER within a factor $O(\log n)$ and an entire set of problems in [ABSS93] — it is only known to imply *quasi-NP-hardness* (a *quasi-NP-hard* problem has a polynomial-time algorithm only if NP $\subseteq$ Time($n^{poly\log(n)}$)).

Plugging our improved analysis of the low degree test into known constructions leads to very efficient *constant-prover* 1-*round proof systems* for NP. Such systems imply the NP-hardness of approximating Set Cover to within a factor of $O(\ln n)$ (see the reduction of [LY93], adapted for more than 2 provers in [BGLR93]). Raz and Safra [RazS96] had before us constructed such systems; so our construction can be viewed as an alternative proof of their result.

In our proof system, a probabilistic polynomial time verifier checks that a given string is in the language by using $O(\log n)$ random bits, and one round of interaction with a constant number of provers during which it receives $O(\log n)$ bit long answers from the provers. If the input is in the language, the provers can answer in a way that makes the verifier accept with probability 1. If the input is not in the language, then regardless of the prover's answers the verifier accepts with probability at most $2^{-\log^{(1-\epsilon)} n}$, for any $\epsilon > 0$. The number of provers in our construction grows as $O(1/\epsilon)$. If we are willing to increase the error probability to $2^{-\log^{1/3} n}$ then the number of provers is 5. (The number of provers can be reduced to 3 using a technique of Tardos [T96]). Whether or not the number of provers can be reduced to to 2 remains an open problem.

Now we briefly describe low degree tests; see Section 2 for more details. Given an m-variate function $f : \mathrm{F}^m \to \mathrm{F}$ over a finite field F, the test wishes to determine whether or not there exists a degree d polynomial that agrees with f in δ fraction of points in F^m. (The function is presented *by value*, and the test has random access into this table of values. Both d and ρ are inputs to the test.) The low degree test is allowed to be probabilistic and it has to read as few values of f as possible.

We are interested in a test described in [RS93] that works roughly as follows: Pick a random "line" in F^m and verify that the restriction of f to this line agrees significantly with some univariate degree d polynomial. If this is the case, accept. This test is similar in flavor to all other known low degree tests, such as the original test in [BFL91] and later ones in [BFLS91, FGLSS91, GLRSW91]. (Many of those tests check the degree of the polynomial in each variable, whereas the test we described checks the total degree.)

Let δ denote the probability with which f passes the low-degree test. Existing analyses of all low degree tests cannot say anything meaningful about f if $\delta < 1/2$; in fact the analyses of [FGLSS91, GLRSW91, RS93, AS92] require $\delta > 1 - O(1/d)$. A crucial ingredient of NP=PCP($\log n, 1$) was an analysis (actually just a combination of the analyses of [AS92, RS93]) of the above test that worked for any $\delta > 1 - \epsilon$, where $\epsilon > 0$ is fixed. This analysis showed that if a function f passes the test with probability $\delta > 1 - \epsilon$, then there exists a degree d polynomial that has agreement $\approx \delta$ with f. (The value of ϵ for which this is true was later improved to 1/8 [FS95].)

In this paper we present an analysis (see Theorem 4) that continues to say something meaningful about f even when δ is fairly close to 0. We show that if $\delta > (md)^c / |F|^\epsilon$ for some fixed $c, \epsilon > 0$, then there exists a degree d polynomial that agrees with f in $\approx \delta$ fraction of the inputs. We remark that a similar statement had earlier been proved for really large fields $|F| > 2^{O(m+d+1/\delta)}$ [A93, T93], but that field size is too large for most applications.

We also prove a related result, Theorem 3, which is more useful for constructing efficient PCP-style verifiers. It says that every function f that passes the low degree test with probability δ has an associated small set of polynomials $P_1, P_2, \ldots$ such that the test fails with high probability if it encounters a point where f does not agree with one of the P_i's. This result is useful because all known verifiers work by checking the properties of some function f provided by the prover. If f is a polynomial, the verifier is extremely unlikely to produce an erroneous answer . Errors creep in when f is not a polynomial but only has significant agreement with some set of polynomials $g_1, g_2, \ldots$. In this case, if the verifier has the bad luck to examine f at a point where f doesn't agree with any of $g_1, g_2, \ldots$, then it could accept erroneously. Our corollary provides the means to combat such errors, since any such $g_1, g_2, \ldots$ turn out to be exactly the set of polynomials $P_1, P_2, \ldots$, mentioned in Theorem 3. The verifier subjects f to a low degree test: at any point where f doesn't agree agree with any of $P_1, P_2, \ldots$, the test fails with high probability, thus averting an erroneous accept.

This intuition is formalized in Section A of the appendix. (We put it in the appendix since it was known to many researchers and is a "folk theorem;" the only part missing until now was our Theorem 3.)

Application to program testing/correcting. Suppose we are given a potentially buggy program that purportedly computes a (unknown) m-variate polynomial over a finite field F. Program testing/correcting [BLR90] concerns the following problems: (i) *testing:* determine δ, the fraction of points at which this program is correct and (ii) *correction:* for each input, correct the output of the program in case it is incorrect. It was open how to do testing if $\delta < 1/2$; our low-degree test (specifically, the version that doesn't use an auxiliary table) closes this open problem when $|F|^\epsilon > poly(md)$. As for correction, note that its meaning is not clear when $\delta < 1/2$, since as many as $O(1/\delta)$ polynomials could have agreement δ with the program. Two notions of correction are possible, as noted in [ALRS92]. The weaker one is that for each input, the corrector outputs $O(1/\delta)$ values, one of which is correct. Such a corrector is known [Su96]. The stronger notion is that the corrector create $O(\frac{1}{\delta})$ programs (polynomials) such that w.h.p. one of them is correct. Finding such a corrector was an open problem. Our analysis leads to such a corrector. We omit details from this abstract, but they are obvious from reading our proofs (specifically, by noting their "algorithmic" nature). We note that the recent techniques of [RazS96] do no seem to provide a tester/corrector.

Past work. The first construction of a nontrivial constant prover 1-round proof system for NP appeared in [LS91]; others appeared in [FL92, BGLR93, T94, FK94, R95]. These systems could not reduce the error probability to below a constant while using $O(\log n)$ random bits (the best construction needs $O(k \log n)$ random bits to make the error probability 2^{-k}; see [R95]). It was also known [FK95] that some obvious ideas (such as "recycling randomness") cannot let us get around this. Earlier this year Raz and Safra [RazS96] found a construction of a proof system achieving subconstant error. Our result, though obtained independently, was completed a couple of months after we had heard of the existence of their result (the missing part at the time was our proof of the bivariate case of Theorem 1). Upon seeing their manuscript in September 1996, it was clear — although their earlier announcement didn't suggest this — that they also rely on a low degree test, albeit a new one and with a very different correctness proof than ours. Also, they extend their ideas to design constant prover systems in which the error probability is $2^{-O(\#\text{ of answer bits})}$; we haven't done that.

Paper organization. We state and explain our main theorem (Theorem 1) and its corollaries (Theorems 3 and 4) in Section 2. We prove the theorem in Section 3. This proof resembles proofs of earlier results [RS93, ALMSS92, A94, FS95], in that it has two parts. First in Section 3.1 we prove the theorem when m is constant (specifically, $m = 2, 3$); this uses algebraic arguments inspired by Sudan's [Su96] work on reconstructing polynomials from very noisy data and Kaltofen's work on "Effective Hilbert Irreducibility" [K85, K95]. Then in Section 3.3 we "bootstrap" to allow larger m. This part uses probabilistic arguments and relies upon the cases $m = 2, 3$ (including Theorems 3 and 4 for the cases $m = 2, 3$). It is inspired by the "symmetry-based" approach of Arora [A94]. Finally, the appendix contains the construction of constant prover 1-round proof systems and proofs of many lemmas.

2 The Low-degree Test

Let F be a finite field and m, d be integers. An *m-variate polynomial over F* is a sum of terms of the form $ax_1^{j_1}x_2^{j_2}\cdots x_m^{j_m}$, where $a \in \mathrm{F}$. The set of such polynomials forms an integral domain, denoted $\mathrm{F}[x_1, \ldots, x_m]$. We will often view such a polynomial as a function from F^m to F. The *degree* of the polynomial is its total degree (thus $j_1 + \cdots + j_m$ is the degree of the above monomial). We will usually reserve the symbol q for $|\mathrm{F}|$, the cardinality of F.

The *distance* between two functions $f, g : \mathrm{F}^m \to \mathrm{F}$, denoted $\Delta(f, g)$, is the fraction of points in F^m they differ on. The *agreement* between the functions is $1 - \Delta(f, g)$.

The low-degree test is given a function $f : \mathrm{F}^m \to \mathrm{F}$. Using randomness, it checks that f looks "locally" like a degree-d polynomial. Magically, it can infer from this that f has significant agreement with a degree-d polynomial. To be more formal we need to define a *line* in F^m.

A *line* in F^m is a set of q points with a parametric representation of the form $\{(u_1 + tv_1, u_2 + tv_2, \ldots, u_m + tv_m) : t \in \mathrm{F}\}$ for some $(u_1, \ldots, u_m), (v_1, \ldots, v_m) \in \mathrm{F}^m$. We refer to the point $(u_1 + av_1, u_2 + av_2, \ldots, u_m + av_m)$ as the *point* $t = a$ of the line.

Note that replacing $(v_1, \ldots, v_m)$ by $c \cdot (v_1, \ldots, v_m)$ for any $c \in \mathrm{F} \setminus \{0\}$ does not change the line. Our convention is to fix one of the representations as canonical.

Definition 1 Let l be the line $\{(u_1 + tv_1, u_2 + tv_2, \ldots, u_m + tv_m) : t \in \mathrm{F}\}$ and $f : \mathrm{F}^m \to \mathrm{F}$ be a function. Let $g(t)$ be a univariate polynomial. Then g *describes* f *at the point* $t = a$ *of* l if

$$g(a) = f(u_1 + av_1, u_2 + av_2, \ldots, u_m + av_m).$$

□

Note that if $f : \mathrm{F}^m \to \mathrm{F}$ is a degree d polynomial, then on every line the restriction of f to that line is described by a univariate degree-d polynomial in the line parameter t. The converse can also be shown to be true: if on every line in F^m, the values of f are described by a univariate degree-d polynomial and F is sufficiently large ($q \geq (d+1)(\frac{p}{p-1})$, where p is the characteristic of the field [FS95]), then f must be a degree-d polynomial.

The low degree test is presented with $f : \mathrm{F}^m \to \mathrm{F}$, and an integer d. It is also presented a table that is meant to be a "proof" that f is a degree d polynomial. This table contains, for each line in F^m, a univariate degree d polynomial that supposedly describes the restriction of f to that line. We will use the term *d-oracle* for any table that contains, for each line in F^m, a univariate degree d polynomial. (The number of variables m can be inferred from the context.)

The Low Degree Test:

Pick a random line l in F^m and read the univariate polynomial, say $p_l(t)$, which the given d-oracle contains for this line. Randomly pick a point x on line l and check whether p_l correctly describes f at x. If so, ACCEPT, else REJECT.

We denote by $\delta_d(f, B)$ the probability that the low degree test accepts a function f and a d-oracle m. We will prove the following result about the low degree test.

Theorem 1 (Main) *There are positive integers c_0, c_1, c_2, and c_3 such that the following are true. Let $f : F^m \to F$ be any function and $d > 0$ be any integer.*

1. *For any $\delta > 0$, if f has agreement δ with some degree d polynomial, then there is a d-oracle B such that $\delta_d(f, B) \geq \delta$.*
2. *If $\delta > 0$ satisfies $q > c_0(dm/\delta)^{c_1}$ and there is a d-oracle B such that $\delta_d(f, B) \geq \delta$, then f has agreement at least δ^{c_3}/c_2 with some degree d polynomial.*

We remark that the first half of this theorem is trivial, since we can just take the degree d polynomial that has agreement δ with f, and construct the d-oracle by using the polynomial's restriction to the line in question. We will prove only the more nontrivial second half. As mentioned earlier, previous results show that the statement in the second half has been shown true for some $\delta < 1$, but much greater than half. This paper shows that the statement is true for $\delta \ll 0.5$, and in fact for δ as small as $dm(c_0/q)^{1/c_1}$, which is tiny if q is $(c_0 dm)^{2c_1}$.

2.1 Two Stronger Forms of Theorem 1

Theorem 1 has two stronger forms, one of which will be useful in constructing proof systems.

We will need the (well-known) fact that there are not "too many" polynomials that have significant agreement with a given function.

Proposition 2 *Let $f : F^m \to F$ be any function. Suppose integer $d > 0$ and fraction ρ satisfy $\rho > 2\sqrt{\frac{d}{q}}$. Then there are at most $2/\rho$ degree d polynomials that have agreement at least ρ with f.*

Proof: Let k be the number of polynomials and $P_1, P_2, \ldots, P_k$ be the polynomials. Then P_1 describes f in at least ρ fraction of the points, P_2 describes f in at least $\rho - d/q$ fraction of the points where $P_1 \neq f$, P_3 describes f in at least $\rho - 2d/q$ fraction of the points where $P_1 \neq f$ and $P_2 \neq f$, etc.

Thus the polynomials together describe f in at least

$$\rho + (\rho - \frac{d}{q}) + (\rho - \frac{2d}{q}) + \cdots$$

fraction of the points. This fraction is at least

$$k\rho - \binom{k}{2}\frac{d}{q}.$$

When $k > 2/\rho$, this fraction is more than 1, which is impossible. □

The first strong form says that "almost all" of the success probability of the low degree test happens at points where f agrees with (one of) a small set of polynomials.

Theorem 3 *Suppose m is an integer such that the statement of Theorem 1 is true for all m-variate functions. Let $f : F^m \to F$ be any function and $d > 0$ be any integer. Let c_0, c_1, c_2 and c_3 refer to the same integers that appeared in Theorem 1 and let $\epsilon > 0$ be any fraction satisfying $q > c_0(d/\epsilon)^{c_1}$. Let $P_1, P_2, \ldots, P_k$ be all the degree d polynomials that have agreement at least ϵ^{c_3}/c_2 with f. Then with probability at least $1 - \epsilon$ one of the following two events happens during the low degree test on f (irrespective of the contents of the d-oracle):*

1. *The test outputs REJECT.*
2. *The test picks a point $x \in F^m$ such that $f(x) = P_i(x)$ for some $i = 1, \ldots, k$.*

Proof: Suppose the probability mentioned in the theorem statement is less than $1 - \epsilon$. We derive a contradiction.

