Perfectly Satisfiable Systems of Linear Equations and Fixed Weight Solutions

Johan Håstad
KTH – Royal Institute of Technology

February 3, 2026

Abstract

We study systems of linear equations modulo two in n variables with three variables in each equation. We assume that the system has a solution with pn variables taking the value 1 for some value $0 < p < 1$. We prove that for any $\delta > 0$ it is hard to find a solution of the same weight that satisfies at least a fraction $c_p + \delta$ of the equations. The constant c_p is upper bounded by .9 for any value of p.

1 Introduction

Systems of linear equations take an almost unique place in mathematics. They are surprisingly expressive and when satisfiable they can be solved efficiently by Gaussian elimination. Gaussian elimination is a very efficient but also a rather fragile algorithm. It is sensitive to errors and even though it efficiently finds some solution it is difficult to use it to find solutions with special structure. It turns out that some computational problems related to systems of linear equations are computationally difficult.

In this paper we are interested in the special case where each equation only contains three variables. This can also be phrased as a Constraint Satisfaction Problem (CSP) and this is the context from which we arrive. In a CSP each variable takes a value in a finite domain and we have a large number of constraints of constant arity over these variables. A particular CSP is defined by the type of constraints allowed. The general study of CSPs is a huge research area covering many aspects. We do not wish to survey the results but let us mention some highlights.

A basic question is which CSPs allow an efficient algorithm to determine whether a given instance is satisfiable. After a long sequence of results, Bulatov [13] and Zhuk [19] independently completed the characterization of when this is possible and in particular proved that the problem is either in P or NP-hard.

Another important area of research is approximability of CSPs. Here one is given an instance in which a fraction c of the constraints are simultaneously satisfiable and the question is whether it is possible to efficiently find an assignment that satisfies at least a fraction s of the constraints. Both the case of $s = c$ is interesting (but usually hard) as well as $s < c$. Here a very important result by Raghavendra [17] shows that, under the unique games conjecture, semidefinite programming can be used to get an algorithm that, within an arbitrarily small ϵ, gives an optimal value of s.

A special case that combines the two views is approximability in the case of satisfiable instances, i.e., $c = 1$. In this situation, proving hardness turns out to be much more difficult. Two of the main problems are that we cannot reduce from unique games as it has non-perfect completeness and we cannot introduce noise in the reductions. For many CSPs it is sometimes difficult to even guess what the best approximation algorithms on satisfiable instances should look like. We refer to the sequence of papers [7, 8, 9, 10, 11] and the references in these papers for a more detailed discussion.

As mentioned initially, we are in this paper interested in satisfiable systems of linear equations with three variables in each equation and only study equations over the field of two elements. Superficially it might seem that there is not much more to say in this case as it is easy to find a solution by Gaussian elimination. As we are in a finite field we do not even have problems with numerical accuracy. What makes the situation interesting, however, is that we are interested in special solutions. In particular we limit the number of variables that take the value 1.

There are several previous paper that study CSPs under the restriction of fixed weight solutions and many have studied CSPs of arity 2. As a prime example one can consider Max-Bisection, the problem of Max-Cut when the cut is required to split the graph in to two pieces of equal size. The hardness results for Max-Cut apply also to Max-Bisection while so far it has not been possible to transfer the algorithmic results. For instance the famous algorithm of Goemans and Williamson [14] gives an approximation ratio of roughly .8786 for Max-Cut while the best approximation algorithm for Max-Bisection by Austrin et el [2] gives a ratio closer to .8776. There is recent work by Brakensiek et al [12] that suggests that the two constants might be different, or at least cannot be proven to be equal without new ideas. It would be amazing if these two very similar problems have optimal approximation ratios that are this close but still distinct.

Austrin and Stanković [4] give detailed results what approximation ratios can be obtained as a function of c and the relative weight p for several CSPs of arity two but do not study the case $c = 1$ in detail. In general this problem of finding fixed weight solutions for satisfiable instances for Boolean CSPs is not so well studied. It is not difficult to see that for Max-Cut (and, more generally linear equations with two variables in each equation) it is possible to find a perfectly satisfying assignment of a given weight if it exists. For 2Sat the problem is NP-hard by a reduction from Vertex Cover. For any edge (i, j) one writes down the clause $(x_i \vee x_j)$ and a satisfying assignment with t ones is now equivalent to a vertex cover of size t. Furthermore, by using the result of Austrin et al [3] that vertex cover on graphs of bounded degree d is hard to approximate within

$2 - o_d(1)$, it follows that fixed weight 2Sat is hard to approximate[1]. While [3] depends on the unique games conjecture using the techniques of [6] it follows that the problem is NP-hard under randomized reductions.

As stated above, in this paper we are interested in instances of 3Lin that have perfect solutions of a given weight and the question is whether it is possible to find a solution of the correct weight that satisfies all or many of the equations. As far as we can tell this question has not been addressed previously.

We prove that there are constants c_p such that given a linear system with the guarantee that there is an assignment with weight pn that satisfies all equations, it is hard to find an assignment of weight pn that satisfies more than a fraction $c_p + \delta$ of all equations for any $\delta > 0$. We can assume that $p \leq \frac{1}{2}$ as we can negate all right-hand sides which negates all values of a solution.

The proof is based on a natural PCP and is very similar to that of [15] but as we cannot introduce noise, modifications are needed. A direct proof yields $c_p = 1 - p + 2p^2$ but this can be improved for $p \neq \frac{1}{4}$ by an easy reduction. One simply adds some new variables and new equations on these new variables that are easy to satisfy. For example, starting with instance constructed for $p = 1/4$ and adding 50% more variables and some equations on these it is possible to get $c_{1/2} = 9/10$.

Our first construction results in a system were we have some few frequent variable and a large number of variables that are less frequent. In this situation, the cardinality constraint only reflects what happens to the many variables. This is not quite satisfactory as a majority (in fact two thirds) of all variable occurrences are not really affected by the cardinality constraints. Motivated by this and by adding a few complications we extend our results to regular systems where each variable appear the same number of times. The bounds are not so different from the non-regular case and for instance for balanced solutions we obtain the constant 10/11.

2 Basic definitions

We call the problem of study 3Lin which is thus given by linear systems of equations modulo 2 where each equation contains three variables. A generic system is denoted L and use n to denote the number of variables in the system and m to be the number of equations. An assignment has *relative weight* p if it gives the value one to pn variables.

Let us comment on one annoying detail. Clearly the relative weight of any assignment is a rational number. Furthermore suppose that $p = p_1/p_2$ where p_1 and p_2 are integers. Then to construct a system with a solution of relative weight p we need the number of variables in the system to be a multiple of p_2 which gives a very annoying condition to address in the construction. We basically sweep this problem under the rug. When claiming that we create

[1] Take any d such that it is hard to distinguish graphs which has a vertex cover of size $3n/5$ and $4n/5$. Then it is NP-hard to find a solution of weight $3n/5$ that satisfies more than a $(1 - \frac{2}{5d})$ fraction of the clauses.

systems where there is a solution of relative weight p, we in fact create a system with a solution of weight that is fixed and known and very close to pn. The error in the relative weight can be made arbitrarily small and is ignored.

When studying systems that can be satisfied by an assignment of relative weight p it turns out to be interesting to study assignments of relative weight s also for $s \neq p$. Our main interest is in $s = p$ but to have good information on a system L formed from two systems L_1 and L_2 on disjoint sets of variables it is not enough to know how each system behaves on assignments of relative weight p.

The first quality measure of solutions that comes to mind is the fraction of equations satisfied. In many situations, however, it is rather simple to satisfy half the equations. Thus we have the following definition.

Definition 2.1 *An assignment to the variables of L has* advantage *δ if it satisfies a fraction $(1+\delta)/2$ of the equations.*

As all our hardness results are through defining a PCP where the verifier accepts depending on the exclusive-or of three bits in the proof, we use the term "verifier accepts with advantage δ" to indicate the the verifier accepts with probability $(1+\delta)/2$.

3 Overview of the proof

Our current proof is very much based on the proof of [15] showing that it hard to distinguish instances of 3Lin where it is possible to satisfy a $(1-\epsilon)$ fraction of the equations from those where you can only satisfy a fraction $(1+\delta)/2$. This overview is easier to read if one has that proof fresh in ones mind and thus a quick glance at [15] before continuing can make the rest of this section easier to follow.

The heart of the argument of [15] is a two-prover protocol turned into a PCP. In the two-prover protocol one prover, P_1, receives k clauses from a 3CNF and returns a satisfying assignment for these clauses. The other prover, P_2, gets one variable from each clause and returns an assignment to these k variables. The verifier checks that the assignment returned by P_1 does satisfy the chosen clauses (this is easy to achieve) and that it is consistent with the values from P_2 (which is the interesting part). It is NP-hard to distinguish the two cases when the provers can make the verifier always accept (as the underlying 3-CNF is satisfiable) from the case where best strategy of the provers makes the verifier accept with probability at most c^k for some $c < 1$.

This is turned into a PCP and we let W be the set of variables sent to P_1 and U the set of variables sent to P_2. We replace the answers of the provers by their long codes. Thus for each set W there is a table, $B(g)$, indexed by functions g mapping $\{0,1\}^W$ to $\{0,1\}$. For a correct proof of a correct statement the value at g is $g(y)$ where y is the answer from P_1. Similarly we have tables $A_U(f)$ replacing the answers of of P_2.

