Electronic Colloquium on Computational Complexity, Report No. 51 (2014)

Hardness of Coloring 2-Colorable 12-Uniform Hypergraphs with $2^{(\log n)^{\Omega(1)}}$ Colors

Subhash Khot* Rishi Saket†

April 12, 2014

Abstract

We show that it is quasi-NP-hard to color 2-colorable 12-uniform hypergraphs with $2^{(\log n)^{\Omega(1)}}$ colors where n is the number of vertices. Previously, Guruswami et al. [GHH+14] showed that it is quasi-NP-hard to color 2-colorable 8-uniform hypergraphs with $2^{2^{\Omega(\sqrt{\log \log n})}}$ colors. Their result is obtained by composing a standard Outer PCP with an Inner PCP based on the Short Code of super-constant degree. Our result is instead obtained by composing a new Outer PCP with an Inner PCP based on the Short Code of degree two.

*New York University, New York, USA. khot@cims.nyu.edu. Research partly supported by NSF Expeditions grant CCF-0832795 and NSF Waterman Award.

†IBM India Research Lab, Banglore, India. rissaket@in.ibm.com.

ISSN 1433-8092

1 Introduction

A k-uniform hypergraph is a collection of vertices and hyperedges where each hyperedge is a subset of k vertices. An independent set in a hypergraph is a subset of vertices that does not contain any hyperedge completely inside it. A hypergraph is said to be q-colorable if the vertices can be partitioned into q disjoint independent sets, or equivalently if the vertices can be colored with q colors so that every edge is non-monochromatic. Coloring a hypergraph using *few* colors is one of the most well studied problems in combinatorics and theoretical computer science.

On graphs (i.e. $k = 2$), there is an efficient algorithm to determine 2-colorability, i.e. bipartiteness. A series of results – [Wig83], [Blu94], [KMS98], [BK97], [ACC06] and [KT12] – give efficient algorithms to color 3-colorable graphs with n^β colors, where the current best value of β is ≈ 0.2038. On the other hand, it is known to be NP-hard to color 3-colorable graphs with 4 colors [KLS00, GK04]. For q-colorable graphs with sufficiently large q, a lower bound of $2^{\Omega(q^{1/3})}$ colors was recently shown by Huang [Hua13], improving upon an earlier bound of $q^{\Omega(\log q)}$ by Khot [Kho01]. Dinur, Mossel and Regev [DMR09] propose a variant of the Unique Games Conjecture referred to as the α-Conjecture and show hardness of coloring 3-colorable graphs with any constant number of colors under this conjecture.

Our understanding is much better for the problem of coloring q-colorable k-uniform hypergraphs with $k \geq 3$ (in this case, even determining 2-colorability is NP-hard). From the algorithmic side, the problem becomes only harder, so the best known algorithms still require $n^{\Omega(1)}$ colors, see Krivelevich et al. [KNS01], Chen and Frieze [CF96] and Kelsen et al. [KMH96]. From the hardness side, there has been steady progress on obtaining stronger and stronger results. We avoid giving a long list of all the known results for different values of q and k and instead refer to the respective papers [Hol02, DRS05, Kho02b, GHH$^+$14, KS14, Sak14]. Here we focus on the case where q and k are allowed to be (preferably small) constants and the concern is obtaining quantitatively strong lower bounds on the number of colors used by efficient algorithms. Guruswami, Håstad and Sudan [GHS02] proved the first superconstant bound, showing hardness of coloring 2-colorable 4-uniform hypergraphs with $\Omega\left(\frac{\log\log n}{\log\log\log n}\right)$ colors. Subsequently, Khot [Kho02a] showed the first poly-logarithmic bound, showing hardness of coloring q-colorable 4-uniform hypergraphs with $(\log n)^{\Omega(q)}$ colors where $q \geq 7$. In recent work, Guruswami et al. [GHH$^+$14] obtained the first super-polylogarithmic bound, showing hardness of coloring 2-colorable 8-uniform hypergraphs with $2^{2^{\Omega(\sqrt{\log\log n})}}$ colors. The main result of this work is a further "exponential" improvement:

Theorem 1.1. *For some absolute constant $c > 0$, it is* quasi-NP-hard[1] *to find an independent set of relative size $2^{-(\log n)^c}$ in an n-vertex 2-colorable 12-uniform hypergraph. Hence, it is* quasi-NP-hard *to color a 2-colorable 12-uniform hypergraph with $2^{(\log n)^c}$ colors. In particular, any $c < \frac{1}{20}$ works.*

We note that all results quoted above, with the exception of [DRS05], also show hardness of finding an independent set of relative size $\delta(n)$ which in turn implies hardness of coloring with $1/\delta(n)$ colors. Our result takes us another step closer to the $n^{\Omega(1)}$ bound, which might perhaps be the truth. We further note that significantly stronger results are known for the case of *almost* coloring: a hypergraph is almost q-colorable if the removal of a small fraction of its vertices and incident hyperedges makes it q-colorable. Given an almost q-colorable graph with $q \geq 3$, it is known to be NP-hard to find an independent set of relative size $q^{-\lfloor \log_2 q \rfloor - 1}$ [DKPS10, KS12] and of relative size $2^{-\frac{q}{2}}$ [Cha13]. Given an almost 2-colorable 4-uniform hypergraph, it is known to be quasi-NP-hard to find an independent set of relative size $2^{-(\log n)^{1-o(1)}}$ [KS14].

Hardness results (including ours) are typically obtained by constructing a *probabilistically checkable proof* (PCP), letting the proof locations be vertices of a hypergraph and letting the tests (or rather the set

[1] A problem is said to be *quasi-NP-hard* if it admits a DTIME($N^{\text{poly}(\log N)}$) reduction from 3SAT.

of proof locations queried in a run of the test) be the hyperedges. Guruswami et al. [GHS02] related the hardness of hypergraph coloring problem to the *covering complexity* of a PCP with the Not-All-Equal predicate. The covering complexity is k if (in the NO case) one needs at least k proofs so that every constraint is satisfied in at least one proof. The number of colors required to color the hypergraph is then 2^k. Dinur and Kol [DK13] study the covering complexity of general predicates. It is easily observed that the covering complexity is at most $O(\log n)$ where the PCP proof has size n and has $\mathsf{poly}(n)$ constraints. This is because if $O(\log n)$ random proofs were constructed, then with high probability over the choice of the proofs, every constraint is satisfied. In terms of covering complexity, ours is the first result to achieve a PCP with covering complexity that is polynomial in $\log n$, specifically $(\log n)^c$. This holds for the Not-All-Equal predicate of arity 12 (optimizing the exponent c and the arity 12 is not the focus of the paper; however the current techniques face a natural barrier of $\frac{1}{2}$ for the exponent c). We consider a new notion called *super-position complexity* of PCPs. Though it resembles the notion of covering complexity, there is no obvious upper bound better than n for the super-position complexity of a PCP. We work with this new notion for most of the paper and in the end show a hardness result for hypergraph coloring problem that amounts to a $(\log n)^c$ covering complexity result.