Let $S \subseteq \mathrm{F}^m$ be the set of points at which f does not agree with any of $P_1, \ldots, P_k$. Then $f|_S$, the restriction of f to S, is a function that passes the low degree test with probability at least ϵ. Let us extend $f|_S$ to a function $g : \mathrm{F}^m \to \mathrm{F}$ by randomly picking values for g at points in $\mathrm{F}^m \setminus S$. Since g passes the low-degree test with probability at least ϵ, Theorem 1 implies that there is a degree d polynomial P that has agreement ϵ^{c_3}/c_2 with g. This agreement must largely be on points in S, since the restriction of g to $\mathrm{F}^m \setminus S$ is a random function. (Note: A simple calculation using Chernoff bounds shows that a random function has agreement approximately $1/q$ with every degree-d polynomial.) Hence we conclude that polynomial P has agreement approximately ϵ^{c_3}/c_2 with $f|_S$. Since none of $P_1, \ldots, P_k$ agrees with f on S, this polynomial must be be different from each P_i. But this contradicts the hypothesis that $\{P_1, \ldots, P_k\}$ is an *exhaustive* listing of the degree d polynomials that have agreement at least ϵ^{c_3}/c_2 with f. $\Box$

The second strong form says, heuristically speaking, that if $q > \mathrm{poly}(\frac{1}{\rho}, \frac{1}{\epsilon}, md)$, then every function that passes the low degree test with probability ρ has agreement at least $\rho - \epsilon$ with some degree d polynomial. (Note:Theorem 1 only guaranteed an agreement ρ^{c_3}/c_2).

Theorem 4 *Suppose m is an integer such that the statement of Theorem 1 is true for all m-variate functions. Let $f : F^m \to F$ be any function and $d > 0$ be any integer. Suppose there is a d-oracle such that* $\Pr[$*low degree test accepts*$] \geq \rho$. *Let $\epsilon > 0$ be any fraction satisfying*

$$q > \frac{64 \cdot 4^{c_3}}{\epsilon^{c_3+3}\rho^{c_3-1}} + c_0(\frac{4dm}{\epsilon\rho})^{c_1},$$

where c_0, c_1, c_2, c_3 refer to the same integers that appeared in Theorem 1.

Then there is a degree d polynomial that has agreement $\rho - \epsilon$ with f.

Proof: Suppose we pick a line l randomly from F^m. An averaging argument using Lemma 17 shows that with probability at least $\epsilon/2$, we pick a line on which the success probability of the low degree test is at least $\rho - \epsilon/2$. In other words,

$$\Pr_l[\text{some univ. deg. } d \text{ polynomial } g_l \text{ describes } f \text{ on } \rho - \epsilon/2 \text{ fraction of points of } l] \geq \frac{\epsilon}{2} \tag{1}$$

Let $\epsilon_1 = \epsilon\rho$ and let $P_1, \ldots, P_k$ be all the degree d polynomials that have agreement at least $\frac{1}{c_2}(\frac{\epsilon_1}{4})^{c_3}$ with f. Let $\rho_1, \ldots, \rho_k$ be their agreements with f. We wish to show that $\rho_i \geq \rho - \epsilon$ for some i. Let us therefore assume that each $\rho_i < \rho - \epsilon$ and show that the probability mentioned in Assertion (1) is less than $\epsilon/2$, thus deriving a contradiction to Assertion (1).

Where could the univariate degree d polynomial mentioned in Assertion (1) come from? There are two cases. *Case (i):* g_l is the restriction of one of the P_i's to the line l. *Case (ii):* g_l is some other polynomial. Note that if case (ii) happens, then l is a line on which the low degree test is succeeding with probability $\rho - \epsilon/2$, and furthermore this success happens on on points where f doesn't equal any of $P_1, P_2, \ldots, P_k$. By Theorem 3, at most $\epsilon_1/4$ of the success probability of the low degree test comes from the points where f doesn't equal any of $P_1, P_2, \ldots, P_k$. By an averaging argument (Lemma 17) it follows that

$$\Pr_l[\text{case (ii) happens}] \leq \epsilon_1/4\rho \leq \epsilon/4.$$

Now we show that $\Pr_l[\text{Case (i) happens}] < \epsilon/4$, thus leading to the desired contradiction.

For $i = 1, 2, \ldots, k$, let γ_i be the fraction of points on l at which polynomial P_i agrees with f. By Lemma 18 it follows that

$$\Pr_l[\gamma_i - \rho_i > \frac{\epsilon}{2}] \leq \frac{4\rho_i}{\epsilon^2 q} \qquad \text{for } i = 1, \ldots, k. \tag{2}$$

Since we assumed that each $\gamma_i < \rho - \epsilon$, we now conclude that

$$\Pr_l[\exists\ i \text{ s.t. } \gamma_i > \rho - \epsilon/2] \leq \frac{4\rho}{\epsilon^2 q} \times k.$$

But Proposition 2 implies that $k \le 2c_2/(\epsilon_1/4)^{c_3}$. Hence

$$\Pr_l[\exists\ i \text{ s.t. } \gamma_i > \rho - \epsilon/2] \le \frac{8\rho c_2}{\epsilon^2 q(\epsilon_1/4)^{c_3}}$$

Note that the probability on the LHS is an upperbound on the $\Pr_l$[Case (i) happens], and that the RHS is less than $\epsilon/4$ for the range in which our parameters lie. Thus $\Pr_l$[Case (i) happens] $< \epsilon/4$. $\Box$

3 Proof of Correctness of Low-degree Test

In this section we prove Theorem 1. For ease of exposition we first restate Theorem 1. From now on we will reserve the symbol f for a function from F^m to F which is the subject of the low degree test.

Definition 2 The *line polynomial for f on line l for degree d*, denoted $P_d^f(l)$, is the univariate degree d polynomial (in the line parameter t) that describes f on more points of l than any other degree d polynomial. (We arbitrarily break ties among different polynomials that describe f equally well on the line.) The *d-success-rate of f on line l*, denoted $\mu_d^f(l)$, is defined as

$$\mu_d^f(l) = \text{fraction of points on } l \text{ where } P_d^f(l) \text{ describes } f.$$

The *d-success-rate of f at point $x \in F^m$* is the fraction of lines through x whose line polynomial describes f at x.

The *d-success rate of f* is the average of its d-success rates on all lines. (Note: By symmetry this is also equal to its average d-success rate at all points.)

Note that the probability that a function $f : \mathrm{F}^2 \to \mathrm{F}$ passes the low degree test is maximised when the accompanying d-oracle contains, for each line l, the polynomial $P_d^f(l)$. Hence it suffices to prove the following.

Theorem 5 (Restatement of Theorem 1 part 2) *There are integers c_0, c_1 such that the following is true. If $f : F^m \to F$ is any function whose d-success rate is at least δ and $q > \frac{1}{c_0}(\frac{dm}{\delta})^{c_1}$, then there exists a degree d polynomial that has agreement at least δ^{c_3}/c_2 with f.*

3.1 The Bivariate Case

In this section we prove Theorem 5 for $m = 2$. Let $f : \mathrm{F}^2 \to \mathrm{F}$ be a function with success-rate at least δ. Our proof goes in two steps.

(Step 1). Show that there is a polynomial $Q \in \mathrm{F}[z, x, y]$ of "not too large degree" such that for a "reasonably large" set of points $S \subseteq \mathrm{F}^2$, the following are true:

$$Q(f(a,b), a, b) = 0 \qquad \forall (a,b) \in S \tag{3}$$

$$\text{the } d\text{-success rate of } f \text{ at every point in } S \text{ is "nontrivial."} \tag{4}$$

(Step 2). Show that any Q that satisfies the conditions in Step 1 has a factor $z - g(x,y)$, such that $g \in \mathrm{F}[x,y]$ is a degree d polynomial and for "many" $(a,b) \in S$

$$(z - g(x,y)) = 0 \qquad \text{at } (z,x,y) = (f(a,b), a, b). \tag{5}$$

By the end of Step 2, we have concluded that f has significant agreement with the degree d bivariate polynomial g. Step 2 depends on a fairly difficult technical fact, Theorem 23, which will be proved separately in Section 4.1. Step 1 is motivated by Sudan's [Su96] technique for reconstructing univariate polynomials from very noisy data.

Sudan makes the following observation.

Proposition 6 *Let $(a_1, y_1), \ldots, (a_n, y_n)$ be any set of n pairs from F^2, and d_z, d_x be any positive integers satisfying $d_x d_x > n$. Then there exists a bivariate polynomial $\Gamma \in F[z,x]$ with $deg_z(\Gamma) \leq d_z$ and $deg_x(\Gamma) \leq d_x$, satisfying*

$$\Gamma(y_i, a_i) = 0 \qquad \text{for } i = 1, \ldots, n \tag{6}$$

Remark: We can view Γ as an implicit description of the sequence $(a_1, y_1), \ldots, (a_n, y_n)$, in the following sense: for each a_i, one of the roots of $\Gamma(z, a_i)$ is y_i.

Proof: If we let γ_{ij} be the coefficient of $z^i x^j$ in Γ, then the contraints in (6) define the following homogeneous linear system with $(1+d_x)(1+d_y)$ unknowns and n constraints. (Note that $a_1, \ldots, a_n, y_1, \ldots, y_n$ are "constants.")

$$\begin{aligned}
\sum_{i=0}^{d_z} \sum_{j=0}^{d_x} \gamma_{ij} y_1^j a_1^i &= 0 \\
\sum_{i=0}^{d_z} \sum_{j=0}^{d_x} \gamma_{ij} y_2^j a_2^i &= 0 \\
\cdots &= 0 \\
\sum_{i=0}^{d_z} \sum_{j=0}^{d_x} \gamma_{ij} y_n^j a_n^i &= 0
\end{aligned}$$

Since $(1+d_x)(1+d_y)$, the number of variables, exceeds n, the number of constraints, a nontrivial solution exists. $\Box$

Then Sudan uses the following lemma from Ar et al. [ALRS92].

Lemma 7 *Let $(a_1, y_1), \ldots, (a_n, y_n) \in F^2$ be any sequence such that for some $\rho > 0$,*

$$\text{there is a degree } d \text{ polynomial } h \in F[x] \text{ s.t. } h(a_i) = y_i \text{ for } \rho n \text{ values of } i. \tag{7}$$

Let $\Gamma \in F[z,x]$ be any polynomial satisfying (6). If $deg_x(\Gamma) + d \cdot deg_z(\Gamma) < \rho n$, then $(z - h(x)) \mid \Gamma$.

Proof: The polynomial $\Gamma(h(x), x)$ has degree at most $deg_x(\Gamma) + d \cdot deg_z(\Gamma)$ and has at least ρn roots. So if $deg_x(\Gamma) + d \cdot deg_z(\Gamma) < \rho n$, this polynomial must be identically 0. $\Box$

Remark: Sudan's observations lead to efficient algorithms because both Lemma 7 and Proposition 6 have "constructive" versions: efficient algorithms exist for polynomial factorization (needed for Lemma 7) and solving linear equations (needed for Proposition 6) over finite fields. The current paper is not about algorithm design, but nevertheless the key algebraic facts used in polynomial factorization and solving linear equations also drive our result. See for example the "effective Hilbert irreducibility" in Section 4.1 and Cramer's Rule in Lemma 8.

Specifically, we need the following generalization of Sudan's ideas to F$[y]$, the ring of univariate polynomials in the formal variable y.

Lemma 8 *Let $S_1, S_2 \subseteq F$ be any subsets of F and $l = |S_1|$. Let $f : S_1 \times S_2 \to F$ be any function and for each $a, b \in F$, let $C_a \in F[y], R_b \in F[x]$ be degree d polynomials. Suppose there is a fraction $\rho > 2d/\sqrt{l}$ such that for all $b \in S_2$,*

$$f(a,b) = C_a(b) = R_b(a) \qquad \text{for at least } \rho l \text{ values of } a \in S_1. \tag{8}$$

Then there is a polynomial $Q \in F[z,x,y]$ satisfying $\deg_z(Q) \leq \sqrt{l}$, $\deg_x(Q) \leq \sqrt{l}$, $\deg_y(Q) \leq dl^{3/2}$ such that

$$\forall a \in S_1, \;\; Q(C_a(y), a, y) = 0 \;\; \text{and} \tag{9}$$

$$\forall b \in S_2, \;\; (z - R_b(x)) \,|\, Q(z, x, b) \tag{10}$$

Proof: Let $\mathrm{F}[y][z,x]$ be the ring of polynomials in the formal variables z and x whose coefficients are from $\mathrm{F}[y]$.

We use the same idea as in Proposition 6, but work over the ring $\mathrm{F}[y]$ instead of over F. Consider the following sequence of $|S_1|$ pairs from $\mathrm{F} \times \mathrm{F}[y]$: $((a, C_a(y)) : a \in S_1)$. Note that there exists a polynomial $Q \in \mathrm{F}[y][z,x]$ with $deg_z(Q), deg_x(Q) \leq \sqrt{l}$ such that

$$Q(C_a(y), a) = 0 \qquad \forall a \in S_1 \tag{11}$$

The reason is that if we let $Q_{ij} \in \mathrm{F}[y]$ be the coefficient of $z^j x^i$ in Q, then the constraints in (11) define a homogeneous system of linear equations over $\mathrm{F}[y]$ with $(1 + deg_x(Q))(1 + deg_z(Q)) > l$ unknowns and l constraints.

$$\sum_{i=0}^{\sqrt{l}} \sum_{j=0}^{\sqrt{l}} Q_{ij} (C_a(y))^j a^i \;\; = \;\; 0 \qquad \forall a \in S_1$$

Since the number of unknowns exceeds the number of constraints, a nontrivial solution exists.

Now we claim that we can find a nontrivial solution Q that in addition is in $\mathrm{F}[y][z,x]$ and satisfies $deg_y(Q) \leq dl^{3/2}$. The reason is that Q is obtained by Cramer's Rule on a system of l constraints, which calls for inverting an $(l-1) \times (l-1)$ matrix. Inverting an $(l-1) \times (l-1)$ matrix involves evaluating polynomials of degree $l - 1$ in the matrix entries. In this case the matrix entries are degree $d\sqrt{l}$ polynomials in $\mathrm{F}[y]$, so matrix inversion produces only polynomials of degree $dl^{3/2}$ in y. Hence $deg_y(Q) \leq dl^{3/2}$.

Finally, the fact that Q satisfies condition (10) follows immediately from Lemma 7 and the condition $\rho > 2d/\sqrt{l} \geq (d+1)/\sqrt{l}$. $\Box$

The following lemma finishes Step 1 of our proof.

Lemma 9 *Let $f : F^2 \to F$ have d-success rate at least δ, let $t = \max\{4 \log q/\delta^3, (\frac{64d}{\delta^3})^2\}$. If $q > 100t^2$, then there is a polynomial $Q \in F[z,x,y]$ of total degree at most $2t^{3/2}d$ and a set of points $S \subseteq F^2$ containing at least $\delta^6/256$ fraction of the points such that*

1. *$Q(f(a,b), a, b) = 0 \qquad \forall (a,b) \in S$.*
2. *The d-success rate of f at each point in S is at least $\delta/2$*

Proof: This proof uses averaging. The main idea is to rotate the coordinate system so that with respect to the new x and y axes, the conditions of Lemma 8 are satisfied for $\rho = \delta^6/256$. The existence of polynomial Q is then implied by the conclusion of that lemma. Note that a rotation of coordinates does not affect the total degree of a polynomial, and we are interested only in the total degree of Q.

Below, when we say "a line in the direction h," we mean a line of the form $\{(u + t \cdot h) : t \in \mathrm{F}\}$. Note that for each point $x \in \mathrm{F}^2$ and direction h, there is exactly one line in direction h that passes through x.

We say that a point $x \in \mathrm{F}^2$ is *good* for a pair of directions (h_1, h_2) if the line polynomials $P_d^f(l_1)$ and $P_d^f(l_2)$ correctly describe f at x, where l_1, l_2 are the lines that pass through x and lie in directions h_1 and h_2 respectively.

