THEORY OF COMPUTING, Volume 9 (28), 2013, pp. 863–887
www.theoryofcomputing.org

SPECIAL ISSUE: ANALYSIS OF BOOLEAN FUNCTIONS

A Two-Prover One-Round Game with Strong Soundness

Subhash Khot* Muli Safra

Received July 2, 2012; Revised May 7, 2013; Published December 6, 2013

Abstract: We show that for any fixed prime $q \geq 5$ and constant $\zeta > 0$, it is NP-hard to distinguish whether a two-prover one-round game with q^6 possible answers has value at least $1-\zeta$ or at most $4/q$. The result is obtained by combining two techniques: (i) An Inner PCP based on the *point versus subspace* test for linear functions. The test is analyzed Fourier analytically. (ii) The Outer/Inner PCP composition that relies on a certain *sub-code covering* property for Hadamard codes. This is a new and essentially black-box method to translate a *codeword test* for Hadamard codes to a *consistency test*, leading to a full PCP construction.

As an application, we show that unless NP has quasi-polynomial time deterministic algorithms, the Quadratic Programming Problem is inapproximable within factor $(\log n)^{1/6-o(1)}$.

ACM Classification: F.1.3

AMS Classification: 68Q17

Key words and phrases: approximation, approximation algorithms, Fourier transform, inapproximability, label cover, Grothendieck inequality, linearity testing, PCP, probabilistically checkable proofs

1 Introduction

It is well known that for many NP-hard problems, even computing approximate solutions is computationally hard. In particular, it is known that given a satisfiable instance of 3SAT, it is NP-hard to find an

A preliminary version of this paper appeared in the Proceedings of the 52nd IEEE FOCS, 2011 [22].

*Research supported by NSF CAREER grant CCF-0833228, NSF Expeditions grant CCF-0832795, NSF Waterman Award and BSF grant 2008059.

 DOI: 10.4086/toc.2013.v009a028

assignment that satisfies a β fraction of the constraints for some absolute constant $\beta < 1$. Indeed, this is precisely the statement of the PCP Theorem [14, 4, 3], which serves as the starting point for many such *inapproximability results* for various NP-hard problems, e. g., [1, 6, 16, 17, 13].

The PCP Theorem may equivalently be stated from a *proof checking viewpoint* (and this viewpoint played a decisive role in its discovery). From the proof checking viewpoint, the PCP Theorem states that NP has *probabilistically checkable proofs* with a very efficient verifier. Specifically, there is a way to write a proof for an NP-statement such that its correctness can be verified by a probabilistic verifier that uses logarithmic randomness and a constant number of queries to the proof. The proof system is *complete* in the sense that a correct proof of a correct statement is accepted with probability 1 and is also *sound* in the sense that any proof of an incorrect statement is accepted with probability strictly bounded away from 1 (this probability can be reduced to an arbitrarily small constant by running the verifier $O(1)$ times). The equivalence between the proof checking viewpoint and the inapproximability viewpoint (i. e., the statement of the PCP Theorem as inapproximability of 3SAT as quoted in the beginning), going back to [11, 14], though not difficult to prove, has led to many illuminating insights and strong inapproximability results over the last two decades. By now, it is a standard practice in PCP literature to freely switch between the proof checking viewpoint and the inapproximability viewpoint and we follow the practice in this paper.

A vast majority of the inapproximability results start by reducing the *hard* 3SAT instance given by the PCP Theorem to a problem called Label Cover, introduced in [1], which is equivalently viewed as a 2-Prover-1-Round Game. A 2P1R Game (see Definition 2.1) has a parameter R that denotes the number of different answers each prover may give on a fixed question. In this paper, all games considered are projection games where the answer of the first prover uniquely determines the answer of the second prover (if the verifier is to accept). For a projection game, the number of answers for the two provers may be different and R denotes the larger of the two numbers. The PCP Theorem combined with Raz's Parallel Repetition Theorem [29] gives the following result.

Theorem 1.1. *There exists an absolute constant $\gamma > 0$ such that for all large constant R, it is NP-hard to distinguish whether the value of a (projection) 2P1R Game with R answers is 1 (called the completeness parameter) or at most $1/R^{\gamma}$ (called the soundness parameter).*

In this paper, we investigate the trade-off between the number of answers R and the soundness parameter. Given the central nature of 2P1R Games, we believe this is a natural pursuit. It is easy to see that the soundness must be at least $\Omega(1/R)$, since the provers may give a random answer and succeed with probability $\Omega(1/R)$. The exponent γ in the above theorem is unspecified in Raz's paper (and the subsequent works of Holenstein [18] and Rao [28]) and even if one were to compute it, it would presumably be very tiny.[1] The main result in this paper is that the above theorem holds essentially with $\gamma = 1/6$, albeit with imperfect completeness.

Theorem 1.2 (Main Theorem)**.** *For any fixed prime $q \geq 5$ and constant $\zeta > 0$, it is NP-hard to distinguish whether a 2P1R Game with $R = q^6$ answers has value at least $1-\zeta$ or at most $4/q$.*

[1] If the value of a game is $1-\alpha$, then the value of the k-wise repeated game is at most $(1-\alpha^{p})^{ck}$ for some absolute constants c and p. We have improvements $p = 32, 3$ and for projection games $p = 2$ from [29, 18, 28] respectively. However, c still remains unspecified and hence the exponent γ remains unspecified in Theorem 1.1.

The exponent γ does play a role in some inapproximability results. For instance, Arora et al. [2] show that the Quadratic Programming Problem is inapproximable within factor $(\log n)^{\gamma}$. This is the problem of maximizing a quadratic form $\sum_{i,j=1}^{n} a_{ij}x_i x_j$ where $a_{ii} = 0$ for $1 \le i \le n$ and the variables x_i are restricted to be $\{-1,1\}$-valued. The problem is known to be approximable within factor $O(\log n)$ [23, 26, 10]. We obtain the following improved inapproximability result. In fact this application was our original motivation.

Theorem 1.3. *For an arbitrarily small constant $\varepsilon > 0$, unless $\mathsf{NP} \subseteq \mathsf{DTIME}(2^{(\log n)^{O(1/\varepsilon)}})$, no polynomial time algorithm can approximate the Quadratic Programming Problem within factor $(\log n)^{1/6-\varepsilon}$.*

This result requires the analogue of Theorem 1.2 when R is a *large enough* growing function of the instance size. We make this statement precise below. Though the statement is a bit cumbersome, it might be useful for future reference. We note that Theorem 1.2 has been recently improved by Chan [9], essentially achieving $\gamma = 1/2$. His result is described in more detail in Section 1.2. His reduction however leads to a large blow-up in the instance size and does not apply for *large enough* values of R and does not lead to improvement to Theorem 1.3.

Theorem 1.4. *For any integer q that is a power of 2, there is a reduction from a SAT instance of size n to a 2P1R Game with $R = q^6$ answers such that:*

- *The size of the 2P1R Game is $2^{O((\log^2 n+\log q)\log^{24} n \cdot q^6 \log^2 q)}$.*
- *If the SAT instance is satisfiable, the value of the 2P1R Game is at least $1 - \dfrac{q^6 \log^2 q}{n^{\Omega(1)}}$.*
- *If the SAT instance is unsatisfiable, the value of the 2P1R Game is at most $4/q$.*

One technical contribution of the paper, perhaps more interesting for future research, is an essentially black-box method to translate a *codeword test* for Hadamard code (i. e., a linearity test) to a *consistency test*, leading to a full PCP construction. We state this as informal Theorem 1.5 towards the end of this section.

1.1 Overview of proofs and techniques

We prove Theorem 1.2 by constructing a PCP (for an NP-complete language) that makes two queries to a proof: one query reads a symbol over an alphabet of size q and the other reads a symbol over an alphabet of size q^6. The PCP has completeness $1-\zeta$ and soundness at most $4/q$. As is standard in the PCP literature, the PCP is obtained by composing the so-called Outer PCP and Inner PCP.

Inner PCP The Inner PCP is essentially a probabilistic testing procedure that tests whether a given function $f : \mathbb{F}_q^m \to \mathbb{F}_q$ satisfies a desired property. Three types of tests have generally been used depending on the desired setting of parameters, e. g., the number of queries, type of the acceptance predicate, completeness and soundness, size of the proof, etc.

- The low degree test [14, 4, 5, 30] that tests whether f is a polynomial of low degree.

- The linearity test that tests whether f is linear (= Hadamard codeword) [3, 19].
- The dictatorship test that tests whether f is a dictatorship, i. e., a function of the form $f(x) = x_i$ for some coordinate $1 \le i \le m$, e. g., [6, 16, 17].

If f satisfies the desired property (i. e., being linear, low degree, or dictatorship), then the test passes with probability (close to) 1 and conversely, if the test passes with *reasonable* probability, f must have a non-trivial agreement with another function g that has the desired property. The low degree test has been analyzed algebraically [14, 4, 5] and also combinatorially [30] whereas the linearity and dictatorship tests are often amenable to Fourier analysis, e. g., [19, 16, 17, 20].

In this paper, we desire an explicit and *good* trade-off between the alphabet size of the PCP and the soundness parameter. At the Inner PCP level, this is achieved by a linearity test that combines the elements of the linearity test and the low degree test. Specifically, given a function $f : \mathbb{F}_q^m \to \mathbb{F}_q$, we wish to test that f is linear. The standard BLR test [8] checks whether $f(x+y) = f(x) + f(y)$ for randomly chosen inputs x and y. We however wish to have a 2-query test and hence we instead do a *point versus subspace* test. We are given the table of values of the function f and in addition, for every subspace $W \subseteq \mathbb{F}_q^m$ of dimension 6, a linear function $\mathcal{T}(W) : W \to \mathbb{F}_q$ that is supposed to be the restriction of f on W (denoted $f|_W$). The test selects a random 6-dimensional subspace W and a random point $w \in W$ and accepts if and only if $f(w) = \mathcal{T}(W)(w)$. Note that the query $f(w)$ is over an alphabet of size q and the query $\mathcal{T}(W)$ is over an alphabet of size q^6.