1.1 Overview of the Proof

Our hardness result follows from a long sequence of reductions, the successive steps presented as Theorems 3.1, 3.2, 3.6, 4.1, 5.2, 6.4, 7.1, 7.2 and 8.1 respectively. It is infeasible to give an overview of all these steps here, so we present only a high level view of some of the steps and emphasize some aspects in which our approach differs from the prior ones, in particular from that of Guruswami et al [GHH$^+$14].

As mentioned before, hardness results (including ours) are typically obtained by constructing a probabilistically checkable proof and letting the proof locations be vertices of a hypergraph and letting the tests be the hyperedges. The PCP is typically viewed as a *composition* of an *outer verifier* with an *inner verifier*. The quantitative strength of the hardness result depends (mainly) on the efficiency of the inner verifier and in particular, on the efficiency (= length) of the *encoding scheme* used by the inner verifier. Several results – such as [GHS02, Hol02, Kho02a, Kho02b, Sak14] – have been obtained using inner verifiers based on the *Long Code*. The Long Code of an m-bit string is a string of length 2^{2^m} and this leads to a large proof (= hypergraph) size, limiting the hardness result to a poly-logarithmic number of colors. At the other end of the spectrum, the *Hadamard Code* of an m-bit string is a string of length 2^m. Using an inner verifier based on the Hadamard Code, Khot and Saket [KS14] obtain a hardness result with $2^{(\log n)^{1-o(1)}}$ colors.[2] However, Hadamard Code can only incorporate (via a technique called *folding*) linear constraints and one is forced to use an underlying NP-hard problem with linear constraints. This forces the PCP to have *imperfect completeness* and one obtains a hardness result only for the *almost* coloring version of the problem. Recently, Barak et al. [BGH$^+$12] proposed a new encoding scheme referred to as the *Short Code* that has length intermediate between the Hadamard Code and the Long Code. To encode an m-bit string u, the Hadamard Code writes down the value of all linear functions on u, whereas the Long Code writes down the value of all functions on u. The Short Code takes an intermediate route and writes down the value of all degree d functions on u for some constant d. The length of the encoding is $\approx 2^{m^d}$ and even though much less than the Long Code, it does increase rapidly for higher degree d. For $d \geq 2$, it allows one to incorporate (via folding) non-linear constraints and hence a PCP with *perfect completeness* is potentially feasible. In a recent work, Dinur and Guruswami [DG13] were indeed able to construct an inner verifier based on the Short Code and

[2] This *almost polynomial factor* is a well-known barrier and should be considered as the best possible bound via the current technology.

obtain hardness results for a variant of the hypergraph coloring problem. Guruswami et al. [GHH$^+$14] were then able to adapt this Short Code based inner verifier for the hypergraph coloring problem, leading to the $2^{2^{\Omega(\sqrt{\log \log n})}}$ bound mentioned before. Their outer verifier is a standard one[3] and its composition with the inner verifier requires using a high degree d, limiting the quantitative bound to $2^{2^{\Omega(\sqrt{\log \log n})}}$ as stated.

Our key idea is to use, at the inner level, a *Quadratic Code* which is same as the Short Code with degree $d = 2$. This leads to a significant saving in the encoding length and we are able to obtain a $2^{(\log n)^{\Omega(1)}}$ bound. However, as we elaborate below, the composition now requires a much stronger guarantee from the outer verifier. The guarantee from the outer verifier is usually in terms of *low soundness*, but we need an additional guarantee that we refer to as the *high super-position complexity* (see below). Much of our effort is then invested in constructing such an outer verifier. We now describe the testing primitive used by the inner verifier and how its analysis motivates (and necessitates) the idea of super-position complexity.

We intend to use the Quadratic Code that encodes an m-bit input $u \in \mathbb{F}[2]^m$ by writing down the values of all quadratic functions on u. This is same as defining an $m \times m$ matrix $M = u \otimes u$ and writing down the values of all linear functions on M (i.e. the Hadamard Code of M). The Quadratic Code is indexed by the set of all $m \times m$ matrices X and the value at location X is given by the entry-wise inner product $\langle M, X \rangle$. We describe a 6-query test to check whether a supposed code is indeed a Quadratic Code (in a loose, *list decoding* sense). It can be adapted, without much additional effort, to a 12-query test of an inner verifier, leading to a hardness result for coloring 12-uniform hypergraphs. This involves reading 6 queries each from two supposed codes and in addition to checking that these are indeed codewords, also checking that these are *consistent*.

The test is as follows. Pick matrices $X, Y, Z \in \mathbb{F}[2]^{m \times m}$ and vectors $a, b \in \mathbb{F}[2]^m$ uniformly and independently at random. Let $\mathsf{Diag}(a)$ be the diagonal matrix with a as the diagonal. Test whether,

$$[C(X) + C(X + \mathsf{Diag}(a))] \cdot [C(Y) + C(Y + \mathsf{Diag}(b))] = C(Z) + C(Z + a \otimes b).$$

It is easy to check that if C is the Quadratic Code of some $u \in \mathbb{F}[2]^m$, then the test always accepts. Indeed, letting $M = u \otimes u$, the right hand side of the equation is ($\langle u, a \rangle$ denotes the inner product over $\mathbb{F}[2]^m$)

$$\langle M, Z \rangle + \langle M, Z + a \otimes b \rangle = \langle M, a \otimes b \rangle = \langle u \otimes u, a \otimes b \rangle = \langle u, a \rangle \cdot \langle u, b \rangle,$$

whereas the left hand side evaluates to the same value:

$$\langle M, \mathsf{Diag}(a) \rangle \cdot \langle M, \mathsf{Diag}(b) \rangle = \langle u, a \rangle \cdot \langle u, b \rangle.$$