Let $G \subseteq \mathrm{F}^2$ denote the set of points at which the success rate of f is at least $\delta/2$. Since the overall success rate is at least δ, averaging shows that at least $\delta/2$ fraction of the points are in G.

Claim 1: *There exist two directions h_1, h_2 and a set of points $H \subseteq G$ with size $|H| \geq \delta^3 |F|^2/8$ such that every point in H is good for (h_1, h_2).*

Proof of Claim 1: We use the probabilistic method. Randomly pick a random point $x \in \mathrm{F}^2$ and a pair of directions (h_1, h_2). (Note: with probability $1 - 1/q$, $h_1 \neq h_2$. We will assume for simplicity that this probability is actually 1.)

$$\begin{aligned}
\Pr_{x,h_1,h_2}[x \in G \bigwedge x \text{ is good for } (h_1,h_2)] &= \Pr_x[x \in G] \times \Pr_{x,h_1,h_2}[x \text{ is good for } (h_1,h_2) \mid x \in G] \\
&\geq \frac{\delta}{2} \times (\frac{\delta}{2})^2. \\
&\geq \frac{\delta^3}{8}.
\end{aligned}$$

In other words, when a pair of directions (h_1, h_2) is picked randomly, then the expected size of the set $\left\{x \in \mathrm{F}^2 : x \in G \bigwedge x \text{ is good for } (h_1, h_2)\right\}$ is at least $\delta^3 |\mathrm{F}|^2/8$. Hence there exists a pair of directions for which this set has size at least $\delta^3 |\mathrm{F}|^2/8$. This finishes the proof of Claim 1. $\Box$

Let h_1, h_2, H be as in Claim 1. Rotate the coordinates so that h_1 becomes the x-axis and h_2 becomes the y-axis. From now on, coordinates are stated in this new system. We use *columns* and *rows* to refer to lines parallel to the y and x axes respectively.

For $a, b \in \mathrm{F}$ let R_b and C_a denote the line polynomials in the row $\{(x, b) : x \in \mathrm{F}\}$ and the column $\{(a, y) : y \in \mathrm{F}\}$ respectively. By the defining property of H, if $(a, b) \in H$, then $C_a(b) = R_b(a) = f(a, b)$.

Let $\gamma = \delta^3/16$. Averaging shows that at least γ of the rows have at least γ fraction of their points in H. Let $S_2 \subseteq \mathrm{F}$ be the set of all such rows. Let $t = 4 \log q/\gamma$. We claim that there exists a set S_1 consisting of t vertical lines such that $\forall b \in S_2$

$$C_a(b) = R_b(a) = f(a, b) \qquad \text{for at least } \gamma t/2 \text{ values of } a \in S_1. \tag{12}$$

The existence of S_1 is proved by the probabilistic method. Pick a set of S_1 randomly by picking t lines with repetition, and show that w.h.p. the resulting set satisfies, for all $b \in S_2$, $|H \cap (S_1 \times \{b\})| \geq \gamma t/2$. (Even though we picked lines with repetition, the probability that any two are the same is at most t^2/q, which is $< 1/100$. Hence w.h.p. the set S_1 has no repeated lines.)

Let $b \in S_2$. The expected fraction of points in $S_1 \times \{b\}$ that lie in H is at least γ. Hence by the Chernoff bound,

$$\Pr_{S_1}[|H \cap (S_1 \times \{b\})| < \gamma t/2] \quad \leq \quad exp(-\frac{\gamma t}{2})$$

$$= \quad exp(-2\log q) \leq \frac{1}{2q}.$$

Thus the probability is at least $1 - |S_2|/2q - 1/100 \geq .49$ that the randomly chosen set S_1 satisfies condition (12).

Thus we have proven the existence of $S_1, S_2 \subseteq \mathrm{F}$ such that they satisfy the hypothesis of Lemma 8 with $\rho = \gamma/2$ and $l = t$. (Notice that by the condition on t, we have that $\rho \geq 2d/\sqrt{l}$.) Let $Q \in \mathrm{F}[z, x, y]$ be the polynomial whose existence is guaranteed by Lemma 8. Then $deg_x(Q), deg_z(Q) \leq \sqrt{t}$ and $deg_y(Q) \leq dt^{3/2}$, and total degree of Q is $2\sqrt{t} + dt^{3/2} < 2dt^{3/2}$.)

To finish we need to define the set S mentioned in the lemma. Let

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

Since each row $b_2 \in S_2$ has at least γ fraction of its points in H and $|S_2| > \gamma\,|\mathrm{F}|$, we have

$$|S| \geq \gamma^2\,|\mathrm{F}|^2 = \frac{\delta^6}{256}\,|\mathrm{F}|^2.$$

Now let $(a, b) \in S$. Since $b \in S_2$, the property of Q implies $(z - R_b(x)) \mid Q(z, x, b)$ and so $Q(R_b(x), x, b) = 0$. Since $(a, b) \in H$, the property of H implies $f(a, b) = C_a(b) = R_b(a)$. Hence $Q(f(a, b), a, b) = 0$. Thus the lemma has been proved. $\Box$

Now we move to Step 2 of our argument.

Lemma 10 *Let $f : F^2 \to F$ be a function, and $Q \in F[z, x, y]$ be a polynomial of total degree D and $S \subseteq F^2$ be a set of points of size at least $\gamma \cdot |F|^2$ such that: (i) $\forall (a, b) \in S$, $Q(f(a, b), a, b) = 0$. (ii) The d-sucess-rate of f at every point in S is at least γ.*

If $|F| > 4D^5/\gamma^2$, then there is a degree D bivariate polynomial $g \in F[x, y]$ that has agreement at least $\gamma^4/8D$ with f and such that $z - g(x, y)$ is a factor of Q.

Proof: The main idea is to use Lemma 7 to show that the restriction of Q on "many" lines has a linear factor that describes f on "many" points of that line. Then we will use Theorem 23 on "effective Hilbert irreducibility" to conclude that Q itself must have a linear factor that describes f in "many" points.

We say a point $(a, b) \in \mathrm{F}^2$ is *nice* for a line l in F^2 if (i) $Q(f(a, b), a, b) = 0$ and (ii) $P_d^f(l)$, the line polynomial of l, describes f at (a, b).

Claim 1: *When a line l is picked randomly, the expected fraction of points on it that are nice for it is at least γ^2.*

Proof: Imagine picking a point (a, b) randomly and then randomly picking a line l that passes through it. The probability that the point is nice for l is at least $\gamma \cdot \gamma = \gamma^2$. The claim now follows by linearity of expectations. $\Box$

Let $Q_1, \ldots, Q_k \in \overline{\mathrm{F}}[z, x, y]$ be all the distinct factors (over the algebraic closure of field F) of Q that involve z. (Note that $k \leq D$.)

Claim 2: *One of the Q_i's is of the form $z - r(x, y)$ where $r \in \overline{F}[x, y]$.*

Proof: For a line l let us denote the restriction of Q to l by $Q|_l \in \mathrm{F}[z, t]$, where t is the line parameter. We define $Q_i|_l$ analogously for $i = 1, .., k$.

Assume for contradiction's sake that no Q_i has the form $z - r(x, y)$ for some $r \in \overline{\mathrm{F}}[x, y]$. Since each Q_i is absolutely irreducible, Theorem 23 implies that the fraction of lines l

such that the restriction $Q_i|_l$ has a factor of the type $z - p(t)$ where $p \in \overline{\mathrm{F}}[t]$, is at most $O(D^3/|\mathrm{F}|)$. Hence the fraction of lines on which either of $Q_1|_l, \ldots Q_k|_l$ has a factor of the type $z - p(t)$ is at most $O(kD^3/|\mathrm{F}|)$. By our assumption on the values of $|\mathrm{F}|$, γ, and D, this fraction is at most $\gamma^2/4$. We show next that this fraction is actually at least $\gamma^2/2$, which is a contradiction.

From the statement of Claim 1 and simple averaging we know that on at least $\gamma^2/2$ fraction of the lines, at least $\gamma^2/2$ fraction of the points are nice for them. Let l be such a line. We show that $Q|_l(z,t)$ has a factor of the form $z - p(t)$ for some $p \in \mathrm{F}[t]$. Let $h \in \mathrm{F}[t]$ be the line polynomial for l (i.e., $h = P_d^f(l)$). Then $Q|_l(h(t),t)$ has $\gamma^2 |\mathrm{F}|/2$ roots and degree only Dd, where $Q|_l(z,t)$ is the restriction of Q to l. But $Dd < \gamma^2 |\mathrm{F}|/2$, so $Q|_l(h(t),t)$ must be identically 0. Hence $z - h(t) \mid Q|_l(z,t)$. □

The following claim finishes the proof of the lemma. Note that the polynomial g in the statement of the claim takes its coefficients from F instead of from the closure field $\overline{\mathrm{F}}$.

Claim 3: *One of the Q_i's is of the form $z - g(x,y)$ where $g \in F[x,y]$ is a degree d polynomial that has agreement at least $\gamma^2/2D$ with f.*

Proof:(of Claim 3) Assume that $l \geq 1$ factors of Q are of the form described in Claim 2, and assume w.l.o.g. that they are $Q_1, \ldots, Q_l$. For $i = 1, \ldots, l$, suppose $Q_i(z,x,y) = z - p_i(x,y)$ where $p_i \in \overline{\mathrm{F}}[x,y]$. From the proof of Claim 2 we know that for at least $\gamma^2/2 - O(D^3k/|\mathrm{F}|)$ fraction of the lines, the following is true (i) the line polynomial $P_d^f(l)$ of the line is the restriction of one of the p_i's to the line, (ii) $P_d^f(l)$ describes f on at least $\gamma^2/2$ fraction of points on l. For simplicity, we use $\gamma^2/4$ as a lowerbound on $\gamma^2/2 - O(D^3k/|\mathrm{F}|)$.

Thus there must exist some $i \in [1,l]$ such that Q_i explains $1/l$ fraction of such lines. We claim that this Q_i is the factor we are looking for (i.e., $g = p_i$). Note that by choice of i, polynomial p_i has agreement $\frac{1}{l} \cdot \frac{\gamma^2}{2} \cdot \frac{\gamma^2}{4}$, with f. This agreement is at least $\frac{\gamma^4}{8D}$.

Note that thus far we only know that $g \in \overline{\mathrm{F}}[x,y]$ and has degree at most D. Now we claim that g actually (i) is a degree d polynomial and (ii) has all its coefficients in F. The reason we claim (ii) is that that the restriction of g on at least $\frac{1}{l} \cdot \frac{\gamma^2}{4}$ fraction of lines is in $\mathrm{F}[t]$ and $\frac{\gamma^2}{4l} > D/|\mathrm{F}|$. (See Lemma 22.) The reason that g has degree at most d instead of D is that its restriction to at least $\frac{\gamma^2}{4l}$ fraction of the lines is a degree d polynomial and $\frac{\gamma^2}{4l} > D/|\mathrm{F}|$. (See Lemma 20.) □

□

Thus we have proved the bivariate case of Theorem 1.

Theorem 11 *Let $F = GF(q)$ and $f : F^2 \to F$ be a function that has d-success rate at least δ. If $q/(\log q)^5 > 2^{105}d^{20}/\delta^{57}$, then there is a bivariate degree d polynomial g that has agreement at least $\delta^{33}/(2^{55}d^4 \log q)$ with f.*

Proof: Follows from Lemmas 9 and 10. □

3.2 The Trivariate Case

We restate Theorem 5 for the case $m = 3$ and prove it. The proof is a minor modification of the proof for $m = 2$.

Lemma 12 *There exist constants c_0, c_1, c_2, c_3 such that for all δ, d and F such that $|F| \geq c_0(d/\delta)^{c_1}$ if $f : F^3 \to F$ has d-success-rate at least δ, then f has agreement at least $\frac{1}{c_2}\delta^{c_3}$ with some degree d polynomial.*

Proof:

As in the proof of the bivariate case we first perform a random transformation of the coordinates. We identify three directions h_1, h_2 and h_3 in F^3 and call all lines of the form $\{a + th_1 | t \in F\}$ as *vertical lines*, all lines of the form $\{a + t(h_2 + bh_3) | t \in F\}$ for any $b \in \mathrm{F}$ as *horizontal* lines.

Claim 1 *Let $\delta_1 = \delta^3/16$. There exist direction h_1, h_2 and h_3 s.t. δ_1 fraction of the points in F^m are in G and are good for h_1, h_2, h_3, have the following three properties:*

- *The vertical line through the point passes the low-degree test.*
- *δ_1 fraction of the horizontal lines through the point pass the low-degree test.*
- *δ_1 fraction of all lines through the point pass the low-degree test.*

Proof: By averaging we know that the d-success-rate of f is at least $\delta/2$ at at least $\delta/2$ fraction of the points. Let G denote this set of points.

We say a point $a \in \mathrm{F}^3$ is *good* for directions h_1, h_2, h_3 if it satisfies conditions (i) and (ii) in the statement of the claim.

We use the probabilistic method to prove the lemma. We randomly pick three directions h_1, h_2, h_3 and a point $a \in \mathrm{F}$ show that

$$\Pr_{a,h_1,h_2,h_3}\left[a \in G \bigwedge a \text{ is good for } h_1, h_2, h_3\right] \geq \delta_1.$$

Note that $\Pr_a[a \in G] \geq \delta/2$, so it suffices to show that

$$\Pr_{a,h_1,h_2,h_3}\left[a \text{ is good for } h_1, h_2, h_3 \mid a \in G\right] \geq 2\delta_1/\delta = \delta^2/8.$$

Consider any $a \in G$. If we pick a random line through a, then it passes the low degree test at a with probability at least $\delta/2$. So if we pick two random directions h_2, h_3 and then a random $b \in \mathrm{F}$, then

$$\Pr_{h_2,h_3,b}\left[\text{the line } \{a + t(h_2 + bh_3) : t \in \mathrm{F}\} \text{ passes the low degree test at } a\right] \geq \delta/2.$$

Hence we conclude by averaging that for at least $\delta/4$ choices of h_2, h_3, the fraction of $b \in \mathrm{F}$ for which this event happens is at least $\delta/4$.

Thus for all $a \in G$,

$$\Pr_{h_1,h_2,h_3}\left[a \text{ is good for } h_1, h_2, h_3\right] \geq \delta/2 \cdot \delta/4 = \delta^2/8.$$

(Notice that with probability $3/|\mathrm{F}|$ the directions h_1, h_2 and h_3 are linearly dependent, in which case our calculation is off by a little. We ignore this factor since it is so close to 0.) □

From now on we assume that h_1, h_2 and h_3 as guaranteed above have been found and that the coordinates have been transformed so that $h_1 = (1, 0, 0)$, $h_2 = (0, 1, 0)$ and $h_3 = (0, 0, 1)$. The set of points of the form $\{(w, x, y) | x, y \in F\}$ will be called the horizontal plane through w. For every w let $(f_{\mathrm{hor}}(w)[y, z]$ denote the bivariate degree d polynomial which has maximum agreement with f on the horizontal plane through w.