We achieve the following soundness guarantee: if f passes the *point versus subspace* test with probability $3/q$, then for some $j \in \mathbb{F}_q, j \neq 0$, the function $j \cdot f$ has a Fourier coefficient with magnitude at least $1/q^2$ (see Lemma 4.4). Interestingly, the test is analyzed by looking at the probability that f (and its restriction $f|_W$) passes the *Gowers Test* $f(x) - f(y) - f(z) + f(-x+y+z) = 0$ for randomly chosen x, y, z (see Section 4.1 for the reason the test is named so). The Gowers Test is considered for the purposes of the analysis only and is not a part of the actual test. The probability of passing the Gowers Test is related to the existence of a *large* Fourier coefficient (see Lemma 4.3) for function $j \cdot f$ for some $j \neq 0$.

The analysis proceeds as follows: assume that f passes the *point versus subspace* test with probability $1/q + \delta$ where $\delta \geq 2/q$. For the sake of simplicity, assume that for every subspace W, the test passes with probability $1/q + \delta$ after selecting W. Thus $f|_W$ has agreement $1/q + \delta$ with a linear function $\mathcal{T}(W)$. We show that this implies $j \cdot f|_W$ has a *large* Fourier coefficient for some $j \neq 0$ and hence $f|_W$ passes the Gowers Test with probability $1/q + \delta^4$. We observe that the probability that f passes the Gowers Test is the average (over W) of the probability that $f|_W$ passes the Gowers Test, up to an additive difference of $e = 3/q^4$. This implies that f passes the Gowers Test with probability $1/q + \delta^4 - e \geq 1/q + 2/q^4$. From this we conclude that for some $j \neq 0$, $j \cdot f$ has a *large* Fourier coefficient, with magnitude at least $1/q^2$.

Thus the Gowers Test serves as a vehicle to pass from the local linearity of f to its global linearity. This is also reminiscent of the *bootstrapping* method that allows the analysis of the low degree test in two (or three) dimensions to carry over to a higher number of dimensions.

Remark It is not clear that we necessarily need a *point versus ℓ-dimensional subspace* test with $\ell = 6$ to achieve a soundness of $O(1/q)$. Here is the limitation of our current analysis. In the Gowers Test on $f|_W$, $dim(W) = \ell$, the inputs x, y, z are linearly dependent with probability $\Theta(1/q^{\ell-2})$. On the other hand, when the Gowers Test is applied to the function f, the inputs x, y, z are linearly dependent with probability

$\Theta(1/q^{m-2})$ which is negligible. Hence we are able to claim that the probability that f passes the Gowers Test is the average (over W) of the probability that $f|_W$ passes the Gowers Test, but only up to an additive difference of $e = \Theta(1/q^{\ell-2})$. As the above calculation shows, we desire that $\delta = O(1/q)$ and that δ^4 dominates $e = \Theta(1/q^{\ell-2})$, which forces us to have $\ell \geq 6$.

Outer PCP, Composed PCP and Sub-Code Covering Property The linearity testing primitive at the Inner PCP level dictates that we use an Outer PCP based on a NP-hard problem with linear constraints. A natural choice is the 3LIN problem over $\mathbb{F}_q$: we are given an instance (X,Φ) where X is a set of variables taking values in $\mathbb{F}_q$ and Φ is a set of linear equations, each equation depending on three variables from X. The goal is to find an assignment $\sigma : X \to \Phi$ that satisfies a good fraction of the equations. A well-known result of Håstad [17] shows that for any constant $\eta > 0$, it is NP-hard to distinguish whether an instance (X,Φ) has an assignment that satisfies $1-\eta$ fraction of the equations (YES Case) or any assignment satisfies at most $1/q+\eta$ fraction of the equations (NO Case).

Starting with the hard instance of 3LIN as above, one builds a 2P1R Game as follows: the first prover is sent a set of k equations at random from Φ. Let V denote the set of $3k$ variables sent to the first prover. The second prover is sent a set $U \subseteq V$ that includes independently for $1 \leq i \leq k$, all three variables in the i^{th} equation with probability $1-\beta$ and exactly one of the three variables in the i^{th} equation with probability $\beta/3$ each (β will be tiny as explained later). The provers answer with assignments to V and U respectively and the verifier accepts if and only if the assignments are consistent on U and moreover all equations on V are satisfied. In the YES Case, the provers have a strategy to make the verifier accept with probability at least $1-k\eta$ whereas in the NO Case, any prover strategy makes the verifier accept with probability at most $2^{-\Omega(\beta k)}$ (see Section 3 for a proof).

The 2P1R Game described is precisely the so-called Outer PCP. The composition of the Outer and Inner PCP amounts to constructing a verifier that behaves as follows: the PCP verifier expects, for each question V (U resp.) to the first (second resp.) prover, a Hadamard encoding of assignment to V (U resp.). The Hadamard code is same as the table of values of a linear function $f_V : \mathbb{F}_q^V \to \mathbb{F}_q$ ($g_U : \mathbb{F}_q^U \to \mathbb{F}_q$ resp.) defined by an assignment to V (U resp.). Moreover, for every V, a table of linear functions on all 6-dimensional subspaces of $\mathbb{F}_q^V$ is expected; these linear functions are supposed to be the restrictions of the global linear function on $\mathbb{F}_q^V$. The verifier now picks a random question V to the first prover and performs the *point versus subspace* test on the table f_V and the corresponding subspaces table.

Note that it appears as if the Hadamard codes on U do not play any role (which would not make sense). We observe that the Hadamard code on U is actually contained in the Hadamard code on V (as a sub-code). This is because $\mathbb{F}_q^U \subseteq \mathbb{F}_q^V$ where we append a vector in $\mathbb{F}_q^U$ with zeroes at positions in $V \setminus U$ and think of it as a vector in $\mathbb{F}_q^V$. Thus there is no need to have a separate Hadamard code on U (though it helps in the analysis to think of these as virtual tables). Moreover if $U \subseteq V \cap V'$ for distinct questions V and V' to the first prover, we can identify the positions in Hadamard codes of V and V' that correspond to the same position in the (virtual) Hadamard code of U.[2]

Now we look carefully at the soundness analysis of the composed PCP. Assume on the contrary that the verifier accepts with probability $4/q$ and for simplicity that for every V, the verifier accepts

[2]The reader might be concerned here that the 3LIN instance has imperfect completeness and whether such identification can be done. This is a subtle and valid point: a short answer is that the identification can be done, but as mentioned in the next paragraph, the standard trick of *folding over equations* is no longer allowed and the *folding* needs to be implemented explicitly.

the *point versus subspace* test on the supposed Hadamard code $f : \mathbb{F}_q^V \to \mathbb{F}_q$ with probability $4/q$. The analysis of the Inner PCP guarantees that for some $j \neq 0$, $j \cdot f$ has a large Fourier coefficient. Thus the function f may be list decoded by making a list of all large Fourier coefficients. Since the sum of squared magnitudes of all Fourier coefficients is 1, the list size is bounded. Our test also incorporates *side-conditions* (see Section 4.3) and ensures that the Fourier coefficients obtained as list decoding satisfy all the equations on V (this is done in a more explicit manner than the standard *folding over equations* trick which is inapplicable in our setting).

Finally we want to infer consistency between f (i. e., supposed Hadamard code on V) and the (virtual) supposed Hadamard codes $g_U : \mathbb{F}_q^U \to \mathbb{F}_q$ on $U \subseteq V$. But as observed $g_U = f|_{\mathbb{F}_q^U}$. We would like to conclude that since $j \cdot f$ has a large Fourier coefficient, so do *many* of the $j \cdot g|_U$ functions. This turns out to be possible if the sub-code spaces $\mathbb{F}_q^U$ over all choices of U (weighted according to the distribution on U for a fixed V) cover the global space $\mathbb{F}_q^V$ almost uniformly. We term this as the *sub-code covering property*. When β is sufficiently small and k is sufficiently large, we note that

$$|U| \approx \left(1 - \frac{2}{3}\beta\right)|V|;$$

thus the size of a typical sub-code space is $q^{-O(\beta k)} = 2^{-O(\beta \log q \cdot k)}$ relative to the size of the code, there are roughly

$$\binom{k}{\beta k} \cdot 3^{\beta k} = 2^{\Omega(\beta \log(1/\beta) \cdot k)}$$

choices of U, and it is not unreasonable to expect that the covering property holds provided $\log(1/\beta) \gg \log q$. We formally prove this as Lemma 3.1.

Once we are able to infer the consistency of tables $f = f_V$ and g_U, as usual, the list decoding and then picking a random Fourier coefficient in the list yields a provers' strategy in the 2P1R Game. This yields a contradiction provided the soundness $2^{-\Omega(\beta k)}$ of the 2P1R Game is low enough, which follows if βk is large enough.

The Codeword Test and the Consistency Test In the PCP literature, the low degree test, linearity test and the dictatorship test at the Inner PCP level are often referred to as the *codeword test* and then the Inner/Outer PCP composition amounts to extending the test to the *consistency test* between two (or more as is necessary in some PCPs) supposed codewords, e. g., f_V and g_U as above. Often this composition presents serious technical challenges and one needs to carefully analyze each PCP construction by itself. Our paper shows how to translate the codeword test for Hadamard code (i. e., the linearity test), essentially in a black-box manner, to a consistency test. In fact the tables g_U are virtual and there is no separate consistency test. As described above, the codeword test for functions f_V automatically serves as a consistency test between f_V and g_U, since the entries in tables f_V and $f_{V'}$ that correspond to the virtual table g_U with $U \subseteq V \cap V'$ are identified together. We state this observation as an informal theorem.[3]

Theorem 1.5 (Informal)**.** *Suppose there is a linearity test for functions $f : \mathbb{F}_q^m \to \mathbb{F}_q$ with perfect completeness such that every function whose all Fourier coefficients are $o(1)$ in magnitude is accepted with*

[3]We remark that a similar informal theorem also holds for the dictatorship test modulo the Unique Games Conjecture, with the notion of Fourier coefficients replaced by influences of co-ordinates.

probability at most s. Then the test can be translated to a PCP with the same predicate, almost perfect completeness and soundness at most $s+o(1)$.