On the other hand, it can be shown, by an elementary Fourier analysis, that if the test passes with probability $\frac{1}{2} + 2^{-O(k)}$, then the given C-table can be decoded (by simply outputting a Fourier coefficient with significant magnitude) to a symmetric rank k matrix $\tilde{M}$. Writing $\tilde{M}$ as a super-position (i.e. sum) of k symmetric rank one matrices $\tilde{M} = \sum_{\ell=1}^{k} u^{(\ell)} \otimes u^{(\ell)}$, this amounts to decoding the C-table to a bounded list $u^{(1)}, \ldots, u^{(k)} \in \mathbb{F}[2]^m$ of inputs.[4]

Typically, the inner verifier also needs to check that the input u satisfies a constraint. In our setting, the constraint will be given as a quadratic equation, say $h(u) = 0$ for some quadratic polynomial h (assume for the ease of this overview that h has no constant term). Let's write the constraint as

$$\sum_{i,j=1}^{m} h_{i,j} u_i u_j = 0.$$

[3] By a standard outer verifier we mean the 2-Prover-1-Round Game, a.k.a. Label Cover, instance obtained by parallel repetition of a *clause versus variable* game constructed from a Gap3SAT instance [ABSS97, BGS98, Hås01].

[4] To express a symmetric rank k matrix as a sum of symmetric rank one matrices needs up to $\frac{3k}{2}$ summands, see Lemma 2.1. We ignore this small issue here.

This amounts to a linear constraint $\langle H, M \rangle$ on the matrix $M = u \otimes u$ where $H = (h_{i,j})$ is also a matrix. If C is the Quadratic Code of u such that $h(u) = 0$, then it satisfies $C(X + H) = C(X)$ for every index X. We can ensure that the supposed code always satisfies this property by identifying the proof locations corresponding to $X + H$ and X for every index X. Also, since $M = u \otimes u$ is symmetric, one expects $C(X) = C(X^{\mathsf{T}})$ and this property can be ensured similarly. This trick is known as folding and its consequence is that the decoded matrix $\tilde{M}$ (as described above, by outputting a significant Fourier coefficient of the given C-table) is symmetric and satisfies the constraint $\langle H, M \rangle = 0$. Since $\tilde{M} = \sum_{\ell=1}^{k} u^{(\ell)} \otimes u^{(\ell)}$, this amounts to saying that

$$\sum_{i,j=1}^{m} h_{i,j} \left(\sum_{\ell=1}^{k} u_i^{(\ell)} u_j^{(\ell)} \right) = 0. \tag{1}$$

We say that the quadratic equation $h = 0$ is satisfied *in super-position* by the k inputs $u^{(1)}, \ldots, u^{(k)}$. In summary, the analysis of the inner verifier furnishes a short list of inputs that together satisfy the quadratic equation $h = 0$ in super-position, in the sense of Equation (1). This is an aspect in which our PCP differs from all earlier ones. In earlier PCPs, the inner verifier furnishes a short list of inputs such that *every* input in the list satisfies the relevant constraint whereas in our case, the constraint is only satisfied in super-position. To accommodate this weaker guarantee furnished by the inner verifier, the outer verifier needs a correspondingly stronger guarantee, which we refer to as the high super-position complexity.

We hope it is now clear why we need the outer verifier to have both the low soundness and high super-position complexity. We elaborate further on the latter property. As is standard, the outer verifier can be viewed as a 2-prover-1-round game where the first prover's answer is $u \in \mathbb{F}[2]^m$ and the second prover's answer is $v \in \mathbb{F}[2]^r$ (where $r \leq m$). The verifier accepts if $\pi(u) = v$ for some *projection map* $\pi : \mathbb{F}[2]^m \mapsto \mathbb{F}[2]^r$ that happens to be linear in our setting. In addition, the answer u must satisfy a quadratic equation $h(u) = 0$ for the verifier to accept. In the YES case, the provers have a strategy that makes the verifier accept with probability 1. In the NO case, the verifier accepts with negligible probability even under a looser criterion for acceptance. The provers are now allowed to furnish a short list $u^{(1)}, \ldots, u^{(k)}$ and $v^{(1)}, \ldots, v^{(k)}$ of answers respectively and the verifier accepts if $\pi(u^{(\ell)}) = v^{(\ell)}$ $\forall \ell \in \{1, \ldots, k\}$ and that $u^{(1)}, \ldots, u^{(k)}$ satisfy the constraint $h = 0$ in super-position. Once we have an outer verifier with such a guarantee, it is straightforward to compose it with the inner verifier described above.

The formal description of the outer verifier appears as Theorem 7.1 and the bulk of our paper is devoted to proving this theorem. It follows via a sequence of reductions (= PCPs), the successive steps presented as Theorems 3.1, 3.2, 3.6, 4.1, 5.2 and 6.4. We focus on constraint satisfaction problems where the constraints are quadratic equations over $\mathbb{F}[2]$. The super-position complexity of a CSP instance is the minimum number of assignments that satisfy every constraint in super-position in the sense of Equation (1). We start by showing that it is NP-hard to distinguish whether a CSP has super-position complexity of 1 or at least k (we choose the parameter k to be poly-logarithmic in the instance size though the result also holds for much higher settings of the parameter). This appears as Theorems 3.1 and 3.2. Interestingly, we do use some of the techniques from Dinur and Guruswami [DG13] here, specifically Lemma 2.3 which in turn is based on techniques from [BKS$^+$10] to test Reed-Muller codes over $\mathbb{F}[2]$. However, we emphasize that Dinur and Guruswami [DG13] employ these techniques in the analysis of the inner verifier whereas for us, these serve as a starting point in a long sequence of reductions.[5] We then use the ingredients used to prove the PCP

[5]One may view Dinur and Guruswami reduction as a sequence of four steps: NP-hardness of 3SAT (= Cook-Levin Theorem), NP-hardness of Gap3SAT (= the PCP Theorem), the Outer PCP and the inner PCP. With this viewpoint, the techniques referred to, are used by Dinur and Guruswami at the inner PCP level whereas we use them to prove the analogue of the Cook-Levin Theorem. We then naturally proceed to prove the analogue of the PCP Theorem.

Theorem (sum-check protocol, low degree test etc) to simultaneously reduce the *arity* of the constraints and to achieve *low soundness*, while preserving the high super-position complexity at every step. In the last step, the constraints are those given by a point-versus-surface low degree test and is naturally viewed as a 2-prover-1-round game, i.e. as the outer verifier. As mentioned before, the inner PCP is then based on the Quadratic Code. Its analysis is elementary and does not use any of the machinery required to analyze the Short Code.

2 Preliminaries

This section describes some useful tools that are used in subsequent sections.