Claim 2 *Let $\delta_2 = \delta_1^4/8$. Then δ_2 fraction of the points (w, x, y) satisfy the following properties:*

- *The vertical line through the point passes the low-degree test.*
- *The value of f at the point agrees with the f_{hor} at that point. the low-degree test.*
- *δ_2 fraction of all lines through the point pass the low-degree test.*

Proof: We use the correctness of the low-degree test for the bivariate case after perturbing the function f as in Theorem 3. We set randomly every point except the δ_1 fraction of points which are guaranteed good by the previous claim. Observe that a random point passes the low-degree test for a random horizontal line through it with probability at least δ_1^2. Say that a horizontal plane is good if the probability that a random point on the plane and a random horizontal line through the plane pass the low-degree test with probability $\delta_1^2/2$. Observe now that $\delta_1^2/2$ fraction of the horizontal planes are good. For a good horizontal plane, the correctness of the bivariate test implies that there exists a low-degree polynomial which has an agreement of $\delta_1^2/4$ with f on the plane. Furthermore, this agreement happens on points which have good success rate on random lines through them as well as on the vertical line through them (since all other points were set randomly.

The final statement of the claim is now obtained as follows. Consider the probability that a random point we pick is on a good horizontal plane and on that plane agrees with the polynomial f_{hor} for that plane. This probability is at least $(\delta_1^2/2)(\delta_1^2/4) = \delta_2$. □

Claim 3 *Let $\delta_3 = \delta_2^2/8$. Then for any $l \geq (\frac{16d}{\delta_2})^2$, there exists a polynomial $Q(w, x, y, z)$ with $\deg_w(Q) = \deg_z(Q) = \sqrt{l}$ and $\deg_x(Q) = \deg_y(Q) = dl^{3/2}$ such that δ_3 fraction of the points (w, x, y) satisfy the following properties:*

- *$Q(w, x, y, f(w, x, y)) = 0$.*
- *δ_3 fraction of all lines through the point pass the low-degree test.*

Proof:[sketch] As in the bivariate analysis, we pick a random set $S \subset F$, $|S| = l$. We construct a polynomial $Q_S(w, x, y, z)$ of degree $\sqrt{l}$ in w and z and degree $dl^{3/2}$ in x and y such that for all $w \in S$, $x, y \in F$, $Q_S(w, x, y, (f_{\text{hor}}(w))[x, y]) = 0$. Call a point *good* if it satisfies the properties listed in the last claim. Call a vertical line good if $\delta_2/2$ fraction of the points on it are good. Observe that $\delta_2/2$ fraction of the vertical lines are good. By applying Chernoff bounds, we find that with probability at least $1 - e^{-\delta_2 l/16}$ (over the choice of S), a good vertical line has $2d\sqrt{l}$ good points from S in it, provided $l \geq (16d/\delta_2)^2$. Now for a vertical line which has at least $2d\sqrt{l}$ good points on it, we observe that the polynomial $(z - (f_{\text{vert}}(x, y))[w])$ divides the polynomial $Q_{x,y}(w, z) \stackrel{\text{def}}{=} Q_S(w, x, y, z)$. Thus with probability $1 - e^{-\delta l/16} > 1/2$ over the choice of S, we find that for a fixed good vertical line, a good point on the line satisfies the conditions of the claim. Thus there must exist a choice of S such that this holds for more than half the good vertical lines. This yields the polynomial Q as guaranteed by the claim. □

Claim 4 *Let $\delta_4 = \delta_3^2$. Then there exists a degree d polynomial $q(w, x, y)$ f and q have agreement $\delta_4 - O(d^3 l^3/|F|)$.*

Proof: For a random point and a random line, the probability that Q is zero and the test succeeds is at least δ_3^2. We can now argue as in Claims 2 and 3 in the proof of Lemma 10 to get the desired polynomial q. $\square$

Thus we have shown that if f has success rate δ, then it has agreement $\delta_4 - O(d^3l^3/|F|) = \frac{1}{c_2}\delta^{c_3}$ with some degree d polynomial provided $|F| \geq 2d^3l^3/\delta_4$ which is a condition of the form $|F| \geq c_0(d/\delta)^{c_1}$. $\square$

Since we have proved the trivariate case of the low degree test, the trivariate cases of Theorem 3 and 4 now follow.

Corollary 13 *There exist constants c_0, c_1 such that if δ, d and F satisfy $|F| \geq c_0(d/\delta)^{c_1}$, then if $f : F^3 \to F$ has d success rate δ, then it has agreement $\delta/2$ with some degree d polynomial.*

Corollary 14 *There exist constants c_0, c_1 such that if γ, d and F satisfy $|F| \geq c_0(d/\gamma)^{c_1}$, then given a function $f : F^3 \to F$ there exists a set of at most $\frac{4}{\gamma}$ polynomials $Q_1, \ldots, Q_k$ such that the success probability of f on points where f does not equal any of the Q_i's is at most γ.*

3.3 The Bootstrapping

This section assumes the truth of Theorem 1 (as well as Theorems 3 and 4) for $m = 2, 3$, and proves it for general m. We rely on symmetry-based arguments similar to those in [A94]. These use the notion of a k-dimensional subspaces of F^m.

Definition 3 Let $m, k \in \mathcal{Z}^+$ and $k < m$. A *k-dimensional subspace* of F^m is a set of points with a parametrization of the form

$$\{\overline{u_0} + t_1 \cdot \overline{u_1} + t_2 \cdot \overline{u_2} + \cdots + t_k \cdot \overline{u_1} : t_1, t_2, \ldots, t_k \in \mathrm{F}\},$$

for some $\overline{u_1}, \overline{u_2}, \ldots, \overline{u_k} \in \mathrm{F}^m$. $\square$

Thus a *line* is a 1-dimensional subspace, for example. We will refer to a 2-dimensional subspace as a *plane* and a 3-dimensional subspace as a *cube*. A function defined on a k-dimensional subspace of F^m is called a degree d polynomial if the function can be expressed as a degree d polynomial in the parameters $t_1, \ldots, t_k$.

Note that each set of $k+1$ distinct points in F^m determines a unique k-dimensional subspace. Likewise, a line and a point outside it determine a unique plane, two lines that are not in the same plane determine a unique cube, and so on. We use the term plane(l, x) to denote the unique plane containing a line l and a point x etc.

Our argument will rely on symmetry, such as the following facts: (i) all points in F^m are part of exactly the same number of k-dimensional subspaces (ii) All lines in F^m are part of exactly the same number of k-dimensional subspaces, etc. We give an illustrative example of a symmetry-based calculation.

Example 1 Suppose $f : \mathrm{F}^m \to \mathrm{F}$ is any function whose d-success-rate is exactly β. For any plane s let t_s be the average d-success-rate of f among lines in l. Then symmetry implies that $E_s[t_s]$, the average of t_s among all planes, is exactly β. The reason is that $\sum_s t_s$ counts every line in F^m an equal number of times.

Now we try to define a function $\hat{f}$ that we hope is "almost" a polynomial and has significant agreement with f.

Definition 4 ($\hat{f}_l$) For any line l we define a function $\hat{f}_l : \mathrm{F}^m \to \mathrm{F}$ as follows. Let $P_d^f(l)$ denote the univariate degree d polynomial that best describes f's restriction to l (see Definition 2). Now consider every plane s that contains l. (Note: since every point $x \notin l$ determines a unique plane with l, the set of planes containing l form a partition of F^m.) Check whether there is a bivariate polynomial, say g, that agrees with $P_d^f(l)$ on line l and that has agreement at least $\delta/4$ with f on plane s. If so, for every point $y \in s$, we define $\hat{f}_l(y)$ to be the value taken by g at y. If no such bivariate polynomial exists, we define $\hat{f}_l(y)$ arbitrarily in this plane.

Lemma 15 *There are constants $r, s > 1$ such that the following is true for each $m > 3$. Let $f : F^m \to F$ have d-success-rate at least δ, and $q = |F| > (\frac{r}{\delta^3})^s$. If a line l is picked randomly, then*

$$E_l[d\text{-}success\text{-}rate\ of\ \hat{f}_l\ in\ F^m] \geq 1 - \frac{\delta^2}{256} \tag{13}$$

$$E_l[agreement\ between\ f\ and\ \hat{f}_l\ in\ F^m] \geq \frac{\delta}{4}. \tag{14}$$

Before proving Lemma 15, we first point out how Theorem 1 follows immediately.

Proof:(of Theorem 1; $m > 3$) We use the probabilistic method: we pick a line l randomly and show that with nonzero probability, we get a line such that the polynomial closest to $\hat{f}_l$ has agreement at least $\delta/24$ with f.

Using an averaging argument along with statement (13) we see that for any $k > 1$,

$$\Pr_l[d\text{-success-rate of } \hat{f}_l \geq 1 - k\frac{\delta^2}{256}] \geq 1 - \frac{1}{k}$$

Using averaging on (14) we see that

$$\Pr_l[\text{agreement between } f \text{ and } \hat{f}_l > \frac{\delta}{8}] > \frac{\delta}{8}.$$

We let $k = 10/\delta$, and conclude that with probability $\delta/8 - \delta/25.6$ the following two events happen (i) d-success-rate of $\hat{f}_l > 1 - \delta/24$ and (ii) the agreement between f and $\hat{f}_l$ is at least $\delta/8$.

In particular, there exists at least one line for which the two events in the preceding paragraph happen. Let l_0 be such a line. The existing analysis of the low degree test [ALMSS92] implies that for each $\delta < 1$, every function with d-success-rate at least $1 - \delta/24$ has agreement at least $1 - \delta/12$ with some degree d polynomial. Let g be this polynomial for f. Since g and f have agreement at least $1 - \delta/12$ and since $\hat{f}_{l_0}$ and f have agreement at least $\delta/8$, we conclude that f and g have agreement at least $\delta/8 - \delta/12 = \delta/24$. $\Box$

Now we prove Lemma 15.

Proof: (Lemma 15) By linearity of expectations it suffices to show that if we pick a pair of lines (l, l') randomly in F^m, then

$$E_{(l,l')}[d\text{-success-rate of } \hat{f}_l \text{ on } l'] \geq 1 - \frac{\delta^2}{256} \tag{15}$$

$$\text{and} \qquad E_{(l,l')}[\text{agreement of } \hat{f}_l \text{ and } f \text{ on } l'] \geq \frac{\delta}{4}. \tag{16}$$

Let $\alpha = 1/32$. The main reason why we can "bootstrap" is the following. The two expectations in statements (15) and (16) are essentially unchanged (except for a "fudge" factor of $1 - 1/\sqrt{q}$, which is negligible) if we change the method of picking (l, l') as follows: instead of picking a random pair of lines in F^m, we pick a pair randomly from all noncoplanar pairs of lines in a *fixed* cube c in which the average d-success-rate of f is at least $\delta(1 - \alpha)$. The reason behind our claim is that when we pick a random pair of lines in F^m, then with probability $1 - q^2/q^m$ they are non-coplanar, in which case they determine a unique cube. Furthermore, this cube is randomly distributed among all cubes, so with a further probability at least $1 - \frac{1}{\alpha^2\delta^2 q}$ the d-success-rate of f in this cube is at least $\delta(1-\alpha)$ (Lemma 19). Thus, if we are willing to ignore a factor $(1 - \frac{1}{q^{m-2}} - \frac{1}{\alpha^2\delta^2 q})$ (which we are, since this is $> 1 - 1/\sqrt{q}$ for a large enough q), it suffices to compute the expectations in (15) and (16) when (l, l') is a random pair of non-coplanar lines in a cube c in which the d-success-rate of f is at least $\delta(1 - \alpha)$. We restrict attention to such (l, l').

By the trivariate case of Theorem 4, there is a degree d trivariate polynomial that has agreement at least $\delta(1 - 2\alpha)$ with f on cube c. Let P_1 be one such polynomial and let $P_2, \ldots P_{k_0}$ be all the other degree d polynomials that have agreement at least $\delta(1 - 6\alpha)$ with f on cube c.

Let c_2, c_3 be the constants of the same name that occured in the statement of Theorem 3 for the case $m = 3$. Let $\epsilon = 1/q^{1/4c_3}$. Let $P_{k_0+1}, \ldots, P_k$ be all the degree d polynomials whose agreement with f on cube c is between ϵ^{c_3}/c_2 and $\delta(1 - 6\alpha)$. By Proposition 2, the set of polynomials we have identified thus far is not too big: $k_0 \le 8/\delta$ and $k \le 4c_2/\epsilon^{c_3}$. Furthermore, we know by the trivariate case of Theorem 3 that if we restrict the low degree test on f to those points of cube c where f doesn't agree with any of $P_1, \ldots, P_k$, then the success probability is at most ϵ. This will be important.

We hope to show ultimately that for "most" lines l, the function $\hat{f}_l$ has high agrement with one of $P_1, P_2, \ldots, P_{k_0}$. For any trivariate polynomial Q and line l, let $Q|_l$ denote its restriction to line l. We likewise define the restriction $Q|_s$ for a plane s. We say that line l is *nice* if the restrictions $P_1|_l, P_2|_l, \ldots, P_k|_l$ are all distinct and $P_d^f(l)$, the univariate degree d polynomial that has the highest agreement with f on l, is one of $P_1|_l, P_2|_l, \ldots, P_{k_0}|_l$.

Let $\gamma = 4\epsilon/\delta = 4/\delta q^{1/4c_3}$.

Claim 1: *At least $1 - \gamma$ fraction of the lines l in cube c are nice.*

Proof of Claim 1: The fraction of lines l for which $P_i|_l = P_j|_l$ for some $i \ne j$ is at most $\binom{k}{2} \times \frac{d}{q}$, since for any fixed i, j, the fraction of lines l for which $P_i|_l = P_j|_l$ is at most d/q. Since $k \le 4c_2/\epsilon^{c_3}$, we have

$$\binom{k}{2} \times \frac{d}{q} \le \frac{8c_2^2 d q^{2c_3/4c_3}}{q} \le \frac{8dc_2^2}{\sqrt{q}}.$$

Now we estimate the fraction of lines for which $P_d^f(l)$ is not one of $P_1|_l, P_2|_l, \ldots, P_{k_0}|_l$. Such a line must satisfy one or more of the following properties.

1. $P_1|_l$ has agreement less than $\delta(1 - 4\alpha)$ with f on line l. By Lemma 18, the fraction of such lines is at most $\frac{1}{4\alpha^2\delta q}$.

2. $P_1|_l$ has agreement $\beta \ge \delta(1 - 4\alpha)$ with f on line l but one of $P_{k_0+1}|_l, \ldots, P_k|_l$ has agreement more than β. By Lemma 18, the fraction of such lines is at most $\frac{1}{4\alpha^2\delta q} \times (k - k_0)$, which is at most $\frac{c_2}{\delta\alpha^2 q^{3/4}}$ since $k \le 4c_2/\epsilon^{c_3} < 4c_2 q^{1/4}$.

3. $P_1|_l$ has agreement $\beta \geq \delta(1-4\alpha)$ with f on line l but some univariate polynomial that is not $P_1|_l, P_2|_l, \ldots, P_k|_l$ has agreement more than β with f on l. Since the success probability of f on points where it does not agree with $P_1|_l, \ldots, P_k|_l$ is at most ϵ, the fraction of lines on which this success probability is more than $\delta(1-4\alpha)$ is at most $\epsilon/\delta(1-4\alpha) < 2\epsilon/\delta < 2/\delta q^{1/4c_3}$.