Our informal theorem applies in particular to the results of Samorodnitsky and Trevisan [31] and Engebretsen and Holmerin [12]. The first paper gives a linearity test over $\mathbb{F}_2^m$ with $k+\binom{k}{2}$ queries and soundness $2^{-\binom{k}{2}}$ which is translated into a PCP in the latter paper, albeit thinking of the test as a dictatorship test with addition of noise (the former paper is able to construct a PCP with slightly worse parameters). Our informal theorem leads to a PCP directly.

1.2 Related work

Our *point versus subspace* test is similar to the *point versus affine line* test [4, 5] and the *point versus affine plane* test [30]. These tests check whether a function has degree d where d is small but still super-constant. We are instead interested in the simplest case, i. e., $d=1$. The tests in [5, 30] have the following soundness guarantee: if a function $f:\mathbb{F}_q^m \to \mathbb{F}_q$ passes with probability δ and $q>\Omega((dm/\delta)^c)$, then f must have agreement at least $\delta^{c'}$ with some degree d polynomial for some positive integers c and c'. We could apply their analysis to the special case $d=1$. However we are interested in the explicit values of c and c' which are not specified in these papers and moreover we cannot afford the dependence on m. In principle, the values of c and c' may be computed by a rigorous examination of analysis therein and perhaps the dependence on m is not necessary. We however skip this arduous task and instead present a self-contained Fourier analysis of the test.

With a more careful (or different, perhaps algebraic or combinatorial along the lines of [5, 30]) analysis, it might be enough to have the *point versus* 2*-dimensional subspace* (or the *point versus affine line*) test instead of the *point versus* 6*-dimensional subspace* test. If so, this would give a PCP with q^2 answers and soundness $O(1/q)$. We do not consider this as the current focus of our paper and hence do not attempt it. Our focus is to demonstrate that it is possible to get an explicit and *good* trade-off between the answer size and the soundness parameter and present the Outer/Inner PCP composition based on the sub-code covering property. Constructing a PCP with q^t answers with $1<t<2$ and soundness $O(1/q)$ however seems much more challenging, if possible at all, and might require an entirely new approach. We note that Moshkovitz and Raz [25] do provide an analysis of the *point versus* 2*-dimensional subspace* test (Theorem 19 therein), in a rather similar way as ours, albeit using the BLR Test as an intermediate vehicle instead of the Gowers Test. The soundness they achieve is $O(1/q^{1/6})$, which would give exponent $\gamma=1/12$ in the trade-off between the alphabet size and the soundness of the test.

A recent result of Chan [9] improves Theorem 1.2 and obtains soundness of $O(\log(R)/\sqrt{R})$ for a 2P1R Game with R answers and almost-perfect completeness. (He does not get improved result for the Quadratic Programming Problem however, due to a blow-up in the instance size.) His techniques are entirely different and his result is more general. We do sketch his result since interestingly there is a (loose) connection to the *point versus affine line* test. His general result is that for any abelian group G and a *pairwise independent* subgroup $H\subseteq G^k$, the constraint satisfaction problem CSP(H) is *approximation resistant* in a very strong sense. This is a k-CSP where the variables take values over G, the set of satisfying assignments is H and *variable translations* are allowed (this is analogue of variable negation in the boolean case). He shows that it is NP-hard to distinguish whether the instance is nearly satisfiable or whether for any global assignment to the instance, the local assignment to a constraint, over a randomly chosen constraint, is nearly uniform in G^k (after accounting for the variable translations). Towards obtaining a result for 2P1R Games, he sets $G = \mathbb{F}_q$ and $H \subseteq \mathbb{F}_q^q$ to be the set of (value tables of) linear functions $f : \mathbb{F}_q \to \mathbb{F}_q$ (i. e., functions of the form $f(x) = ax + b$ and hence $|H| = q^2$). The CSP is then viewed as a variable versus constraint 2P1R Game which resembles the *point versus affine line* test for a linear function. The strong soundness guarantee of the CSP implies the soundness of the 2P1R Game. Roughly speaking, for any assignment to the CSP variables (i. e., the first prover's answers), the second prover cannot do significantly better than picking a random answer (i. e., a random assignment from H). The second prover has q^2 possible answers and for a random answer, his expected agreement with the first prover is $1/q$. The soundness of the 2P1R Game is $O(\log q/q)$: one loses a logarithmic factor since the second prover, instead of picking a random answer, can do slightly better by picking the best possible answer, i. e., answer from H that has the best agreement with the first prover's answers.

Finally, we remark that the focus of our results is incomparable to that of the Sliding Scale Conjecture of Bellare et al. [7] and the Projection Games Conjecture of Moshkovitz [24]. We focus on the trade-off between the soundness and the answer-size R of a 2P1R Game while thinking of R as a constant (or slowly growing with the instance size) whereas these conjectures focus on the case where the (inverse of) soundness and R are polynomial in the instance size.

2 Preliminaries

In this section, we briefly describe preliminary background and the tools used in this paper.

2.1 2-Prover 1-Round Games

Definition 2.1. A 2P1R Game $\mathcal{G}(\mathcal{V},\mathcal{U},\mu,\mathcal{R},\mathcal{S},\{\pi_{VU}\})$ consists of sets of questions $\mathcal{V},\mathcal{U}$ and sets of answers $\mathcal{R},\mathcal{S}$ for the two provers respectively, a distribution μ on the set of question pairs $\mathcal{V}\times\mathcal{U}$ and for every question pair (V,U) in the support of μ, a predicate $\pi_{VU} : \mathcal{R}\times\mathcal{S} \to \{0,1\}$ that defines the pairs of accepting answers. A strategy of the provers is a map $\phi : \mathcal{V} \to \mathcal{R}, \phi : \mathcal{U} \to \mathcal{S}$. The value of the strategy ϕ is:

$$\mathrm{val}(\phi,\mathcal{G}) := \Pr_{(V,U)\sim\mu}\left[\pi_{VU}(\phi(V),\phi(U)) = 1\right].$$

The value of the game $\mathrm{val}(\mathcal{G})$ is the maximum value of any prover strategy. A Projection Game is one where for every answer of the first prover, there is exactly one accepting answer of the second prover. For a Projection Game, the predicate π_{VU} can be thought of as a map $\pi_{VU} : \mathcal{R} \to \mathcal{S}$ and the accepting answers are of the form $(r, \pi_{VU}(r))$ for $r \in \mathcal{R}$. For a Projection Game, $|\mathcal{S}| \le |\mathcal{R}|$.

A 2P1R Game is best viewed as a game between the two provers and a verifier. The verifier picks a random question pair (V,U) from the distribution μ, asks one question each to the two provers respectively, and accepts if and only if the provers' answers satisfy the predicate π_{VU}. The probability of acceptance by the verifier is same as the value of the provers' strategy.

Definition 2.2. Given a 2P1R Game $\mathcal{G}(\mathcal{V},\mathcal{U},\mu,\mathcal{R},\mathcal{S},\{\pi_{VU}\})$, the k-wise repeated game is

$$\mathcal{G}^{\otimes k}(\mathcal{V}^k,\mathcal{U}^k,\mu^k,\mathcal{R}^k,\mathcal{S}^k,\{\pi^k_{V'U'}\}),$$

where for $V' = (V_1,\dots,V_k)$ and $U' = (U_1,\dots,U_k)$, $\pi^k_{V'U'} := \bigwedge_{i=1}^k \pi_{V_iU_i}$.

We state below Raz's Parallel Repetition Theorem along with the recent improvements (and simplifications) by Holenstein and Rao.

Theorem 2.3 ([29, 18, 28]). *There exists an absolute constant $c > 0$ such that for a 2P1R Game $\mathcal{G}$ with* $\mathrm{val}(\mathcal{G}) = 1 - \varepsilon$,

$$\mathrm{val}(\mathcal{G}^k) \leq (1 - \varepsilon^3)^{ck}.$$

For a Projection Game, the bound of $(1 - \varepsilon^2)^{ck}$ holds.

2.2 Hardness of 3LIN

Our reduction is from the 3LIN problem over a finite field. For the proof of Theorem 1.2, we use Håstad's well-known hardness result for 3LIN [17]. For the proof of Theorem 1.4, we need a hardness result for 3LIN with very good completeness and we use a result of Khot and Ponnuswami [21].

Definition 2.4. For a prime power q, an instance (X, Φ) of 3LIN consists of a set of variables X over $\mathbb{F}_q$ and a set of linear constraints Φ such that each constraint depends on exactly three variables. Let $\mathsf{OPT}(X, \Phi)$ denote the maximum fraction of constraints satisfied by any assignment.

Theorem 2.5 ([17]). *For every constant $\eta > 0$ and a prime q, it is NP-hard to distinguish whether a 3LIN instance (X, Φ) over $\mathbb{F}_q$ has $\mathsf{OPT}(X, \Phi) \geq 1 - \eta$ or $\mathsf{OPT}(X, \Phi) \leq 1/q + \eta$.*

Theorem 2.6 ([21]). *There is a reduction from a SAT instance Γ with size n to a 3LIN instance (X, Φ) with size N over a field $\mathbb{F}_q$ of characteristic 2 such that:*

- *N and the running time of the reduction are bounded by $2^{O(\log^2 n)}$.*
- *If Γ is satisfiable, then $\mathsf{OPT}(X, \Phi) \geq 1 - 2^{-\Omega(\sqrt{\log N})}$.*
- *If Γ is unsatisfiable, then $\mathsf{OPT}(X, \Phi) \leq 1 - \Omega\left(\frac{1}{\log^3 N}\right)$.*

Remark 2.7. Theorem 2.6 is stated in [21] over the field $\mathbb{F}_2$. It also holds for any extension field $\mathbb{F}_q$ since in the YES Case, the good assignment is over $\mathbb{F}_2$ and in the NO Case, any assignment over $\mathbb{F}_q$ can be mapped to an equally good assignment over $\mathbb{F}_2$, e. g., by taking the last bit of every $\mathbb{F}_q$-element.