2.1 Tensor Decomposition of Symmetric Matrices

The following lemma shows a canonical way to write a symmetric matrix as a sum of symmetric rank one matrices. We only consider matrices over a field $\mathbb{F}[q]$ of characteristic 2.

Lemma 2.1. *Given a symmetric matrix $A \in \mathbb{F}[q]^{m\times m}$ of rank k over a field $\mathbb{F}[q]$ of characteristic 2, there are k linearly independent vectors $\overline{z}_1, \ldots, \overline{z}_k \in \mathbb{F}[q]^m$ from the column space of A such that,*

$$A = \sum_{i=1}^{s} \overline{z}_i \otimes \overline{z}_i + \sum_{j=1}^{t} \left(\overline{z}_{s+2j-1} \otimes \overline{z}_{s+2j} + \overline{z}_{s+2j} \otimes \overline{z}_{s+2j-1} \right) \tag{2}$$

$$= \sum_{i=1}^{s} \overline{z}_i \otimes \overline{z}_i + \sum_{j=1}^{t} \Big(\overline{z}_{s+2j-1} \otimes \overline{z}_{s+2j-1} + \overline{z}_{s+2j} \otimes \overline{z}_{s+2j} + (\overline{z}_{s+2j-1} + \overline{z}_{s+2j}) \otimes (\overline{z}_{s+2j-1} + \overline{z}_{s+2j}) \Big), \tag{3}$$

where $k = s + 2t$ for some non-negative integers s and t. In particular, A is a sum of at most $\frac{3k}{2}$ symmetric rank one matrices.

Proof. Note that the second equation in the statement of the lemma follows from the first by observing that $\overline{a} \otimes \overline{b} + \overline{b} \otimes \overline{a} = \overline{a} \otimes \overline{a} + \overline{b} \otimes \overline{b} + (\overline{a} + \overline{b}) \otimes (\overline{a} + \overline{b})$. So we focus on obtaining the decomposition as in the first equation. If $A = 0$, there is nothing to prove. If $A = (a_{ij}) \neq 0$, then we consider two cases and in each case, we give a decomposition of A into a single term in Equation (2) and a matrix of lower rank A'. The lemma then follows by an inductive argument on A'. We use a crucial fact that in a field $\mathbb{F}[q]$ of characteristic 2, every element is a square. In particular, for any $a \in \mathbb{F}[q]$, $a \neq 0$, the element $\frac{1}{\sqrt{a}}$ exists.

Case (i): Consider the case when A has a non-zero diagonal element, i.e. $a_{ii} \neq 0$ for some $i \in \{1, \ldots, m\}$. Let $\overline{a}_i$ be the i^{th} column of A and let $\overline{b}_i = \frac{1}{\sqrt{a_{ii}}} \cdot \overline{a}_i$. Consider the symmetric matrix,

$$A' = A + \overline{b}_i \otimes \overline{b}_i.$$

It is easy to see that the i^{th} column as well as row of A' is zero. This implies that $\overline{b}_i$ is linearly independent of the columns of A' and $\mathsf{rank}(A') = \mathsf{rank}(A) - 1$. We can then inductively decompose A' keeping in mind that the decomposition will involve vectors that are linearly independent of $\overline{b}_i$.

Case (ii): Now consider the case when all diagonal elements of A are zero, but there are indices $i \neq j$ such that $a_{ij} = a_{ji} \neq 0$. As before, let $\overline{b}_i = \frac{1}{\sqrt{a_{ij}}} \cdot \overline{a}_i$ and $\overline{b}_j = \frac{1}{\sqrt{a_{ij}}} \cdot \overline{a}_j$. Since $a_{ii} = a_{jj} = 0$, we have $\overline{b}_i \neq \overline{b}_j$. Consider the symmetric matrix

$$A' = A + \overline{b}_i \otimes \overline{b}_j + \overline{b}_j \otimes \overline{b}_i.$$

The i^{th} and the j^{th} columns as well as rows of A' are zero. This implies that $\overline{b}_i$ and $\overline{b}_j$ are linearly independent of the columns of A' and $\mathsf{rank}(A') = \mathsf{rank}(A) - 2$. We can then inductively decompose A' keeping in mind that the decomposition will involve vectors that are linearly independent of $\overline{b}_i$ and $\overline{b}_j$. □

2.2 Representations of Monomial Assignments

This and the next section describe the basic setup used by Dinur and Guruswami [DG13] for analyzing their inner verifier. Their verifier relies on the Short Code (we do not define it here since we won't be using it) that was proposed and analyzed by Barak et al. [BGH$^+$12].

Let $x_1, \ldots, x_m$ be variables over $\mathbb{F}[2]$. Fix a degree parameter $d \geq 1$ and let $\mathcal{S}_d$ be the set of all monomials $\prod_{i \in S} x_i$ corresponding to non-empty subsets $S \subseteq [m]$ of size at most d. An assignment $\sigma : \mathcal{S}_d \mapsto \mathbb{F}[2]$ is referred to as a *monomial assignment*. One can naturally extend assignment σ to all polynomials of degree at most d by linearity, i.e. if $q(x) = c + \sum_{S \subseteq [m], 1 \leq |S| \leq d} c_S \prod_{i \in S} x_i$ is a polynomial, then

$$\sigma(q) = c + \sum_{S \subseteq [m], 1 \leq |S| \leq d} c_S \cdot \sigma\left(\prod_{i \in S} x_i\right).$$

Lemma 2.2. *For any monomial assignment $\sigma : \mathcal{S}_d \mapsto \mathbb{F}[2]$, there is a subset $\beta \subseteq \mathbb{F}[2]^m$ such that for all polynomials $q(x)$ of degree at most d,*

$$\sigma(q) = \sum_{a \in \beta} q(a). \tag{4}$$

Proof. Let $\mathcal{P}_d$ be the linear vector space of all polynomials $q(x)$ of degree at most d. The dimension of this space equals the number of monomials (including the empty monomial), i.e. $\sum_{i=0}^{d} \binom{m}{i}$. Let $\mathcal{A}$ be the set of all inputs $a \in \mathbb{F}[2]^m$ with Hamming weight at most d so that $|\mathcal{A}| = \sum_{i=0}^{d} \binom{m}{i}$ is same as the dimension of $\mathcal{P}_d$. For every fixed $a \in \mathbb{F}[2]^m$, the map $q(x) \mapsto q(a)$ is a linear map on $\mathcal{P}_d$. We will show that these maps are linearly independent and hence form a basis for the space of all linear maps on $\mathcal{P}_d$ and in particular, the linear map σ can be expressed as their linear combination, proving the lemma. In order to show the linear independence of the maps $\{q(x) \mapsto q(a) \mid a \in \mathcal{A}\}$, it suffices to show that if a degree (at most) d polynomial $q(x)$ vanishes on all inputs in $\mathcal{A}$, then it vanishes identically. Indeed, if on the contrary, $q(x) \neq 0$, then $q(x) = \prod_{i \in S} x_i + \sum_{S' \neq S} c_{S'} \prod_{j \in S'} x_j$ where $\prod_{i \in S} x_i$ is a monomial of highest degree that has a non-zero coefficient in $q(x)$. Clearly, for the input $a \in \mathbb{F}[2]^m$ whose non-zero co-ordinates are precisely on the set S, we have $q(a) \neq 0$ reaching a contradiction. □