Hence the fraction of lines that are not nice is at most

$$\frac{8dc_2^2}{\sqrt{q}} + \frac{1}{4\alpha^2\delta q} + \frac{2}{\delta^3\alpha^2 q^{3/4}} + \frac{2}{\delta q^{1/4c_3}}.$$

The last term dominates when q is large enough, so this fraction is at most $4/\delta q^{1/4c_3}$. $\square$

We say that a plane s in c is *well-behaved* if (i) each of $P_1|_s, P_2|_s, \ldots, P_{k_0}|_s$ has agreement at least $\delta(1-8\alpha)$ with f on s (ii) no bivariate polynomial that is not one of $P_1|_s, \ldots, P_k|_s$ has agreement more than $\delta(1-8\alpha)$ with f on plane s.

Claim 2: *At least* $1-\gamma$ *fraction of planes in* c *are well-behaved.*

Proof of Claim 2: Each of $P_1, \ldots, P_{k_0}$ has agreement at least $\delta(1-6\alpha)$ with f on cube c. Picking a random plane involves picking three points at random from the cube. Hence we can use pairwise independence (i.e., Chebyshev's inequality) to conclude

$$\Pr_s[\text{agreement between } P_i|_s \text{ and } f \text{ on } s \text{ is} < \delta(1-8\alpha)] \leq \frac{4}{\alpha^2\delta q^2}.$$

Next, we bound the fraction of planes s such that some bivariate polynomial different from $P_1|_s, \ldots, P_k|_s$ has agreement at least $\delta(1-8\alpha) > \delta/2$ with f on plane s. Note that in such a plane the restriction of f to points where it doesn't agree with $P_1, \ldots, P_k$ passes the low degree test with probability at least $\delta/2$. But the case $m=3$ of Theorem 3 and symmetry implies that the average of this rate over the entire cube is at most ϵ. Hence the fraction of such planes is at most $2\epsilon/\delta < 2/\delta q^{1/4c_3}$.

Thus the fraction of planes that are not well-behaved is at most $4/\alpha^2\delta q^2 + 2/\delta q^{1/4c_3}$, which for large enough q is at most $4/\delta q^{1/4c_3}$. $\square$

Claim 3: *For at least* $1-\sqrt{\gamma}$ *fraction of lines in cube* c*, at least* $1-\sqrt{\gamma}$ *fraction of the planes containing that line are well-behaved.*

Proof of Claim 3: Among all planes that contain any line l, let σ_l denote the fraction that are well-behaved. Then by symmetry we know that $E_l[\sigma_l]$ is exactly the fraction of well-behaved planes in cube c, which is at least $1-\gamma$ by Claim 2. Averaging implies that $\sigma_l \geq 1-\sqrt{\gamma}$ for at least $1-\sqrt{\gamma}$ fraction of l. $\square$

Now call a line l *super* if it is nice and if at least $1-\sqrt{\gamma}$ fraction of the planes containing l are well-behaved. By Claims 1 and 3, at least $1-\gamma-\sqrt{\gamma}$ fraction of lines in cube c are super.

Claim 4: *If line* l *is super, then for every line* l' *that is non-coplanar with* l*,*

$$d\textit{-success-rate of } \hat{f}_l \textit{ on } l' \geq 1-\sqrt{\gamma} \tag{17}$$

and for a random line l' *noncoplanar with* l*,*

$$E_{l'}[\textit{agreement between } \hat{f}_l \textit{ and } f \textit{ on cube } c] \geq \delta(1-\sqrt{\gamma})(1-8\alpha). \tag{18}$$

Proof of Claim 4: Recall that the set of planes containing l is a partition of cube c. Since l is nice, $P_d^f(l)$ is $P_i|_l$ for some $i \in [1, k_0]$. In any plane s containing l, the bivariate polynomial used to define $\hat{f}_l$ in that plane must agree with $P_i|_l$ on l and must have agreement at least $\delta/2$ with f on s. If s is well-behaved for l, then only $P_i|_s$ qualifies. Hence the agreement between $\hat{f}_l$ and f on this plane is at least $\delta(1-8\alpha)$. Summing over all planes containing l, we see that the agreement between $\hat{f}_l$ and f on the cube c is at least $(1-\sqrt{\gamma})\cdot\delta(1-8\alpha)$. Now the claim in (18) follows.

Now we prove the claim in (17). Consider any line l' non-coplanar with l. Every plane s containing l meets l' in exactly one point, say x. If s is well-behaved, then $\hat{f}_l(x) = P_i(x)$, as already argued. Hence $P_i|_{l'}$, the restriction of P_i to l', has agreement at least $1-\sqrt{\gamma}$ with f on l'. $\Box$

By examining Claim 4 we realize that the Lemma is more or less proved, since at least $1-\gamma-\sqrt{\gamma}$ fraction of lines in c are super. We make $q > (2^{32}/\delta^4)^{4c_3}$, which makes $1-\sqrt{\gamma} > 1-\delta^2/512$. Now the first expectation is $\delta(1-\sqrt{\gamma})(1-\gamma-\sqrt{\gamma})(1-8\alpha)$ which is at least $\delta/4$. The second expectation is $(1-\sqrt{\gamma})(1-\gamma-\sqrt{\gamma}) > 1-\delta^2/256$.
$\Box$

4 Some technical lemmas

We prove some of the lemmas used elsewhere in the paper.

Lemma 16 (Schwartz) *An m-variate degree D polynomial that is nonzero can be zero at at most $D/|Q|$ fraction of points in F^m.*

Proof: Simple induction on degree. $\Box$

Lemma 17 *Let $r_1, r_2, \ldots, \in [0,1]$ be some real numbers whose average is α. Then (i) at least $\frac{\alpha-\rho}{1-\rho}$ fraction of them are greater than ρ (ii) at most $1/k$ fraction of them are more than $k\cdot\alpha$.*

Proof: (i) If the desired fraction is s, then s satisfies $s+(1-s)\rho \geq \alpha$.(ii) If the desired fraction is t then t satisfies $t\cdot k\alpha \leq \alpha$. $\Box$

In the following lemmas, $\mathrm{F} = \mathrm{GF}(q)$ is any finite field. For the next lemma, we remind the reader that a line in F^m has q points.

Lemma 18 (Well-distribution lemma for lines) *Let $S \subseteq F^m$ be a set whose size is $\mu\cdot q^m$. For every $K > 0$, at least $1-\frac{1}{K^2}$ fraction of lines in F^m have between $\mu q(1-\frac{K}{\sqrt{\mu q}})$ and $\mu q(1+\frac{K}{\sqrt{\mu q}})$ points from S.*

Proof: Imagine picking a line $l = \{\overline{u}+t\overline{v} : t \in \mathrm{F}\}$ randomly, that is, by picking $\overline{u}, \overline{v}$ randomly from F^m. For $a \in \mathrm{F}$ let the random variable X_a be 1 if $\overline{u}+a\overline{v} \in S$ and 0 otherwise. Then $E[X_a] = \mu$ and the X_a's are pairwise independent. By the Chebyshev inequality,

$$\Pr[|\sum_{a\in\mathrm{F}} X_a - \mu q| \geq K\sqrt{\mu q}] \leq \frac{1}{K^2}.$$

$\Box$

We choose to state the next lemma in terms of the d-success-rate, but it is also true if instead of d-success-rate we associate any set of positive fractions with the lines of F^m and look at their average value in a random cube.

Lemma 19 (Well-distribution lemma for cubes) *For any $\alpha > 0$ and $m > 3$, if any function $f : F^m \to F$ has d-success-rate δ, then in a random cube C,*

$$\Pr_{\text{cube } C}[\textit{Average d-Success-rate of } f \textit{ on lines in } C \leq (1-\alpha)\delta] \leq \frac{2}{\alpha^2\delta^2|F|}.$$

Proof: Let the random variable Y_C denote the average d-success-rate of f on lines in cube C. By symmetry (see the note on symmetry in Example 1), $E_C[Y_C] = \delta$. Let X_C be the random variable $Y_C - \delta$, so $E_C[X_C] = 0$. By the Chebyshev inequality,

$$\Pr_C[|X_C| \geq \alpha\delta] \leq \frac{V[X_C]}{\alpha^2\delta^2},$$

where $V[X_C]$ denotes the variance of X_C. We show next (using an argument from [BGS95]) that $V[X_C] \leq 2/|\mathrm{F}|$, thus proving the lemma.

Let $q = |\mathrm{F}|$ and let K denote the number of lines in F^3 (thus $K = \binom{q^3}{2}/\binom{q}{2}$). Let us number these lines in some arbitrary way, so that we can refer to the "ith line of F^3." We similarly talk about the "ith line of cube C." Now for $1 \leq i \leq K$, let X_i be the r.v. X_i = success rate of f on ith line of $C - \delta$. Note that $X_C = E_{i \leq K}[X_i]$. Hence

$$V[X_C] \;=\; E_C[\, E_{i,j \leq K}[X_i X_j]\,] \tag{19}$$

To upperbound the expression in (19), we note that $1 - 1/q$ fraction of all pairs (i,j) correspond to a pair of non-coplanar lines in F^3. We claim that $E[X_i X_j] \leq 1/q$ when i, j are non-coplanar. The reason is that we can pick a random cube C in F^m by picking a random pair of noncoplanar lines l_1, l_2 in F^m and making l_1 the ith line of C and l_2 the jth line of C; this completely determines C. Once we have fixed l_1, the fraction of lines that are non-coplanar with it is $1 - q^2/q^m$, so the average d-success-rate of f among lines that are non-coplanar with l_1 is within $[\delta - 1/q, \delta + 1/q]$. Hence $E[|X_j||X_i|] \leq 1/q$.

Thus

$$V[X_C] \leq \frac{1}{q}\cdot 1 + (1 - \frac{1}{q})\cdot\frac{1}{q} \leq \frac{2}{q}.$$

□

Now we state two lemmas that are related to Schwartz's lemma.

Lemma 20 *Let $p \in F[x_1, x_2, \ldots, x_m]$ be a polynomial of degree exactly D. Then on at least $1 - D/|F|$ fraction of lines in F^m, the restriction of p has degree no less than D*

Proof: Let $\{(a_1 + b_1 t, \ldots, a_m + b_m t) : t \in \mathrm{F}\}$ denote a generic line (i.e., think of $a_1, \ldots, a_m$, and $b_1, \ldots, b_m$ as variables).

Then p can be expressed as a degree D polynomial in $\mathrm{F}[a_1, \ldots, a_m, b_1, \ldots, b_m, t]$. View it as a univariate degree D polynomial in t whose coefficients are in the ring $\mathrm{F}[a_1, \ldots, a_m, b_1, \ldots, b_m]$. Since the leading coefficient is a polynomial of degree exactly D, Schwartz's lemma implies that this coefficient cannot vanish for more than $D/|\mathrm{F}|$ values of $(a_1, \ldots, a_m, b_1, \ldots, b_m) \in \mathrm{F}^{2m}$. □

Lemma 21 *Let F be a finite field and $\overline{F}$ be its closure. Let $p \in \overline{F}[x_1, x_2, \ldots, x_m]$ be a polynomial of degree exactly D. If on more than $D/|F|$ fraction of the points in F^m, the value of p at the point is in F, then p must be a polynomial in $F[x_1, \ldots, x_m]$.*

Proof: The proof imitates the proof of Schwartz's lemma. It uses the observation that a univariate degree D polynomial that has at least one coefficient in $\overline{F} \setminus \mathrm{F}$ must take values in $\overline{F} \setminus \mathrm{F}$ on at least $|\mathrm{F}| - D$ points in F. The reason is that $D+1$ values are sufficient to recover the polynomial by interpolation.

(By contrast, the proof of Schwartz's lemma uses the fact that a nonzero degree D univariate polynomial can be 0 only at D points in F.) □

Lemma 22 *Let F be a finite field and $\overline{F}$ be its closure. Let $p \in \overline{F}[x_1, x_2, \ldots, x_m]$ be a polynomial of degree exactly D. If on more than $D/|F|$ fraction of the lines in F^m, the restriction of p to the line is in $F[t]$ (where t is the line parameter), then p must be a polynomial in $F[x_1, \ldots, x_m]$.*

Proof: Follows from Lemma 21 in a way analogous to the proof of Lemma 20 followed from Schwartz's Lemma. □

4.1 A version of Hilbert Irreducibility

In order to prove the cases $m = 2, 3$ we needed something like the following: if a polynomial is irreducible, then its restriction on most lines does not have a "linear" factor. Now we state and prove this fact. It is a simpler version of Kaltofen's "Effective Hilbert Irreducibility" [K85]. in that it focusses only factors that are monic and linear in one of the variables. The proof essentially follows from Kaltofen [K95], and is included here for completeness.

A polynomial (in this section, "polynomial" means a formal polynomial) $Q \in \mathrm{F}[z, y_1, \ldots, y_m]$ is said to be *monic with respect to z* if the leading coefficient of z is a constant (i.e., an element of F). It is *absolutely irreducible* if it does not factor over $\overline{\mathrm{F}}$, the algebraic closure of F.

Theorem 23 *Let $Q \in F[z, y_1, y_2, \ldots, y_m]$ be a degree l polynomial that is absolutely irreducible and monic in z. Then the fraction of $(a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_m) \in F^{2m}$ for which*

$$Q(z,\ a_1 t + b_1,\ \ldots, a_m t + b_m) \in F[z,t] \text{ has a factor of the form } z - p(t) \text{ in } \overline{F}[z,t]$$

is at most $1 - O(l^3/q)$.

Remark: The monicity condition is not strictly necessary. It suffices instead that if $Q = \sum_i Q_i(y_1, \ldots, y_m) z^i$ then $\gcd(Q_1, \ldots, Q_l) = 1$. If this condition is satisfied then we can apply the lemma to the polynomial $Q_l(y_1, \ldots, y_m)^l Q(\frac{z}{Q_l}, y_1, \ldots, y_m)$.

In the proofs that follow we will use $\mathrm{F}(y_1, y_2, \ldots, y_k)$ and $\overline{F}(y_1, y_2, \ldots, y_k)$ to denote the quotient fields of $\mathrm{F}[y_1, y_2, \ldots, y_k]$ and $\overline{F}[y_1, y_2, \ldots, y_k]$ respectively.

In the following lemma, congruence (i.e., $\equiv$) modulo $[y_1, \ldots, y_m]^{l+1}$ means that the polynomials on the two sides of the $\equiv$ are identical once we throw away all terms of total degree $l+1$ or higher.

Lemma 24 *Let $Q \in F[z, y_1, \ldots, y_m]$ and $g \in F[y_1, \ldots, y_m]$ be any polynomials. Then for any $l \geq 0$,*

$$Q(g(y_1, \ldots, y_m), y_1, \ldots, y_m) \equiv 0 \ (\bmod\ [y_1, \ldots, y_m]^{l+1}) \tag{20}$$

iff there is a polynomial $h \in F[z, y_1, \ldots, y_m]$ such that

$$Q(z, y_1, \ldots, y_m) \equiv (z - g(y_1, \ldots, y_m)) h(z, y_1, \ldots, y_m) \ (\bmod\ [y_1, \ldots, y_m]^{l+1}). \tag{21}$$

Furthermore, if such an h exists, then it is unique modulo $[y_1, \ldots, y_m]^{l+1}$.