Remark 2.8. We present the proof of Theorem 1.2 by constructing a PCP over the prime field $\mathbb{F}_q$ and the Fourier analysis is over this field. For the proof of Theorem 1.4, we use the same PCP but over a field of characteristic 2. We will omit the straightforward extension of our (Fourier) analysis to this case.

2.3 Hadamard code and Fourier analysis

Let q be a prime, $\omega := e^{2\pi i/q}$ be the complex q^{th} root of unity and $\Omega := \{1, \omega, \ldots, \omega^{q-1}\}$.

Definition 2.9. The Hadamard Code of $\alpha \in \mathbb{F}_q^n$ is defined as the table of values of the linear function $\chi_\alpha : \mathbb{F}_q^n \to \Omega$ where

$$\forall x \in \mathbb{F}_q^n, \qquad \chi_\alpha(x) = \omega^{\alpha \cdot x}.$$

The vector space of all functions $f : \mathbb{F}_q^n \to \mathbb{C}$ has an orthonormal basis $\{\chi_\alpha \mid \alpha \in \mathbb{F}_q^n\}$ where the inner product between two functions $f, g : \mathbb{F}_q^n \to \mathbb{C}$ is defined as (the expectation is over x uniformly distributed in $\mathbb{F}_q^n$)

$$\langle f, g\rangle := \mathop{\mathbb{E}}_{x}\left[f(x)\overline{g(x)}\right].$$

Hence, every $f : \mathbb{F}_q^n \to \mathbb{C}$ can be expressed uniquely as

$$f = \sum_{\alpha\in\mathbb{F}_q^n} \widehat{f}(\alpha)\,\chi_\alpha\,.$$

The coefficients $\widehat{f}(\alpha) \in \mathbb{C}$ are called Fourier coefficients. These are defined by:

$$\widehat{f}(\alpha) = \langle f, \chi_\alpha\rangle = \mathop{\mathbb{E}}_{x}\left[f(x)\overline{\chi_\alpha(x)}\right].$$

By Parseval's identity, $\sum_\alpha |\widehat{f}(\alpha)|^2 = \|f\|_2^2 = \mathbb{E}_x[|f(x)|^2]$. In particular, for a function taking values in Ω, the sum of squared absolute values of all its Fourier coefficients equals 1.

2.4 Quadratic Programming Problem

Definition 2.10. Given a real symmetric matrix $A = \{a_{ij}\}_{i=1}^n$ with zero diagonal entries, the Quadratic Programming Problem seeks to maximize $\sum_{i,j=1}^n a_{ij}x_ix_j$ where $\forall i, x_i \in \{-1,1\}$. Let $\mathsf{OPT}(A)$ denote this maximum (which is non-negative since the quadratic form is multi-linear and its maximum remains the same if the variables are allowed to be in $[-1,1]$ and then $\{\forall i, x_i = 0\}$ is a feasible solution).

It is important that the diagonal entries are zero as otherwise it is NP-hard to even decide whether there is a $\{-1,1\}$-valued assignment for which the quadratic form is non-negative. Given a graph $G(V = \{1,\ldots,n\}, E)$, the maximum of the quadratic form

$$\sum_{(i,j)\in E} \frac{1-x_ix_j}{2} - m$$

is non-negative if and only if the graph has a cut with m edges and finding a maximum cut is NP-hard.

Theorem 1.3 follows directly from Theorem 1.4 and reducing the hard instance of 2P1R Game to the Quadratic Programming Problem via Arora et al.'s reduction [2] below. Theorem 1.4 is proved using the same PCP used to prove Theorem 1.2, but we need to use Theorem 2.6 as a starting point of the PCP instead of Theorem 2.5.

Theorem 2.11 ([2]). *There is a reduction from a Projection Game $\mathcal{G}(\mathcal{V}, \mathcal{U}, \mu, \mathcal{R}, \mathcal{S}, \{\pi_{VU}\})$ to a Quadratic Programming Problem instance A such that:*

- *The running time of the reduction and the number of variables of the instance A is polynomial in $|\mathcal{V}| + |\mathcal{U}|$ and $2^{|\mathcal{R}|}$.*
- *If* $\mathrm{val}(\mathcal{G}) \geq 1-\eta$*, then* $\mathsf{OPT}(A) \geq 1 - \eta - \dfrac{1}{|\mathcal{V}|+|\mathcal{U}|}$.
- *If* $\mathrm{val}(\mathcal{G}) \leq \varepsilon$*, then* $\mathsf{OPT}(A) \leq O(\varepsilon)$.

2.5 Hellinger and statistical distance

The squared Hellinger distance between distributions D_1 and D_2 over a discrete probability space $\mathcal{A}$ is

$$H^2(D_1,D_2) := \frac{1}{2}\sum_{a\in\mathcal{A}}\left(\sqrt{D_1(a)}-\sqrt{D_2(a)}\right)^2 = 1-\sum_{a\in\mathcal{A}}\sqrt{D_1(a)D_2(a)}\,.$$

It is clear that $1-H^2(\cdot,\cdot)$ is multiplicative for product distributions D_1^k, D_2^k on space $\mathcal{A}^k$, i. e.,

$$1-H^2(D_1^k,D_2^k) = (1-H^2(D_1,D_2))^k\,.$$

The statistical distance between D_1 and D_2 is:

$$\Delta(D_1,D_2) := \frac{1}{2}\sum_{a\in\mathcal{A}}|D_1(a)-D_2(a)|\,.$$

We have the standard inequality:

Lemma 2.12 ([27])**.** $H^2(D_1,D_2) \le \Delta(D_1,D_2) \le \sqrt{2}\cdot H(D_1,D_2)\,.$

3 The Outer PCP

In this section, we describe our Outer PCP. For the ease of exposition, we do it in three stages, starting with the standard *Variable Versus Equation* Game. We also prove the *sub-code covering property*. Let (X,Φ) be an instance of the 3LIN problem over $\mathbb{F}_q$ given by Theorem 2.5 or Theorem 2.6.

3.1 The Variable Versus Equation Game

Consider the 2P1R Game where the verifier picks a random equation $E\in\Phi$ and then picks one of the three variables $x\in E$. The first prover is sent the question E (i. e., the three variables appearing in E) and the second prover is sent the question x. The provers answer with the $\mathbb{F}_q$-values of all the variables they receive. The verifier accepts if and only if the two provers agree on the variable x and moreover the values given by the first prover satisfy the equation E. It is well-known and easy to check that if $\mathsf{OPT}(X,\Phi) = 1-\varepsilon$, then the value of the game is $1-\varepsilon/3$.

3.2 The Basic 2P1R Game

We now slightly modify the Variable Versus Equation Game and call it the *Basic 2P1R Game* (think of β as small):

- The verifier picks an equation $E\in\Phi$ at random. Let $x,y,z\in X$ be the variables in E.
- The first prover is sent the equation E.
- The second prover is sent the equation E with probability $1-\beta$ and one of the three variables x,y,z with probability $\beta/3$ each.

- The provers answer with the values of all the variables they receive.
- The verifier accepts if and only if the two provers agree on the values of the variables and moreover the values given by the first prover satisfy the equation.

Assume that $\mathsf{OPT}(X,\Phi) = 1-\varepsilon$. The value of the game is at least $1-\varepsilon$ since the provers may stick to a $(1-\varepsilon)$-satisfying assignment to (X,Φ) and in that case, the verifier may reject only when the equation E is not satisfied.

On the other hand, the value of the game is at most $1-\Omega(\varepsilon\beta)$ since with probability β, the second prover receives exactly one variable and the variable versus equation game is played.

3.3 The Final 2P1R Game (Outer PCP)

The Outer PCP is now obtained as the k-wise repetition applied to the Basic 2P1R Game. Specifically:

- The verifier picks equations $E_1,\ldots,E_k \in \Phi$ at random. Let V denote the set of $3k$ variables appearing in these equations.
- The question V is sent to the first prover.
- Let $U \subseteq V$ be chosen by including independently for each $1 \le i \le k$, all three variables in equation E_i with probability $1-\beta$ and one of the three variables in equation E_i with probability $\beta/3$ each. The question U is sent to the second prover.
- The provers answer with the values of all the variables they receive.
- The verifier accepts if and only if the two provers agree on the values of the variables in U and moreover the values given by the first prover satisfy all the k equations $E_1,\ldots,E_k$.

Assume again that $\mathsf{OPT}(X,\Phi) = 1-\varepsilon$. The value of the game is at least $1-\varepsilon k$ since the provers may stick to a $(1-\varepsilon)$-satisfying assignment to (X,Φ) and in that case, the verifier may reject only when at least one equation E_i is not satisfied.

On the other hand, we observe that the value of the game is at most $(1-\Omega(\varepsilon^2))^{\Omega(\beta k)}$. As noted before, the value of the Basic Game is $1-\Omega(\varepsilon\beta)$ and if we apply the parallel repetition theorem (Theorem 2.3) directly, we end up with an upper bound of $(1-\Omega(\varepsilon^2\beta^2))^{\Omega(k)}$. This bound is not good enough for us and we get the better bound as follows. Call a coordinate useful if the second prover receives a single variable, i. e., the standard Variable Versus Equation Game is played. The expected number of useful coordinates is βk. Hence the probability that less than $\beta k/2$ coordinates are useful is at most $2^{-\Omega(\beta k)}$ and may be ignored in comparison to the desired bound. The repeated game restricted to the question pairs where at least $\beta k/2$ coordinates are useful can be thought of as a convex combination of sub-games, each sub-game being the standard Variable Versus Equation Game repeated at least $\beta k/2$ times. Each such sub-game has value at most $(1-\Omega(\varepsilon^2))^{\Omega(\beta k)}$ and hence so does the overall repeated game.

