A Sub-Constant Error-Probability Low-Degree-Test and a Sub-Constant Error-Probability PCP Characterization of NP *

Ran Raz † Shmuel Safra ‡

Abstract

We introduce a new low-degree–test, a one that uses the restriction of low-degree polynomials to planes (i.e., affine sub-spaces of dimension 2), rather than the restriction to lines (i.e., affine sub-spaces of dimension 1). We prove the new test to be of a very small error-probability (in particular, much smaller than a constant).

The new test enables us to prove a low-error characterization of NP in terms of PCP. Specifically, our theorem states that, for any given $\epsilon > 0$, membership in any NP language can be verified with $O(1)$ accesses, each reading logarithmic number of bits, and such that the error-probability is $2^{-\log^{1-\epsilon} n}$. Our results are in fact stronger, as stated below.

One application of the new characterization of NP is that approximating SET-COVER to within logarithmic factors is NP-hard.

Previous analysis for low-degree–tests, as well as previous characterizations of NP in terms of PCP, have managed to achieve, with constant number of accesses, error-probability of, at best, a constant. The proof for the small error-probability of our new low-degree–test is, nevertheless, significantly simpler than previous proofs. In particular, it is combinatorial and geometrical in nature, rather than algebraic.

*To be found at URL http://www.math.tau.ac.il/school/courses/PCP

†Weizmann Inst., ISRAEL. ranraz@wisdom.weizmann.ac.il

‡Tel-Aviv University, ISRAEL.

1 Characterizations of NP in terms of PCP

Characterizing the class NP, by itself or with respect to other computational-complexity classes, is perhaps one of the most fundamental avenues of research in theory of computer-science.

Since the early days, when the classes P and NP were defined, and the question was posed as to whether they are the same or do they differ, many problems were shown to be NP-complete, thereby increasing the weight on finding stricter characterization for the class NP.

NP has since been given a few alternative characterizations. The one most commonly applied being Cook's [Coo71], which characterizes NP in terms of efficient verification of proofs (or nondeterministic computations).

A new perspective, by which improved characterizations of NP can be obtained, has been recently proposed. The motivation for which stems from questions regarding the hardness of approximation versions, for problems whose exact computation is known to be NP-hard.

This avenue of research was initiated by [FGL+91], which introduced a new methodology for proving hardness results for approximation problems. The method takes advantage of results in a seemingly unrelated area — that of interactive proofs [GMR89, Bab85, BGKW88, LFKN92, Sha92, BFL91] — however interprets those results with quite a different perspective in mind.

Much effort has been invested since towards a better understanding of this methodology, and the class NP has consequently gained stricter characterizations [AS92, ALM+92, BGS95], which are referred to as *characterizations of NP in terms of PCP* (or, in short, PCP characterizations of NP).

The PCP characterization of NP — though has taken around 20 years to be formulated — seems now as the most natural extension of the old characterization of NP [Coo71], if one has in mind proving hardness results for approximation problems. This characterization has already been used to obtain quite a few hardness results for approximation problems [FGL+91, AS92, ALM+92, PY91, LY94, BGLR93, KLS93, BGS95, Hås96a, Hås96b, Hås97].

The previous characterization of NP in terms of the

PCP hierarchy [AS92, ALM$^+$92], seemed at first as the best possible up to constant factors. A stronger characterization was later conjectured in [BGLR93]; one that, as an immediate outcome, implies NP-hardness of approximating SET-COVER to within logarithmic factors [LY94, BGLR93]. The conjecture itself, nonetheless, seems to be of independent interest, allowing a deeper understanding into probabilistic checking of proofs (PCP) and the class NP.

1.1 Cook's Characterization of NP

The known PCP characterization of NP and the one obtained herein are surveyed next. Let us first present (a slight variation of) Cook's characterization of NP.

Theorem 1 ([Coo71]) *For every $L \in NP$ there exists a polynomial time procedure that, on input $I \in \{0,1\}^n$, constructs a set of Boolean functions $\Phi_{L,I} = \{\varphi_1, ..., \varphi_l\}$ over Boolean variables $\mathcal{Y} = \{y_1, ..., y_l\}$ — representing the guess bits (or witness) — such that*

1. *Each $\varphi_i \in \Phi_{L,I}$ depends on $O(1)$ variables in $\mathcal{Y}$*
2. *$I \in L$ if and only if there exists an assignment $A: \mathcal{Y} \to \{0,1\}$ such that all $\varphi_i \in \Phi_{L,I}$ evaluate to* `true` *on A.*

1.2 PCP Characterization of NP

The characterization of NP in terms of PCP differs from Cook's characterization only in regards to how many of the Boolean functions can be simultaneously satisfied in case the input I is not in L (see condition 3). The following theorem was proved by [ALM$^+$92], based on methods of [AS92]:

Theorem 2 ([AS92, ALM$^+$92]) *For every $L \in NP$ there exists a polynomial time procedure that, on input $I \in \{0,1\}^n$, constructs a set of Boolean functions $\Phi_{L,I} = \{\varphi_1, ..., \varphi_l\}$ over Boolean variables $\mathcal{Y} = \{y_1, ..., y_l\}$ such that*

1. *Each $\varphi_i \in \Phi_{L,I}$ depends on $O(1)$ variables in $\mathcal{Y}$*
2. *$I \in L$ if and only if there exists an assignment A to the variables $\mathcal{Y}$ such that all $\varphi_i \in \Phi_{L,I}$ evaluate to* `true` *on A.*
3. *If $I \notin L$ then for any assignment A to the variables $\mathcal{Y}$, more than $\frac{l}{2}$ of $\varphi_i \in \Phi_{L,I}$ evaluate to* `false` *on A.*

In contrast to Cook's characterization, here, if the input I is not in L, at least half of the Boolean functions are unsatisfied, no matter which assignment is chosen. One should note that this variation, of the new characterization of NP from the old one, however syntactically small, is in fact not at all minor. It causes both a great increase in how difficult it is to prove the theorem, and, in return, enables quite stronger implications deduced.

1.3 Stronger PCP Characterizations of NP

The characterization of NP described above seemed, at first, as the best possible. The PCP hierarchy [AS92] was defined so as to consider the number of random bits (which corresponds to the logarithm of the total number of functions in $\Phi_{L,I}$, and is always $O(\log n)$ in the formalism used herein, because we require the generating procedure to halt in polynomial time) and the total number of answer bits (which is, according to our formalism, the number of bits each Boolean function depends on). The possibility of letting variables range over sets containing more than two elements was not considered in [AS92].