Proof: By the division theorem, there exists a unique pair of polynomials $h(z, y_1, \ldots, y_m)$ and $r(y_1, \ldots, y_m)$ such that

$$Q(z, y_1, \ldots, y_m) = (z - g(y_1, \ldots, y_m))h(z, y_1, \ldots, y_m) + r(y_1, \ldots, y_m).$$

Thus

$$Q(g(y_1, \ldots, y_m), y_1, \ldots, y_m) \equiv 0 \ (\bmod\ [y_1, \ldots, y_m]^{l+1})$$

iff $r(y_1, \ldots, y_m) \equiv 0 \ (\bmod\ [y_1, \ldots, y_m]^{l+1})$, which happens iff

$$Q(z, y_1, \ldots, y_m) \equiv (z - g(y_1, \ldots, y_m))h(z, y_1, \ldots, y_m) \ (\bmod\ [y_1, \ldots, y_m]^{l+1}).$$

Lastly, uniqueness of h follows by the uniqueness of the solution to the division theorem. $\square$

The following lemma will be important.

Lemma 25 (A Version of Hensel Lifting) *Let $Q \in F[z, y_1, \ldots, y_m]$ be any polynomial and $\alpha \in \overline{F}$ be a root of multiplicity 1 of the polynomial $p(z) \stackrel{\text{def}}{=} Q(z, 0, 0, \ldots, 0)$. Then for each $l \geq 1$ there exists a unique polynomial $q_l \in \overline{F}[y_1, \ldots, y_m]$ of total degree l such that*

$$Q(q_l(y_1, \ldots, y_m), y_1, \ldots, y_m) \equiv 0 \ (\bmod\ [y_1, \ldots, y_m]^{l+1}) \ \textit{ and } q_l(0, 0, \ldots, 0) = \alpha.$$

Proof: We simplify notation by using the symbol $\hat{y}$ as a shorthand for $y_1, \ldots, y_m$ throughout the proof.

By Lemma 24, it suffices to prove that for each $l \geq 0$, there exists a unique pair of polynomials $h_l \in \overline{\mathrm{F}}[z, \hat{y}], q_l \in \overline{\mathrm{F}}[\hat{y}]$ whose total degree in $\hat{y}$ is at most l and which satisfy:

$$Q(z, \hat{y}) \equiv (z - q_l(z, \hat{y})) \cdot h_l(z, \hat{y}) \ (\bmod\ [\hat{y}]^{l+1}) \ \text{ and } q_l(0, 0, \ldots, 0) = \alpha. \tag{22}$$

We use induction on l. The base case $l = 0$ is trivial, since $q_0 = \alpha$ and $h_0(z) = Q(z, 0, \ldots, 0)/(z - \alpha)$ are the only such polynomials.

Assume the statement is true up to $l \leq k$. The uniqueness property implies that $q_0, q_1, \ldots, q_k, h_0, \ldots, h_k$ satisfy for all $1 \leq i \leq k, 0 \leq j \leq i$:

$$q_i(\hat{y}) \equiv q_{i-j}(\hat{y}) \ (\bmod\ [\hat{y}]^{i-j+1})$$

and

$$h_i(z, \hat{y}) \equiv h_{i-j}(z, \hat{y}) \ (\bmod\ [\hat{y}]^{i-j+1}).$$

This means, for example, that each q_i is expressible as

$$q_i(\hat{y}) = q_j(\hat{y}) + \ \text{ terms whose degree in } \hat{y} \text{ is between } j+1 \text{ and } i.$$

A similar fact holds for the h_i's. Thus if any polynomials q_{k+1}, h_{k+1} satisfy condition (22) for $l = k + 1$, then they must necessarily be of the form

$$q_{k+1}(\hat{y}) = q_k(\hat{y}) + \sum_{d_1, \ldots, d_m : \sum_i d_i = k+1} c_{d_1, \ldots, d_m} \prod_i y_i^{d_i}, \tag{23}$$

$$h_{k+1}(z, \hat{y}) = h_k(z, \hat{y}) + \sum_{d_1, \ldots, d_m : \sum_i d_i = k+1} e_{d_1, \ldots, d_m}(z) \prod_i y_i^{d_i}, \tag{24}$$

where each $c_{d_1,\ldots,d_m} \in \overline{\mathrm{F}}$ and each $e_{d_1,\ldots,d_m} \in \overline{\mathrm{F}}[z]$.

Now we show that there exist unique values for the $c_{d_1,\ldots,d_m}$'s and the $e_{d_1,\ldots,d_m}$'s such that q_{k+1}, h_{k+1} have the required properties. Let $R \in \overline{\mathrm{F}}[z, y_1, \ldots, y_m]$ be defined as

$$R(z, \hat{y}) = Q(z, \hat{y}) - (z - q_k(\hat{y}))h(z, y_1, \ldots, y_m) \ (\bmod\ [\hat{y}]^{k+2}). \tag{25}$$

Note that each term in R has degree $k+1$ in $\hat{y}$. Express it as

$$R(z, \hat{y}) = \sum_{d_1,\ldots,d_m : \sum_i d_i = k+1} r_{d_1,\ldots,d_m}(z) \prod_i y_i^{d_i}.$$

Since we desire q_{k+1}, h_{k+1} to satisfy

$$Q(z, \hat{y}) \equiv (z - q_{k+1}(\hat{y}))h_{k+1}(\hat{y}) \ \ (\bmod\ [\hat{y}]^{k+2}]),$$

we replace this in (25) to get

$$R(z, \hat{y}) = (z - q_{k+1}(\hat{y}))h_{k+1}(\hat{y}) - (z - q_k(\hat{y}))h(z, y_1, \ldots, y_m) \ (\bmod\ [\hat{y}]^{k+2}).$$

Now replacing the expressions from (23) and (24) in (25) and equating coefficients of like terms, we get the following for every tuple of degrees $(d_1, \ldots, d_m)$ satisfying $\sum_i d_i = k+1$:

$$(z - q_0)e_{d_1,\ldots,d_m}(z) + h_0(z)c_{d_1,\ldots,d_m} = r_{d_1,\ldots,d_m}(z).$$

Since α is a root of multiplicity 1 of $Q(z, 0, \ldots, 0)$, $h_0(z) = Q(z, 0, \ldots, 0)/(z - \alpha)$ does not have a root at $z = \alpha$. Furthermore, $q_0 = \alpha$. So this system solves as

$$c_{d_1,\ldots,d_m} = r_{d_1,\ldots,d_m}(\alpha)/h_0(\alpha)$$

and

$$e_{d_1,\ldots,d_m} = \frac{1}{z - \alpha}(r_{d_1,\ldots,d_m}(z) - h_0(z)c_{d_1,\ldots,d_m}).$$

Thus we have proved both the existence and uniqueness of q_{k+1}, h_{k+1}, thus completing the induction.

□

We say that a univariate polynomial $p \in \mathrm{F}[x]$ is *square-free* if it doesn't have a repeated root. For such a polynomial if $f \cdot g = p$ is any factorization of p in $\overline{\mathrm{F}}[x]$, then f, g have no common factor. The following lemma about *discriminants* gives a necessary and sufficient condition for square-freeness.

Lemma 26 *A degree k polynomial $p = \sum_{i=0}^{d} p_i x^i \in F[x]$ is square-free iff the determinant of the following $(2d-1) \times (2d-1)$ matrix (the so-called* discriminant*) is nonzero, where $g_i = (i+1)p_{i+1}$ for $i \leq d-1$ and $g_d = 0$.*

$$\begin{pmatrix}
p_0 & 0 & 0 & \cdots & 0 & g_0 & 0 & 0 & \cdots & 0 \\
p_1 & p_0 & 0 & \cdots & 0 & g_1 & g_0 & 0 & \cdots & 0 \\
p_2 & p_1 & p_0 & \cdots & 0 & g_2 & g_1 & g_0 & \cdots & 0 \\
p_3 & p_2 & p_1 & \cdots & 0 & g_3 & g_2 & g_1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
p_d & p_{d-1} & p_{d-2} & \cdots & p_0 & g_d & g_{d-1} & g_{d-2} & \cdots & g_0 \\
0 & p_d & p_{d-1} & \cdots & p_1 & 0 & g_d & g_{d-1} & \cdots & g_1 \\
0 & 0 & p_d & \cdots & p_2 & 0 & 0 & g_d & \cdots & g_2 \\
0 & 0 & 0 & \cdots & p_3 & 0 & 0 & 0 & \cdots & g_3 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & \cdots & p_d & 0 & 0 & 0 & \cdots & g_d
\end{pmatrix}$$

Proof: p has a repeated factor iff p and p' (the derivative of p) share a factor, which happens iff there are nonzero degree $d-1$ polynomials A and B such that $A \cdot p + B \cdot p' = 0$. We can try to find such polynomials by representing their coefficients as unknowns and writing a (homogeneous) linear system of equations. The determinant we have written is the determinant of this system; it is zero iff a nontrivial solution (i.e., A and B) exists. □

Corollary 27 *Let $Q \in F[z, y_1, \ldots, y_m]$ have degree l and be absolutely irreducible. Then there is a nonzero polynomial $\Phi_Q \in F[y_1, \ldots, y_l]$ of degree $O(l^2)$ such that for all $b_1, \ldots, b_m \in F$,*

$$\Phi_Q(b_1, \ldots, b_m) \neq 0 \;\Rightarrow\; Q(z, b_1, \ldots, b_m) \text{ is square-free in } F[z].$$

Proof: Write Q as $\sum_{i=0}^{l} z^i p_i(y_1, \ldots, y_m)$, that is, as a polynomial in $\mathrm{F}[y_1, \ldots, y_m][z]$. Let Φ_Q be the discriminant of this polynomial. Since the discriminant is a polynomial of degree $(2l+1)$ in the coeeficients, and each coeeficient in this case is itself a degree l polynomial in $\mathrm{F}[y_1, \ldots, y_m]$, we conclude that $\Phi_Q \in \mathrm{F}[y_1, \ldots, y_m]$ has degree $l(2l+1)$. Furthermore, since Q is irreducible in $\mathrm{F}[y_1, \ldots, y_m][z]$, Gauss' Lemma implies that Q is irreducible in $\overline{F}(y_1, \ldots, y_m)[z]$, so the discriminant Φ_Q is non-zero. □

Lemma 28 *Let $Q \in F[z, y_1, y_2, \ldots, y_m]$ be a monic, degree l polynomial that is absolutely irreducible. Suppose $b_1, \ldots, b_m$ satisfy $\Phi_Q(b_1, \ldots, b_m) \neq 0$, where Φ_Q is the polynomial in Corollary 27. Then there exists a nonzero polynomial $\Psi_Q \in \overline{F}[v_1, \ldots, v_m]$ of degree l^3 such that for all $a_1, \ldots, a_m \in F$,*

$$\Psi_Q(a_1, a_2, \ldots, a_m) \neq 0 \implies f(z, a_1t+b_1, \ldots, a_mt+b_m) \text{ has no factor like } z - p(t) \text{ in } \overline{F}[z,t].$$

Proof: By the hypothesis, $Q(z, b_1, \ldots, b_m)$ is square-free. Define $T \in \mathrm{F}[z, y_1, \ldots, y_m]$ as

$$T(z, y_1, \ldots, y_m) = Q(z, y_1 + b_1, \ldots, y_m + b_m).$$

Clearly, T is absolutely irreducible and $T(z, 0, \ldots, 0)$ is square-free. For each $a_1, \ldots, a_m \in \mathrm{F}$, let $T_{a_1, \ldots, a_m} \in \overline{\mathrm{F}}[z, t]$ be defined as

$$T_{a_1, \ldots, a_m}(z, t) \;=\; T(z, a_1 t, a_2 t, \ldots, a_m t). \tag{26}$$

We wish to give a "nice" description (namely, as roots of a low-degree polynomial Ψ_Q) of those tuples $(a_1, \ldots, a_m)$ for which

$$T_{a_1, \ldots, a_m} \text{ has a factor of the form } z - p(t) \text{, where } p \in \overline{\mathrm{F}}[t]. \tag{27}$$

Let $\alpha_1, \ldots, \alpha_k$ be all the roots of $T(z, 0, \ldots, 0)$. Thus $k \leq l$ and the α_i's are distinct. By Lemma 25, for each $i = 1, \ldots, k$, there is a unique degree l polynomial $g_i \in \overline{\mathrm{F}}[y_1, \ldots, y_m]$ such that

$$T(g_i(y_1, \ldots, y_m), y_1, \ldots, y_m) \equiv 0 \;(\bmod\; [y_1, \ldots, y_m]^{l+1}) \qquad \text{and } g_i(0, \ldots, 0) = \alpha_i. \tag{28}$$

Note that $g_i \neq g_j$ for $i \neq j$, since g_i and g_j differ at $(0, \ldots, 0)$. Further, for each i

$$T(g_i(y_1, \ldots, y_m), y_1, \ldots, y_m) \neq 0, \tag{29}$$

since otherwise $z - g_i(y_1, \ldots, y_m)$ would be a factor of T and T is known to be absolutely irreducible.

Now let us identify tuples $(a_1, \ldots, a_m)$ for which $T_{a_1,\ldots,a_m}$ has a linear factor. For each $i = 1, \ldots, k$, think of the polynomial $g_i(a_1 t, \ldots, a_m t)$ as a univariate polynomial in t. By examining (28) and the definition of $T_{a_1,\ldots,a_m}$, we see that for each $a_1, \ldots, a_m \in \mathrm{F}$,

$$T_{a_1,\ldots,a_m}(g_i(a_1 t, \ldots, a_m t), t) \equiv 0 \pmod{[t]^{l+1}} \text{ and } g_i(a_1 t, \ldots, a_m t) \text{ is } \alpha_i \text{ at } t = 0.$$

The degree of $g_i(a_1 t, \ldots, a_m t) \in \overline{\mathrm{F}}[t]$ is at most l. So we conclude from the uniqueness condition in the conclusion of Lemma 24 that $T_{a_1,\ldots,a_m}$ has a factor of the form $z - p(t)$ for $p \in \overline{\mathrm{F}}[t]$ iff that factor is $z - g_i(a_1 t, \ldots, a_m t)$ for some $i \in [1..k]$. In other words, iff $T_{a_1,\ldots,a_m}(g_i(a_1 t, \ldots, a_m t), t)$ is the zero polynomial. Now we show that the set of $(a_1, \ldots, a_m)$ for which $T_{a_1,\ldots,a_m}(g_i(a_1 t, \ldots, a_m t), t)$ is the zero polynomial have a nice description as the roots of some polynomial Ψ_Q.

When $v_1, \ldots, v_n$ are indeterminates, then polynomial $T(g_i(v_1 t, \ldots, v_m t), v_1 t, \ldots, v_m t)$ is nonzero (see (29)). Write this polynomial as $\sum_j p_{ij}(v_1, \ldots, v_m) t^i$, where each $p_{ij} \in \overline{\mathrm{F}}[v_1, \ldots, v_m]$ is a degree l^2 polynomial. For each i pick a j_i such that p_{i,j_i} is nonzero. Then define Ψ_Q as

$$\Psi_Q(v_1, \ldots, v_m) = \prod_i p_{i,j_i}(v_1, \ldots, v_m). \tag{30}$$

Now consider any $(a_1, \ldots, a_m)$ such that $\Psi_Q(a_1, \ldots, a_m) \neq 0$. Then $T(g_i(a_1 t, \ldots, a_m t), a_1 t, \ldots, a_m t)$ is a nonzero polynomial in $\overline{\mathrm{F}}[t]$ for $i = 1, \ldots, k$. As already argued, $T_{a_1,\ldots,a_m}$ has no linear factor for such an $(a_1, \ldots, a_m)$.
$\square$

Now we are ready to prove Theorem 23.