The proof is checked by picking a random function, f, on $\{0,1\}^U$, and random function, g_1, on $\{0,1\}^W$, and setting $g_2(y) = g_1(y) + f(\pi(y)) + \mu(y)$ where μ is random noise, addition is exclusive-or, and π is the projection operator from W to U. The noise μ takes the value one randomly and independently for each y with probability ϵ and is otherwise zero. After expanding both A and B by the Fourier Transform and and doing some simple calculations one concludes that if the verifier accepts with advantage δ then

$$\delta = \sum_\beta \hat{A}_{\pi_2(\beta)} \hat{B}_\beta^2 (1-2\epsilon)^{|\beta|}, \tag{1}$$

where π_2 is a "mod 2 projection". From this relationship it is possible to extract strategies in the two-prover game. For instance, P_1 picks β with probability $\hat{B}_\beta^2$ and then answers with a random element from β. To analyze this strategy it is important to bound the expected size of β and the factor $(1-2\epsilon)^{|\beta|}$ is crucial for this. The origin of this factor is the noise function μ.

In the current paper we follow the same approach but as we want perfect completeness we cannot allow any noise and we end up with (essentially) the expression

$$\sum_\beta \hat{A}_{\pi_2(\beta)} \hat{B}_\beta^2 \tag{2}$$

and we need to worry about large size β. In the current situation we prove a weaker theorem than [15] and we only need to analyze the case when (2) is close to one. It is easy to see that to make it exactly one, A and B have to be matching exclusive-ors. It is also not difficult to construct explicit solutions where the verifier accepts with probability one and we give examples in Section 7.1 below. This is unavoidable as determining whether there is a proof in the PCP that is accepted with probability one can be done in polynomial time.

We are, however, interested in solutions where only a fraction p of the variables take the value 1 and to achieve this we change f to take the value one with (roughly) probability p. In this situation a correct long code also has a fraction p of values that are one while large exclusive-ors are essentially unbiased. Modulo some technicalities this implies that we can focus on sets β such that $\pi_2(\beta)$ is small with high probability. One could have hoped that such β are themselves small, but only something weaker is true. There is a small set, S^β, such that if $\pi_2(\beta)$ is small, it is, with high probability, contained in $\pi(S^\beta)$ (which equals $\pi_2(S^\beta)$ with high probability). This set, S^β, is not in general a subset of β and might not even give possible answers for P_1. It is true, however, that elements of $\pi_2(\beta)$ are possible answers for P_2. By adding the variables of tk clauses to both U and W for a large parameter t we can prove that contributions from terms that do not satisfy the clauses sent to P_1 is small. Let us turn to the formal argument and start with some preliminaries.

4 Preliminaries

In the current paper we have both bits (which we view as elements of the finite field $\mathbb{F}_2$) and real numbers. To avoid having two different variants of "+" we, in the technical part of the paper, use ±1 to denote the elements of $\mathbb{F}_2$ and addition in $\mathbb{F}_2$ turns into multiplication of real numbers. In other words, a linear equation which in $\mathbb{F}_2$-notation reads as

$$x + y + z = 1$$

turns into an equation

$$xyz = -1.$$

This change turns 1 into -1 and hence if x is of relative weight p then

$$\sum_{i=1}^{n} x_i = (1-2p)n.$$

As stated earlier, as we can negate variables, we focus on $p \leq \frac{1}{2}$ and hence this sum is positive. The distribution μ_p on n bits picks each bit, independently of the other bits, to be 1 with probability $1-p$ and to be -1 otherwise. We have the basic characters $\chi_i^p(x_i)$, depending on a single input which take the value $\sqrt{p/(1-p)}$ at 1 and $-\sqrt{(1-p)/p}$ at -1. For general sets $\alpha \subseteq [n]$ we have

$$\chi_\alpha^p(x) = \prod_{i \in \alpha} \chi_i^p(x_i).$$

It is well known and easy to check that these form a complete orthonormal basis of $\{-1,1\}^n$ under μ_p. Of particular interest to us is the case $p = 1/2$ where we drop the superscript and simply use μ for the probability measure and χ_α for the characters.

Any function $\{-1,1\}^n$ can be expanded as the Fourier transform in any such basis and

$$f(x) = \sum_{\alpha \subseteq n} \hat{f}_\alpha^p \chi_\alpha^p(x)$$

where

$$\hat{f}_\alpha^p = E_x[f(x)\chi_\alpha^p(x)]$$

in which x is selected according to μ_p. It is easy to go from one basis to another and in particular $\chi_i(x) = x$ and this can be expressed in the p-biased basis as

$$\chi_i(x) = (1-2p)\chi_\emptyset^p(x) + 2\sqrt{p(1-p)}\chi_i^p(x).$$

We use the long code, introduced in [5], to code an element, x, from a set S. Let F_S denote the set of functions f mapping S to $\{-1,1\}$.

Definition 4.1 *The* long code *of $x \in S$ is a table A mapping F_S to $\{-1,1\}$ where $A(f) = f(x)$.*

Note that the definition depends both on the element x and the set, S, from which it is chosen. In our PCPs the proof contains tables that are supposed to be long codes of certain strings. In some cases such a table is of size $2^{|S|}$ with one entry for each f. In some other cases the table contains only $2^{|S|-1}$ bits with one entry for each pair $f, \neg f$ giving the value for $A(f)$. This is a *folded* table. If the value for $\neg f$ is desired this is defined as $\neg A(f)$. We use folded tables that come in complementary pairs, A^0 and A^1. If A^0 contains the value for f then A^1 contains the value for $\neg f$. This ensures that for any assignment correctly coded by the two tables A^0 and A^1, we have exactly half ones when looking at the union of the tables. This step might seem unnecessary as one normally thinks of a folded long code as unbiased. It is true that the resulting logical long code is unbiased. We want, however, that the actually stored bits of the long code are equally often 1 and -1 which is a different property. Suppose for instance, that we, for some input y, choose f to represent that pair $(f, -f)$ iff $f(y) = 1$. Then with this folding the long code for y consists of all ones. The coding of any other input is unbiased but we want it to be unbiased for many input. There are many ways to achieve this and complimentary pairs is just one possibility that is convenient for us.

Suppose $U \subseteq W$, then for $y \in \{-1, 1\}^W$, we let $\pi_U(y)$ be the string $x \in \{-1, 1\}^U$, such that $x_i = y_i$ for $i \in U$. When U is clear from the context we drop it for readability and simply write $\pi(y)$.

5 The basic protocol to show hardness

As is standard we construct a PCP where the Verifier flips $O(\log n)$ random coins, reads three bits, and accepts if the exclusive-or of the three bits have a prescribed value. While most modern hardness results start with arbitrary label cover, we start with the specific[2] label cover used in [15]. To be more precise, we start with a 3Sat instance $\varphi = \vee_{i=1}^{m} C_i$ where each clause C_i is of size exactly three and each variables appears in exactly 5 clauses. It has n variables and hence $m = 5n/3$.

By [1], one can construct such CNFs where it is NP-hard to distinguish whether φ is satisfiable or any assignment falsifies a fraction γ of the clauses. Here γ is an absolute constant strictly greater than 0 and we do not need its actual value as it is absorbed in other constants. We use φ to create a two-prover game and let k and t be two constants to be specified later. The protocol of [15] is the same but uses $t = 0$.

Basic two-Prover protocol

- The verifier V uniformly at random selects $(t+1)k$ clauses C_{i_j} from φ.

[2] We use the particulars of this system when proving Lemma 6.5 and Lemma 6.6 below. It is possible that one could work with a general label cover, but this would require a different argument.

- V sends the $(t+1)k$ clauses to P_1 which returns values to all $3(t+1)k$ variables in these clauses.
- V randomly selects k of the chosen clauses and a random variable in each of these clauses. It sends these k variables jointly with the remaining tk clauses to P_2 which returns values to the k chosen variables as well as the $3tk$ variables in the clauses.
- V accepts iff the assignment sent by the two variables are consistent on the common variables and satisfy all the chosen clauses.

We have the following, by now standard, theorem.

Theorem 5.1 *If φ is satisfiable then there is a strategy for P_1 and P_2 that makes V always accept. If any assignment of the variables of φ falsifies a fraction $\gamma > 0$ of the clauses, then there is a constant $c_\gamma < 1$ such that for any strategy of P_1 and P_2, V accepts with probability at most c_γ^k.*

The first part of the theorem is obvious as P_1 and P_2 can just agree to answer according to a fixed satisfying assignment. The second part follows by the parallel repetition theorem for two-prover games by Raz [18].

We denote a typical set of variables sent to P_1 by W and a typical set of variables sent to P_2 by U. Rather than adding more notation we assume in some places that each of these sets also carries the information of the identity of the picked clauses. Thus W is sometimes a set of variables and sometimes a set of clauses. As the notions are very close hopefully this is not confusing. Most of the time W contains $3(t+1)k$ variables and U contains $(3t+1)k$ variables while always $U \subseteq W$. A technical point is that we pick the $(t+1)k$ clauses with repetition. This implies that the k single variables in U are picked with the uniform probability independent of the other tk clauses. It might be the case that we have duplicated variables in W or U, but this happens only with probability $O(k^2t^2/n) = O(1/n)$ and, for notational convenience, we ignore this possibility in our analysis of soundness. This small probability can be absorbed in the error terms.

Let S_U be the set of assignments on U that satisfy the clauses sent to P_2 and similarly we have sets S_W. For any answer, $y \in S_W$ there is a unique answer, $\pi(y) \in S_U$ that makes V accept. It is simply the restriction of the assignment y on W to the subset U. Of course π depends on the identity of W and U but we suppress this dependence.

We now turn this basic two-prover protocol into a PCP. This written proof contains a number of tables. In a correct proof for a satisfiable φ these tables are long codes of local views of a satisfying assignment. Of course when analyzing soundness we do not have any control over these tables as they might not contain any systematic information. We use the term "supposed long codes" to hopefully give the reader the correct intuition. To be able to adjust the relative weight of a correct proof we use multiple copies of some tables.