Once a larger range is allowed, the number of variables each Boolean function depends on may be considerably smaller than the total number of bits the function takes as input. One may therefore contemplate stronger characterizations of NP in terms of PCP. Namely, NP can possibly be characterized as in the above except that, while the number of variables would remain constant, their range could become larger, thereby achieving smaller, hopefully sub-constant, error-probability (the error-probability is defined to be the fraction of functions that can be satisfied simultaneously in case the original input is not in L).

A stronger characterization of NP along these lines, taking these parameters into consideration, was conjectured in [BGLR93] (formulated there for a somewhat different model). These two parameters — the number of variables each function depends on and the probability of error — play a crucial role in some applications of PCP characterizations of NP. The number of variables preferably remains constant while the probability of error would hopefully become exponentially small in the number of bits required to represent the value for each variable.

Thus, the two differences, between the current PCP characterization of NP and the conjectured one, are in regards to the range of values for the variables, and the probability of error. The conjecture proceeds as follows:

Conjecture 3 ([BGLR93]) *For every $L \in NP$, and at most logarithmic m, there exists a polynomial time procedure that, on input $I \in \{0,1\}^n$, constructs a set of Boolean functions $\Phi_{L,I} = \{\varphi_1, ..., \varphi_l\}$ over variables $\mathcal{Y} = \{y_1, ..., y_l\}$ which range over $[1, ..., 2^m]$ such that*

1. *Each $\varphi_i \in \Phi_{L,I}$ depends on $O(1)$ variables in $\mathcal{Y}$*
2. *$I \in L$ if and only if there exists an assignment A to the variables $\mathcal{Y}$ such that all $\varphi_i \in \Phi_{L,I}$ evaluate to* `true` *on A.*
3. *If $I \notin L$ then for any assignment A to the variables $\mathcal{Y}$, at most $\frac{l}{2^m}$ of $\varphi_i \in \Phi_{L,I}$ evaluate to* `true` *on A.*

The conjecture takes an extreme view, expecting every bit the functions depend on to decrease the probability of error by a factor of 2. The barrier of constant error-probability (for a constant number of accesses whatever m is), however, seemed to have been unbreakable, so far.

A construction of a function-system of super-polynomial ($n^{O(\log n)}$ to be precise) size, which, however, otherwise complies with the conditions of the [BGLR93] conjecture, is implied as an application of [Ra95]. An exact characterization of NP, however, cannot be deduce from the analysis there.

Proving the above conjecture would directly imply NP-hardness for the problem of approximating SET-COVER to within logarithmic factors (see [BGLR93]). Nevertheless, the conjecture has the potential of an impact quite larger in scope, deepening our understanding with regards to the class NP.

According to a different perspective of the PCP theorem the above conjecture suggests that any mathematical proof (presented in the right format) can be verified by reading only a *constant* number of its lines, each consisting of m bits, and such that the probability of not detecting an error in the proof would be *exponentially small in m.*

1.4 Our New Characterization of NP in Terms of PCP

Our results, concerning characterization of NP in terms of PCP, approaches the above conjecture, and prove it true for a sizable fraction of the appropriate range. For logarithmic m (recall that m denotes the number of bits required to represent each variable), the error-probability obtained is $2^{-m^{1-\epsilon}}$ for any $\epsilon > 0$. If a similar result for ϵ equal 0 were shown, the conjecture would follow. Our results in fact bring us closer to the above conjecture:

Theorem 4 *There exists a constant $\beta > 0$, such that if $M(n)$ is any function satisfying $M(n) \leq (\log n)^\beta$ then for every $L \in NP$, there exists a polynomial time procedure that, on input $I \in \{0,1\}^n$, constructs a set of Boolean-functions $\Phi_{L,I} = \{\varphi_1, ..., \varphi_l\}$ over variables $\mathcal{Y} = \{y_1, ..., y_l\}$ which range over $[1, ..., 2^{M(n)}]$, such that*

1. *Each $\varphi_i \in \Phi_{L,I}$ depends on $O(1)$ variables in $\mathcal{Y}$.*
2. *$I \in L$ if and only if there exists an assignment A to the variables $\mathcal{Y}$ such that all $\varphi_i \in \Phi_{L,I}$ evaluate to* `true` *on A.*
3. *If $I \notin L$ then for any assignment A to the variables $\mathcal{Y}$, at most $\frac{l}{2^{M(n)}}$ of $\varphi_i \in \Phi_{L,I}$ evaluate to* `true` *on A.*

In addition, if one is interested in $M(n)$ closer to $\log n$, the following holds as well:

Theorem 5 *Let $M(n)$ be any function which is at most $(\log n)^{1-\beta}$ for any constant $\beta > 0$. Then for every $L \in NP$, there exists a polynomial time procedure that, on input $I \in \{0,1\}^n$, constructs a set of Boolean-functions $\Phi_{L,I} = \{\varphi_1, ..., \varphi_l\}$ over variables $\mathcal{Y} = \{y_1, ..., y_l\}$ which range over $[1, ..., 2^{M(n)\cdot(\log M(n))^c}]$, for some constant c, such that*

1. *Each $\varphi_i \in \Phi_{L,I}$ depends on a constant C_M variables in $\mathcal{Y}$.*
2. *$I \in L$ if and only if there exists an assignment A to the variables $\mathcal{Y}$ such that all $\varphi_i \in \Phi_{L,I}$ evaluate to* `true` *on A.*
3. *if $I \notin L$ then for any assignment A to the variables $\mathcal{Y}$, at most $\frac{l}{2^{M(n)}}$ of $\varphi_i \in \Phi_{L,I}$ evaluate to* `true` *on A.*

[BGLR93] formulated the above conjecture, for a somewhat different model, in order to be applied to obtain NP-hardness result for approximating SET-COVER to within logarithmic factors. Indeed, combined with some observations, regarding the relation between the two models, and other properties the characterization should comply with, NP-hardness for SET-COVER can be deduced. This is to be elaborated in the full version of the paper.

1.5 The Proof

A critical section of previous proofs for the old PCP theorem is the so called "low-degree–test". The analysis of the error-probability of that test is the most technical part of those proofs. This analysis, however, is not tight. An improved analysis (for possibly a different test) may imply a stronger PCP characterization of NP, and may at the same time simplify the entire proof.