We note that the k equations chosen by the verifier in the first step are disjoint (i. e., do not share variables) with high probability and we make this assumption in all that follows.

3.4 Sub-code covering property

Fix a question $(E_1,\dots,E_k)$ to the first prover in the Outer PCP and let V denote the set of variables $\{x_1,\dots,x_{3k}\}$ in these equations. The question to the second prover is a subset $U \subseteq V$ where independently for each $1 \le i \le k$, all three variables $x_{3(i-1)+1}, x_{3(i-1)+2}, x_{3i}$ are included in U with probability $1-\beta$ and exactly one of the three variables is included with probability $\beta/3$ each. Consider two distributions on $\mathbb{F}_q^V = \mathbb{F}_q^{3k}$:

- Distribution $\mathcal{D}$ is uniform on $\mathbb{F}_q^V$.
- Distribution $\mathcal{D}'$ is obtained by picking a random question $U \subseteq V$ to the second prover (as described above), picking a string in $\mathbb{F}_q^U$ uniformly at random, and then "pulling up" this string to $\mathbb{F}_q^V$ by inserting 0 at positions in $V \setminus U$.

Lemma 3.1. *The statistical distance $\Delta(\mathcal{D},\mathcal{D}')$ is upper bounded by $O(\sqrt{k}\cdot q\beta)$.*

Proof. We will bound the squared Hellinger distance on one coordinate, then use the multiplicativity of the squared Hellinger distance for product distribution, and then upper bound the statistical distance in terms of the Hellinger distance.

Let $\mathcal{Q}$ denote the uniform distribution on $\mathbb{F}_q$ and $(\mathcal{Q},\mathcal{Q},\mathcal{Q})$ denote its three independent copies. On any single coordinate $1 \le j \le k$, note that $\mathcal{D}_j$ is same as the distribution $(\mathcal{Q},\mathcal{Q},\mathcal{Q})$ whereas $\mathcal{D}'_j$ can be written as:

$$\mathcal{D}'_j = (1-\beta)\cdot(\mathcal{Q},\mathcal{Q},\mathcal{Q}) + \frac{\beta}{3}\cdot(\mathcal{Q},0,0) + \frac{\beta}{3}\cdot(0,\mathcal{Q},0) + \frac{\beta}{3}\cdot(0,0,\mathcal{Q})\,.$$

The total probability mass attached by $\mathcal{D}_j$ to triples in $\mathbb{F}_q^3$ where the number of non-zero entries is three, two, one, and zero respectively is:

$$p_3 = \frac{(q-1)^3}{q^3},\ p_2 = \frac{3(q-1)^2}{q^3},\ p_1 = \frac{3(q-1)}{q^3},\ p_0 = \frac{1}{q^3}\,.$$

Similarly, the total probability mass attached by $\mathcal{D}'_j$ to triples in $\mathbb{F}_q^3$ where the number of non-zero entries is three, two, one, and zero respectively is:

$$p'_3 = (1-\beta)\frac{(q-1)^3}{q^3},\ p'_2 = (1-\beta)\frac{3(q-1)^2}{q^3},\ p'_1 = (1-\beta)\frac{3(q-1)}{q^3} + \beta\frac{q-1}{q},\ p'_0 = (1-\beta)\frac{1}{q^3} + \beta\frac{1}{q}\,.$$

Hence

$$\begin{aligned}
1 - H^2(\mathcal{D}_j,\mathcal{D}'_j) &= \sqrt{p_3p'_3} + \sqrt{p_2p'_2} + \sqrt{p_1p'_1} + \sqrt{p_0p'_0} \\
&= \frac{(q-1)^3}{q^3}\sqrt{1-\beta} + \frac{3(q-1)^2}{q^3}\sqrt{1-\beta} \\
&\qquad + \frac{3(q-1)}{q^3}\sqrt{1+\left(\frac{q^2}{3}-1\right)\beta} + \frac{1}{q^3}\sqrt{1+(q^2-1)\beta} \\
&\ge 1 - O(\beta^2q^2)\,,
\end{aligned}$$

where all the square roots are estimated using

$$\sqrt{1+x} \geq 1 + \frac{x}{2} - x^2$$

for $x \in [-1/2, 1/2]$ and $q^2\beta \ll 1/2$ (as will be the case). Thus the expression is lower bounded by $a_0 + a_1\beta + a_2\beta^2$ where it is easily checked that $a_0 = 1$ and $|a_2| = O(\beta^2 q^2)$. The main point is that the coefficient a_1 vanishes. Indeed, a_1 equals

$$\frac{1}{2}\left(-\frac{(q-1)^3}{q^3} - \frac{3(q-1)^2}{q^3} + \frac{3(q-1)}{q^3} \cdot \left(\frac{q^2}{3} - 1\right) + \frac{1}{q^3} \cdot (q^2 - 1)\right) = 0\,.$$

By multiplicativity of the squared Hellinger distance, we have

$$1 - H^2(\mathcal{D}, \mathcal{D}') \;=\; (1 - H^2(\mathcal{D}_j, \mathcal{D}'_j))^k \;\geq\; (1 - O(\beta^2 q^2))^k \;\geq\; 1 - O(\beta^2 q^2 k)\,.$$

Finally $\Delta(\mathcal{D}, \mathcal{D}') \leq \sqrt{2} \cdot H(\mathcal{D}, \mathcal{D}') \leq O(\sqrt{k} \cdot q\beta)$. □

3.5 The choice of parameters

Let $\mathsf{OPT}(X, \Phi) = 1 - \varepsilon$ where (X, Φ) is a NO instance of 3LIN given by Theorem 2.5 or Theorem 2.6.

- In Theorem 2.5, $\mathsf{OPT}(X, \Phi)$ is close to $1/q$ and hence $\varepsilon = \Omega(1)$.
- In Theorem 2.6, $\mathsf{OPT}(X, \Phi) \leq 1 - \Omega(1/\log^3 N)$ where N is the size of the 3LIN instance and hence $\varepsilon = \varepsilon(N) = \Omega(1/\log^3 N)$.

The parameters will be chosen so that for a large enough constant C,

- The soundness of the Outer PCP, which is at most $(1 - \Omega(\varepsilon^2))^{-\Omega(\beta k)}$, is at most $1/q^C$.
- $\Delta(\mathcal{D}, \mathcal{D}') \leq \dfrac{1}{Cq^2}$.

Using Lemma 3.1, it suffices to choose

$$k = \frac{C_*^3 q^6 \log^2 q}{\varepsilon^4} \quad \text{and} \quad \beta = \frac{\varepsilon^2}{C_*^2 \cdot q^6 \cdot \log q}$$

for a large enough constant C_*.

4 The Inner PCP

In this section, we describe our Inner PCP. We first analyze a test that we call Gowers Test, which is then used to analyze the actual Inner PCP presented in Section 4.3. We begin with some notation and a simple lemma.

Let $\omega := e^{2\pi i/q}$ be the complex q^{th} root of unity and $\Omega := \{1, \omega, \ldots, \omega^{q-1}\}$. For $f, g : \mathbb{F}_q^m \to \Omega$, let $\mathrm{AGR}(f, g)$ denote the agreement between the two functions, i. e., the fraction of points on which they agree. Let f_i denote the function $f_i(x) := f(x)^i$. Note that for $z \in \Omega$, the expression $(1 + z + z^2 + \cdots + z^{q-1})/q$ equals 1 if $z = 1$ and 0 otherwise.

Lemma 4.1. *If $f : \mathbb{F}_q^m \to \Omega$ is a function such that for some linear function χ_α, $\mathrm{AGR}(f, \chi_\alpha) \geq 1/q + \delta$. Then $\sum_{j=1}^{q-1} \left|\widehat{f}_j(j\alpha)\right| \geq q\delta$.*

Proof. The lemma follows by noting that:

$$\begin{aligned}
\mathrm{AGR}(f, \chi_\alpha) &= \mathop{\mathbb{E}}_x\left[\frac{1}{q}\sum_{j=0}^{q-1}(f(x)\overline{\chi_\alpha(x)})^j\right] \\
&= \frac{1}{q} + \frac{1}{q}\sum_{j=1}^{q-1}\mathop{\mathbb{E}}_x\left[f_j(x)\overline{\chi_{j\alpha}(x)}\right] \\
&= \frac{1}{q} + \frac{1}{q}\sum_{j=1}^{q-1}\widehat{f}_j(j\alpha)\,. \qquad \square
\end{aligned}$$

4.1 The Gowers test

For a function $f : \mathbb{F}_q^m \to \mathbb{C}$, the Gowers Uniformity Norm U_2 [15] is defined as

$$\|U_2(f)\|^4 := \mathop{\mathbb{E}}_{x,y,z}\left[f(x)\overline{f(x+y)f(x+z)}f(x+y+z)\right]\,.$$

We will study the probability that a function $f : \mathbb{F}_q^m \to \Omega$ passes the test

$$f(x)\overline{f(x+y)f(x+z)}f(x+y+z) = 1\,,$$

for a random choice of x, y, z. It is thus natural to name the test as the Gowers Test. It will be more convenient for us to think of the test equivalently as

$$\text{Gowers Test :} \qquad f(x)\overline{f(y)f(z)}f(-x+y+z) = 1\,.$$

Lemma 4.2. *Let $f : \mathbb{F}_q^m \to \Omega$ be a function. Then the acceptance probability of the Gowers Test is:*

$$\Pr_{x,y,z}\left[f(x)\overline{f(y)f(z)}f(-x+y+z) = 1\right] = \frac{1}{q} + \frac{1}{q}\sum_{j=1}^{q-1}\sum_\alpha\left|\widehat{f}_j(j\alpha)\right|^4\,.$$