Proof:(of Theorem 23) Pick $(b_1, \ldots, b_m)$ randomly from F^m. With probability $1 - O(l^2/q)$, the polynomial Φ_Q Corollary 27 becomes nonzero. Pick $(a_1, \ldots, a_m)$ randomly from F^m. With a further probability $1 - l^3/q$, the polynomial Ψ_Q from Lemma 28 becomes nonzero and so $Q(z, a_1 t + b_1, \ldots, a_m t + b_m)$ has no linear factor. Thus we have shown that with probability $(1 - O(l^2/q)(1 - O(l^3/q)) = (1 - O(l^3/q))$, $Q(z, a_1 t + b_1, \ldots, a_m t + b_m)$ has no linear factor. $\square$

5 Conclusions

We do not know how to reduce the number of provers in our constructions to 2. So long as we use the verifier composition idea of [AS92], 3 provers appears to be the best possible. Reducing the number of provers to 2 would imply the NP-hardness of approximation problems dealt with in [ABSS93].

Thanks

Sanjeev Arora thanks Laci Babai and Kati Friedl for introducing him to "symmetry-based" arguments for the low degree test in summer 1993. We thank Dick Lipton for saving us from fruitless labor on an incorrect conjecture on irreducibility (he provided a counterexample). We also thank Erich Kaltofen for providing pointers to his work.

References

[ALRS92] S. Ar, R. Lipton, R. Rubinfeld and M. Sudan. Reconstructing algebraic functions from noisy data. *Proceedings of the* 33rd *Symposium on Foundations of Computer Science*, IEEE, 1992.

[A93] S. Arora *Unpublished, 1993.*

[A94] S. Arora. *Probabilistic Checking of Proofs and Hardness of Approximation Problems.* PhD thesis, U.C. Berkeley, 1994. Available from `http://www.cs.princeton.edu/~arora` .

[ABSS93] S. Arora, L. Babai, J. Stern and Z. Sweedyk. The hardness of approximating problems defined by linear constraints. *Proceedings of the* 34th *Symposium on Foundations of Computer Science*, IEEE, 1993.

[ALMSS92] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy. Proof verification and the hardness of approximation problems. *Proceedings of the* 33rd *Symposium on Foundations of Computer Science*, IEEE, 1992.

[AS92] S. Arora and S. Safra. Probabilistic checking of proofs: a new characterization of NP. *Proceedings of the* 33rd *Symposium on Foundations of Computer Science*, IEEE, 1992.

[BFL91] L. Babai, L. Fortnow, and C. Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1:3–40, 1991.

[BFLS91] L. Babai, L. Fortnow, L. Levin, and M. Szegedy. Checking computations in polylogarithmic time. *Proceedings of the* 23rd *Annual Symposium on Theory of Computing*, ACM, 1991.

[BGS95] M. Bellare, O. Goldreich and M. Sudan. Free bits, PCPs and non-approximability — towards tight results. *Proceedings of the* 36th *Symposium on Foundations of Computer Science*, IEEE, 1995. TR95-024 of ECCC, the *Electronic Colloquium on Computational Complexity*, `http://www.eccc.uni-trier.de/eccc/`.

[BGLR93] M. Bellare, S. Goldwasser, C. Lund, and A. Russell. Efficient probabilistically checkable proofs. *Proceedings of the* 25th *Annual Symposium on Theory of Computing*, ACM, 1993. (See also Errata sheet in *Proceedings of the* 26th *Annual Symposium on Theory of Computing*, ACM, 1994).

[BLR90] M. Blum, M. Luby, and R. Rubinfeld. Self-testing/correcting with applications to numerical problems. In *Proc. 22nd ACM Symp. on Theory of Computing*, pages 73–83, 1990.

[F96] U. Feige. A threshold of $\ln n$ for Set Cover. *Proceedings of the* 28th *Annual Symposium on Theory of Computing*, ACM, 1996.

[FGLSS91] . U. Feige, S. Goldwasser, L. Lóvasz, S. Safra and M. Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM*, 43(2):268-292, 1996.

[FK94] U. Feige and J. Kilian. Two prover protocols – Low error at affordable rates. *Proceedings of the* 26th *Annual Symposium on Theory of Computing*, ACM, 1994.

[FK95] U. Feige and J. Kilian. Impossibility results for recycling random bits in two-prover proof systems. *Proceedings of the* 27th *Annual Symposium on Theory of Computing*, ACM, 1995.

[FL92] U. Feige and L. Lóvasz. Two-prover one-round proof systems: Their power and their problems. *Proceedings of the* 24th *Annual Symposium on Theory of Computing*, ACM, 1992.

[FS95] K. FRIEDL AND M. SUDAN. Some improvements to low-degree tests. *Proceedings of the Third Israel Symposium on Theory and Computing Systems*, IEEE, 1995.

[GLRSW91] P. GEMMELL, R. LIPTON, R. RUBINFELD, M. SUDAN AND A. WIGDERSON. Self-testing/correcting for polynomials and for approximate functions. *Proceedings of the* 23rd *Annual Symposium on Theory of Computing*, ACM, 1991.

[GRS] O. GOLDREICH, R. RUBINFELD AND M. SUDAN. Learning polynomials with queries: The highly noisy case. *Proceedings of the* 36th *Symposium on Foundations of Computer Science*, IEEE, 1995.

[K85] E. KALTOFEN. Polynomial-time reductions from multivariate to bi- and univariate integral polynomial factorization. *SIAM Journal on Computing*, 14(2):469–489, 1985.

[K85] E. KALTOFEN. Effective Hilbert irreducibility. *Information and Control*, 66:123–137, 1985.

[K95] E. KALTOFEN. Effective Noether irreducibility forms and applications. *Journal of Computer and System Sciences*, 50(2):274-295, 1995.

[LS91] D. LAPIDOT AND A. SHAMIR. Fully Parallelized Multi-prover protocols for NEXP-time. *Proceedings of the* 32nd *Symposium on Foundations of Computer Science*, IEEE, 1991.

[LFKN92] C. LUND, L. FORTNOW, H. KARLOFF, AND N. NISAN. Algebraic methods for interactive proof systems. *Journal of the ACM*, 39(4):859-868, 1992.

[LY93] C. LUND AND M. YANNAKAKIS. On the hardness of approximating minimization problems. *Proceedings of the* 25th *Annual Symposium on Theory of Computing*, ACM, 1993.

[PS94] A. POLISHCHUK AND D. SPIELMAN. Nearly Linear Sized Holographic Proofs. *Proceedings of the* 26th *Annual Symposium on Theory of Computing*, ACM, 1994.

[R95] R. RAZ. A parallel repetition theorem. *Proceedings of the* 27th *Annual Symposium on Theory of Computing*, ACM, 1995.

[RazS96] R. RAZ AND S. SAFRA. Personal communication. March 1996. Manuscript. September 1996.

[RS93] R. RUBINFELD AND M. SUDAN. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing* 25:2, pp. 252–271, 1996.

[Sch80] J. T. SCHWARTZ. Probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27:701–717, 1980.

[Sh92] A. SHAMIR. IP = PSPACE. *Journal of the ACM*, 39(4):869-877, 1992.

[Su96] M. SUDAN. Maximum likelihood decoding of Reed Solomon codes. *Proceedings of the* 37th *Symposium on Foundations of Computer Science*, IEEE, 1996.

[T93] G. TARDOS. *Personal Communication*, 1993.

[T94] G. TARDOS. Multi-prover encoding schemes and three-prover proof systems. *Proceedings of the* 9th *Annual Conference on Structure in Complexity Theory*, IEEE, 1994.

[T96] G. TARDOS. *Personal Communication*, 1996.

A Construction of Constant-Prover 1-Round systems

A p-prover 1-round proof system for a language L consists of a verifier that checks membership proofs for L (a proof that a given input x is in L) in the following way. The proof consists of p oracles. (An oracle is a table that, for some $a, b > 0$, contains 2^b strings from $\{0,1\}^a$. When we supply this oracle a b-bit address, it returns the a-bit string stored at the corresponding location. We call a the *answer size* of the oracle.) The verifier is probabilistic. It uses its randomness to compute one address in each of the p oracles, reads the strings in those locations, and then computes an ACCEPT or REJECT decision. (The name "p-prover 1-round system" is a holdover from the past; we could also use "p-oracle 1-round systems.")

To construct very efficient $O(1)$-prover 1-round proof systems for SAT we use two standard techniques. First we plug our low degree test into a construction of [ALMSS92] to get a proof system with 3 provers that uses $O(\log n)$ random bits but the oracles in the proof have answer size $2^{\log^\beta n}$ for some $\beta < 1$. Then we use "verifier composition," a technique from [AS92], to reduce the answer size to $O(\log n)$ (the number of stays $O(1)$). Our technique for verifier composition will also rely on the low degree test.

To use verifier composition we also need to ensure that the verifier's ACCEPT/REJECT decision is computed in a very simple way, by evaluating a small circuit. At the start, the verifier uses its random string and the input to compute a circuit C and one location in each of the p oracles. After reading the oracles, the verifier outputs ACCEPT iff the concatenation of the strings it just read is a satisfying assignment to C. The size of C (= number of wires in it) is called the *circuit size* of the verifier.

Now we define $\mathrm{MIP}[p, r, a, e]$, the class of languages that have such verifiers.

Definition 5 (**MIP**$[p, r, a, e]$) *For a positive integer p and functions $r, a, e : \mathcal{Z}^+ \to \mathcal{Z}^+$, a language L is said to belong to MIP$[p, r, a, e]$ if there exists a probabilistic polynomial-time verifier V that on any input $x \in \{0,1\}^n$ uses $r(n)$ random bits, expects the membership proof to contain p oracles of answer size $a(n)$, and has the following behavior:*

1. *If $x \in L$, then there exist oracles $\pi_1, \ldots, \pi_p$ such that V always outputs ACCEPT (i.e., outputs ACCEPT with probability 1).*

2. *If $x \notin L$, then there for every set of oracles $\pi_1, \ldots, \pi_p$, verifier V outputs ACCEPT with probability at most $e(n)$.*

Furthermore, the circuit size of the verifier is polynomial in $a(n)$.

A.1 A basic primitive

Our constructions of verifiers rely on an algebraic procedure that allows them to reconstruct "many" values of a polynomial using $O(1)$ queries. In describing this procedure we closely follow the exposition in [A94], chapter 3. The only difference is a tremendous performance gain due to our new analysis of the low degree test.

As already mentioned, our verifiers rely on the fact that a satisfying assignment can be encoded as a degree d polynomial, for some appropriate d. The verifier expects the proof of satisfiability to contain such a polynomial, represented *by value*. This means that the proof contains some oracle $f : \mathrm{F}^m \to \mathrm{F}$ (the encoding is such that $|\mathrm{F}|^m$, the size of this oracle, is polynomial in the size of the assignment being encoded). Using the low degree test the the verifier checks that f has reasonable agreement with a degree d polynomial. Next, to check satisfiability, the verifier picks in some way (note: we're omitting many details here) k points $z_1, z_2, \ldots, z_k \in \mathrm{F}^m$ and then has to reconstruct the values of P at those points, where P is any polynomial that has significant agreement with f. Now we describe a procedure from [ALMSS92] (who essentially borrowed it from [LS91]) that allows the verifier to do this reconstruction, provided the proof contains additional information. The most important property of this procedure is that the verifier reads only 3 entries from the oracles provided to it, even though k might be pretty large (and not a constant).

Reconstruction Procedure:

Given: A function $f : \mathrm{F}^m \to \mathrm{F}$, an integer d, a fraction $\delta > 0$, and k points $z_1, \ldots, z_k \in \mathrm{F}^m$.

Desired: One of the k-tuples $(P_1(z_1), \ldots, P_1(z_k))$, $(P_2(z_1), \ldots, P_2(z_k)), \ldots$ or $(P_r(z_1), \ldots, P_r(z_k))$, where $P_1, \ldots, P_r$ are all the the polynomials that have agreement at least δ with f.

Auxiliary Information: Two tables T and T_1 in which each entry has $\mathrm{poly}(dk \log q$ bits.

Procedure's Properties: The procedure outputs either REJECT or a k-tuple.

- Let $q > (c_0 d)^{4c_3}$ and $\epsilon = q^{-1/4c_1c_3}$. Then
$$\Pr[\text{procedure doesn't output one of the desired tuples}] \leq \frac{kd}{\sqrt{q}} + \epsilon.$$
- If f is a degree d polynomial, then there exist tables T, T_1 such that
$$\Pr[\text{procedure outputs } f(z_1), \ldots, f(z_k)] = 1.$$

Procedure's Complexity: Uses $O(m \log q)$ random bits. Reads 1 entry from each of T and T_1 and one value of f.

Now we describe the procedure. Recall that a degree-k curve is a set of points with a parametric representation like $\{(c_1(t), \ldots, c_m(t)) : t \in \mathrm{F}\}$, where each c_i is a degree k univariate polynomial. Note that the restriction of a degree d polynomial to this curve is a univariate polynomial of degree kd.

Below, we talk about a *random* degree k curve that passes through $z_1, \ldots, z_k$. We can pick such a curve by choosing a random point $y \in \mathrm{F}_m$ and identifying (using interpolation) m degree k univariate polynomials $c_1(t), \ldots, c_m(t)$ such that

$$\forall 1 \leq i \leq k \qquad (c_1(i), c_2(i), \ldots, c_m(i)) = z_i \tag{31}$$
$$(c_1(k+1), c_2(k+1), \ldots, c_m(k+1)) = y \tag{32}$$

Here we are using the integers $1, 2, \ldots, |\mathrm{F}|$ to also denote field elements. Note that by choosing the $k+1$th point of the curve randomly from F^m, we have ensured that the the last $|\mathrm{F}| - k$ points on the curve are randomly (though not independently) distributed in F^m. This will be important.

Now we describe the procedure. Note that part of the procedure just consists in doing the low degree test at a point $C(a)$ on the curve C.

Description of Reconstruction Procedure:

INPUTS: Function $f : \mathrm{F}^m \to \mathrm{F}$, two oracles T_1, T_2, and k points $z_1, \ldots, z_k \in \mathrm{F}^m$.
T_1 contains a sequence of univariate degree d polynomials, one for each line in F^m. Oracle T_2 contains, for each degree-$(k+1)$ curve in F^m that passes through $z_1, \ldots, z_k$, a univariate degree $(k+1)d$ polynomial.