Note that the subset β guaranteed by Lemma 2.2 need not be unique. For a monomial assignment $\sigma : \mathcal{S}_d \mapsto \mathbb{F}[2]$, let β_σ denote a minimum sized subset β satisfying the conclusion of the lemma (i.e. Equation (4)).

2.3 A Useful Tool from Dinur and Guruswami [DG13]

We now state (a minor variant of) the main tool we borrow from Dinur and Guruswami [DG13] paper. Let $\mathcal{F}_m$ be the space of all functions $f : \mathbb{F}[2]^m \mapsto \mathbb{F}[2]$. For a subset $\beta \subseteq \mathbb{F}[2]^m$, define the character

$\chi_\beta : \mathcal{F}_m \mapsto \{-1, 1\}$ as:

$$\chi_\beta(f) = (-1)^{\sum_{x\in\beta} f(x)} = (-1)^{\sum_{x\in\mathbb{F}[2]^m} \mathbb{1}_\beta(x)f(x)} = (-1)^{\langle \mathbb{1}_\beta, f\rangle},$$

where $\mathbb{1}_\beta$ denotes the indicator function of that subset. If β_σ is a (minimum sized) subset corresponding to a monomial assignment σ as defined earlier, then for any polynomial g of degree at most d,

$$\chi_{\beta_\sigma}(g) = (-1)^{\sum_{x\in\beta_\sigma} g(x)} = (-1)^{\sigma(g)}.$$

The following is a minor variant of a theorem proved in [DG13]. The ideas in its proof go back to the analysis of testing Reed-Muller codes in [BKS+10].

Lemma 2.3. *Let $\beta = \beta_\sigma$ be a (minimum sized) subset corresponding to some monomial assignment σ such that $|\beta| \geq 2^{d/2}$ and $\alpha, \gamma \subseteq \mathbb{F}[2]^m$ are arbitrary. Then*

$$\left|\mathbb{E}_{g,h}\left[\chi_\beta(gh)\chi_\gamma(g)\chi_\alpha(h)\right]\right| \leq 2^{-2^{d/4-2}+1},$$

where g is a uniformly random polynomial of degree at most $3d/4$ and h is a uniformly random polynomial of degree at most $d/4$ with no constant term.

Proof. The expectation can be upper bounded by

$$\mathbb{E}_h\left[\left|\mathbb{E}_g\left[\chi_\beta(gh)\chi_\gamma(g)\right]\right|\right]. \tag{5}$$

The inner expectation is same as

$$\mathbb{E}_g\left[\chi_\beta(gh)\chi_\gamma(g)\right] = \mathbb{E}_g\left[(-1)^{\langle \mathbb{1}_\beta\cdot h+\mathbb{1}_\gamma, g\rangle}\right]. \tag{6}$$

We use the fact that the space of polynomials of degree at most $m - 3d/4 - 1$ is precisely the orthogonal space of the space of polynomials of degree at most $3d/4$. Thus the expectation in Equation (6) is 1 if $\mathbb{1}_\beta \cdot h + \mathbb{1}_\gamma$ is a polynomial of degree at most $m - 3d/4 - 1$ and zero otherwise. Hence the expression in Equation (5) is same as

$$\Pr_h\left[\mathbb{1}_\beta \cdot h + \mathbb{1}_\gamma \text{ is a polynomial of degree at most } m - 3d/4 - 1\right],$$

where h is a random polynomial of degree at most $d/4$ with no constant term. By Lemma 2.5, this probability is upper bounded by $2^{-2^{d/4-2}+1}$. □

Lemma 2.5 is an immediate consequence of a similar lemma in [DG13].

Lemma 2.4. *For a uniformly random polynomial h of degree at most $d/4$ and β such that $|\beta| \geq 2^{d/2}$,*

$$\Pr_h\left[\mathbb{1}_\beta \cdot h \text{ is a polynomial of degree at most } m - 3d/4 - 1\right] \leq 2^{-2^{d/4-2}}.$$

Lemma 2.5. *For a uniformly random polynomial h of degree at most $d/4$ with no constant term and any γ, β such that $|\beta| \geq 2^{d/2}$,*

$$\Pr_h\left[\mathbb{1}_\beta \cdot h + \mathbb{1}_\gamma \text{ is a polynomial of degree at most } m - 3d/4 - 1\right] \leq 2^{-2^{d/4-2}+1}.$$

Proof. If there is no h such that $\mathbb{1}_\beta \cdot h + \mathbb{1}_\gamma$ is a polynomial of degree at most $m - 3d/4 - 1$ then we are done. Otherwise the set of all such h is an affine subspace and translating it to include the origin yields the subspace (of the same size) of h' such that $\mathbb{1}_\beta \cdot h'$ is a polynomial of degree at most $m - 3d/4 - 1$. An appeal to Lemma 2.4 completes the proof. We may lose a factor of 2 in the probability bound due to conditioning on only those h that have no constant term. □

2.4 Arora-Sudan Analysis of the Low Degree Test

Let $\mathbb{F}[q]$ be a field and d, m be positive integers. Suppose we are given a table of values of a function $f : \mathbb{F}[q]^m \mapsto \mathbb{F}[q]$ that is supposed to be a degree d polynomial. Suppose, in addition, we are given, for every line ℓ in the space $\mathbb{F}[q]^m$, a univariate degree d polynomial f_ℓ that is supposed to be the restriction of the supposed global polynomial f to that line.[6] For a point $\overline{v}$ on the line ℓ, we denote by $f_\ell(\overline{v})$ the value given by f_ℓ at the point $\overline{v}$. The following theorem was proved by Arora and Sudan [AS03].