Proof. The acceptance probability can be expressed as

$$\begin{aligned}
&\frac{1}{q} + \frac{1}{q}\sum_{j=1}^{q-1}\mathbb{E}\left[(f(x)\overline{f(y)f(z)}f(-x+y+z))^j\right] \\
&= \frac{1}{q} + \frac{1}{q}\sum_{j=1}^{q-1}\sum_{\alpha,\phi,\psi,\gamma}\widehat{f}_j(\alpha)\overline{\widehat{f}_j(\phi)\widehat{f}_j(\psi)}\widehat{f}_j(\gamma)\,\mathbb{E}\left[\chi_{\alpha-\gamma}(x)\chi_{-\phi+\gamma}(y)\chi_{-\psi+\gamma}(z)\right] \\
&= \frac{1}{q} + \frac{1}{q}\sum_{j=1}^{q-1}\sum_\alpha\left|\widehat{f}_j(\alpha)\right|^4,
\end{aligned}$$

noting that the expectation vanishes unless $\alpha = \phi = \psi = \gamma$. We can replace α by $j\alpha$ without changing the summation. $\square$

4.2 The Gowers test with side conditions

At the Inner PCP level, we need to check not only that a function f is linear, i. e., $f=\chi_\alpha$ for some α, but also that α itself satisfies a given set of linear constraints. We call these as *side conditions* and modify the Gowers Test so as to incorporate these side conditions (and henceforth Gowers Test refers to one with side conditions incorporated).

Given:

- A function $f:\mathbb{F}_q^m\to\Omega$.
- Side conditions $h_i\cdot x=b_i, i=1,\dots,k$ where $h_i\in\mathbb{F}_q^m,\ b_i\in\mathbb{F}_q$. Assume that $\{h_i\}_{i=1}^k$ are linearly independent and H is their linear span.

The Test:

- Pick $x,y,z\in\mathbb{F}_q^m$ at random.
- Pick $a=(a_1,\dots,a_k)\in\mathbb{F}_q^k$ at random and let $h=\sum_{i=1}^k a_ih_i$ and $a\cdot b:=\sum_{i=1}^k a_ib_i$.
- Accept if and only if
$$f(x)\overline{f(y)f(z)}f(-x+y+z+h)=\omega^{a\cdot b}.$$

Lemma 4.3. *The following hold:*

1. *The Gowers Test always passes with probability at least $1/q$.*
2. *If the Gowers Test passes with probability at least $1/q+\delta$, then there exists $1\le j\le q-1$ and $\alpha\in\mathbb{F}_q^m$ that respects the side conditions (i. e., $\forall i,\alpha\cdot h_i=b_i$) and $|\widehat{f_j}(j\alpha)|\ge\sqrt{\delta}$.*
3. *If f has agreement $1/q+\delta$ with a linear function χ_α that respects the side conditions (i. e., $\forall i,\alpha\cdot h_i=b_i$), then f passes the Gowers Test with probability at least $1/q+\delta^4$.*

Proof. The acceptance probability of the Gowers Test with side conditions is:

$$\begin{aligned}
&\frac{1}{q}+\frac{1}{q}\sum_{j=1}^{q-1}\mathbb{E}\left[(f(x)\overline{f(y)f(z)}f(-x+y+z+h))^j\omega^{-ja\cdot b}\right]\\
&=\frac{1}{q}+\frac{1}{q}\sum_{j=1}^{q-1}\sum_{\alpha,\phi,\psi,\gamma}\widehat{f_j}(\alpha)\overline{\widehat{f_j}(\phi)\widehat{f_j}(\psi)}\widehat{f_j}(\gamma)\,\mathbb{E}\left[\chi_{\alpha-\gamma}(x)\chi_{-\phi+\gamma}(y)\chi_{-\psi+\gamma}(z)\right]\mathop{\mathbb{E}}_a\left[\chi_\gamma(h)\omega^{-ja\cdot b}\right]\\
&=\frac{1}{q}+\frac{1}{q}\sum_{j=1}^{q-1}\sum_{\alpha}\left|\widehat{f_j}(\alpha)\right|^4\cdot\mathop{\mathbb{E}}_a\left[\omega^{\sum_{i=1}^k a_i(\alpha\cdot h_i-jb_i)}\right]\\
&=\frac{1}{q}+\frac{1}{q}\sum_{j=1}^{q-1}\sum_{\alpha|\forall i,\alpha\cdot h_i=jb_i}\left|\widehat{f_j}(\alpha)\right|^4\\
&=\frac{1}{q}+\frac{1}{q}\sum_{j=1}^{q-1}\sum_{\alpha|\forall i,\alpha\cdot h_i=b_i}\left|\widehat{f_j}(j\alpha)\right|^4.
\end{aligned}$$

The first conclusion is immediate. The second uses the fact that the sum of squared absolute values of Fourier coefficients of f_j is 1. The third follows in conjunction with Lemma 4.1 and convexity of the function x^4. $\square$

4.3 The Point-Subspace test with side conditions (Inner PCP)

Given:

- A function $f : \mathbb{F}_q^n \to \Omega$.
- Side conditions $h_i \cdot x = b_i, i = 1, \ldots, k$. Assume that $\{h_i\}_{i=1}^k$ are linearly independent and H is their linear span.
- A table $\{\mathcal{T}(W) \mid W \in \mathcal{C}\}$ where $\mathcal{C}$ denotes the class of $k+6$ dimensional subspaces of the form $W = D \oplus H$ for a 6-dimensional subspace D of $\mathbb{F}_q^n$ such that $D \cap H = \{\mathbf{0}\}$. The entry $\mathcal{T}(W)$ is a linear function $\chi : W \to \Omega$ that respects the side conditions, i. e.,

$$\chi(x+y) = \chi(x) \cdot \chi(y) \qquad \text{and} \qquad \chi(x+h_i) = \chi(x) \cdot \omega^{b_i} \quad \forall x, y \in W, i = 1, \ldots, k\,.$$

The Test:

1. Pick a random $W \in \mathcal{C}$ and let $\mathcal{T}(W)$ be the linear function on W.
2. Pick a random $w \in W$.
3. Accept if and only if $f(w) = \mathcal{T}(W)(w)$.

4.3.1 Completeness

Suppose $f : \mathbb{F}_q^n \to \Omega$ is linear that respects the side conditions, i. e.,

$$f(x+y) = f(x) \cdot f(y) \qquad \text{and} \qquad f(x+h_i) = f(x) \cdot \omega^{b_i} \quad \forall x, y \in \mathbb{F}_q^n, i = 1, \ldots, k\,.$$

Then letting $\mathcal{T}(W)$ to be the restriction $f|_W$, the point-subspace test passes with probability 1.

4.3.2 Soundness

Lemma 4.4. *If the point-subspace test passes with probability $3/q$ then there exists $1 \le j \le q-1$ and $\alpha \in \mathbb{F}_q^n$ that respects the side conditions (i. e., $\forall i, \alpha \cdot h_i = b_i$) and $|\widehat{f_j}(j\alpha)| \ge 1/q^2$.*

Proof. Assume that the point-subspace test passes with probability $1/q + \delta$ where $\delta \ge 2/q$. This means that:

$$\mathop{\mathbb{E}}_W [\mathrm{AGR}\,(f|_W, \mathcal{T}(W))] = 1/q + \delta\,.$$

By Lemma 4.3(1,3) and using convexity of the function x^4, we conclude that

$$\mathop{\mathbb{E}}_W [\Pr[f|_W \text{ passes Gowers Test}]] \ge 1/q + \delta^4. \tag{4.1}$$

Now let us consider the probability that f passes the Gowers Test. The Gowers Test picks three points x,y,z independently (from the global space $\mathbb{F}_q^n$). Let $\perp$ be the event that $\operatorname{span}(x,y,z,H)$ has dimension $k+3$ and let Γ be the set of all such triples (x,y,z). Conditional on $\perp$ happening, the Gowers Test picks a triple in Γ uniformly at random. An alternate way of picking a triple in Γ uniformly at random is to pick a subspace $W \in \mathcal{C}$ and then pick $(x,y,z) \in W^3$ conditional on the event that $\operatorname{span}(x,y,z,H)$ has dimension $k+3$. Let $\perp(W)$ be the event that $\operatorname{span}(x,y,z,H)$ has dimension $k+3$ when $x,y,z \in W$ are picked at random. Thus:

$$\begin{aligned}
\Pr[f \text{ passes Gowers Test} \mid \perp] &= \mathop{\mathbb{E}}_{W}\big[\Pr[f|_W \text{ passes Gowers Test} \mid \perp(W)]\big] \\
&\geq \mathop{\mathbb{E}}_{W}\big[\Pr[f|_W \text{ passes Gowers Test}] - \Pr[\neg\perp(W)]\big] \\
&\geq \mathop{\mathbb{E}}_{W}\big[\Pr[f|_W \text{ passes Gowers Test}]\big] - \frac{3}{q^4} \\
&\geq \frac{1}{q} + \delta^4 - \frac{3}{q^4} \;\geq\; \frac{1}{q} + \frac{2}{q^4},
\end{aligned}$$

where we used $\Pr[\neg\perp(W)] \leq 3/q^4$ and $\delta \geq 2/q$. Also, noting that

$$\Pr[\perp] \geq 1 - \frac{3}{q^{n-k-2}},$$

we get that

$$\Pr[f \text{ passes Gowers Test}] \geq \frac{1}{q} + \frac{1}{q^4}.$$

It follows now from Lemma 4.3(2) that there exists $1 \leq j \leq q-1$ and α such that $\forall i, \alpha \cdot h_i = b_i$ and $|\widehat{f_j}(j\alpha)| \geq 1/q^2$. □

5 The composed PCP

We now describe the composed PCP and prove Theorem 1.2. The Outer PCP is constructed from a 3LIN instance (X,Φ) as described in Section 3. The 3LIN instance is either $(1-\eta)$-satisfiable or at most $(1-\varepsilon)$-satisfiable as per Theorem 2.5 or Theorem 2.6. The various parameters are chosen as in Section 3.5.