- For each possible set U sent to P_2 we have M supposed long codes A_U^i of elements in S_U for $1 \leq i \leq M$.
- For each possible set W sent to P_1 we have a pair of complementary supposed long codes, B_W^i, of elements in S_W for $i \in \{0, 1\}$.

We turn to the basic PCP. We have a parameter q which controls the bias of the function f. This is a number close to p and we give details in Lemma 5.3 below. The algorithm for the verifier is as follows.

- Pick U and W as in the two-prover protocol.
- Pick a random $i \in [M]$ and random $j_1, j_2 \in \{0, 1\}$, all three independently.
- Define $f \in F_{S_U}$ by setting $f(x) = -1$ with probability q and 1 otherwise for each $x \in S_U$ independently and uniformly.
- Define $g_1 \in F_{S_W}$ by setting $g_1(y)$ to an unbiased bit independently for all $y \in S_W$.
- Set $g_2(y) = g_1(y)f(\pi(y))$ for each $y \in S_W$.
- Accept iff $B_W^{j_1}(g_2)B_W^{j_2}(g_1)A_U^i(f) = 1$.

We emphasize that A_U^i is not folded while B_W^0 and B_W^1 is a complementary pair of folded tables.

5.1 Analyzing completeness

If φ is satisfiable then for each U we let x_U be the restriction of a fixed satisfying assignment to U and similarly y_W is the restriction to W. We let A_U^i and B_W^j be the long codes of x_U and y_W, respectively. In other words, $A_U^i(f) = f(x_U)$ and $B_W^j(g) = g(y_W)$ for any i, j, f, and g. As $\pi(y_W) = x_U$, the definition of g_2 implies that the verifier always accepts.

We want to calculate the relative weight of this proof, i.e. the fraction of bits that take the value -1. In order to do this we add an extra step as follows.

- If the number of x such that $f(x) = -1$ is not contained in the range $[(q - \epsilon)|S_U|, (q + \epsilon)|S_U|]$, reject this choice for f and make a new choice. Call this a *biased* choice of f.

This extra step has the consequence that we can remove a big part of A_U^i that is never used. The step does not affect the completeness and, as we see below, it hardly affects soundness, but it changes the relative weight of the proof. As written originally, A_U^i contains entries for all possible f and if it is a correct long code then half of the entries are -1. By removing entries corresponding to very unlikely f we make the relative weight close to q. Let us first state a standard lemma.

Lemma 5.2 *The probability that f is biased is at most* $\exp(-c\epsilon^2 7^{tk})$ *for an absolute constant c.*

Proof: (Sketch) The function f is picked by at least 7^{tk} independent bits. The expected fraction of ones is q. The probability that we get a deviation of at least ϵ is, by standard Chernoff bounds, as stated in the lemma. ■

Suppose that for the setting $M = 1$, the proof has N bits of which $\delta_A N$ come from A-tables and $\delta_B N = (1 - \delta_A)N$ come from the B-tables. Here δ_A is a constant depending only on ϵ, q, t and k that can be calculated but, as this is not needed, let us not make an explicit formula for this constant.

The table A contains an entry for each f such that the number of x with $f(x) = -1$ is in the given range. By symmetry, the number of -1s in such a table does not depend on which input it codes and it easy to see that this number is between $(q - \epsilon)\delta_A N$ and $(q + \epsilon)\delta_A N$. Suppose that the true number is $q'\delta_A N$. Here q' is a number that can be calculated from ϵ, t, k and q and does not depend on the assignment that satisfies φ. The number of -1s in the B-tables is, by the complementarity property, $\frac{1}{2}\delta_B N$. With M copies of each A-table the total number of bits in the proof becomes $M\delta_A N + \delta_B N$ and the number of -1s becomes $Mq'\delta_A N + \frac{1}{2}\delta_B N$. This implies that the fraction of -1s is

$$\frac{q' M\delta_A + \frac{1}{2}\delta_B}{M\delta_a + \delta_B} \tag{3}$$

which tends to q' as M increases. We summarize the argument in the current section as a lemma.

Lemma 5.3 *If φ is satisfiable then there is a proof that makes V always accept. The relative weight of this proof does not depend on the assignment that satisfies φ and can be made arbitrarily close to q by making ϵ small and M large.*

For a given value of p and given values of ϵ and M we get a value $q = q(p, \epsilon, M, k, t)$ such that the relative weight of the given proof is, essentially, p. As discussed earlier we cannot make it exactly p. This happens when p is irrational or has an awkward denominator that does not fit well with the PCP. We can, however, make it very close to p and we ignore this point and $q(p, \epsilon, M, k, t)$ is a number such that the relative weight of the solution is as close to p as we desire. The below observation follows directly from (3) and that $|q' - q| \le \epsilon$.

Lemma 5.4 *We have*

$$\lim_{M\to\infty} |p - q(p, \epsilon, M, k, t)| \le \epsilon.$$

6 Analyzing soundness

Let us analyze the probability that V accepts a given proof. Let us fix the values of U and W and drop these as indices of A and B. Define

$$A(f) = \frac{1}{M}\sum_{i=0}^{M-1} A^i(f),$$

and

$$B(g) = \frac{1}{2}(B^0(g) + B^1(g)).$$

These functions are not Boolean but take values in $[-1, 1]$. If a verifier accepts with advantage γ for this fixed value of U and W, then

$$\gamma = E_{i,j_1,j_2,f,g_1}[A^i(f)B(g_1^{j_1})B(g_2^{j_2})] = E_{f,g_1}[A(f)B(g_1)B(g_2)]. \qquad (4)$$

Set $C(f) = E_{g_1}[B(g_1)B(g_2)]$ and expanding B by the Fourier transform we get.

$$C(f) = E_{g_1}[\sum_{\beta_1} \hat{B}_{\beta_1}\chi_{\beta_1}(g_1)\sum_{\beta_1} \hat{B}_{\beta_2}\chi_{\beta_2}(g_1(f\circ\pi))] = \sum_{\beta} \hat{B}_\beta^2 \chi_\beta(f\circ\pi),$$

as terms with $\beta_1 \neq \beta_2$ have expectation zero. Now

$$\chi_{\beta(f\circ\pi)} = \prod_{y\in\beta} f(\pi(y)) = \chi_{\pi_2(\beta)}(f),$$

where $x \in \pi_2(\beta)$ iff there is an odd number of y such that $\pi(y) = x$. We conclude that

$$\hat{C}_\alpha = \sum_{\beta\ |\pi_2(\beta)=\alpha} \hat{B}_\beta^2,$$

and we note that $\hat{C}_\alpha$ are positive numbers whose sum is $\|B\|_2^2 \leq 1$. Plugging this into (4) we get

$$\gamma = E_f[A(f)C(f)] = E_f[A(f)\sum_\alpha \hat{C}_\alpha \chi_\alpha(f)]. \qquad (5)$$

In the analysis of [15] one extracts a good strategy for the provers in the two-prover game as soon as (5) is strictly greater than zero. Currently we are trying to do less and only extract a strategy when the expectation is at least c for some constant c which is quite close to, but strictly smaller than, 1. Before going into the details how to do this, let us give some intuition.

As all the Fourier coefficients of C are positive and they sum to at most one, the only way that (5) can equal one is that $C_\alpha = 1$ for some α and $A = \chi_\alpha$. If we on top of this require that A only contains a fraction q of values that are -1 the only alternative is that α is a singleton. As the size of α increases, the bias of $\chi_\alpha(f)$ tends to 0 and the following simple lemma is useful.

Lemma 6.1 *Let X and Y be two discrete random variables taking values in $[-1,1]$ and such that $E[X]=e_x$ and $E[Y]=e_y$. Then $E[XY] \le 1-|e_x-e_y|$.*

Proof: First we claim that we can assume that the two variables only take values ± 1. Suppose a pair of values (x_i,y_i) appears with probability s and $|x_i| \ne 1$. We can replace this by $(-1,y_i)$ appearing with probability $(1-x_i)s/2$ and $(1,y_i)$ appearing with probability $(1+x_i)s/2$. This does not change any of $E[X]$, $E[Y]$, and $E[XY]$. We can repeat this to make also Y take values ± 1.

Now we can observe that pairs with $x=-y$ must appear with probability at least $\frac{1}{2}|e_x-e_y|$ to give the assumed difference in absolute value. ∎

We split C in to its high and low degree parts. To be more precise we let ℓ be an odd integer and set set

$$C^\ell(f) = \sum_{|\alpha|\le \ell} \hat{C}_\alpha \chi_\alpha(f).$$

and $\tilde{C}(f) = C(f) - C^\ell(f)$. Let $c^\ell = \sum_{|\alpha|\le\ell} \hat{C}_\alpha$ be the sum of the Fourier coefficients of C^ℓ. Note that $|C^\ell(f)| \le c^\ell$ for any f and $|\tilde{C}(f)| \le 1-c^\ell$. Remember that B is folded and hence any α such that $\hat{C}_\alpha \ne 0$ is of odd size. For notational convenience define $e_A = E_f[A(f)]$.

Lemma 6.2 *We have*

$$E_f[A(f)\tilde{C}(f)] \le (1-c^\ell)\min(1, 1+(1-2q)^{\ell+2}-e_A, 1+e_A).$$

Proof: First note that if $0 \le e_A \le (1-2q)^{\ell+2}$, the lemma true as $|\tilde{C}(f)| \le 1-c^\ell$ and $|A(f)| \le 1$. Assume first that $e_A > (1-2q)^{\ell+2}$. We have, by Lemma 6.1 and using that $E[\chi_\alpha(f)] = (1-2q)^{|\alpha|}$ and that $|\alpha| \ge \ell+2$,

$$E[A(f)\chi_\alpha(f)] \le 1-|e_A-(1-2q)^{|\alpha|}| \le 1+(1-2q)^{|\alpha|}-e_A \le 1+(1-2q)^{\ell+2}-e_A.$$

We conclude that

$$E_f[A(f)\tilde{C}(f)] = \sum_{|\alpha|\ge\ell+2} \hat{C}_\alpha E[A(f)\chi_\alpha(f)] \le (1-c^\ell)(1+(1-2q)^{\ell+2}-e_A).$$

The case when $e_A < 0$ is similar using that $E[\chi_\alpha(f)] \ge 0$ for any α as $q \le \frac{1}{2}$. ∎

The key lemma to establish soundness is the following.