Theorem 2.6. *There are constants $c_0, c_1, c_2, c_3 > 0$ such that the following holds. For any parameter $\delta > 0$ such that $q \geq c_0(dm/\delta)^{c_1}$, let $\{P_1, \ldots, P_t\}$ be the set of degree d polynomials that agree with f at δ^{c_2}/c_3 fraction of the points. Then, taking the probability over a random line ℓ and random point $\overline{v}$ on the line,*

$$\Pr_{\ell, \overline{v}} [f(\overline{v}) \notin \{P_1(\overline{v}), \ldots, P_t(\overline{v})\} \text{ and } f_\ell(\overline{v}) = f(\overline{v})] \leq \delta.$$

Also, by coding theoretic bounds $t \leq 2c_3/\delta^{c_2}$.

2.5 Super-position Complexity

Definition 2.7. *Let $a^{(1)}, \ldots, a^{(t)} \in \mathbb{F}[2]^m$ be t assignments and $q(x) = 0$ be a quadratic equation in m boolean variables with $q(x) = c + \sum_{i=1}^m c_i x_i + \sum_{1 \leq i < j \leq m} c_{ij} x_i x_j$. We say that the t assignments satisfy the equation $q(x) = 0$ in super-position if*

$$c + \sum_{i=1}^m c_i \left(\sum_{\ell=1}^t a_i^{(\ell)} \right) + \sum_{1 \leq i < j \leq m} c_{ij} \left(\sum_{\ell=1}^t a_i^{(\ell)} a_j^{(\ell)} \right) = 0.$$

Note that for $t = 1$, this is same as saying that $q(a^{(1)}) = 0$, i.e. that $a^{(1)}$ satisfies the equation (in the standard sense). Also, if $q(x)$ is linear, this is same as saying that the assignment $a = \sum_{\ell=1}^t a^{(\ell)}$ satisfies the equation (in the standard sense).

Definition 2.8. *Given a system of quadratic equations $\{q_i(x) = 0\}_{i=1}^L$, its super-position complexity is the minimum number t, if it exists, such that there are t assignments $a^{(1)}, \ldots, a^{(t)} \in \mathbb{F}[2]^m$ that satisfy every equation $q_i(x) = 0,\ i \in \{1, \ldots, L\}$ in super-position. Otherwise, one may define the super-position complexity to be ∞ (but we will not encounter this scenario).*

3 Starting Point for Our PCPs

In this section, we describe a set of results that serve as the starting point for our PCPs. The main theorem is Theorem 3.2 that provides a *super-position gap* for constraint satisfaction problems with constraints that are quadratic equations over $\mathbb{F}[2]$. The theorem states that given an instance of such a CSP, it is NP-hard to distinguish whether it has a satisfying assignment (i.e. has super-position complexity of 1) or has high super-position complexity. Theorem 3.1 is a preparatory step towards the main Theorem 3.2. For subsequent applications, we need certain strengthenings of the main theorem stated as Theorem 3.4 and 3.6.

[6] A line is a set $\ell(t) = \overline{\alpha} + t\overline{\beta}$ parameterized by $t \in \mathbb{F}[q]$ for some $\overline{\alpha}, \overline{\beta} \in \mathbb{F}[q]^m$.

3.1 CSPs with High Degree Equations

Recall that given n boolean variables $x_1, \ldots, x_n$ and the degree parameter d, $\mathcal{S}_d$ denotes the set of all (non-empty) monomials of size at most d over the n variables. Given a monomial assignment $\sigma : \mathcal{S}_d \mapsto \mathbb{F}[2]$, one can extend it naturally to all polynomials of degree at most d by linearity. Moreover there exists a set $\beta_\sigma \subseteq \mathbb{F}[2]^n$ (of minimal size, by definition) such that for all polynomials $q(x)$ of degree at most d, $\sigma(q) = \sum_{s \in \beta_\sigma} q(s)$. A monomial assignment σ is said to satisfy a system of degree d polynomial equations $\{q_i(x) = 0\}_{i=1}^m$ if $\sigma(q_i) = 0$ for every $i \in \{1, \ldots, m\}$. We prove the following theorem in this section.

Theorem 3.1. *For any $d \geq 3$, there is a DTIME($n^{O(d)}$) reduction from 3SAT to a system $\mathcal{B}$ of degree d equations over $\mathbb{F}[2]$ such that,*

YES Case: If the 3SAT instance is satisfiable then there is an assignment that satisfies (all equations in) $\mathcal{B}$.

NO Case: If the 3SAT instance is unsatisfiable then for any monomial assignment $\sigma : \mathcal{S}_d \mapsto \mathbb{F}[2]$ that satisfies (all equations in) $\mathcal{B}$, one must have $|\beta_\sigma| \geq 2^{d-3}$.

Proof. Suppose the 3SAT instance consists of n boolean variables $x_1, \ldots, x_n$ and m clauses. For $i = 1, \ldots, m$, the i^{th} clause can be written as an equation $p_i(x) = 0$ where $p_i(x)$ is a polynomial of degree at most 3. It depends on at most 3 variables, but this will not be relevant to us. Let $d \geq 3$ be as in the statement of the theorem. We construct a system $\mathcal{B}$ of equations as desired by adding the equation

$$\left(\prod_{i \in S} x_i\right) p_i(x) = 0,$$

for all monomials $\prod_{i \in S} x_i$ of degree at most $d - 3$ and $i = 1, \ldots, m$. Note that every equation in $\mathcal{B}$ has degree at most d. In the YES case, if the 3SAT instance has a satisfying assignment, then clearly the same assignment satisfies all equations in $\mathcal{B}$. So we focus on the NO case. Let a monomial assignment $\sigma : \mathcal{S}_d \mapsto \mathbb{F}[2]$ be given that satisfies all equations in $\mathcal{B}$ and let $\beta_\sigma \subseteq \mathbb{F}[2]^n$ be the corresponding set. Note that for any polynomial $q(x)$ of degree at most $d - 3$ and any $i \in \{1, \ldots, m\}$, the equation $q(x)p_i(x) = 0$ is a linear combination of equations in $\mathcal{B}$ and hence must be satisfied by σ, i.e. $\sigma(q \cdot p_i) = 0$.