The original proof for the PCP theorem uses the low-degree–test of [RS92]. As mentioned above, our main new tool is a new low-degree–test, and a new analysis for that test. We prove the new test to be of a very small error-probability. In fact, our analysis for the error-probability of the new test is almost tight. Previous low-degree–tests achieved error-probability of, at best, a constant. [1]

The new low-degree–test uses the restriction of low-degree polynomials to planes (i.e., affine sub-spaces of dimension 2), rather than the standard use of restriction to lines (i.e., affine sub-spaces of dimension 1). Functions over geometric spaces, represented by their values over planes, are a special case of the format used in [GS].

Thus, the proof for the new PCP characterization of NP has two parts. The first is the low-degree–test, and

[1] Subsequent to the results reported herein, Arora and Sudan [AS97] presented a new analysis for the low-degree–test of [RS92], showing it to exhibit sub-constant error-probability, similar to the one achieved by our analysis for the new low-degree–test.

the second is the use of the low-degree–test to achieve the new characterization. The second part is not at all trivial. There are two possible strategies in trying to achieve the second step. The first is relying on previous works like [ALM+92, BGLR93, Ra95], and trying to combine them together with the improved low-degree–test (using some new ideas) to achieve the final characterization. The strategy we have chosen, in presenting the results reported herein, however, is to not rely on any previous work, rather rework the proof, extending when necessary, making unclear issues precise, and simplifying where we could. There are several advantages for this strategy. First, if one wants to understand the entire argument one has to understand only the relevant material, and not every single paper in the area. Second, the results we get this way are stronger. And finally, we believe that this way the proof is more precise, and easier to verify (although it is still very hard).

Our proof starts off with the general scheme of [AS92], as amended (following the technique of [LS91], improved in [FL92]) by [ALM+92]. This scheme, however, is insufficient when sub-constant error-probability is sought after. Further alterations are therefore required, in order to allow undisturbed flow of such small error-probability between various parts of the proof.

The specific paradigm used in our analysis is syntactically different than previous presentations. It should be viewed however simply as an alternative structure, which in essence is not much different to previous ones, however allows us to define precisely the various necessary parameters. Let us emphasize that our exposition by no means turn the proof into a simple one, nevertheless, makes it precise and certainly, despite proving a stronger claim, not more complex than previous presentations, perhaps simpler.

In this extended abstract we discuss only the first part of the proof, that is, the low-degree–test. The second part is omitted.

2 Low-Degree–Tests

Let $\mathcal{D}$ be a vector-space of dimension d over a field $\mathcal{F}$, that is,

$$\mathcal{D} = \mathcal{F}^d$$

and let $T: \mathcal{D} \to \mathcal{F}$ be an arbitrary function, which can be thought of as a table of values that has one entry for each element of $\mathcal{D}$. Such a table T can fall into one of the following categories:

1. T is a polynomial of low-degree, i.e. it can be described as a polynomial of d variables, and such that the total-degree of the polynomial is small — how small is determined by a parameter of the test, denoted r. In most cases the ratio $r/|\mathcal{F}|$ is taken to be very small. Denote this ratio by α.
2. T is not a polynomial of low degree, however, there exists a low-degree polynomial $g: \mathcal{D} \to \mathcal{F}$, such that T and g agree on a large fraction of the domain $\mathcal{D}$ — how large a fraction is again given as a parameter of the test, denoted δ. T, in that case, is said to be *close* to g.
3. T is not a polynomial of low-degree, and is not close to any polynomial of low-degree.

Informally, a low degree test is a game between a verifier and a prover, where the prover tries to convince the verifier that the table T falls into the first category, that is, T is a polynomial of low-degree. It should be the case that if T falls into the first category, and the prover is honest, then the verifier accepts with probability 1 (i.e., if T falls into the first category the prover has a strategy that causes the verifier to always accept). If T, however, falls into the third category then the verifier should reject with probability of at least $1-\gamma$ (where γ is again a parameter), no matter what the strategy of the prover is.

2.1 The Line-Point Test

Let us first describe the standard low-degree–test, first introduced by [RS92]. We call this test; the Line-Point test, as it compares values on lines to values on points. This test is the one used in the original proof for the PCP theorem.

A line in $\mathcal{D}$ is an affine subspace of $\mathcal{D}$ of dimension 1. Denote by $\mathcal{S}^1$ the set of all lines in $\mathcal{D}$. For every such line ℓ, let us consider the restriction of the function T to ℓ. If T is a degree-r polynomial[2] then, obviously, its restriction to each line is a uni-variate–polynomial of degree r.

Let now T be a table, as above, supposedly containing the values for each point of a certain low-degree polynomial. Let T' be a table, supposedly containing the restriction of T to every line of $\mathcal{D}$.

For $\ell \in \mathcal{S}^1$, denote by $T'(\ell)$ the ℓ^{th} element of T' (which is supposedly the restriction of T to ℓ). Given T and T', let us say that T' is *honest* if indeed, for every ℓ, $T'(\ell)$ is the restriction of T to ℓ.

The aim of the Line-Point low-degree–test is to output `true` with probability 1 if T falls into the first category **and** T' is honest, however to output `false` with probability at least $1-\gamma$ if T falls into the third category (no matter what T' is in that case). If T falls into the second category, or alternatively, if T falls into the first one, however T' is not honest, then the output of the test may be arbitrary.

By a contrapositive perspective, if the test outputs `false` with probability greater than 0, then either T' is not honest, or T is not a polynomial of low-degree. On

[2] A polynomial of degree $< r$ is nonetheless of degree r as well

the other hand, if the test outputs true with probability greater than γ, then T is either a polynomial of low-degree or is close to a certain polynomial of low-degree.

Note that the smaller γ is, the stronger is the test.

The Line-Point test proceeds as follows: The verifier picks a random line $\ell \in \mathcal{S}^1$, hence ℓ contains $|\mathcal{F}|$ elements of $\mathcal{D}$. The verifier chooses a random element $x \in \ell$, i.e., an element of $\mathcal{D}$ contained in ℓ. If $T'(\ell)$ is indeed a polynomial of low-degree, and $T'(\ell)$ agrees with $T(x)$, then the verifier outputs true. Otherwise, the verifier outputs false.

Clearly, if the verifier outputs false then either T' is not honest, or T is not a polynomial of low-degree. Therefore, to prove the correctness of the test, one just needs to show that if the verifier outputs true with probability greater than γ, then T is close to a polynomial of low-degree g (or, even better, T is exactly a polynomial of low-degree).

