Improved low-degree testing and its applications

Sanjeev Arora*
Princeton University

Madhu Sudan†
IBM T. J. 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 [29]. 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 used in the factoring of multivariate polynomials.

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 [28]. 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 $\delta \ll 0.5$, then the tester/corrector determines δ and generates $O(\frac{1}{\delta})$ randomized programs, such that one of the programs is correct on every input, with high probability.

*Supported by an NSF CAREER award and an Alfred P. Sloan Fellowship.arora@cs.princeton.edu.

†P.O. Box 218, Yorktown Heights, NY 10598, U.S.A. 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 [7], and NP = PCP$(\log n, 1)$ [6, 5] 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 [24, 31] 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 [24]. 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$(\log n, 1)$. 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 [13, 6], CHROMATIC NUMBER and SET COVER [25], AND MAX-3SAT [5]. 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 [4] — it is only known to imply *quasi-NP-hardness* (a *quasi-NP-hard* problem is one that has a polynomial-time algorithm only if NP $\subseteq$ Time$(n^{\mathrm{poly}\log(n)})$).

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 [25], adapted for more than 2 provers in [10]). Raz and Safra [28] 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 us-

ing $O(\log n)$ 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 probably be reduced further using a technique of Tardos [35], but that still does not lead to a 2-prover proof system. Getting a 2-prover proof system with $O(\log n)$ randomness and answer size but subconstant error probability remains an open question.

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 [29] that works roughly as follows: Pick a random "line" in $\mathrm{F}^m$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 [7] and later ones in [8, 13, 18]. (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 [13, 18, 29, 6] 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 [6, 29]) of the above test that worked for $\delta > 1-\epsilon$ for some fixed $\epsilon > 0$. 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 $1/8$ [17].)

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 $\delta > (md)^c / |\mathrm{F}|^\epsilon$ 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 $|\mathrm{F}| > 2^{O(m+d+1/\delta)}$ [2, 33]. (However, 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 only when f is not a polynomial but 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 therefore 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. A formal proof of this "folklore result" is included in the full paper, and some pointers appear in Section 4.

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 [11] deals with 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, a version slightly different from the one described in the next section in that it doesn't use a d-oracle) closes this open problem when $|\mathrm{F}|^\epsilon > \mathrm{poly}(md)$. As for correction, note that its meaning is unclear 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 [1]. 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 [32]. 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. Details of the proof are omitted from this abstract, but they are obvious from reading our proofs (specifically, by noting their "algorithmic" nature).

Past work on constant-prover proof systems. The first construction of a nontrivial constant prover 1-round proof system for NP appeared in [23]; others appeared in [16, 10, 34, 14, 27]. 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 [27]). It was also known [15] that some obvious ideas (such as "recycling randomness") cannot let us get around this. Earlier this year Raz and Safra [28] found a construction of a proof system achiev-

ing 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.

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 [29, 5, 3, 17], 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 [32] work on reconstructing polynomials from very noisy data and Kaltofen's work on "Effective Hilbert Irreducibility" [20, 21, 22]. Then in Section 3.2 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 [3].

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 $a \in$ 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 |F|, the cardinality of F.

The *distance* between two functions $f, g : \mathrm{F}^m \to$ 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 $f : \mathrm{F}^m \to$ 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 $l = \{(u_1 + tv_1, \ldots, u_m + tv_m) : t \in \mathrm{F}\}$ be a line, $f : \mathrm{F}^m \to$ F be a function and $g(t)$ 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 $f : \mathrm{F}^m \to$ 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 [17]), then f must be a degree-d polynomial.

The low degree test is presented with $f : \mathrm{F}^m \to$ 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 B. 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 $f : \mathrm{F}^m \to \mathrm{F}$ 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 is true for some $0.5 < \delta < 1$. 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 state the stronger forms here.