Assume for the sake of contradiction that $|\beta_\sigma| < 2^{d-3}$. Fix an arbitrary $a \in \beta_\sigma$. By Lemma 2.13 of [GHH$^+$14], there exists a polynomial $q(x)$ of degree at most $d - 3$ such that $q(a) = 1$ and $\forall b \in \beta_\sigma, b \neq a, q(b) = 0$. Since the 3SAT instance is unsatisfiable, the assignment a fails on some, say j^{th}, clause, i.e. $p_j(a) = 1$. We reach a contradiction by observing that

$$\sigma(q \cdot p_j) = \sum_{s \in \beta_\sigma} q(s)p_j(s) = q(a)p_j(a) + \sum_{s \in \beta_\sigma, s \neq a} q(s)p_j(s) = p_j(a) = 1.$$

□

3.2 Quadratic CSP with Superposition Gap

We recall Definition 2.7 and prove our main theorem in this section.

Theorem 3.2. *There is a reduction from* 3SAT *to an instance $\mathcal{A}$ of quadratic equations such that,*

YES Case. *If the* 3SAT *instance is satisfiable then there is an assignment to $\mathcal{A}$ that satisfies all the equations.*

NO Case. *If the* 3SAT *instance is unsatisfiable then there are no t assignments to $\mathcal{A}$ that satisfy all the equations simultaneously in super-position for any $1 \leq t \leq k$. Here k is a parameter and the reduction runs in time $N^{O(\log k)}$ where N is the size of the* 3SAT *instance.*

Proof. We first reduce 3SAT to a system of degree d equations $\mathcal{B}$ as given by Theorem 3.1. The size of instance $\mathcal{B}$ is $N^{O(d)}$ where N is the size of the 3SAT instance. Let $x_1, \ldots, x_n$ be the variables of the instance $\mathcal{B}$ and the degree parameter d will be set later. Note that in the YES case, the instance $\mathcal{B}$ has a satisfying assignment $a \in \mathbb{F}[2]^n$ whereas in the NO case, for any assignment $\sigma : \mathcal{S}_d \mapsto \mathbb{F}[2]$ that satisfies $\mathcal{B}$, it must be that $|\beta_\sigma| \geq 2^{d-3}$ (to recall again, $\mathcal{S}_d$ is the set of all (non-empty) monomials over variables $x_1, \ldots, x_n$ of degree at most d). We construct the desired system $\mathcal{A}$ of quadratic equations as follows.

- For every $A \subseteq [n]$, $1 \leq |A| \leq d$ we have a variable y_A. This variable is supposed to represent the monomial $\prod_{i \in A} x_i$ and in the YES case, it takes the same value as this monomial under a satisfying assignment to $\mathcal{B}$.
- Add all the equations of $\mathcal{B}$ replacing each monomial $\prod_{i \in A} x_i$ by the corresponding variable y_A. These equations are linear in the variables $\{y_A \mid 1 \leq |A| \leq d\}$ (so this is simply a linearization of $\mathcal{B}$).
- For every pair $A, B \subseteq [n]$ such that $1 \leq |A|, |B|, |A \cup B| \leq d$, add the quadratic equation $y_A y_B = y_{A \cup B}$. Note that this quadratic equation is indeed satisfied in the YES case since the variables y_A have values same as the corresponding monomials under an assignment to the variables $x_1, \ldots, x_n$.

This completes the construction of the instance $\mathcal{A}$. In the YES case, taking the satisfying assignment $a \in \mathbb{F}[2]^n$ to instance $\mathcal{B}$ and assigning to every variable y_A the value $\prod_{i \in A} a_i$ satisfies all equations of instance $\mathcal{A}$. In the NO case, we wish to show that no t assignments $\sigma_1, \ldots, \sigma_t : \{y_A \mid 1 \leq |A| \leq d\} \mapsto \mathbb{F}[2]$ can satisfy all equations of $\mathcal{A}$ in super-position for any $1 \leq t \leq k$. Assume on the contrary that this is the case. Note that any assignment σ_i is naturally also an assignment $\sigma_i : \mathcal{S}_d \mapsto \mathbb{F}[2]$. Hence there exists a corresponding set $\beta_{\sigma_i} \subseteq \mathbb{F}[2]^n$ as in Lemma 2.2. Let $\sigma = \sum_{i=1}^t \sigma_i$ so that $\beta_\sigma = \bigoplus_{i=1}^t \beta_{\sigma_i}$ (here $\oplus$ is the symmetric difference operator on sets; strictly speaking, β_σ is the set of minimal size that is *equivalent to* $\bigoplus_{i=1}^t \beta_{\sigma_i}$).

The equations of $\mathcal{A}$ are either those obtained by linearization of $\mathcal{B}$ or those of the type $y_A y_B = y_{A \cup B}$. The equations of the first kind are linear in the variables $\{y_A\}$ and since $\sigma_1, \ldots, \sigma_t$ satisfy these equations in super-position, the assignment $\sigma = \sum_{i=1}^t \sigma_i$ also satisfies all these equations. In other words, $\sigma : \mathcal{S}_d \mapsto \mathbb{F}[2]$ satisfies the instance $\mathcal{B}$ and hence by the guarantee offered by the NO case of Theorem 3.1, we have $|\beta_\sigma| \geq 2^{d-3}$. We now show that $|\beta_\sigma|$ being large implies that $\sigma_1, \ldots, \sigma_t : \{y_A \mid 1 \leq |A| \leq d\} \mapsto \mathbb{F}[2]$ cannot simultaneously satisfy the equations $y_A y_B = y_{A \cup B}$ in super-position. Assume on the contrary that this is the case, i.e. for all A, B such that $1 \leq |A|, |B|, |A \cup B| \leq d$,

$$\sum_{i=1}^{t} \sigma_i(y_{A \cup B}) = \sum_{i=1}^{t} \sigma_i(y_A)\sigma_i(y_B).$$