Lemma 6.3 *If the provers can win the two-prover game with probability at most ϵ_1 then $E[A(f)C^\ell(f)] \le c^\ell(1-2q)\max(0,e_A) + \sqrt{\ell\epsilon_1} + \sqrt{\ell/t}$.*

Proof: To estimate $E[A(f)C^\ell(f)]$ we expand both factors in the basis χ^q_α. As noted above, for singletons α we have

$$\chi_\alpha(f) = (1-2q)\chi^q_\emptyset(f) + 2\sqrt{q(1-q)}\chi^q_\alpha(f),$$

and by multiplicativity

$$\chi_\alpha(f) = (1-2q)^{|\alpha|}\chi_\emptyset^q(f) + \sum_{\alpha' \subseteq \alpha} c_{q,|\alpha'|,|\alpha|}\chi_{\alpha'}^q(f), \tag{6}$$

where the sum is over non-empty α' and, by Plancherel,

$$\sum_{\alpha' \subseteq \alpha} c_{q,|\alpha'|,|\alpha|}^2 \leq 1. \tag{7}$$

We are interested in, possibly large, β such that $\pi_2(\beta)$ is of size at most ℓ, and start with a preliminary observation useful for studying projections.

Lemma 6.4 *Suppose $\alpha \subseteq \{-1,1\}^U$ is a set of size ℓ_0, then there is a $U' \subseteq U$ of size at most $\ell_0 - 1$ such that for any $x \neq x'$, both in α, $\pi_{U'}(x) \neq \pi_{U'}(x')$.*

Proof: We prove the lemma by induction of ℓ_0. It is clearly true for $\ell_0 = 1$. For the general case take any coordinate, i, on which all elements in α do not agree. Split α into a disjoint union of α^0 and α^1 where the elements with $x_i = b$ are included in α^b. By induction we have coordinates $U^{b,'}$ that split these sets. Setting U' to be the union of $U^{0,'}$ and $U^{1,'}$ together with the element i gives a set with the requested property. ■

Let us fix W and study what happens when we pick a random partner U and first look at the case when $\pi_2(\beta)$ is of size one.

Lemma 6.5 *Suppose β is a set of odd size. Then there is an assignment z such that if $|\pi_2(\beta)| = 1$, then $\pi_2(\beta) = \pi(z)$.*

Proof: For each $i \in W$ divide β into two sets, β_0^i and β_1^i, depending on the value of y_i. Define z_i to the bit, b, such that β_b^i is of odd size.

Now if $i \in U$ then $\pi_2(\beta)$ must contain an element with ith coordinate z_i. This follows as there is an odd number of elements with $y_i = z_i$ in β. If $\pi_2(\beta)$ is a singleton then it must hence equal $\pi(z)$. ■

Note that we do not claim that z belongs to β and it is easy to come up with examples where it does not. For instance β might be a sub-hypercube with one element, w, removed and in this case $z = w$. We proceed to study that case when $\pi_2(\beta)$ is small but of size larger than one.

Lemma 6.6 *Suppose β is a set of odd size and that there is some U such that $\pi_2(\beta)$ is of size at most ℓ. Then there is a set, S^β, of size at most ℓ such that*

$$Pr[|\pi_2(\beta)| \leq \ell \wedge \pi_2(\beta) \neq \pi(S^\beta)] \leq O(\ell/t).$$

Proof: Fix a value U^0 of U such that $|\pi_2(\beta)| \leq \ell$ and such that $|\pi_2(\beta)|$ is as large as possible given that the size is at most ℓ. Suppose $\pi_2(\beta) = \alpha$ and let U' be the set of size at most $|\alpha| - 1$ found by Lemma 6.4. For each element $x \in \pi_{U'}(\alpha)$, β contains an odd number of elements, y, such that $\pi_{U'}(y) = x$.

For each such x we can find an assignment z^x as in the proof of Lemma 6.5. Namely look at all $y \in \beta$ such that $\pi_{U'}(y) = x$. As this is an odd number, for each i there is a value z^x_i such that there is an odd number of such y with ith coordinate z^x_i. Let S^β be the set of all such z^x.

Now for any U that contains U' and any $x \in \alpha$ we must have at least an element y in $\pi_2(\beta)$ such that $\pi_{U'}(y) = \pi_{U'}(x)$. Thus $\pi_2(\beta)$ contains at least ℓ_0 elements. By definition, $\pi_2(\beta)$ either contains exactly ℓ_0 elements or more than ℓ elements. In the former case, $\pi_2(\beta)$ must equal $\pi(S^\beta)$. As the probability that U contains U' is $1 - O(\ell/t)$, Lemma 6.6 follows. ■

We return to the proof of Lemma 6.3 and expand A by its q-biased Fourier transform and using (6) we derive the q-biased expansion of C^ℓ resulting in the equality

$$E[A(f)C^\ell] = \sum_\alpha \hat{C}_\alpha \left(((1-2q)^{|\alpha|})e_A + \sum_{\alpha' \subseteq \alpha} \hat{A}^q_{\alpha'} c_{q,|\alpha'|,|\alpha|} \right), \tag{8}$$

where we the last sum is over non-empty α'. As $\sum_\alpha \hat{C}_\alpha = c^\ell$ where each number is positive and $|\alpha| \geq 1$ the first term in all the summands is bounded by the first term of the lemma and we need to look at the inner sum.

Say that a term is unusual if α' is not a subset of $\pi_2(S^\beta)$. By Lemma 6.6, the expectation of contribution from unusual terms is at most $O(\ell/t)$. Summing only over usual terms and using Cauchy-Schwarz and Plancherel each twice we get

$$\begin{aligned}
\sum_\beta \hat{B}^2_\beta \sum_{\alpha' \subseteq \pi_2(S^\beta)} \hat{A}^q_{\alpha'} c_{q,|\alpha'|,|\alpha|} &\leq \\
\sum_\beta \hat{B}^2_\beta \left(\sum_{\alpha' \subseteq \pi_2(S^\beta)} (\hat{A}^q_{\alpha'})^2 \right)^{1/2} \left(\sum_{\alpha' \subseteq \pi_2(S^\beta)} c^2_{q,|\alpha'|,|\alpha|} \right)^{1/2} &\leq \\
\sum_\beta \hat{B}^2_\beta \left(\sum_{\alpha' \subseteq \pi_2(S^\beta)} (\hat{A}^q_{\alpha'})^2 \right)^{1/2} &\leq \\
\left(\sum_\beta \hat{B}^2_\beta \right)^{1/2} \left(\sum_\beta \hat{B}^2_\beta \sum_{\alpha' \subseteq \pi_2(S^\beta)} (\hat{A}^q_{\alpha'})^2 \right)^{1/2} &\leq \\
\left(\sum_\beta \hat{B}^2_\beta \sum_{\alpha' \subseteq \pi_2(S^\beta)} (\hat{A}^q_{\alpha'})^2 \right)^{1/2} &.
\end{aligned} \tag{9}$$

Let us look at the following strategy in the two-prover game.

- P_1, upon receiving W looks at $B = B_W$ and picks β with probability $\hat{B}^2_\beta$. If S^β is a set of size at most ℓ it returns a random element from this set and otherwise it returns any default message.

- P_2, upon receiving U, looks at $A = A_U$ and picks α with probability $(\hat{A}^q_\alpha)^2$. If α is a non-empty, P_2 returns a random string from this set and otherwise any default message.

The probability that the two strings returned are consistent is at least

$$\frac{1}{\ell}\sum_\beta \hat{B}^2_\beta \sum_{\alpha\subseteq\pi_2(S^\beta)} (\hat{A}^q_\alpha)^2. \tag{10}$$

Indeed if the two provers pick α an β, where $\alpha \subseteq \pi_2(S^\beta)$ then the probability the two strings are consistent is at least $1/\ell$. The verifier also checks that the answers satisfy the corresponding clauses. This is always true for the answers from P_2. We claim that the expected contribution to (10) from answers where the answer from P_1 does not satisfy the clauses picked is at most $1/t$. To see this pick any such answer by P_1. The probability that it contributes anything to (10) is $1/t$ as any violated clause is sent also to P_2 with probability $(1-1/t)$. We conclude that the expectation of (10) is at most $\epsilon_1 + 1/t$. By convexity we have that $E[\sqrt{X}] \leq \sqrt{E[X]}$ and hence the expectation of (9) is bounded by

$$\sqrt{\ell(\epsilon_1 + \frac{1}{t})} \leq \sqrt{\ell\epsilon_1} + \sqrt{\ell/t}.$$

Collecting terms we have established Lemma 6.3. ■

Having established both soundness and completeness we can finally state our first theorem. Let $F_p(e) = 1 - 2pe$ be the line through the points $(0,1)$ and $(1, 1-2p)$.

Theorem 6.7 *For any $\delta > 0$, $p < \frac{1}{2}$, it is NP-hard to distinguish instances of 3Lin such that*

- *There is an assignment of relative weight p that satisfies all equations.*
  - *For any s, such that $0 \leq s < 1/2$, any assignment of relative weight s has advantage at most $F_p(1-2s) + \delta$.*
  - *For any s, such that $1/2 \leq s < 1$, any assignment of relative weight s has advantage at most $2 - 2s + \delta$.*

Proof: Assume first that $s \leq \frac{1}{2}$. In the proof we set ℓ (the analysis parameter), t (controlling the number of clauses sent to both players), k (the the number of clauses actually used in the two-prover protocol), and M (the number of copies of the A-tables) large enough and ϵ (the tolerance in the bias in the selection of the f input) small enough. We later check that we do not get conflicts among the different conditions but let us for the moment suppose not. Then to make the arguments more transparent we, in this proof, drop terms that can be made arbitrarily small as follows.