Such theorems were proven in [BLR90, Lip91, FSH94, GLR+91, RS92, She91, Rub90, Sud92, FS91] for some values of δ and γ. These proofs, however, do not generalize to the cases where γ is small, say $\gamma < 1/2$. In order to reduce the error-probability in the PCP theorem, one should prove the correctness of the test for much smaller values.

Theorem 6 ([ALM+92]) *There exist constants $1 > \gamma, \delta > 1/2$, such that for every T, T' (as above), if the Line-Point test outputs* true *with probability greater than γ then there exists a degree-r polynomial $g: \mathcal{D} \rightarrow \mathcal{F}$, such that T and g agree on a fraction of at least δ of the domain $\mathcal{D}$.*

2.2 The Plane-Point Test

A plane in $\mathcal{D}$ is an affine subspace of $\mathcal{D}$ of dimension 2. Denote by $\mathcal{S}^2$ the set of all planes in $\mathcal{D}$. For every such plane $\wp$, let us consider the restriction of the function T to $\wp$. If T is a degree-r polynomial then, obviously, its restriction to each plane is a bi-variate–polynomial of degree r.

Let now T be a table, as above, supposedly containing the values for each point of a certain low-degree polynomial. Let T' be a table, supposedly containing the restriction of T to every plane of $\mathcal{D}$. As before, given T and T', let us say that T' is *honest* if for every $\wp$, $T'(\wp)$ is the restriction of T to $\wp$.

The Plane-Point test proceeds as follows: Pick a random plane $\wp \in \mathcal{S}^2$, hence $\wp$ contains $|\mathcal{F}|^2$ elements of $\mathcal{D}$. Now choose a random element $x \in \wp$. If $T'(\wp)$ is indeed a polynomial of low-degree, and $T'(\wp)$ agrees with $T(x)$, then the test outputs true. Otherwise, the test outputs false.

As before, to prove the correctness of the test, one just needs to show that if the test outputs true with probability greater than γ, then T is close to a polynomial of low-degree g.

Our analysis proves the correctness of the Plane-Point test, for very small values of γ. This will follow as an application of our analysis for the new low-degree–test.

Theorem 7 *There exists a constant c, such that for every $\gamma > cd\alpha^{1/c}$ (recall that $\alpha = r/|\mathcal{F}|$.), and every T, T' (as above), if the Plane-Point test outputs* true *with probability greater than γ then there exists a degree-r polynomial $g: \mathcal{D} \rightarrow \mathcal{F}$, such that T and g agree on a fraction of at least δ of the domain $\mathcal{D}$, where $\delta = \Omega(\gamma)$.*

2.3 Our New Low-Degree–Test

Thus, our argument can be used to achieve an almost tight analysis for the Plane-Point low-degree–test. We prefer, however, to introduce a new low-degree–test, one that doesn't use the table of points at all, and uses only the consistencies among restrictions to planes. Besides being more elegant, more general and a bit stronger, the new test gives a new insight into the problem, and, in fact, gives the main intuition for the proof.

Consider a table T consisting of one entry, $T(\wp)$, for each $\wp \in \mathcal{S}^2$; assume that that entry contains an r-degree polynomial, which is supposedly the restriction of g to $\wp$ (where as before, g is a polynomial of total-degree r). The new test proceeds as follows: Pick two random planes $\wp_1, \wp_2 \in \mathcal{S}^2$, such that $\wp_1 \cap \wp_2$ is a line $\ell \in \mathcal{S}^1$. If $T(\wp_1)$ agrees with $T(\wp_2)$ on their intersection ℓ then the test outputs true. Otherwise, the test outputs false.

The power of the new test is that the two polynomials, corresponding to $\wp_1, \wp_2$, have to agree on every single point of the line of intersection ℓ. Our analysis is based on the fact that if they disagree, and they are both polynomials of low-degree then they may agree on only a very small fraction of the points in ℓ. Using this fact, we are able to give an argument which is combinatorial (and geometric) in nature, rather than the previous algebraic arguments. At least in the case of small dimension (i.e. $d = 3$), the entire argument is significantly simpler than the previous algebraic proofs.

Theorem 8 *There exists a constant c, such that for every $\gamma > cd\alpha^{1/c}$, and every T (as above), if the new test outputs* true *with probability greater than γ then there exists a degree-r polynomial $g: \mathcal{D} \rightarrow \mathcal{F}$, such that for at least δ fraction of planes $\wp \in \mathcal{S}^2$, $T(\wp)$ is the restriction of g to $\wp$, (where $\delta = \Omega(\gamma)$).*

Theorem 7 now follows as a simple consequence of Theorem 8: Given tables T, T' such that the Plane-Point test outputs true with a non-negligible probability, it is not hard to show that on the table T' the new test outputs true with a non-negligible probability as well. Contrapositively, if the new test outputs true with a very small probability (on a table T') then the Plane-Point test outputs true with a very small probability on T, T' (for any T). The precise claim (and proof) is omitted in this version of the paper.

In order to use the new low-degree–test to prove low-error PCP characterizations of NP, one needs a stronger version of Theorem 8. Section 3 shortly explains why the stronger theorem is needed. The stronger theorem is then stated in Section 4.

3 Consistency with Low Error-Probability

Assume now a low-degree test proven to satisfy the above requirements (say, the requirements of the Plane-Point test). Consider a two-phase process that first apply the test, and then, if the value returned is true, proceeds under the assumption that T represents a certain polynomial g of low-degree. One can then sample g on a random element x of $\mathcal{D}$ just by reading the value $T(x)$. By the correctness of the test, one can assume that $T(x) = g(x)$. This, of course, is not always true. The larger $1 - \gamma$ and δ are, the safer it is to trust that assumption.

In previous applications of low-degree–tests for proving PCP characterizations of NP, this two-phase process is applied. The table is first checked so that serious inconsistencies are detected with high probability. It is then assumed that the values obtained for the requested entries indeed correspond to a single global low-degree polynomial.

When trying to achieve much smaller probabilities of error for the PCP theorem, however, such a two-phase process cannot do. E.g., it is possible to assign values to one subset of the entries to the table, that would cause the first phase (that of checking consistency) to succeed with small, however non-negligible, probability. At the same time, however, a disjoint subset of the entries can be assigned values that have nothing to do with those of the first subset. Values which the second phase would return with non-negligible probability. Hence, with tests that perform only a constant number of accesses, the barrier of constant error-probability cannot be broken.