PROCEDURE:

1. `Randomly pick a degree` k `curve` $C(t)$ `in` F^m `whose first` k `points are` $z_1, \ldots, z_k$`. Pick a random` $a \in \mathrm{F}$`, and compute the point` $C(a) \in \mathrm{F}^m$`. Pick a random line` l `that passes through` $C(a)$`.`

2. `Read the value of` f `at` $C(a)$`. Read the polynomial given for curve` C `in oracle` T_2`; say it is` $g_C(t)$`. Read the polynomial given for line` l `in oracle` T_1`; say it is` $h_l(t)$`.`

3. `If` $g_C(t)$ `and` $h_l(t)$ `produce the value` $f(C(a))$ `at point` $C(a)$`, then output` $(g_C(1), g_C(2), \ldots, g_C(k))$`, the values of` g_C `at` $1, 2, \ldots, k \in$ `F. Otherwise output REJECT.`

Complexity: The procedure runs in time $\mathrm{poly}(m + d + \log|\mathrm{F}| + k)$. Randomness is required only to generate $O(1)$ elements of F^m, so only $O(m \log|F|)$ random bits are needed. Whenever we use this procedure, the function f is supposed to represent an assignment to n variables. The field size, the degree and the number of variables have been carefully chosen so that $|\mathrm{F}|^m = \mathrm{poly}(n)$. Thus the procedure require $O(m \log|\mathrm{F}|) = O(\log n)$ random bits. Also, $d > m$, so the running time and the size of the oracle entries are $\mathrm{poly}(d + k)$.

Now we prove the correctness of the procedure. We are only interested in two cases. In the first case, the oracle-constructor is trying to help the verifier. Then it is clear that by just taking f to be a degree d polynomial and constructing oracles T_1, T_2 appropriately, it can make the verifier accept with probability 1. Now suppose the oracle constructor is malicious. Let c_1, c_2, c_3 be constants of the same name that appeared in Theorem 3. Let $\epsilon = q^{-1/4c_1c_3}$, Let $P_1, \ldots P_r$ be all degree d polynomials that have agreement at least ϵ^{c_3}/c_2 with f. We say that the procedure *makes a mistake* if it outputs a k-tuple that isn't one of $(P_1(z_1), \ldots, P_1(z_k))$, $(P_2(z_1), \ldots, P_2(z_k)), \ldots$ or $(P_r(z_1), \ldots, P_r(z_k))$.

Lemma 29 *Suppose ϵ is as described in the previous paragraph and $q > (c_0 d)^{4c_3}$. Then*

$$\Pr[\textit{procedure makes a mistake}] \leq \frac{kd}{\sqrt{q}} + \epsilon.$$

Proof:(sketch) Let us try to identify characteristics of any curve C, point $C(a)$ and line l that causes the procedure to make a mistake. It must be that (i) For each polynomial P_i there is some point among $z_1, \ldots, z_k$ at which P_i and g_C disagree (since otherwise the procedure would output $(P_i(z_1), \ldots, P_i(z_k))$, and thus not make a mistake). In other words, the univariate polynomial g_C differs from each of the restrictions $P_1|_C, \ldots P_r|_C$. (ii) f passes the low degree test using line l (iii) The curve polynomial g_C produces the value $f(C(a))$ at $C(a)$ (since otherwise the procedure would output REJECT).

We upperbound the probability of making a mistake as follows. Suppose curve C satisfies condition (i). Since two univariate degree kd polynomials can agree at at most kd points, we conclude that on such a curve, $1 - dkr/q$ fraction of $a \in \mathrm{F}$ are such that g_C does not agree with any of $P_1|_C, \ldots P_r|_C$ at $C(a)$. Thus conditions (ii) and (iii) become difficult to satisfy: if $g_C(a) = f(C(a))$ for "many" a — as required by condition (iii) — then on most such points f must disagree with all of $P_1, \ldots, P_r$, in which case the low degree test is very unlikely to succeed.

Now we formalize this. Let $S \subseteq \mathrm{F}^m$ be the set of points where f doesn't agree with any of $P_1, P_2, \ldots, P_r$. With each point $x \in \mathrm{F}^m$ let us associate a number ρ_x as follows: if $x \notin S$ then $\rho_x = 0$ and otherwise ρ_x is the success probability of the low degree test at x. By Theorem 3,

$$E_{x \in \mathrm{F}^m}[\rho_x] \leq \epsilon.$$

Now if C is a curve, we denote by $Y_C \in [0,1]$ the average of ρ_x among all points $x \in C$. When the test picks a random curve C, then the last $|\mathrm{F}| - k$ points of the curve are randomly distributed in F^m. Hence by linearity of expectations $E_C[Y_C] \leq \epsilon$ (we are assuming $k \ll q$, so the first k points don't affect the expectation by much).

On any curve C that satisfies condition (i),

$$\Pr_{a,l}[\text{made a mistake on } C(a) \text{ using line } l] \leq \frac{dkr}{q} + E_{a \in \mathrm{F}}[\rho_{C(a)}] = \frac{dkr}{q} + Y_C.$$

Hence

$$\Pr_{C,a,l}[\text{made a mistake on } C(a) \text{ using line } l] \leq \frac{dkr}{q} + E_C[Y_C] \leq \frac{dkr}{q} + \epsilon.$$

Now the lemma follows by noticing that r, the number of polynomials with agreement at least ϵ^{c_3}/c_2 with f, is at most $4\epsilon^{c_3}/c_2$ by Proposition 2. $\Box$

A.2 A 3-Prover Proof System

The reconstruction procedure described above can be used in conjunction with any efficient PCP system to obtain efficient constant prover proof systems. For instance, we could start with an amplified version of the proof system of [ALMSS92]: In this proof system the verifier, V_1, queries a proof of length n in k places and accepts valid proofs, while accepting proofs of incorrect theorems with probability at most $\gamma = \exp(-k)$. Furthermore the randomness complexity of this verifier is $O(\log n) + O(k)$. To turn this verifier into a verifier for a constant prover proof system (in particular a 3-prover proof system), we extend the proof π used by verifier V_1 into a low-degree polynomial and then use the reconstruction procedure described above to reconstruct the responses to all k queries making only 3 queries to three tables. Details follow:

Let π be a table of n bits. Let F be a field of order q and let $H \subset F$. Then if m is such that $|H|^m \geq n$, then we can view π as a function from $H^m \to \{0,1\}$. Further we can extend π to obtain a polynomial $\hat{\pi} : F^m \to F$ so that $\hat{\pi}$ restricted to H^m is π and the degree d of $\hat{\pi}$ is at most $m|H|$. The new verifier works with 3 tables, f (which is supposed to be $\hat{\pi}$), T and T_1, where the latter two are supposed to be the auxiliary tables of the reconstruction procedure. To simulate the action of the verifier on queries $z_1, \ldots, z_k \in H^m$, we use the reconstruction procedure on three tables f, T and T_1 to reconstruct a tuple $(a_1, \ldots, a_k)$ such that, with high probability, this is the output of $(P_i(z_1), \ldots, P_i(z_k))$ for some $i \in \{1, \ldots, r\}$ where $P_1, \ldots, P_r$ are all the polynomials which agree with f in δ fraction of the places. The new verifier then accepts the proof only if the verifier of the reconstruction procedure does not reject and only if the k tuple returned by the reconstruction procedure is accepted by V_1 as responses to queries $z_1, \ldots, z_k$.

To analyze the complexity of this verifier, first observe that the number of provers used is 3. Also the length of the provers responses (entry size in the tables) is at most $md \log q$. The randomness complexity is the sum of the randomness complexity of the two verifiers which is at most $O(\log n) + O(k) + O(m \log q)$. Finally the error of the verifier is at most the sum of the errors of the two components and is thus at most $\frac{kd}{\sqrt{q}} + \epsilon + \gamma$. By picking F and H appropriately, we can now obtain efficient 3-prover proof systems. In particular, given any $\beta < 1$, by setting $q = 2^{\log^\beta n}$ and $|H| = q^\alpha$ for some sufficiently small but positive α, we obtain the following lemma.

Lemma 30 *For every $\beta < 1$, $NP \in MIP[3, O(\log n), 2^{O(\log^\beta n)}, 2^{-\log^\beta n}]$.*

The verifier described above can be used as a starting point for applying the recursive proof checking technique of Arora and Safra [AS92]. In particular we can use the following lemma from Bellare et al. [BGLR93] proved using recursion.

Lemma 31 ([BGLR93]) *$MIP[p, r, c, a, 2^{-k}] \subset MIP[p + 2, r + (\log a + k)^3, (\log a + k)^3, 2^{-(k(n)/p)+3}]$.*

Combining Lemma 31 with Lemma 30 we get the following theorem.

Theorem 32 *$NP \subset MIP[5, O(\log n), O(\log n), 2^{-\log^{1/3} n}]$.*

In order to reduce the error probability above to $2^{-\log^\beta n}$ for any $\beta < 1$, we need to replace Lemma 31 with our own protocol to be used for recursive application. The protocol involves simple modifications of the protocol used to obtain Lemma 30. We list the differences here:

- First the lemma is used to prove that a verifier for a p-prover proof system would accept a given set of answers. Hence the recursive verifier works in the encoded theorems model of [BFLS91] and uses further the concatenated input model of [AS92]. Specifically the protocol is designed to prove that $x_1, \ldots, x_p$ would form a satisfying input to a circuit C, given an encoding of $x_1, \ldots, x_p$ in the form of low-degree polynomials.

- This procedure uses the reconstruction procedure to verify that an efficient PCP system would have accepted the encodings of $x_1, \ldots, x_p$. In order to do so, it generates k points used to check that the concatenation of $x_1, \ldots, x_p$ would be accepted by C, and in addition generates p points, one from each x_i which is used to verify that the claimed concatenation of the x_i's is really consistent with the individual encodings of x_i's.

- Last point of difference is that we don't use the protocol of [ALMSS92] to generate the test for the reconstruction procedure. Instead we go back to the protocol of Babai et al. [BFLS91] and use this protocol and use a fresh analysis of the protocol keeping the improved low-degree test in mind. This allows us to push the parameter β in the exponent of the error term arbitrarily close to 1.

We omit the details of the protocol and simply state the lemma that can be obtained from the above.

Lemma 33 *There exist $\epsilon > 0$ and $\alpha < \infty$ such that for every r, p, a, e the following holds: $MIP[p, r, a, e] \subset MIP[p+3, r + O(m \log |F|), O((poly\log a) d \log |F|), e^{\frac{1}{2p+2}}]$. where d, m are any positive integers and F is a any finite field satisfying the following conditions:*

- $e^{\frac{1}{2p+2}} \geq d^{\alpha}/|F|^{\epsilon}$.
- $(d/m)^m \geq a^{O(1)}$.

Using Lemma 33 we now get the following theorem.

Theorem 34 *For every $\beta < 1$, there exists a $p < \infty$ such that*

$$NP \subset MIP[p, O(\log n), O(\log n), 2^{-\Omega(\log^{\beta} n)}].$$

Proof: We start with a 3-prover proof system with $\beta = 1/2$ as given by Lemma 30 We then recurse $(\frac{1}{1-\beta})$-times using Lemma 33 with $|F| = 2^{\log^{\beta} n}$, $m = \log^{1-\beta} n$ for all applications. The choice of d is set to satisfy the condition $(d/m)^m \geq a^{O(1)}$ and we pick $d = m 2^{\log^{1-i(1-\beta)} n}$ in the ith application. This yields NP $\subset$ MIP$[3(\frac{1}{1-\beta} + 1), O(\log n), O(\log^{\beta} n \log\log n), 2^{-\Omega(\log^{\beta} n)}]$. $\square$

B Self-correction of programs

Consider a program $\mathcal{P}$ that is supposed to be computing an unknown polynomial g. Suppose $\mathcal{P}$ is correct on only some tiny δ fraction of the inputs. Our testing procedure allows us to estimate the largest δ for which the program's output agrees with the output of some polynomial, to within an additive error of $O(1/q^{\epsilon})$ over a field of size q.

The task of self-correcting this program needs to be defined carefully. For starters, there can be more than one polynomial agreeing with the program $\mathcal{P}$ in δ fraction of the inputs. In fact, we can have $O(\frac{1}{\delta})$ such polynomials. However we can be expected to reconstruct $O(\frac{1}{\delta})$ (randomized) "programs", each of which computes a polynomial (and is correct on every input with high probability), such that every polynomial that has $\frac{1}{\delta}$ agreement with $\mathcal{P}$ is computed by one of the programs. This task was left as an open problem in Ar et al. [ALRS92], and no polynomial (in m, d and $\frac{1}{\delta}$) time algorithm is known for this problem. Goldreich et al. [GRS] solve this problem when $\delta \geq 2\sqrt{d/q}$ in time exponential in d. We now describe our solution that works when $\delta \geq (md/q)^{\epsilon}$, for some positive ϵ, and is the first polynomial time-bounded solution for any $\delta < 1/2$.

Given a program $\mathcal{P}$, our algorithm works in two phases: First, a preprocessing phase, where we instantiate $k \leq O(\frac{1}{\delta})$ programs $\mathcal{P}_1, \ldots, \mathcal{P}_k$. In the second phase a program $\mathcal{P}_i$ takes an input $x \in F^m$ and computes its output $\mathcal{P}_i(x)$. The guarantee is that at the end of the first phase, with high probability, we create k randomized programs, such that the output of each is (with high probability) a polynomial; furthermore, for every polynomial g which agrees with $\mathcal{P}$ on δ fraction of the input, one of the programs $\mathcal{P}_i$ computes g correctly with high probability on every input.

The two phases are based on the analysis of the bootstrapping method described in Section 3.3 which is in turn based on the work of Arora [A94]. In the preprocessing stage, we pick a random line l from the space F^m and find all degree d univariate polynomials describing $\mathcal{P}$ restricted to the line on $\delta/2$ fraction of the places. We claim (without proof) that no polynomial which describes $\mathcal{P}$ on δ fraction of the places will be unrepresented. Further, no two such polynomials will turn out to be identical on this line. Lastly we claim (based on the strong version of the low-degree test) that no "spurious" polynomials will be discovered by this procedure. Let $p_1, \ldots, p_k$ be the polynomials found this way. We create k programs with $\mathcal{P}_i$ containing p_i and l as its identifying polynomial.

In the second phase, given an input $x \in F^m$, the program $\mathcal{P}_i$ picks a random line l_1 passing through x and computes all trivariate degree d polynomials $h_1, \ldots, h_{k'}$ agreeing with $\mathcal{P}$ in $\delta/4$ fraction of the points on the cube containing l and l_1. Again, with high probability, we assert that no two of the polynomials h_{j_1}, h_{j_2} turn out to be identical on l. We then pick the unique polynomial h_j such that h_j restricted to l is p_i. We then return the value of h_j evaluated at x as the output of $\mathcal{P}_i$ on input x. It can be argued that if all the assertions (claimed to hold with high probability earlier) hold, then the output of $\mathcal{P}_i$ is always according to some fixed polynomial g_i which agrees with $\mathcal{P}$ on at least $\delta/2$ fraction of the input. Furthermore, g_i is the unique polynomial (among all such) such that g_i restricted to l is p_i. Thus we get the following theorem:

Theorem 35 *There exists a randomized polynomial time algorithm that, when given oracle access to a program $\mathcal{P}$ and parameters d, δ, F and m, can create $O(\frac{1}{\delta})$ randomized programs $\mathcal{P}_1, \ldots, \mathcal{P}_k$ such that for every degree d polynomial $g : F^m \to F$ which has δ agreement with $\mathcal{P}$, there exists a program $\mathcal{P}_i$ that computes g correctly on every input with high probability.*