- Since M is large and ϵ is small, we, by Lemma 5.4, set $p = q$.

- Since ℓ is large we replace the term $(1-2q)^{\ell+2}$ in Lemma 6.2 by 0.
- Since k is large, ϵ_1 in Lemma 6.3 is very small and is replaced by 0.
- Since t is large, ℓ/t in Lemma 6.3 is very small and is replaced by 0.

Firstly, from Lemma 6.2 we have

$$E_f[A(f)\tilde{C}(f)] \leq (1-c^\ell)\min(1,(1-e_A),1+e_A) \leq (1-c^\ell)F_p(e_A) \qquad (11)$$

and, from Lemma 6.3,

$$E[A(f)C^\ell(f)] \leq c^\ell(1-2p)\max(0,e_A) \leq c^\ell F_p(e_A). \qquad (12)$$

As the A-tables give almost all the variables, again up to small error we know that $E_U[e_A] = 1-2s$. We do not have control over c^ℓ but clearly from (11) and (12) we can conclude that

$$E[A(f)C(f)] \leq F(e_A) \qquad (13)$$

and as F is linear this gives the bound of the lemma (without the error term δ which comes from the dropped error terms on the way). To look at errors more closely let us first recall where we dropped terms.

- The ignored case of a biased f as discussed in Lemma 5.2.
- The term $(1-2q)^{\ell+2}$ in Lemma 6.2.
- The terms $\sqrt{\ell/t}$ and $\sqrt{\ell\epsilon_1}$ in Lemma 6.3.
- The difference between p and q as implicitly discussed in Lemma 5.3.

To make them all small we fix parameters in the following order.

- Pick ϵ small enough so that $p+3\epsilon < \frac{1}{2}$ and $\epsilon < \delta/7$.
- Determine ℓ such that $(1-2q)^{\ell+2} \leq \delta/7$ for any $|p-q| \leq 2\epsilon$.
- Determine t and k such that $\sqrt{\ell/t}$ and $\sqrt{\ell\epsilon_1}$ are both bounded by $\delta/7$ and such that the error term in Lemma 5.2 is bounded by $\delta/7$.
- Find a values of M such that $|F_q(t) - F_p(t)| \leq \delta/6$ for any $t \in [0,1]$.

We get no conflicts and hence for each constant value of δ we can make the sum of the error terms stay below δ. This completes the proof when $s \leq \frac{1}{2}$.

If $s \geq \frac{1}{2}$ we replace $F(e_A)$ in the above bounds by $1+e_A$ which is a valid upper bound in both Lemma 6.2 and Lemma 6.3 (after we have dropped the error terms). We conclude that $E[A(f)C(f)] \leq 2-2s$ yielding the claimed bound. ■

Let us point out that the given theorem is the best that can be proved without using information from the structure of pairs (U, W) that show up in the proof. Let us consider the interesting case $s \leq \frac{1}{2}$. It can be that for a fraction $(1-2s)$ of all pairs (U, W) we have that A_U is the constant 1 and B_W is the long code of some assignment satisfying the local constraints. For the remaining $2s$ fraction of the proof, A_U and B_W give a perfect (but unbiased) solution to that part of the proof, as described in Section 7.1 below. This situation cannot appear exactly in a real proof as the bipartite graph of appearing pairs (U, W) is connected but no such property was used in the the current proof. Let us look at the given proof more closely.

The fraction of all entries in the A-tables that take the value -1 is s and as these are most of the bits of the PCP, the relative weight of the given proof is close to s. In the part of the proof where the A-tables are the constant 1 we have that $B(g_1)B(g_2)$ is equal to f at a single input which is 1 with probability $1-q$ and thus gives an advantage of $1-2q$. In the other part of the proof the verifier always accepts and thus has advantage one. Thus the overall advantage is

$$(1-2s)(1-2q) + 2s = 1 - 2q(1-2s) = F_q(1-2s),$$

matching the bound of the theorem.

The case of $s = p$ is possibly the most natural to analyze which results in the following corollary.

Corollary 6.8 *Suppose $0 < p < \frac{1}{2}$ and $\delta > 0$. Then its NP-hard to distinguish*

- *Instances of 3Lin that has a satisfiable assignment of relative weight p.*
- *Instances of 3Lin where any assignment of relative weight p has advantage at most $1 - 2p + 4p^2 + \delta$.*

7 Hardness for $s = 1/2$ and non-degenerate systems

Theorem 6.7 gives a non-trivial bound for any $s \neq 1/2$ but this is easy to obtain by a padding argument. We add n' new variables and m' equations of the form $x_i + x_j + x_k = 1$. We want the property that any assignment with tn' ones on the new variables satisfy $(t^3 + 3t(1-t)^2 + o(1))m'$ of the added equations. Such a system can either be constructed randomly (assuming that $m = \omega(n)$) or by choosing all possible combinations of (i, j, k). The parameters n' and m' are derived from parameters of the system to pad, but as we can duplicate equations in this original system we can get a suitable value for m'. To make calculations simpler note that satisfying a fraction $t^3 + 3t(1-t)^2$ of equations is the same as having advantage $(2t-1)^3$.

Take the system as given by Theorem 6.7 for $p = 1/4$ and suppose it has n variables and m equations. Add a system as described in the previous paragraph with $n' = n/2$ new variables and $m' = m/4$ equations on these new variables.

This creates a system with $3n/2$ variables which, in the positive case, can be perfectly satisfied by an assignment of relative weight $1/2$. Namely the solution of relative weight $1/4$ on the old variables combined with the all ones solution on the new variables.

Let us look at soundness and let us again drop error terms. Take any solution of relative weight $1/2$ and suppose it assigns $tn' = tn/2$ ones on the new variables. As the total number of ones in the solution is $3n/4$ the relative weight on the old variables is $\frac{3}{4} - \frac{t}{2}$. If we let $D(s)$ give the upper bounds found in Theorem 6.7, then the maximal advantage, up to error terms, is

$$\frac{4}{5}\left(D(\frac{3}{4} - \frac{t}{2}) + \frac{1}{4}(2t-1)^3\right). \tag{14}$$

Let us maximize this over t. If $t \leq \frac{1}{2}$ then $\frac{3}{4} - \frac{t}{2} \geq \frac{1}{2}$ and both terms (14) are increasing with t. For $\frac{1}{2} \leq t \leq 1$, as D is linear in this range, the second derivative of (14) is non-negative and hence it is maximized either by $t = \frac{1}{2}$ or $t = 1$. The value at $t = 1/2$ is $\frac{4}{5}(1+0) = \frac{4}{5}$ while the value at $t = 1$ is $\frac{4}{5}(D(\frac{1}{4}) + \frac{1}{4}) = \frac{4}{5}$. We summarize this analysis in a theorem.

Theorem 7.1 *Suppose $\delta > 0$. Then its NP-hard to distinguish*

- *Instances of 3Lin that has a satisfiable assignment of relative weight $\frac{1}{2}$.*
- *Instances of 3Lin where any assignment of relative weight $\frac{1}{2}$ has advantage at most $\frac{4}{5} + \delta$.*

This theorem takes care of the case $p = 1/2$ but of course we can get improved results also for other values of p. For suitable parameters q, a, and b we take the system obtained from Theorem 6.7, add an new variables and bm' equations on these new variables and optimize over a, b and q such that $q + a = p(1+a)$. This optimization is not very difficult but let us not do it explicitly and state only the big picture.

Theorem 7.2 *Suppose $\delta > 0$ and $0 \leq p \leq \frac{1}{2}$. Then its NP-hard to distinguish*

- *Instances of 3Lin that has a satisfiable assignment of relative weight p.*
- *Instances of 3Lin where any assignment of relative weight p has advantage at most $D(p)$, where $D(p) = 3/4$ for $0 < p \leq \frac{1}{4}$ and then monotonically increases to $\frac{4}{5}$ at $p = 1/2$.*

Proof: Suppose first that $p \leq \frac{1}{4}$ and start with the system given by Theorem 6.7 with $q = 1/4$. Now add $(\frac{1}{4p} - 1)n$ new variables, but no equations. The total number of variables is now $n/4p$ and thus a solution that has relative weight $\frac{1}{4}$ on the old variables and is all zero outside has relative weight p. Thus completeness is clear.

As for soundness take any assignment of relative weight p to all variables. It has relative weight as most $1/4$ on the old variables and as the bound of

Theorem 6.7 is increasing in s for $s \leq \frac{1}{4}$ and $s = \frac{1}{4}$ gives the bound of the theorem, the theorem is true in this case.

It is a bit disappointing to have the new variables not appear in any equations but as we have no lower bound on the number of times a variable appears this is, formally speaking, OK.

Now let us turn to the case $p \geq \frac{1}{4}$. Let us give the argument that gives some function D with the given property but not the best. Set $q = 1/4$ and add $n' = an$ new variables that we in the completeness case all set to one. To get a relative weight of p we need

$$\frac{1}{4} + a = p(1 + a)$$

which is satisfied by $a = (4p - 1)/(4(1 - p))$. When analyzing soundness we assume again that relative weight of a solution is t on the new variables. The situation is quite similar to the case when $p = 1/2$ and it easy to see that it is non-optimal to have relative weight above one half on the original variables. For $p \geq 2/5$ it is possible to get relative weight one half with $t_p = (5p - 2)/(4p - 1)$. We add bm equations to make this case and $t = 1$ equally good. The objective value at $t = 1$ is $\frac{3}{4} + b$ and for $t = t_p$ it is $1 + b(2t_p - 1)^3$ and it is easy to solve for b. When $p < 2/5$ we instead compare $t = 0$ and $t = 1$ and do a similar optimization. It is not difficult to see that the result is monotone in p and the values at the end-points of the interval are already calculated. ■

It is not difficult to, by a similar analysis, derive statements similar to Theorem 7.2 when we bound solutions of relative weight $s \neq p$. We leave the details to the reader.