The verifier in the composed PCP expects the proof to contain, for every question V to the first prover, two tables $\mathcal{L}_V$ and $\mathcal{T}_V$.

The tables $\mathcal{L}_V$ The table $\mathcal{L}_V$ gives the Hadamard code of the assignment to V. Concretely, let $\{x_1,\ldots,x_{3k}\}$ be the variables in V and $\sigma : X \to \mathbb{F}_q$ be the global assignment that is supposed to be an almost satisfying assignment to (X,Φ). The verifier expects, for the question V, the table of values of the linear function $\mathcal{L}_V : \mathbb{F}_q^V = \mathbb{F}_q^{3k} \to \Omega$ where

$$\mathcal{L}_V(y) = \omega^{\sum_{i=1}^{3k} y_i \sigma(x_i)}.$$

Note that for every question U to the second prover, a table $\mathcal{L}_U$ that gives the Hadamard code of the assignment to U may be expected as well. However if $U \subseteq V$, then the table $\mathcal{L}_U$ is contained in the table $\mathcal{L}_V$. Specifically, for any $z \in \mathbb{F}_q^U$, let $z^{\uparrow}$ denote the vector obtained by extending z to a vector in $\mathbb{F}_q^V$ be inserting zeroes at the positions in $V \setminus U$. Then

$$\mathcal{L}_U(z) = \omega^{\sum_{i:x_i \in U} z_i \sigma(x_i)} = \omega^{\sum_{i=1}^{3k} z_i^{\uparrow} \sigma(x_i)} = \mathcal{L}_V(z^{\uparrow})\,.$$

Thus there is no need to have separate $\mathcal{L}_U$ tables; we do however think of these as virtual tables. Whenever there are questions V, V' to the first prover such that $U \subseteq V \cap V'$, the table $\mathcal{L}_U$ is contained in both the tables $\mathcal{L}_V$ and $\mathcal{L}_{V'}$. We identify the locations in these two tables that correspond to the same location in the (virtual) $\mathcal{L}_U$ table.

The tables $\mathcal{T}_V$ Let V be a question to the first prover that consists of equations (i. e., side conditions) $\{h_i \cdot x = b_i\}_{i=1}^k$, the i^{th} equation depending only on the variables $(x_{3(i-1)+1}, x_{3(i-1)+2}, x_{3i})$. Except with a negligible probability (that we may ignore) over the choice of V, the k equations are disjoint (i. e., do not share variables). Thus the vectors h_i are linearly independent and let H be their linear span. The table $\mathcal{T}_V$ contains, for every $k+6$ dimensional subspace $W \subseteq \mathbb{F}_q^V$ such that $H \subseteq W$, a linear function $\mathcal{T}_V(W) : W \to \Omega$ that respects the side conditions.

The PCP verifier The verifier picks a random question V to the first prover in the Outer PCP. Let $\{h_i \cdot x = b_i\}_{i=1}^k$ be the side conditions and H be the linear span of vectors $\{h_i\}_{i=1}^k$. Let $\mathcal{L}_V : \mathbb{F}_q^V \to \Omega$ and $\mathcal{T}_V$ be the associated tables. The verifier picks a random $k+6$ dimensional subspace W, $H \subseteq W \subseteq \mathbb{F}_q^V$, and a random point $w \in W$. The verifier accepts if and only if

$$\mathcal{L}_V(w) = \mathcal{T}_V(W)(w)\,.$$

5.1 Completeness

Let $\sigma : X \to \mathbb{F}_q$ be a global assignment that satisfies $1-\eta$ fraction of equations. The table $\mathcal{L}_V$ is the Hadamard code (i. e., linear function) of the assignment σ restricted to V. The table $\mathcal{T}_V$ gives, for every subspace W, the linear function $\mathcal{T}_V(W) = \mathcal{L}_V|_W$. The test may fail only when there is some equation in V that is not satisfied by σ. This happens with probability at most ηk.

5.2 Soundness

Assume on the contrary that the test accepts with probability $4/q$. By an averaging argument, for at least $1/q$ fraction of the questions V, the test accepts with probability at least $3/q$. Fix any such *good* V. By the analysis of the Inner PCP, Lemma 4.4, it follows that there exists $1 \le j \le q-1$ such that for $f := \mathcal{L}_V$, there is a Fourier coefficient $|\widehat{f_j}(j\alpha)| \ge 1/q^2$ and α respects the side conditions. Note that for the uniform distribution $\mathcal{D}$ on $\mathbb{F}_q^V$,

$$\widehat{f_j}(j\alpha) = \mathop{\mathbb{E}}_{x \in \mathcal{D}}\left[f_j(x)\overline{\chi_{j\alpha}(x)}\right]\,.$$

By the sub-code covering property, Section 3.5, it follows that for some error parameter $e, |e| \leq 1/(Cq^2)$,

$$
\begin{aligned}
\widehat{f_j}(j\alpha) &= \mathop{\mathbb{E}}_{x\in\mathcal{D}'}\left[f_j(x)\overline{\chi_{j\alpha}(x)}\right] + e \\
&= \mathop{\mathbb{E}}_{U}\left[\mathop{\mathbb{E}}_{y\in\mathbb{F}_q^U}\left[f_j(y^{\uparrow})\overline{\chi_{j\alpha}(y^{\uparrow})}\right]\right] + e \\
&= \mathop{\mathbb{E}}_{U}\left[\mathop{\mathbb{E}}_{y\in\mathbb{F}_q^U}\left[f_j|_U(y)\overline{\chi_{j\alpha_{\downarrow}}(y)}\right]\right] + e \\
&= \mathop{\mathbb{E}}_{U}\left[\widehat{f_j|_U}(j\alpha_{\downarrow})\right] + e\,,
\end{aligned}
$$

where $f_j|_U$ denotes the restriction of f_j to $\mathbb{F}_q^U \subseteq \mathbb{F}_q^V$ and $\alpha_{\downarrow}$ denotes the vector obtaining by dropping coordinates of α in $V \setminus U$. It follows that

$$
\mathop{\mathbb{E}}_{U}\left[\,\left|\widehat{f_j|_U}(j\alpha_{\downarrow})\right|\right] \geq |\widehat{f_j}(j\alpha)| - |e| \geq \frac{1}{q^2} - \frac{1}{Cq^2} \geq \frac{1}{2q^2}\,.
$$

Thus, with probability at least $1/(4q^2)$ over the choice of U (for the fixed good V; call such (V,U) a *good* pair),

$$
\left|\widehat{f_j|_U}(j\alpha_{\downarrow})\right| \geq \frac{1}{4q^2}\,.
$$

Note also that $f|_U = g = \mathcal{L}_U$ which is the supposed (virtual) Hadamard code for an assignment to U and $f_j|_U = g_j$.

Now we derive strategies for the provers in the Outer PCP as follows: the first prover, on receiving a question V, considers the function $f = \mathcal{L}_V$, lists all α such that for some $1 \leq j \leq q-1$, $|\widehat{f_j}(j\alpha)| \geq 1/q^2$ and α respects the side conditions, and then outputs a random element from the list. The list size is bounded by q^5. The second prover, on receiving a question U, considers the function $g = \mathcal{L}_U$, lists all γ such that for some $1 \leq j \leq q-1$, $|\widehat{g_j}(j\gamma)| \geq 1/(4q^2)$, and outputs a random element from the list. The list size is bounded by $16q^5$. The analysis above shows that when (V,U) is a good pair, there are α and $\alpha_{\downarrow} = \gamma$ in the lists for V and U respectively and α respects the side conditions. This and the bound on the list sizes shows that with probability at least

$$
\frac{1}{q}\cdot\frac{1}{4q^2}\cdot\frac{1}{q^5}\cdot\frac{1}{16q^5}\,,
$$

the provers succeed. This is a contradiction since the soundness of the Outer PCP is at most $1/q^C$ for a large constant C as in Section 3.5.

6 Proofs of Theorem 1.4 and Theorem 1.3

Now we prove Theorem 1.4. It directly implies Theorem 1.3 using the reduction from 2P1R Game to Quadratic Programming Problem as in Theorem 2.11 (one sets $q = (\log n)^{O(1/\varepsilon)}$).

To prove Theorem 1.4, we use the same composed PCP as in Section 5, except that the 3LIN instance from Theorem 2.6 is used as a starting point. A minor issue is that we presented our entire PCP for a prime field $\mathbb{F}_q$ whereas now we need a PCP over a field of characteristic 2. We omit the straightforward extension of our analysis to this case. Along with the setting of parameters in Section 3.5, we get a reduction from a 3LIN instance (with the field $\mathbb{F}_q$ of characteristic 2) to a 2P1R Game $\mathcal{G}$ such that:

- The provers have q^6 answers.
- The size of the 2P1R Game is at most $q^{O(k)}N^k$ if the 3LIN instance has size (i. e., number of equations) N.
- If the 3LIN instance is $(1-\eta)$-satisfiable, then the value of the game is at least $1-\eta k$.
- If the 3LIN instance is at most $(1-\varepsilon)$-satisfiable, then the value of the game is at most $4/q$.

The 3LIN instance from Theorem 2.6 has instance size $N = 2^{O(\log^2 n)}$ where n is the size of the starting SAT instance. We have

$$\eta = \eta(N) \leq 2^{-\Omega(\sqrt{\log N})} \quad \text{and} \quad \varepsilon = \varepsilon(N) \geq \Omega\left(\frac{1}{\log^3 N}\right).$$

As noted in Section 3.5, it suffices to have

$$k = \frac{C_*^3 q^6 \log^2 q}{\varepsilon^4} \quad \text{and} \quad \beta = \frac{\varepsilon^2}{C_*^2 \cdot q^6 \cdot \log q}.$$

The soundness parameter of the 2P1R Game is $4/q$ as required. The completeness parameter is $1-\eta k$ with $k = O(q^6 \log^2 q \log^{24} n)$ and $\eta = n^{-\Omega(1)}$. The size of the 2P1R Game is at most $q^{O(k)}N^k$ which is $2^{O((\log^2 n + \log q)\log^{24} n \cdot q^6 \log^2 q)}$ as stated.