Of critical importance, in the development of the ideas leading so far, is the work of [GS]. The work reported here is in fact an extension of the line of research pursued there. [GS] shows how to avoid the all familiar low-degree test, and obtain instead a *consistency test.* Namely, a test that, assuming an arbitrary function f, enables consistent reading of tuples of several values of f. A consistency-test of constant error-probability is presented there. This is in contrast to previous proofs for the old PCP theorem that assumed f to be a low-degree polynomial. The theorem of [GS], although implying a low-degree test, is proven there using only combinatorial properties of the representation of f.

It follows that, for constant error-probability, one needs not assume any properties regarding the function which is to be read consistently. Our work here proceeds to show that in order to obtain sub-constant error-probability, one can incorporate some restrictions regarding f. Namely, requiring it to comply with some conditions akin to properties of error-correcting–codes.

It seems hence to be the case that low-degree polynomials were previously taken advantage of for two different type of properties, which were however not distinguished. The first being the geometrical nature of their domain. The other being extended error-correcting properties.

Consistent Readers

Let us shift to a more general paradigm by which to approach the issue. The procedure in quest is to be given a table, whose values supposedly comply with some global consistency condition (being of low degree is just one possible consistency condition). Given any entry to the table, the procedure is required to return, with very high probability, a value for the requested entry that would correspond to a function that conforms to the consistency condition. The procedure is to achieve that, while performing only a very small number of accesses to the table.

Let us dwell a little on what should be considered a faulty outcome for a random test of such a procedure. If the procedure just happens to step on values that cannot possibly agree with any globally consistent function, it might as well announce it. The table is supposedly everywhere consistent, hence any such inconsistency disqualifies it, and the procedure might as well halt.

Our attention therefore should be focused on the values the procedure returns in case no inconsistency is detected, and how safe it is to assume that the values returned correspond to a globally consistent function.

Requiring, as in the old test, that the values returned for most random tests correspond to a single consistent function, cannot be guaranteed when such small probabilities of error are at aim. E.g., it is possible to partition the set of entries to the table into several subsets, where entries in each subset are assigned values corresponding to a distinct consistent function. This can be carried out while the probability of all accesses of a random test to fall within one subset is non-negligible. No inconsistency can be detected in those cases. The values returned, however, clearly do not all (even most) correspond to a single function.

Limited Pluralism

Our requirements regarding the values returned by the procedure are therefore relaxed so as to allow several permissible functions. [3] The functions allowed are those

[3] We have learnt that the same idea was found independently and before us by several other researchers. In particular, it was

whose values agree with a non-negligible fraction of the entries in the table. The procedure is required to return *impermissible* values — i.e., values that do not correspond to any permissible function — only with a very small probability.

If, as in the case with low-degree polynomials, two globally consistent functions can agree on at most a small fraction of the entries, there can be only a small number of functions that would be considered permissible.

A short summary of the requirements of our test is as follows. Given a table supposedly describing a globally consistent function, the outcome for a particular random choice of the test is considered non-faulty in one of the following cases:

1. The values the test returns correspond to one of the permissible functions.

2. An inconsistency is announced, and the table is indeed inconsistent.

It should be the case that all other possibilities, where the outcome can be labeled faulty, occur only with a very small probability.

4 The New Low-degree–test; Strong Version

Denote by e the set of all polynomials $g: \mathcal{D} \to \mathcal{F}$ of total degree r, and as before denote $\alpha = r/|\mathcal{F}|$. As before, consider one polynomial $g \in \mathsf{e}$. Consider a table T consisting of one entry, $T(\wp)$, for each $\wp \in \mathcal{S}^2$; assume that that entry contains an r-degree polynomial, which is supposedly the restriction of g to $\wp$.

T is said to be *globally-consistent* if indeed there exists $g \in \mathsf{e}$, such that for every $\wp$, $T(\wp)$ is the restriction of g to $\wp$. $g' \in \mathsf{e}$ is said to be T-permissible if T agrees with g' on a non-negligible, (determined by a parameter denoted δ), fraction of the planes.

Denote by $\mathcal{P}$ the set of all pairs of planes of $\mathcal{D}$ that intersect by a line:

$$\mathcal{P} = \left\{ \langle \wp_1, \wp_2 \rangle \,|\wp_1, \wp_2 \in \mathcal{S}^2 \text{ and } \wp_1 \cap \wp_2 \in \mathcal{S}^1 \right\}.$$

Consider a pair $\langle \wp_1, \wp_2 \rangle \in \mathcal{P}$ of planes of $\mathcal{D}$ that intersect by a line $\ell = \wp_1 \cap \wp_2$. The pair $\langle \wp_1, \wp_2 \rangle$ is said to be *T-pairwise-consistent* if $T(\wp_1)$ agrees with $T(\wp_2)$ on ℓ.

The following theorem is the stronger version of Theorem 8. It will follow as a consequence of Theorem 10 proven below:

---

discussed in a talk by Sanjeev Arora in a conference at Weizmann Institute of Science (Jan. 1994).

Theorem 9 *Given any table T, the fraction of pairs of planes $\langle \wp_1, \wp_2 \rangle \in \mathcal{P}$ that are T-pairwise-consistent however do not agree with any T-permissible polynomial g' is at most $O\left(d \cdot \alpha^{1/c} + \delta\right)$, where c is a (large) universal constant.*

It follows that, given a random point $x \in \mathcal{D}$, one may choose a random pair of planes $\langle \wp_1, \wp_2 \rangle \in \mathcal{P}$, one of which containing x, and verify that $\langle \wp_1, \wp_2 \rangle$ is T-pairwise-consistent. If successful, one proceeds while trusting the value thus obtained for x to agree with one of the T-permissible low-degree polynomials. Only a small fraction of the random tests would lead to trust impermissible values. Furthermore, by a slight extension of this test it is possible to obtain a similar consistent-reader for an arbitrary point x (rather than just for a random one).

4.1 Local Consistency versus Global Consistency

Before getting to the proof, let us introduce some notations and present the previous theorem in slightly different form.

Local Consistency

Let $\mathrm{CONS}_T(\mathcal{P})$ be the set of all pairs of planes $\langle \wp_1, \wp_2 \rangle \in \mathcal{P}$ that are T-pairwise-consistent. Local consistency is now defined as follows:

Definition 1 *T is* γ–pairwise–consistent *if*

$$\Pr_{\langle \wp_1, \wp_2 \rangle \,\in_{\mathrm{R}}\, \mathcal{P}} \left[\langle \wp_1, \wp_2 \rangle \in \mathrm{CONS}_T(\mathcal{P}) \right] \geq \gamma$$

where " $\in_{\mathrm{R}}$ " denotes a choice of an element according to the uniform distribution.