The constructed system splits into two parts, one interesting (the one from Theorem 6.7) and one trivial with a known solution and we feel this is artificial and let us discuss one definition that rules out such behavior.

Definition 7.3 *A system L is* non-redundant *if it does not imply any equation of the forms $x_i = x_j + c$ or $x_i = c$ for any values of i, j and c.*

It is not difficult to prove that we can find good almost balanced solutions to non-redundant systems.

Theorem 7.4 *Suppose L is a non-redundant and satisfiable system with n variables and m equations. Then it is possible to efficiently find an assignment of relative weight $\frac{1}{2}$ that satisfies $m(1 - O(1/\sqrt{n}))$ equations.*

Proof: By Gaussian elimination we can find a basis of over $\mathbb{F}_2$ of the linear space of solutions to L. This basis allows us to pick a uniformly random point in this space. From the assumption that L is non-redundant it follows that the coordinates of such a random point are unbiased and pairwise independent. This implies that for a random point

$$E\left[\left(\sum_{i=1}^{n} x_i\right)^2\right] = n \tag{15}$$

and by using conditional expectations we can, in polynomial time, find a solution to L that satisfies $(\sum_{i=1}^{n} x_i)^2 \leq n$. Suppose for concreteness that the sum is positive and equals r. Of all variables x_i that equal 1 we choose the $r/2$ variables that occur in the fewest number of equations. As we have more than $n/2$ variables that equal one, the total number of equations in which the chosen variables appear is at most $3rm/n$. Changing the values on the chosen variables produces the solution needed to prove the lemma. ■

It is clear that we could allow a few variables to be fixed by L and some pairs of variables being equal and prove a slightly weaker version of Theorem 7.4. All we need is that the space of all solutions to L is nice enough that we get a good bound in (15). Let us check that the systems underlying Theorem 6.7 do have a fairly nice set of solutions.

7.1 Solutions satisfying all equations in system from Theorem 6.7

Take any global assignment, τ and for each clause C, let us associate one or three assignments. If C is satisfied by τ then we associate the restriction of τ to the variables of C. If it does not, we instead associate the three assignments that, on the variables of C, equal to τ in exactly one point. The construction implies that if we let β the the set of assignments associated with C then if U is any singleton set containing a variables in C we have that $\pi_{U,2}(\beta)$ is the singleton set with the assignment given by τ to the variable in U.

Now for a set of clauses sent to P_1 let T_W be the direct product of the sets of assignments assigned to the clauses in W. If τ falsifies r of these clauses in W, then T_W is a set of 3^r assignments. Similarly we let T_U be the set of assignments constructed in the same way from U where each singleton variable is given the value according to τ. We set

$$A_U(f) = \prod_{x \in T_U} f(x)$$

and

$$B_W(g) = \prod_{y \in T_W} g(y).$$

It is now easy to check that this proof is accepted with probability one. This follows from the fact that $\pi_2(T_W) = T_U$ which in its turn follows by construction.

As τ was arbitrary this gives a rich family of solutions. It is not quite non-redundant as we have values in B-tables corresponding to g being identically 1 and also in complementary tables we have pairs of variables that are each others negations. It is, however, non-redundant enough to explain the fact that we do not get any hardness for $s = 1/2$ in Theorem 6.7.

8 The regular case

In the proof of Theorem 6.7 we constructed a system of equations where each equation contains one variable from an A-table and two variables from a B-table. We made many copies of each A-table and as a consequence the variables from the B-tables are much more frequent in the resulting linear system. As a result, a fraction $2/3$ of all variables occurrences are not really affected by our global relative weight constraints, a fact that can be seen as aesthetically not very pleasing. One can argue that rather than restricting the fraction of variables that are true one should weight this by the number of occurrences to make frequent variables be more important. One very structured case where the two notions agree is given by regular instances where each variable appears the same number of times. The purpose of this section is to modify the construction to make the resulting system regular. Let us first ignore the possibility that W might contain intersecting clauses.

There are two major reasons for non-regularity of the current PCP. One, as pointed out above, is the unbalance between the A-tables and B-tables. Each query in the B-tables is uniformly random and hence all these variables appear the same number of times. The functions f used to query the A-table are however biased and this is a major cause for non-uniformity. If you look at binary strings of length N then the number of strings with $N/3$ ones are only about half as many as those with $1 + N/3$ ones and this is a major source of non-regularity. To address these problems we make the following two modifications.

- For suitable values M_A and M_B, we have M_A copies of each A-table and M_B copies of each B-table. The latter in the form of complementary pairs.
- We pick a random f such that $f(x) = -1$ for exactly a fraction q of all x.

Since each equation contains two variables from the B-tables and one from the A-tables we balance M_A and M_B such that the total size of the B-tables is twice that of the total size of the A-tables. As each question inside an A-table or a B-table is uniformly random this is enough to get regularity. Let us give some details.

If all the selected clauses are disjoint then the size of S_W is $7^{(t+1)k}$ and hence the size of each such B-table is $2^{7^{(t+1)k}}$ and we have $m^{(t+1)k}$ different B-tables. The size of S_U is $7^{tk}2^k$ and if we denote number by N the size of each A-table is $\binom{N}{qN}$. There are $m^{tk}n^k$ different A-tables and using $m = (5n/3)$ it is easy to determine M_A/M_B to make the total size of the B-tables be twice the total size of the A-tables. To make the instance exactly regular we need some additional modifications and let us only sketch these.

Suppose some clauses in W intersect resulting in a S_W of a different size. Let S be the maximal possible size of S_W (which certainly is less than $2^{3(t+1)k}$ and possibly equals $7^{(t+1)k}$), then we make $2^{S-|S_W|}$ copies of this B_W making the total size of all these tables independent of W. The duplication of A tables due to intersecting clauses gets slightly more involved as numbers do not divide as nicely, but there is only a constant number of different table sizes and hence we hope that the reader is convinced that this can be done. From now on we skip this detail and analyze the given PCP.

The analysis of the completeness remains as before and the only notable change is the final calculation of the relative weight of the resulting proof. The A-tables have relative weight q and the B-tables have relative weight $1/2$ and as the latter make up for two thirds of the proof, the relative weight of the overall proof is

$$\frac{2}{3} \cdot \frac{1}{2} + q \cdot \frac{1}{3} = (1+q)/3.$$

and we summarize this in a lemma.

Lemma 8.1 *If φ is satisfiable then there is a proof in the modified regularized PCP that makes V always accept. The relative weight of this proof is $(1+q)/3$.*

Let us look at soundness and fix a pair (U, W). Naturally we define

$$B(g) = \frac{1}{2M_B} \sum_{i=1}^{M_B} \left(B_i^0(g) + B_i^1(g) \right). \tag{16}$$

and as we are equally interested in the bias of A and B we define $e_A = E_f[A(f)]$ and $e_B = E_g[B(f)]$. It is not difficult to see that biasing B comes with an immediate cost.

Lemma 8.2 *We have $\|B\|_2^2 \leq 1 - |e_B|$.*

Proof: Look at the pairs $(B_i^0(g), B_i^1(g))$. As the two tables are complementary, these two bits are represented in different ways. If $B_i^0(g)$ is given by the stored bit in B_i^0 then $B_i^1(g)$ is given by the complement of the bit stored in B_i^1. If the two bits stored are equal then in fact the two values cancel each other in the sum (16). This must happen for a fraction $|e_B|$ of all pairs and hence

$$E_g[|B(g)|] \leq 1 - |e_B|,$$

and as $|B(g)| \leq 1$ we have

$$E_g[B(g)^2] \leq E_g[|B(g)|] \leq 1 - |e_B|.$$

■

An important fact used in the soundness proof is that $\chi_\alpha(f)$ is close to unbiased when $|\alpha|$ is large. This is no longer true as for instance when $\alpha = S_U$ the value of $\chi_\alpha(f)$ is constant and in general, α and its complement give the same value for $|\chi_\alpha(f)|$. This proves that very large α give large expectations but also tells us that it is enough to analyze α that contains at most half the elements. To study the relevant expectations we recall the Krawtchouk polynomials.

Suppose we have N Boolean variables, x, then we have let

$$K_\ell^N(s) = \sum_{|\alpha|=\ell} \chi_\alpha(x)$$

where x is any input with s coordinates equal to -1. As we sum over all α of size ℓ, the value only depends on the number of ones in x and hence this is well defined. Note that

$$\binom{N}{\ell}^{-1} K_\ell^N(s) = \binom{N}{s}^{-1} K_s^N(\ell).$$

The left hand side is the expectation of $\chi_\alpha(x)$ when α is of size ℓ and x has weight s. On the right hand side ℓ and s trade places but as $\chi_\alpha(x)$ is symmetric with respect to α and x this gives the same value. From the definitions we get the following lemma.