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.*

□

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. □

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 simple averaging argument 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,

$$\Pr_l \left[\begin{array}{l} \text{some univ. deg. } d \text{ poly. } g_l \text{ describes } f \\ \text{on } \rho - \epsilon/2 \text{ fraction of points of } l \end{array} \right] \geq \frac{\epsilon}{2} \quad (1)$$

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 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. The following bound on γ_i follows from a simple application of Chebychev's inequality (proof omitted from this abstract):

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

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

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

A simple inclusion-exclusion based counting argument shows that $k \leq 2c_2/(\epsilon_1/4)^{c_3}$. Hence

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

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] $< \epsilon/4$. □

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" and a "reasonably large" set of points $S \subseteq \mathrm{F}^2$ such that for every $(a, b) \in S$, the following are true:

$$Q(f(a, b), a, b) = 0 \qquad (3)$$

$$d\text{-success rate of } f \text{ at } (a, b) \text{ is "non-negligible."} \qquad (4)$$

(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" $(a, b) \in S$,

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

By the end of Step 2, we have concluded that f has significant agreement with the degree d bivariate polynomial g. Step 2 uses Theorem 6 which is a version of a family of results known as *Hilbert irreducibility* theorems. They study the preservation of the irreducibility of a multivariate polynomial, when values of most variables are substituted with constants or linear forms in one or two new variables. We will need a version which leaves one variable unsubstituted and all other variables get substituted with a linear form in one new variable. This specific substitution has been studied by Kaltofen [21], who bounds the probability with which the polynomial may factor after the substitution, if the substitution is performed "randomly". The bound presented in [21] is too weak for our purposes. Fortunately, in a later work Kaltofen [22] presents improved bounds. The bounds in [22] are presented for a different substitution, but the analysis easily extends to the substitution studied in [21]. We summarize this theorem below.

Theorem 6 ([20]) *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 the polynomial $Q(z,\ a_1t+b_1,\ \ldots, a_mt+b_m)$ in z and t has a factor of the form $z - p(t)$ is at most $1 - O(l^3/q)$.*

Step 1 is motivated by Sudan's [32] technique for reconstructing univariate polynomials from very noisy data. Sudan makes the following observation.

Proposition 7 *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 $deg_z(\Gamma) \le d_z$ and $deg_x(\Gamma) \le d_x$, satisfying*

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

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.") For $k = 1, \ldots, n$,

$$\sum_{i=0}^{d_z} \sum_{j=0}^{d_x} \gamma_{ij} y_k^j a_1^i \quad = \quad 0$$

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

Then Sudan uses the following lemma from Ar et al. [1].

Lemma 8 *Let $(a_1, y_1), \ldots, (a_n, y_n) \in F^2$ be any sequence such that for some $\rho > 0$,*

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

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

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. □

We need the following generalization of the ideas described above.

Lemma 9 *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 $b \in S_2$, there exist at least ρl values of $a \in S_1$ s.t.*

$$f(a, b) = C_a(b) = R_b(a).$$

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

$$\forall a \in S_1, \quad Q(C_a(y), a, y) = 0 \text{ and} \tag{8}$$

$$\forall b \in S_2, \quad (z - R_b(x)) \,|\, Q(z, x, b) \tag{9}$$

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 [32], 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

$$Q(C_a(y), a) = 0 \qquad \forall a \in S_1 \tag{10}$$

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 (10) 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.

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

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 (9) follows immediately from Lemma 8 and the condition $\rho > 2d/\sqrt{l} \geq (d+1)/\sqrt{l}$. □

The following lemma finishes Step 1 of our proof.

Lemma 10 *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 9 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: Omitted; involves picking two random directions h_1, h_2 and computing the expectation. □

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 $\forall b \in S_2$

$$\exists \gamma t/2 \text{ values of } a \in S_1 \text{ s.t. } C_a(b) = R_b(a) = f(a, b). \tag{11}$$

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,

$$\begin{aligned} \Pr_{S_1}[|H \cap (S_1 \times \{b\})| < \gamma t/2] &\leq exp(-\frac{\gamma t}{2}) \\ &= exp(-2\log q) \leq \frac{1}{2q}. \end{aligned}$$

Thus the probability is at least $1 - |S_2|/2q - 1/100 \geq .49$ that the randomly chosen set S_1 satisfies condition (11).

Thus we have proven the existence of $S_1, S_2 \subseteq \mathrm{F}$ such that they satisfy the hypothesis of Lemma 9 with $\rho = \gamma/2$ and $l = t$. (Notice that by the definition of 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 9. 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. □

Now we move to Step 2 of our argument.

Lemma 11 *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 8 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 6 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. □

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 6 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 $\frac{\gamma^2}{4l} > D/|\mathrm{F}|$. 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 $\frac{\gamma^2}{4l} > D/|\mathrm{F}|$. □

□

Thus we have proved the bivariate case of Theorem 1.