It is not hard to prove that if all pairs of planes in $\mathcal{P}$ are pairwise consistent (that is, T is 1-pairwise-consistent) then there exists $g \in \mathsf{e}$ whose restriction to every plane $\wp$ is $T(\wp)$. We are interested however in the case where γ is much smaller, sub-constant in fact. In that case, perfect global consistency does not necessarily hold.

Global Consistency

The *support* of a low degree polynomial $g \in \mathsf{e}$ is the set of all planes $\wp \in \mathcal{S}^2$ that are assigned by T values that agree with g, i.e., where $T(\wp) = g_{\downarrow \wp}$. A $[T, \mathsf{e}]$-*extension* is a function

$$A: \mathcal{S}^2 \to \mathsf{e}$$

labeling every plane $\wp$ by $g \in \mathsf{e}$ so that $\wp$ is in the support of g, that is,

$$\forall \wp \in \mathcal{S}^2, A(\wp)_{\downarrow \wp} = T(\wp).$$

A is *γ-consistent* if at least γ fraction of the pairs of planes intersecting by a line are labeled the same by A, that is, if

$$\Pr_{\langle \wp_1, \wp_2 \rangle \in_R \mathcal{P}} [A(\wp_1) = A(\wp_2)] \geq \gamma.$$

Definition 2 *T is* γ-consistent *if there exists a γ-consistent $[T, \mathsf{e}]$-extension A.*

T being γ-consistent implies that T is γ–pairwise–consistent, as if two planes are assigned the same global encoding by A they must agree on their intersection. The other direction is nonetheless not necessarily true.

From Local to Global

The reader might note that γ-consistency, even for a rather small γ, implies that there is only a small number of global encodings whose support is not negligible. Our Ultimate Claim can now be stated:

Theorem 10 *There exists a global constant $c \geq 2$ so that for any γ, if T is γ–pairwise–consistent then T is $(\gamma - \epsilon \cdot d)$-consistent, for $\epsilon = c \cdot \alpha^{1/c}$.*

In other words, Theorem 10 implies that only a small $(\leq \epsilon \cdot d)$ fraction of the pairs of planes in $\mathcal{P}$ can be given pairwise consistent values by T, however not belong to the support of one, out of a small number of, global encodings. That is, the probability of error of T is $\leq \epsilon \cdot d$.

5 Proof of the Theorem for $d = 3$

Proof: Our proof for Theorem 10 is by induction on d, the dimension of $\mathcal{D}$. For $d = 2$ there is only one plane and nothing to prove. Thus, the first interesting case is $d = 3$. In this version of the paper we will prove the theorem for $d = 3$, where one can see the main idea of the proof, and get the main intuition. The case $d > 3$ follows by induction, using the same ideas (but is a bit more complicated).

Alternatively, as was recently shown by [AS97], once proving the case $d = 3$, one can use the so called "bootstrapping" method of [RS92] (see also [ALM$^+$92]) to get the result for any d. The bootstrapping arguments in [RS92, ALM$^+$92, AS97] are for the Line-Point test. A variation of this argument, however, works for the Plane-Point test, and for the new test as well.

Consistency Graph

In order to isolate and clearly identify the few places in which properties related to low-degree–polynomials are used, the relationship between the entries of the table T are represented by a special type of graph G_T.

Let $G_T = \langle V, E \rangle$ be a graph, whose vertex set has one vertex for every plane, hence $V = \mathcal{S}^2$.

For two vertexes $\wp_1$ and $\wp_2$ in G_T, either $\wp_1$ and $\wp_2$ are parallel, or they intersect by a line. In the former case, the two vertexes are connected by an edge; in the latter the two vertexes are connected if they are T-pairwise-consistent.

Let ψ be the uniform probability distribution, defined over V. For ease of reading, let us extend the notation ψ to sets of vertexes of G and to sets of pairs of vertexes, in the natural way, i.e., for $U \subseteq V$, $\psi(U) = \Pr_{v \in_\psi V} [v \in U]$ and for $M \subseteq V^2$, $\psi(M) = \Pr_{v_1, v_2 \in_\psi V} [\langle v_1, v_2 \rangle \in M]$ (where $\in_\psi$ corresponds to a choice of an element according to ψ).

The *edge-probability* of a graph $G = \langle V, E \rangle$, — denoted $e(G)$ — is defined to be

$$e(G) = \psi(E).$$

By the definition of G_T, the edge-probability of G_T can be bounded from below as follows:

Claim 1 $e(G_T) \geq \gamma$.

Transitivity

Let $G = \langle V, E \rangle$ be a graph; a *non-transitive triplet* in G are three vertexes, $v_1, v_2, v_3 \in V$ such that out of the three inner pairs (v_1, v_2), (v_1, v_3) and (v_2, v_3), exactly two of the edges exist in G (i.e., belong to E). G is *transitive* if it has no non-transitive triplets.

G_T is not necessarily transitive; nevertheless, it turns out to be "almost transitive" in the sense that, it is possible to remove edges from G_T such that the edge-probability of G_T remains almost the same, and so that G_T becomes transitive. For that purpose, let us introduce the following parameter:

Given any graph $G = \langle V, E \rangle$, consider a non-edge $\langle v_1, v_2 \rangle \in \binom{V}{2} \setminus E$ in G; let $\beta(v_1, v_2)$ be the probability, over $v_3 \in_\psi V$, that the triplet v_1, v_2, v_3 is non-transitive (due to the non-edge $\langle v_1, v_2 \rangle$), that is

$$\beta(v_1, v_2) = \Pr_{v_3 \in_\psi V} [\langle v_1, v_3 \rangle, \langle v_3, v_2 \rangle \in E].$$

Let $\beta(G)$ be the maximum, over $\langle v_1, v_2 \rangle \notin E$, of $\beta(v_1, v_2)$.