7 Acknowledgement

We thank Dana Moshkovitz and Ran Raz for a discussion about the sub-code covering property (in a different context). Thanks to both for pointing us to their analysis of the point versus 2-dimensional subspace test. We also thank Preyas Popat for pointing out a flaw in an earlier version of the paper. Many thanks to anonymous referees for suggesting several improvements to the presentation.

References

[1] SANJEEV ARORA, LÁSZLÓ BABAI, JACQUES STERN, AND ELIZABETH Z. SWEEDYK: The hardness of approximate optima in lattices, codes, and systems of linear equations. *J. Comput. System Sci.*, 54(2):317–331, 1997. Preliminary version in FOCS'93. [doi:10.1006/jcss.1997.1472] 864

[2] SANJEEV ARORA, ELI BERGER, ELAD HAZAN, GUY KINDLER, AND MULI SAFRA: On non-approximability for quadratic programs. In *Proc. 46th FOCS*, pp. 206–215. IEEE Comp. Soc. Press, 2005. See also at ECCC. [doi:10.1109/SFCS.2005.57] 865, 872

[3] SANJEEV ARORA, CARSTEN LUND, RAJEEV MOTWANI, MADHU SUDAN, AND MARIO SZEGEDY: Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998. Preliminary version in FOCS'92. See also at ECCC. [doi:10.1145/278298.278306] 864, 866

[4] SANJEEV ARORA AND SHMUEL SAFRA: Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998. Preliminary version in FOCS'92. [doi:10.1145/273865.273901] 864, 865, 866, 869

[5] SANJEEV ARORA AND MADHU SUDAN: Improved low-degree testing and its applications. *Combinatorica*, 23(3):365–426, 2003. Preliminary version in STOC'97. [doi:10.1007/s00493-003-0025-0] 865, 866, 869

[6] MIHIR BELLARE, ODED GOLDREICH, AND MADHU SUDAN: Free bits, PCPs, and nonapproximability—towards tight results. *SIAM J. Comput.*, 27(3):804–915, 1998. Preliminary version in FOCS'95. See also at ECCC. [doi:10.1137/S0097539796302531] 864, 866

[7] MIHIR BELLARE, SHAFI GOLDWASSER, CARSTEN LUND, AND A. RUSSELL: Efficient probabilistically checkable proofs and applications to approximations. In *Proc. 25th STOC*, pp. 294–304. ACM Press, 1993. See correction to abstract in STOC'94. [doi:10.1145/167088.167174] 870

[8] MANUEL BLUM, MICHAEL LUBY, AND RONITT RUBINFELD: Self-testing/correcting with applications to numerical problems. *J. Comput. System Sci.*, 47(3):549–595, 1993. Preliminary version in STOC'90. [doi:10.1016/0022-0000(93)90044-W] 866

[9] SIU ON CHAN: Approximation resistance from pairwise independent subgroups. In *Proc. 45th STOC*, pp. 447–456. ACM Press, 2013. See also at ECCC. [doi:10.1145/2488608.2488665] 865, 869

[10] MOSES CHARIKAR AND ANTHONY WIRTH: Maximizing quadratic programs: extending Grothendieck's inequality. In *Proc. 45th FOCS*, pp. 54–60. IEEE Comp. Soc. Press, 2004. [doi:10.1109/FOCS.2004.39] 865

[11] ANNE CONDON: The complexity of the max word problem and the power of one-way interactive proof systems. *Comput. Complexity*, 3(3):292–305, 1993. Preliminary version in STACS'91. [doi:10.1007/BF01271372] 864

[12] LARS ENGEBRETSEN AND JONAS HOLMERIN: More efficient queries in PCPs for NP and improved approximation hardness of maximum CSP. *Random Structures & Algorithms*, 33(4):497–514, 2008. Preliminary version in STACS'05. [doi:10.1002/rsa.20226] 869

[13] URIEL FEIGE: A threshold of $\ln n$ for approximating set cover. *J. ACM*, 45(4):634–652, 1998. Preliminary version in STOC'96. [doi:10.1145/285055.285059] 864

[14] URIEL FEIGE, SHAFI GOLDWASSER, LÁSZLÓ LOVÁSZ, SHMUEL SAFRA, AND MARIO SZEGEDY: Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996. Preliminary version in FOCS'91. [doi:10.1145/226643.226652] 864, 865, 866

[15] TIMOTHY GOWERS: A new proof of Szemerédi's theorem for progressions of length four. *Geometric and Functional Analysis*, 8(3):529–551, 1998. [doi:10.1007/s000390050065] 877

[16] JOHAN HÅSTAD: Clique is hard to approximate within $n^{1-\varepsilon}$. *Acta Mathematica*, 182(1):105–142, 1999. Preliminary version in FOCS'96. [doi:10.1007/BF02392825] 864, 866

[17] JOHAN HÅSTAD: Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. Preliminary version in STOC'97. [doi:10.1145/502090.502098] 864, 866, 867, 871

[18] THOMAS HOLENSTEIN: Parallel repetition: Simplification and the no-signaling case. *Theory of Computing*, 5(8):141–172, 2009. Preliminary version in STOC'07. [doi:10.4086/toc.2009.v005a008] 864, 871

[19] SUBHASH KHOT: Improved inapproximability results for MaxClique, chromatic number and approximate graph coloring. In *Proc. 42nd FOCS*, pp. 600–609. IEEE Comp. Soc. Press, 2001. [doi:10.1109/SFCS.2001.959936] 866

[20] SUBHASH KHOT, GUY KINDLER, ELCHANAN MOSSEL, AND RYAN O'DONNELL: Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007. Preliminary version in FOCS'04. See also at ECCC. [doi:10.1137/S0097539705447372] 866

[21] SUBHASH KHOT AND ASHOK KUMAR PONNUSWAMI: Better inapproximability results for maxclique, chromatic number and Min-3Lin-Deletion. In *Proc. 33th Internat. Colloq. on Automata, Languages and Programming (ICALP'06)*, pp. 226–237. Springer, 2006. [doi:10.1007/11786986_21] 871

[22] SUBHASH KHOT AND MULI SAFRA: A two prover one round game with strong soundness. In *Proc. 52nd FOCS*, pp. 648–657. IEEE Comp. Soc. Press, 2011. [doi:10.1109/FOCS.2011.62] 863

[23] ALEXANDRE MEGRETSKI: Relaxation of quadratic programs in operator theory and system analysis. In *Proc. International Workshop on Operator Theory and Applications (IWOTA'00)*, pp. 365–392. Springer, 2000. [doi:10.1007/978-3-0348-8362-7_15] 865

[24] DANA MOSHKOVITZ: The projection games conjecture and the NP-hardness of $\ln n$-approximating set-cover. In *Proc. 15th Internat. Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'12)*, pp. 276–287. Springer, 2012. See also at ECCC. [doi:10.1007/978-3-642-32512-0_24] 870

[25] DANA MOSHKOVITZ AND RAN RAZ: Two-query PCP with subconstant error. *J. ACM*, 57(5):29:1–29:29, 2010. Preliminary version in FOCS'08. See also at ECCC. [doi:10.1145/1754399.1754402] 869

[26] ARKADI NEMIROVSKI, KEES ROOS, AND TAMÁS TERLAKY: On maximization of quadratic form over intersection of ellipsoids with common center. *Mathematical Programming*, 86(3):463–473, 1999. [doi:10.1007/s101070050100] 865

[27] DAVID POLLARD: *A user's guide to measure theoretic probability*. Cambridge Univ. Press, 2002. CUP. 873

[28] ANUP RAO: Parallel repetition in projection games and a concentration bound. *SIAM J. Comput.*, 40(6):1871–1891, 2011. Preliminary version in STOC'08. [doi:10.1137/080734042] 864, 871

[29] RAN RAZ: A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. Preliminary version in STOC'95. [doi:10.1137/S0097539795280895] 864, 871

[30] RAN RAZ AND SHMUEL SAFRA: A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proc. 29th STOC*, pp. 475–484. ACM Press, 1997. [doi:10.1145/258533.258641] 865, 866, 869

[31] ALEX SAMORODNITSKY AND LUCA TREVISAN: A PCP characterization of NP with optimal amortized query complexity. In *Proc. 32nd STOC*, pp. 191–199. ACM Press, 2000. [doi:10.1145/335305.335329] 869

AUTHORS

Subhash Khot
Professor
New York University, New York, NY
khot@cs.nyu.edu
http://www.cs.nyu.edu/~khot

Muli Safra
Professor
Tel Aviv University, Tel Aviv, Israel
safra@tau.ac.il
http://www.tau.ac.il/~safra

ABOUT THE AUTHORS

SUBHASH KHOT is a Professor in the Computer Science Department at New York University. He has a Ph. D. from Princeton University, completed in 2003 under the supervision of Sanjeev Arora.

His teacher and mentor at Vyankatrao Highschool, Mr. Vaman G. Gogate played a decisive role in directing his attention to mathematics. If not for Mr. Gogate's guidance and some fortuitous turn of events, the chance of someone from the remote town of Ichalkaranji pursuing mathematical research was strictly nil. Subsequently, Mr. Gogate also mentored Amit Deshpande and Raghav Kulkarni (CS theorists) and Abhijit Gadde (string theorist). He is in touch with all his past students and continues to provide guidance on all aspects of life.

MULI SAFRA is a Professor of Computer Science at Tel Aviv University. He has a Ph. D. from the Weizmann Institute, completed in 1990 under the supervision of Amir Pnueli. He is one of the recipients of the 2001 Gödel Prize.