Theorem 12 *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 10 and 11. □

3.2 The Bootstrapping

Section 3.1 showed the correctness of Theorem 1 for the case of $m = 2$. An easy generalization of the proof (whose details we omit here) carries over for larger m, except the "constants" c_1, c_2 and c_3 then depend on m. In order to avoid any dependence on m, some other idea is needed. We describe this now.

This section assumes the truth of Theorem 1 (as well as Theorems 3 and 4) for $m = 2, 3$, and proves Theorem 1 for general m. The proof relies on symmetry-based arguments similar to those in [3]. 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 $\overline{u_1}, \overline{u_2}, \ldots, \overline{u_k} \in \mathrm{F}^m$. □

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 $\mathrm{plane}(l, x)$ 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 s. 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.

The following two lemmas are both conseqeunces of symmetry-based arguments that we will need. Both can be shown using straightforward application of Chebychev's inequality. We omit the proofs here.

Lemma 13 (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.*

Lemma 14 (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,*

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

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*

$$E_l[d\textit{-success-rate of } \hat{f}_l \textit{ in } F^m] \ge 1 - \frac{\delta^2}{256} \quad (12)$$

$$E_l[\textit{agreement between } f \textit{ and } \hat{f}_l \textit{ in } F^m] \ge \frac{\delta}{4}. \quad (13)$$

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 (12) 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 (13) 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 [5] 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 $\delta/8 - \delta/12 = \delta/24$. □

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

$$E_{(l,l')}[d\text{-success-rate of } \hat{f}_l \text{ on } l'] \geq 1 - \tfrac{\delta^2}{256} \quad (14)$$

$$E_{(l,l')}[\text{agreement of } \hat{f}_l \text{ and } f \text{ on } l'] \geq \tfrac{\delta}{4}. \quad (15)$$

Let $\alpha = 1/32$. The main idea why we can "bootstrap" (i.e., reduce the m-variate case to the trivariate case) is that the two expectations in statements (14) and (15) 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 why this doesn't change the expectation 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 14). 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 (14) and (15) 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 $\delta(1 - 6\alpha)$. Proposition 2 shows that the set of polynomials we have identified thus far is not too big: $k_0 \leq 8/\delta$ and $k \leq 4c_2/\epsilon^{c_3}$. 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 $i \neq j$ 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 $k \leq 4c_2/\epsilon^{c_3}$, we have

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

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 13, the fraction of such lines is at most $\frac{1}{4\alpha^2\delta q}$.

2. $P_1|_l$ has agreement $\beta \geq \delta(1 - 4\alpha)$ with f on line l but one of $P_{k_0+1}|_l, \ldots, P_k|_l$ has agreement more than β. By Lemma 13, 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 $k \leq 4c_2/\epsilon^{c_3} < 4c_2 q^{1/4}$.

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 $4/\delta q^{1/4c_3}$. □

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) Every bivariate polynomial besides $P_1|_s, \ldots, P_k|_s$ has have agreement less 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

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

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 $4/\delta q^{1/4c_3}$. □

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. □

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,*

$$\textit{d-success-rate of } \hat{f}_l \textit{ on } l' \geq 1-\sqrt{\gamma} \qquad (16)$$

and for a random line l' noncoplanar with l,

$$E_{l'}\left[\begin{array}{l} \textit{agreement between } \hat{f}_l \\ \textit{and } f \textit{ on cube } c \end{array} \right] \geq \delta(1-\sqrt{\gamma})(1-8\alpha). \quad (17)$$

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 (17) follows.

Now we prove the claim in (16). 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'. □

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 $(1-\sqrt{\gamma})(1-\gamma-\sqrt{\gamma}) > 1-\delta^2/256$. □

4 Construction of constant prover protocols

The construction is in two steps. *Step 1:* Construct a 3 prover protocol in which the number of random bits is $O(\log n)$ and the provers' answer size is $2^{\log^\beta n}$ for some $\beta < 1$. *Step 2:* Use "verifier composition," a technique from [6], to compose the verifier in Step 1 with itself. Doing this enough times reduces the answer size to $O(\log n)$, while keeping the number of provers at $O(1)$.

Both steps rely on a procedure of [5], which uses the low degree test to reconstruct "many" values of a polynomial using $O(1)$ provers (this procedure is similar in spirit to many others that preceded it in literature). This procedure is described in Section 4.4 of [3]. The analysis given there relies on the old result about the low degree test, and therefore only shows that the procedure fails with probability less than 1/2. Using our Theorem 3 in the analysis shows that the failure probability of the procedure is below $1/\sqrt{q}$ or so.