The consistency graph G_T is in fact such that $\beta(G_T)$ is rather small — as proven below. Let us, at this point, show a general upper bound — for any graph G — of the weight of edges to be removed from G, as a function of $\beta(G)$, in order for G to become transitive:

Lemma 2 *Any graph $G = \langle V, E \rangle$ has a transitive subgraph $G' = \langle V, E' \rangle$ (where $E' \subseteq E$) such that $\psi(E') \geq \psi(E) - 3\sqrt{\beta(G)}$.*

Proof: We will give an algorithm that on input G, outputs such a graph G'. The algorithm identifies a sequence of graphs $G = G_0, G_1, ..., G_l = G'$, where G' satisfies the necessary conditions. First, for a vertex

$v \in V$, denote by $\mathcal{N}_i(v)$ the neighbor set of v in G_i. For $i = 0, 1, 2, ...$, if G_i is transitive then fix $l = i$ and $G_i = G'$.

Otherwise, $G_{i+1} = \langle V, E_{i+1} \rangle$ is constructed from $G_i = \langle V, E_i \rangle$ by applying one of the following two steps:

1. If there exists a vertex $v_i \in V$ such that
$$\psi(\mathcal{N}_i(v_i)) \leq \sqrt{\beta(G)}$$
then E_{i+1} is E_i except all edges containing v_i are removed.

2. Otherwise, unless G_i is transitive there must exist a vertex $v_i \in V$ that is disconnected from at least one vertex in its connected component C_i in G_i, and such that
$$\psi(\mathcal{N}_i(v_i)) > \sqrt{\beta(G)}.$$
In that case, all edges between $\mathcal{N}_i(v_i)$ and $C_i \setminus \mathcal{N}_i(v_i)$ are removed from E_i in order to obtain E_{i+1}.

Let us now show that the total edge-probability removed going from G to G' is small, that is,
$$\psi(E \setminus E') \leq 3\sqrt{\beta(G)}.$$
For that purpose, define I_1 to be the set of indexes in which step 1 is applied, and I_2 be the rest; we show that the two totals of edge probability removed in steps of I_1 and of I_2 are both small.

As far as for I_1, for each i, at most $2\sqrt{\beta(G)} \cdot \psi(v_i)$ edge probability is removed (the factor 2 is due to the fact that each edge contributes twice to the edge-probability). Since for distinct $i \neq i'$, $v_i \neq v_{i'}$, the total edge-probability removed due to enforcement of step 1 is $\leq 2\sqrt{\beta(G)}$.

Now let us turn to I_2, and show that
$$\sum_{i \in I_2} \psi(E_i \setminus E_{i+1}) \leq \sqrt{\beta(G)}$$
Denote
$$T_i = (C_i \setminus \mathcal{N}_i(v_i)) \times \mathcal{N}_i(v_i)$$
and note that it is clearly the case that for any distinct $i \neq i' \in I_2$, T_i and $T_{i'}$ are disjoint. Therefore, it is enough to prove that, for every $i \in I_2$,
$$\frac{\psi(E_i \setminus E_{i+1})}{\psi(T_i)} < \sqrt{\beta(G)}$$
That is, the weight of edges removed in each application of step 2 is bounded from above by a value smaller than $\sqrt{\beta(G)}$ time the weight of T_i. Since the total weight of all T_i's is at most 1, the total edge-probability removed in steps of I_2 must be $\leq \sqrt{\beta(G)}$. The remainder of the proof of Lemma 2 is devoted to the proof of the above statement.

Let $v \in C_i \setminus \mathcal{N}_i(v_i)$ and $v' \in \mathcal{N}_i(v_i)$, such that $\langle v, v' \rangle \in E_i$, the triplet v_i, v, v' is non-transitive due to the non-edge $\langle v_i, v \rangle$. The weight of those non-transitive triangles is bounded from above by $\beta(G)$, hence for every $v \in C_i \setminus \mathcal{N}_i(v_i)$,
$$\Pr_{v' \in_\psi V}[v' \in \mathcal{N}_i(v_i) \text{ and } \langle v, v' \rangle \in E_i] \leq \beta(G)$$
and therefore
$$\psi(E_i \setminus E_{i+1}) \leq \beta(G) \cdot \psi(C_i \setminus \mathcal{N}_i(v_i)).$$
On the other hand, $\psi(\mathcal{N}_i(v_i)) > \sqrt{\beta(G)}$ hence
$$\psi(T_i) = \psi(\mathcal{N}_i(v_i)) \cdot \psi(C_i \setminus \mathcal{N}_i(v_i))$$
$$> \sqrt{\beta(G)} \cdot \psi(C_i \setminus \mathcal{N}_i(v_i))$$
Lemma 2 follows. □

Non-Transitivity Limited

Let us now consider, for every non-edge in G_T, the weight of non-transitive triangle due to that non-edge:

Claim 3 $\beta(G_T) \leq \alpha^+ \stackrel{\text{def}}{=} \alpha + (1/|\mathcal{F}|)$.

Proof: Let $\wp_1$ and $\wp_2$ be disconnected vertexes of G_T. It must then be the case that $\wp_1$ and $\wp_2$ intersect by a line $\ell = \wp_1 \cap \wp_2 \in \mathcal{S}^1$, and $T(\wp_1)$ disagrees with $T(\wp_2)$ on ℓ. $T(\wp_1)$ and $T(\wp_2)$ therefore may agree on at most an α fraction of ℓ. Denote by ℓ' the set of disagreement between $T(\wp_1)$ and $T(\wp_2)$. ℓ' is of $\geq 1 - \alpha$ fraction in ℓ.

Given a vertex $\wp_3$ of G_T, it is either the case that $\wp_3$ does not intersect ℓ, which occurs with probability $|\mathcal{F}|^{-1}$. Otherwise, $\wp_3$ does intersects ℓ, and $\wp_3$ would then be connected to both $\wp_1$ and $\wp_2$ only if $\wp_3$ does not intersect ℓ'. Since only an $\leq \alpha$ fraction of $\wp_3 \in \mathcal{S}^2$ does intersect ℓ however does not intersect ℓ', the claim follows. □

By Lemma 2, G_T has a transitive sub-graph G'_T such that
$$e(G'_T) \geq e(G_T) - 3\sqrt{\beta(G_T)}$$
hence, by Claim 3, and Claim 1, we have
$$e(G'_T) \geq \gamma - 3\sqrt{\alpha^+}.$$
The graph G'_T is transitive, and therefore it is a union of disjoint cliques. Consider one large clique (of fraction more than α^+). I.e. a set Y of vertexes of G_T, such that $\psi(Y) \geq \alpha^+$, and every two vertexes of Y are connected in G_T. By Lemma 4 (below), there exists $g \in \mathsf{e}$, such that for every $\wp \in Y$, g agrees with $T(\wp)$ on S. In other words; every plane $\wp \in Y$ is in the support of g.