Lemma 8.3 *Suppose $|\alpha| = \ell$, then*

$$E_f[\chi_\alpha(f)]| = \binom{|S_U|}{\ell}^{-1} K_\ell^{|S_U|}(q|S_U|) = \binom{|S_U|}{q|S_U|}^{-1} K_{q|S_U|}^{|S_U|}(\ell).$$

The size of the Krawtchouk polynomials is well studied and one source is [16]. Let us only give the main points. If q and ℓ are fixed then

$$\lim_{N\to\infty} \binom{N}{\ell}^{-1} K_\ell^N(qN) = (1-2q)^\ell.$$

Furthermore

$$\binom{N}{qN}^{-1} K_{qN}^N(\ell)$$

is a decreasing function from $\ell = 0$ to the first zero of K_{qN} that appears at $\ell = \frac{N}{2} - (1+o(1))N\sqrt{q(1-q)}$ after which it remains $o(1)$ (as function of N) until $\ell = \frac{N}{2}$. We note that the stated properties are, up to $o(1)$, the same properties as used of $E[\chi_\alpha(f)]$ in the proof of Lemma 6.2.

We return to the case of estimating $E[A(f)C(f)]$ in the modified PCP. We use a similar split $C = C^\ell + \tilde{C}$ as in the previous proof but let C^ℓ also contain the Fourier terms of size at least $|S_U| - \ell$. We let $\tilde{c}$ be sum of the Fourier coefficients of $\tilde{C}$. Note that, by Lemma 8.2, $c^\ell + \tilde{c} \le 1 - |e_B|$. We now have the following Lemma 6.2 using the above stated properties of the Krawtchouk polynomials.

Lemma 8.4 *In the modified regular case, we have*

$$E_f[A(f)\tilde{C}(f)] \le \tilde{c}\min(1, 1 + (1-2q)^{\ell+2} + o(1) - e_A, 1 + o(1) + e_A).$$

To estimate $E[A(f)C^\ell(f)]$ we proceed as in the previous proof and one complication we need to address is to estimate the probability that $\pi_2(\beta)$ is very large and in particular when it contains all but at most ℓ elements of S_U. This is only a small extension of Lemma 6.6.

Lemma 8.5 *Suppose β is a set of odd size and that there is some U such that $\pi_2(\beta)$ is of size at least $|S_U| - \ell$ where $\ell < k$. Then there is a set, S^β, of size at most ℓ such that*

$$Pr[|\pi_2(\beta)| \ge |S_U| - \ell \wedge \pi_2(\beta) \ne S_U/\pi(S^\beta)] \le O(\ell/t).$$

Proof: The proof is very similar to the proof of Lemma 6.6. We use that some sets are of even size and then a subset is of odd size iff its complement is off odd size.

Take any U such that $\pi_2(\beta)$ is of minimal size conditioned on it being of size at least $|S_U| - \ell$. Suppose that $\pi_2(\beta) = S_U - \alpha'$ where $|\alpha'| \leq \ell$. Take a set U' splitting α' as given by Lemma 6.4. Suppose the ith element of α' takes the value x_i on U' where, by construction, $x_i \neq x_j$ for $i \neq j$. Let us look at set the $y \in \beta$ such that $\pi_{U'}(y) = x_i$ and let z denote the coordinates of y on $U \setminus U'$ writing $y = (x, z)$. Define z_i such that (x_i, y_i) is the element in α' with $\pi_{U'}(y) = x_i$. Since $\pi_2(\beta) = S_U - \alpha'$ for $z \neq z_i$, we have an odd number of $y \in \beta$ such that $\pi_U(y) = (x_i, z)$ while there is an even number of y such that $\pi_U(y) = (x_i, z_i)$. As $\ell < k$, we have that $U \setminus U'$ contains at least one of the singleton variables of U. This implies that the number of possible values of z is even and hence the number of z with $z \neq z_i$ is odd and hence the number $y \in \beta$ with $\pi_{U'}(y) = x_i$ is odd. Define z^i as in the proof of Lemma 6.6, namely z^i_j is such that there is an odd number of $y \in \beta$ with $\pi_{U'}(y) = x_i$ and jth coordinate equal to z^i_j.

Now consider any other partner U_1 of W that contains U' and suppose that $|\pi_2(\beta)| \geq |S_U| - \ell$. For any i there is an odd number of elements in $\pi_2(\beta)$ such that $\pi_{U'}(y) = x_i$. As there is an even number of elements x in S_{U_1} with $\pi_{U'}(x) = x_i$ there is some value z such that the number of $y \in \beta$ such that $\pi_{U_1}(y) = (x_i, z)$ is even. In particular there is some element outside $\pi_2(\beta)$ that projects onto x_i. By our choice of U there must be a unique such element. By our choice of z^i_j there is an odd number of elements in $\pi(\beta)$ that project onto x_i and has j coordinate equal to z^i_j. As there is an even number of elements in S_{U_1} that project onto x_i there is also an odd number of elements outside $\pi_2(\beta)$ that project onto x_i and have jth coordinate z^i_j. As there is a unique element outside $\pi_2(\beta)$ that project onto x_i we conclude that the missing element is (x_i, z^i).

As in the proof of Lemma 6.6 the lemma now follows from the fact that the probability that U^1 does not contain U' is $O(\ell/t)$. ■

The other crucial lemma for the soundness is Lemma 6.3 where we show that if $E[A(f)C^\ell(f)]$ is too large then it is possible to extract a good strategy for the two provers. One difference is that since A is only defined for f of relative weight exactly q we do not immediately have the Fourier transform of A in the q-biased basis and hence it is not clear how to extract a strategy for P_2.

For any f not of relative weight q, define $\tilde{f}$ to be a uniformly random element of relative weight q as close as f as possible. In other words if f has more than fraction q of coordinates that are -1 we randomly select a subset of these coordinates to change to 1 in order to make the relative weight exactly q. If it has relative weight below q we instead change a random fraction of the coordinates that are 1. Now define

$$\tilde{A}(f) = E_{\tilde{f}}[A(\tilde{f})].$$

In particular for f of relative weight q we have $\tilde{A}(f) = A(f)$ while for other f it is the average over some nearby points. We claim that $E_f[A(f)C^\ell(f)]$

where f has relative weight exactly q is close to $E_f[\tilde{A}(f)C^\ell(f)]$ where f has each coordinate equal to -1 with probability q. As we know how to extract a strategy in the latter case this is enough. Let us first see that the C^ℓ is likely to be almost the same for f and $\tilde{f}$.

Lemma 8.6 *Suppose f according to μ_q, then*

$$E[|C^\ell(f) - C^\ell(\tilde{f})|] \leq O(\ell/\sqrt{tk}).$$

Proof: By the triangle inequality it is enough to prove this inequality for χ_α where $|\alpha| \leq \ell$. In expectation when we change f to $\tilde{f}$ we modify $O(\sqrt{tk})$ coordinates where each is uniformly chosen. As the probability that any single of these coordinate belongs to α is ℓ/tk we have

$$E_f[|\chi_\alpha(f) - \chi_\alpha(\tilde{f})|] \leq O(\ell/\sqrt{tk}),$$

and the lemma follows. ∎

By definition

$$E_{f\in\mu_q}[\tilde{A}(f)C^\ell(f)]$$

equals

$$E_{f\in\mu_q}[A(\tilde{f})C^\ell(f)],$$

which, by Lemma 8.6, is within $O(\ell/\sqrt{tk})$ of

$$E_{f\in\mu_q}[A(\tilde{f})C^\ell(\tilde{f})],$$

which equals

$$E_f[A(f)C^\ell(f)],$$

when f is chosen to take the value exactly at fraction q of the points. The strategy of P_1 and P_2 based on the A and B tables can now be defined in a very similar way to the proof of Lemma 6.3. One very small difference is that we possibly have to look at sets S^β both in the case when $\pi_2(\beta)$ is of size at most ℓ and when it is of size at least $|S_U| - \ell$. It is not obvious whether the same β can give both small and large projections, and we did not investigate this. The possibility, however, only results in a factor of two as if both are possible we simply choose from each of the two sets with probability $1/2$. Repeating the previous proof we now get the following lemma.

Lemma 8.7 *If the provers can win the two-prover game with probability at most ϵ_1, then in the modified, regular PCP we have $E[A(f)C^\ell(f)] \leq c^\ell(1-2q)\max(0, e_A)(1+o(1)) + \sqrt{2\ell\epsilon_1} + O(\ell/\sqrt{tk})$, where the $o(1)$ is true for fixed ℓ as t and k increases.*

Having established the key lemmas for completeness and soundness we get a hardness result. We only state the result when we are looking for a proof with at most half ones.

Theorem 8.8 *Suppose $\frac{1}{3} < p < \frac{1}{2}$ and set $q = 3p - 1$. Then for any $\delta > 0$, it is NP-hard to distinguish regular instances of 3Lin such that*

- *There is an assignment of relative weight p that satisfies all equations.*
- *For any s, such that $0 \leq s < 1/2$, any assignment of relative weight s gives an advantage as follows.*
  - *If $q \geq \frac{1}{6}$ then the advantage is at most $2s + \delta$*
  - *If $q \leq \frac{1}{6}$ and $s \geq \frac{1}{3}$ then the advantage is at most $1 - 6q(1 - 2s) + \delta$.*
  - *If $q \leq \frac{1}{6}$ and $s \leq \frac{1}{3}$ then the advantage is at most $3s(1 - 2q) + \delta$.*

Proof: The completeness is clear by Lemma 8.1 and let us turn to soundness. As we need to average over all pairs (U, W) let us set $e_B^W = E_g[B_W(g)]$ and define e_A^U similarly. We have a number of error terms that, as in the previous, proof can be incorporated in the error bound δ and let us ignore them. Thus in all calculations from now on we drop any term that goes to 0 when ℓ, k and t tends to infinity. First note that it is never optimal to have a negative e_A^U or e_B^W. This follows as all our bounds are increasing for negative value of the densities and decreasing for positive values. Thus we can always increase the bounds if there are terms of both signs. As the total expectation is positive we can hence assume that all expectations are non-negative.