As a consequence of this procedure, Step 1 is easy: Just repeat the ALMSS protocol $O((\log n)^\beta)$ times using standard pseudorandomness techniques. Instead of making queries to $O((\log n)^\beta)$ independent provers, use the reconstruction procedure to "Aggregate Queries" (see Section 4.1.2 in [3]) and thus end up with 3 provers. Step 2 is also standard and follows the general idea of [6] of making the provers encode their answers using low degree polynomials.

We note that the simple ideas above yield a proof system with error probability $2^{-\log^{0.5-\epsilon} n}$. Reducing error to $2^{-\log^{1-\epsilon} n}$ requires other ideas.

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 [6], 3 provers appears to be the best possible. Reducing the number of provers to 2 would imply the NP-hardness of approximation problems studied in [4].

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

[1] S. AR, R. LIPTON, R. RUBINFELD AND M. SUDAN. Reconstructing algebraic functions from noisy data. IEEE FOCS 1992.

[2] S. ARORA *Unpublished, 1993*.

[3] 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` .

[4] S. ARORA, L. BABAI, J. STERN AND Z. SWEEDYK. The hardness of approximating problems defined by linear constraints. IEEE FOCS 1993.

[5] S. ARORA, C. LUND, R. MOTWANI, M. SUDAN AND M. SZEGEDY. Proof verification and the hardness of approximation problems. IEEE FOCS 1992.

[6] S. ARORA AND S. SAFRA. Probabilistic checking of proofs: a new characterization of NP. IEEE FOCS 1992.

[7] L. BABAI, L. FORTNOW, AND C. LUND. Non-deterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1:3–40, 1991.

[8] L. BABAI, L. FORTNOW, L. LEVIN, AND M. SZEGEDY. Checking computations in polylogarithmic time. ACM STOC 1991.

[9] M. BELLARE, O. GOLDREICH AND M. SUDAN. Free bits, PCPs and non-approximability — towards tight results. IEEE FOCS 1995 TR95-024 of ECCC, the *Electronic Colloquium on Computational Complexity*, `http://www.eccc.uni-trier.de/eccc/`.

[10] M. BELLARE, S. GOLDWASSER, C. LUND, AND A. RUSSELL. Efficient probabilistically checkable proofs. ACM STOC 1993. (See also Errata sheet in *Proceedings of the* 26th ACM STOC, ACM, 1994).

[11] M. BLUM, M. LUBY, AND R. RUBINFELD. Self-testing/correcting with applications to numerical problems. In ACM STOC 1990.

[12] U. FEIGE. A threshold of $\ln n$ for Set Cover. ACM STOC 1996.

[13] . 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.

[14] U. FEIGE AND J. KILIAN. Two prover protocols – Low error at affordable rates. ACM STOC 1994.

[15] U. FEIGE AND J. KILIAN. Impossibility results for recycling random bits in two-prover proof systems. ACM STOC 1995.

[16] U. FEIGE AND L. LÓVASZ. Two-prover one-round proof systems: Their power and their problems. ACM STOC 1992.

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

[18] P. GEMMELL, R. LIPTON, R. RUBINFELD, M. SUDAN AND A. WIGDERSON. Self-testing/correcting for polynomials and for approximate functions. ACM STOC 1991.

[19] O. GOLDREICH, R. RUBINFELD AND M. SUDAN. Learning polynomials with queries: The highly noisy case. IEEE FOCS 1995

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

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

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

[23] D. LAPIDOT AND A. SHAMIR. Fully Parallelized Multi-prover protocols for NEXP-time. IEEE FOCS 1991.

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

[25] C. LUND AND M. YANNAKAKIS. On the hardness of approximating minimization problems. ACM STOC 1993.

[26] A. POLISHCHUK AND D. SPIELMAN. Nearly Linear Sized Holographic Proofs. ACM STOC 1994.

[27] R. RAZ. A parallel repetition theorem. ACM STOC 1995.

[28] R. RAZ AND S. SAFRA. A sub-constant error-probability low-degree test and a sub-constant error-probability PCP characterization of NP. To appear ACM STOC 1997.

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

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

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

[32] M. SUDAN. Maximum likelihood decoding of Reed Solomon codes. IEEE FOCS 1996.

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

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

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