Thus, for every large clique there is a corresponding $g \in \mathsf{e}$. Since the number of edges in small cliques is small, we can ignore small cliques, and the theorem follows. □

Lemma 4 (Interpolation) *Consider $Y \subset \mathcal{S}^2$ such that $|Y| \geq \alpha^+ \cdot |\mathcal{S}^2|$, and such that every two planes $\wp_1, \wp_2 \in Y$ with $\wp_1 \cap \wp_2 = \ell \in \mathcal{S}^1$ are T-pairwise-consistent. Then, there exists (exactly one) encoding $g \in \mathbf{e}$ such that for every $\wp \in Y$, $T(\wp)$ is the restriction of g to $\wp$.*

□

Acknowledgment

Johan Håstad had a small nevertheless critical contribution to this work, however persistently refuses to co-author the paper. Oded Goldreich deserves as good as grade for his contributions and support.

References

[ALM+92] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and intractability of approximation problems. In *Proc. 33rd IEEE Symp. on Foundations of Computer Science*, pages 13–22, 1992.

[AS92] S. Arora and S. Safra. Probabilistic checking of proofs: A new characterization of NP. In *Proc. 33rd IEEE Symp. on Foundations of Computer Science*, pages 2–13, 1992.

[AS97] S. Arora and M. Sudan. Improved Low Degree Testing and its Applications. In *Proc. 29th ACM Symp. on Theory of Computing*, 1997 (this proceeding).

[Bab85] L. Babai. Trading group theory for randomness. In *Proc. 17th ACM Symp. on Theory of Computing*, pages 421–429, 1985.

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

[BGKW88] M. Ben-or, S. Goldwasser, J. Kilian, and A. Wigderson. Multi prover interactive proofs: How to remove intractability assumptions. In *Proc. 20th ACM Symp. on Theory of Computing*, pages 113–121, 1988.

[BGLR93] M. Bellare, S. Goldwasser, C. Lund, and A. Russell. Efficient multi-prover interactive proofs with applications to approximation problems. In *Proc. 25th ACM Symp. on Theory of Computing*, pages 113–131, 1993.

[BGS95] M. Bellare, O. Goldreich, and M. Sudan. Free bits and nonapproximability. In *Proc. 27th ACM Symp. on Theory of Computing*, 1995. to appear.

[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.

[Coo71] S. Cook. The complexity of theorem-proving procedures. In *Proc. 3rd ACM Symp. on Theory of Computing*, pages 151–158, 1971.

[FGL+91] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Approximating clique is almost NP-complete. In *Proc. 32nd IEEE Symp. on Foundations of Computer Science*, pages 2–12, 1991.

[FL92] U. Feige and L. Lovász. Two-prover one-round proof systems: Their power and their problems. In *Proc. 24th ACM Symp. on Theory of Computing*, pages 733–741, 1992.

[FS91] K. Friedl and M. Sudan. Some improvements to low-degree–tests. In *Proc. of the Third Israel Symposium on Theory of Computing*, ACM 1991.

[FSH94] K. Friedl, A. Shen, and Z. Hatsagi. The low-degree test. In *Proc. 5th SIAM Symposium on Discrete Algorithms*, 1994.

[GLR+91] P. Gemmell, R. Lipton, R. Rubinfeld, M. Sudan, and A. Wigderson. Self-testing/correcting for polynomials and for approximate functions. In *Proc. 23rd ACM Symp. on Theory of Computing*, pages 32–42, 1991.

[GMR89] S. Goldwasser, S. Micali, and C. Rackoff. The knowledge complexity of interactive proofs. *SIAM J. of Computation*, 18:186–208, 1989.

[GS] O. Goldreich and S .Safra. On the probabilistic checkable proofs of n. In preparation.

[Hås96a] Johan Håstad. Testing of the long code and hardness for clique. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, pages 11–19, Philadelphia, Pennsylvania, 22–24 May 1996.

[Hås96b] Johan Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. In *Proc. 37th IEEE Symp. on Foundations of Computer Science*, 14–16 October 1996.

[Hås97] Johan Håstad. Some optimal inapproximability results. In *Proc. 29th ACM Symp. on Theory of Computing*, 1997 (this proceeding).

[KLS93] S. Khanna, N. Linial, and S. Safra. On the hardness of approximating the chromatic number. In *Proceedings of the 2nd Israel Symposium on Theory and Computing Systems, ISTCS*, pages 250–260. IEEE Computer Society Press, 1993.

[LFKN92] C. Lund, L. Fortnow, H. Karloff, and N. Nisan. Algebraic methods for interactive proof systems. *Journal of the ACM*, 39(4):859–868, October 1992.

[Lip91] R. Lipton. New directions in testing. In J. Feigenbaum and M. Merritt, editors, *Distributed Computing and Cryptography*. American Mathematical Society, 1991. Dimacs Series in Discrete Mathematics and Theoretical Computer Science, volume 2.

[LS91] D. Lapidot and A. Shamir. Fully parallelized multi prover protocols for NEXPTIME. In *Proc. 32nd IEEE Symp. on Foundations of Computer Science*, pages 13–18, 1991.

[LY94] Carsten Lund and Mihalis Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41(5):960–981, 1994.

[PY91] C. Papadimitriou and M. Yannakakis. Optimization, approximation and complexity classes. *Journal of Computer and System Sciences*, 43:425–440, 1991.

[Ra95] R. Raz, A Parallel Repetition Theorem. To appear in *SIAM Journal on Computing.* Prelim. version in *27th STOC*, pp. 447-456, 1995.

[RS92] R. Rubinfeld and M. Sudan. Testing polynomial functions efficiently and over rational domains. In *Proc. 3rd Annual ACM-SIAM Symp. on Discrete Algorithms*, pages 23–32, 1992.

[Rub90] R. Rubinfeld. *A mathematical theory of self-checking, self-testing and self-correcting Programs.* PhD thesis, U.C. Berkeley, 1990.

[Sha92] A. Shamir. IP = PSPACE. *Journal of the ACM*, 39(4):869–877, October 1992. Prelim. version in 1990 FOCS, pages 11–15.

[She91] A. Shen. Multilinearity test made easy. Manuscript, 1991.

[Sud92] M. Sudan. *Efficient checking of polynomials and proofs and the hardness of approximation problems.* PhD thesis, U.C. Berkeley, 1992.