With the convention of dropping the terms that can be made arbitrarily small, the advantage over a random assignment as given by

$$E_{U,W,f,g_1}[A_U(f)B_W(g_1)B_W(g_2)],$$

and is bounded by

$$E_{U,W}[(1 - |e_B^W|)F_q(e_A^U)] = E_{U,W}[(1 - e_B^W)F_q(e_A^U)], \tag{17}$$

as we assume that e_B^B is non-negative. We bound this using the assumption that the proof has relative weight s which is the same as

$$E_{U,W}[2e_B^W + e_A^U] = 3(1 - 2s). \tag{18}$$

Let us find the solution to this by inspection rather than heavy calculation. We are picking random pairs of table A_U and B_W resulting in random pairs (e_A^U, e_B^W). The pairs have some additional structure as the same number appears in many pairs but let us ignore this and allow arbitrary pairs that satisfy (18) and get an upper bound of (17) valid for all such probability distributions of pairs.

For any fixed choice of the values e_A^U, as the bound is linear in e_B^W, we can assume that (except for one W that we ignore as it gives a small error term) that e_B^W equals 0 or 1.

For any W such that $e_B^W = 1$, and any U that is paired with such a W it is optimal to make $e_A^U = 1$. Indeed increasing any such e_A^U to 1 and decreasing other values usually increases, but certainly does not decrease, (17). By the same argument as above (and using that e_A^U is non-negative) we can assume that e_A^U only takes the values 0 and 1. Assume that $e_B^W = e_A^U = 1$ happens probability e and $e_B^W = 0$ and $e_A^U = 1$ happens with probability e_0. With this notation (18) turns into

$$3e + e_0 = 3(1 - 2s) \tag{19}$$

while the expectation of (18) becomes

$$e_0 F_q(1) + (1 - e_0 - e) = 1 - e - 2qe_0 = 2s + e_0(\frac{1}{3} - 2q),$$

where we used (19) in the last step. If $q \geq \frac{1}{6}$ this is optimized when $e_0 = 0$ giving the first part of the lemma. When $q \leq 1/6$ we want to make e_0 as large as possible but we need to satisfy the constraints $e \geq 0$ and $e + e_0 \leq 1$. When $s \geq \frac{1}{3}$ the optimal solution is $e_0 = 3(1-2s)$ and $e = 0$ giving the optimal value $1 - 6q(1-2s)$. For $s \leq \frac{1}{3}$ the optimal value is $e_0 = 3s$ and $e = 1 - 3s$ giving the optimal value $3s(1-2q)$.

■

Setting $s = p$ we get a corollary.

Corollary 8.9 *Suppose $\frac{7}{18} \leq p < \frac{1}{2}$ and $\delta > 0$. Then its NP-hard to distinguish*

- *Regular instances of 3Lin that has a satisfiable assignment of relative weight p.*
- *Regular instances of 3Lin where any assignment of relative weight p as advantage at most $2p + \delta$.*

For $\frac{1}{3} \leq p \leq \frac{7}{18}$ a similar statement holds with the bound in the soundness replaced by $7 - 30p + 36p^2 + \delta$.

As in non-regular case we can improved bounds for some values of p by padding. A difference here is that if add an variables when we must also add am equations to keep the system regular. The case $p = \frac{1}{2}$ gives nice rational numbers also in this situation.

Theorem 8.10 *Suppose $\delta > 0$. Then its NP-hard to distinguish*

- *Regular instances of 3Lin that has a satisfiable assignment of relative weight $\frac{1}{2}$.*
- *Regular instances of 3Lin where any assignment of relative weight $\frac{1}{2}$ has advantage at most $\frac{9}{11} + \delta$.*

Proof: We start with the system with a solution of relative weight $\frac{7}{18}$ and add $2n/9$ new variables and $2m/9$ new equations on this variables with right hand sides one. This gives a system with $11m/9$ equations which has a solution with relative weight $\frac{1}{2}$.

In the soundness, the two interesting cases are all new variables getting the value one and when half the old variables are one. It is easy to check that in either case the objective value is at most $m(1 + \delta)$.

■

The case for general p is slightly less nice compared to the non-regular case.

Theorem 8.11 *Suppose $\delta > 0$ and $0 \leq p \leq \frac{1}{2}$. Then its NP-hard to distinguish*

- *Regular instances of 3Lin that has a satisfiable assignment of relative weight p.*
- *Regular instances of 3Lin where any assignment of relative weight p has advantage at most $D(p) + \delta$, where $D(p) = 1 - \frac{4p}{7}$ for $0 < p \leq \frac{7}{18}$ and and is upper bounded by $.86$ when $\frac{7}{18} \leq p \leq \frac{1}{2}$.*

Proof: In all cases we start with a system with n variables and m equations and a perfect solution with $7n/18$ variables being true. In the soundness case we only have advantage $\frac{7}{9} + \delta$ and hence falsify $(\frac{1}{9} - \delta/2)m$ equations.

When $p \leq 7/18$, we add $(\frac{7}{18p} - 1)n$ variables and $(\frac{7}{18p} - 1)m$ equations with right hand sides zero. In the soundness case the best solution is to make all these new variables 0 and thus we falsify the same number of equations. This means that we falsify a fraction slightly smaller than $\frac{2p}{7}$ and this gives the advantage stated.

For $\frac{7}{18} \leq p \leq \frac{1}{2}$ we do not have a clean argument and the bound has been verified by computer. The function $D(p)$ that one obtains does not seem to be monotone. ∎

9 Final remarks

Most of the current paper deals with hardness results and hence it would seem natural to complement this by some more algorithmic results. There are two natural and simple algorithms. One was used already in Theorem 7.4 and picks a random element from the set of satisfying assignment and then (greedily or randomly) modifies it to get the correct weight. This is straightforward to analyze in the non-redundant situation. It is not hard to analyze this algorithm in the general case but it does not turn into a nice to state theorem.

Another natural algorithm is to give each variable the value true with probability p to create a solution of relative weight close to p (and later make a small adjustment). Also this is easy to analyze but the result depends on the distribution on the right hand sides.

The instances produced by our hardness reductions are a bit non-symmetric. We have two types of variables, those from A-tables and those from B-tables. While the former are biased the latter are not. The right hand sides of these instances are unbiased due to complementarity property of the B-tables.

Thus apart from the lack of algorithmic results also in hardness analysis there are many potential cases (regular or not, redundant or not, biased right hand sides or not). Maybe we still do not have the most natural class of instances of perfectly satisfiable instances of 3Lin with solution sets of fixed relative weight. Possibly there is no such class but it is always nice to dream of a nice class with matching hardness and algorithmic results.

Looking at other problems, it would be interesting to study the approximability of fixed weight 2Sat on satisfiable instances. Here there is also room for both algorithmic results as well as more explicit hardness reductions.

Acknowledgment. I am grateful to Per Austrin and Pasin Manurangsi for pointing out that it is NP-hard to approximate fixed weight 2Sat on Satisfiable instances. I also thank Per Austrin for many very helpful comments on an early version of the current paper as well as pointing out that [6] can use to eliminate the unique games conjecture in [3].

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan, and M.Szegedy. Proof verification and intractability of approximation problems. *Journal of the ACM*, 45:501–555, 1998.

[2] P. Austrin, S. Benabbas, and K. Georgiou. Better balance by being biased: A 0.8776-approximation for max bisection. *ACM Trans. Algorithms*, 13(1), 2016.

[3] P. Austrin, S. Khot, and M. Safra. Inapproximability of vertex cover and independent set in bounded degree graphs. *Theory of Computing*, 7(1):27–43, 2011.

[4] P. Austrin and A. Stanković. Global Cardinality Constraints Make Approximating Some Max-2-CSPs Harder. In Dimitris Achlioptas and László A. Végh, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*, volume 145 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 24:1–24:17, Dagstuhl, Germany, 2019.

[5] M. Bellare, O. Goldreich, and M. Sudan. Free bits, PCPs and non-approximability—towards tight results. *SIAM Journal on Computing*, 27:804–915, 1998.

[6] A. Bhangale and S. Khot. UG-hardness to NP-hardness by losing half. *Theory of Computing*, 18(5):1–28, 2022.

[7] A. Bhangale, S. Khot, and D. Minzer. On approximability of satisfiable k-CSPs: I. In *Proceedings of the 54th Annual ACM Symposium on Theory of Computing*, pages 976–988, 2022.

[8] A. Bhangale, S. Khot, and D. Minzer. On approximability of satisfiable k-CSPs: II. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 632–6442, 2023.

[9] A. Bhangale, S. Khot, and D. Minzer. On approximability of satisfiable k-CSPs: III. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 643–655, 2023.

[10] A. Bhangale, S. Khot, and D. Minzer. On approximability of satisfiable k-CSPs: IV. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1423–1434, 2024.

[11] A. Bhangale, S. Khot, and D. Minzer. On approximability of satisfiable k-CSPs: V. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 62–71, 2025.

[12] J. Brakensiek, N. Huang, A. Potechin, and U. Zwick. MAX BISECTION might be harder to approximate than MAX CUT, 2025.

[13] A. Bulatov. A dichotomy theorem for nonuniform CSPs. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 319–330, 2017.

[14] M. Goemans and D. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42:1115–1145, 1995.

[15] J. Håstad. Some optimal inapproximability results. *Journal of ACM*, 48:798–859, 2001.

[16] G. Kalai and N. Linial. On the distance distribution of codes. *IEEE Transactions on Information Theory*, 41(5):1467–1472, 2002.

[17] P. Raghavendra. Optimal algorithms and inapproximability results for every csp? In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, page 245–254, New York, NY, USA, 2008. Association for Computing Machinery.

[18] R. Raz. A parallel repetition theorem. *SIAM J. on computing*, 27:763–803, 1998.

[19] D. Zhuk. A proof of the CSP dichotomy conjecture. *Journal of the ACM*, 67(5), 2020.