Since σ_i are also thought of as monomial assignments $\sigma_i : \mathcal{S}_d \mapsto \mathbb{F}[2]$, the above amounts to saying that

$$\sum_{i=1}^{t} \sigma_i(gh) = \sum_{i=1}^{t} \sigma_i(g)\sigma_i(h), \tag{7}$$

where g, h are non-empty monomials in variables $x_1, \ldots, x_n$ such that the sizes of g, h, gh are all upper bounded by d. In particular, this holds whenever g and h are non-empty monomials of degree at most $3d/4$ and $d/4$ respectively (assume d is divisible by 4). We observe that by linearity, Equation (7) holds also when g is a polynomial of degree at most $3d/4$ and h is a polynomial of degree at most $d/4$ with no constant

term. Indeed, suppose $g = c + \sum_A c_A \prod_{j \in A} x_j$ where $1 \leq |A| \leq 3d/4$ and $h = \sum_B c'_B \prod_{j \in B} x_j$ where $1 \leq |B| \leq d/4$. Then

$$\sum_{i=1}^{t} \sigma_i(gh) = \sum_B c \cdot c'_B \sum_{i=1}^{t} \sigma_i \left(\prod_{j \in B} x_j \right) + \sum_{A,B} c_A c'_B \sum_{i=1}^{t} \sigma_i \left(\prod_{j \in A \cup B} x_j \right),$$

which is same as

$$\sum_B c \cdot c'_B \sum_{i=1}^{t} \sigma_i \left(\prod_{j \in B} x_j \right) + \sum_{A,B} c_A c'_B \sum_{i=1}^{t} \sigma_i \left(\prod_{j \in A} x_j \right) \sigma_i \left(\prod_{j \in B} x_j \right),$$

which in turn can be re-written as

$$\sum_{i=1}^{t} \left(c + \sum_A c_A \, \sigma_i \left(\prod_{j \in A} x_j \right) \right) \cdot \left(\sum_B c'_B \, \sigma_i \left(\prod_{j \in B} x_j \right) \right),$$

which equals $\sum_{i=1}^{t} \sigma_i(g)\sigma_i(h)$ as desired. We get back to Equation (7) and switch from values over $\mathbb{F}[2]$ to real values in $\{-1, 1\}$, i.e. replace $\sigma_i(g)$ by $(-1)^{\sigma_i(g)}$. Noting that $\sigma = \sum_{i=1}^{t} \sigma_i$, we get

$$(-1)^{\sigma(gh)} = \prod_{i=1}^{t} \left((-1)^{\sigma_i(g)} \wedge (-1)^{\sigma_i(h)} \right).$$

Note that addition over $\mathbb{F}[2]$ now becomes multiplication over signs $\{-1, 1\}$ and multiplication over $\mathbb{F}[2]$ now becomes the operation $a \wedge b = (1 + a + b - ab)/2$ over signs $\{-1, 1\}$. Since $(-1)^{\sigma_i(g)} = \chi_{\beta_{\sigma_i}}(g)$, we get that

$$\chi_{\beta_\sigma}(gh) \left[\prod_{i=1}^{t} (\chi_{\beta_{\sigma_i}}(g) \wedge \chi_{\beta_{\sigma_i}}(h)) \right] = 1, \tag{8}$$

whenever g is a polynomial of degree at most $3d/4$ and h is a polynomial of degree at most $d/4$ with no constant term. We reach a contradiction by showing that if g and h are chosen as random polynomials of the kind prescribed, the expectation of the left hand side of Equation (8) is nearly zero. Indeed, replacing each expression $a \wedge b$ by $(1 + a + b - ab)/2$ and expanding the product into a sum of 4^t terms, the left hand side of Equation (8) is a sum of 4^t terms of type

$$\left(\frac{1}{2^t} \right) \chi_{\beta_\sigma}(gh) \chi_\gamma(g) \chi_\alpha(h),$$

for some $\gamma, \alpha \subseteq \mathbb{F}[2]^n$. The sets γ, α are related to the sets β_{σ_i}, but this is not relevant for the argument. We finish the proof by showing that the expectation of the term above is negligible and hence the sum of the expectations of the 4^t terms is negligible too. The claim follows by Lemma 3.3 below. It is enough to take $d = O(\log k)$. □

Lemma 3.3. *For $\beta_\sigma \subseteq \mathbb{F}[2]^n$, $|\beta_\sigma| \geq 2^{d-3}$, $d \geq 6$ and arbitrary $\gamma, \alpha \subseteq \mathbb{F}[2]^n$, we have*

$$|\mathbb{E}_{g,h}[\chi_{\beta_\sigma}(gh)\chi_\gamma(g)\chi_\alpha(h)]| \leq 2^{-2^{d/4-2}+1},$$

where g is a random polynomial of degree at most $3d/4$ and h is a random polynomial of degree at most $d/4$ with no constant term.

Proof. For $d \geq 6$, we have $2^{d-3} \geq 2^{d/2}$. The proof follows from Lemma 2.3. □

3.3 Strengthening of Theorem 3.2

We will need to consider quadratic equations over $\mathbb{F}[q]$ that is an extension field of $\mathbb{F}[2]$. In particular we need analogue of Theorem 3.2 where the conclusion holds even for $\mathbb{F}[q]$-valued assignments. In this section, while considering quadratic equations over $\mathbb{F}[q]$, we only consider equations that have $\mathbb{F}[2]$ coefficients and no linear terms, i.e. equations of the form $c + \sum_{1 \le i \le j \le m} c_{ij} x_i x_j = 0$ where $c, c_{ij} \in \mathbb{F}[2]$. The notion of satisfying an equation in super-position is similar as before. Assignments $a^{(1)}, \ldots, a^{(t)} \in \mathbb{F}[q]^m$ are said to satisfy an equation $c + \sum_{1 \le i \le j \le m} c_{ij} x_i x_j = 0$ in super-position if,

$$c + \sum_{1 \le i \le j \le m} c_{ij} \left(\sum_{\ell=1}^{t} a_i^{(\ell)} a_j^{(\ell)} \right) = 0.$$

Theorem 3.2 easily implies the theorem below.

Theorem 3.4. *Let $\mathbb{F}[q]$ be an extension field of $\mathbb{F}[2]$ with $q = 2^r$. There is a reduction from* 3SAT *to an instance $\mathcal{C}$ of quadratic equations over $\mathbb{F}[q]$ such that*

- *The equations have $\mathbb{F}[2]$ coefficients and no linear terms.*
- YES Case. *If the* 3SAT *instance is satisfiable then there is an assignment to $\mathcal{C}$ that satisfies all the equations. In fact there is such an assignment that is $\mathbb{F}[2]$ valued.*
- NO Case. *If the* 3SAT *instance is unsatisfiable then there are no t assignments to $\mathcal{C}$ that are $\mathbb{F}[q]$ valued and satisfy all the equations simultaneously in super-position for any $1 \le t \le k$. Here k is a parameter and the reduction runs in time $N^{O(r \log k)}$ where N is the size of the* 3SAT *instance.*

Proof. The instance $\mathcal{C}$ is essentially the same as the instance $\mathcal{A}$ given by Theorem 3.2. The only difference is that every linear term x_i is replaced by a quadratic term x_i^2. Specifically, an equation $c + \sum_{i=1}^{m} c_i x_i + \sum_{1 \le i < j \le m} c_{ij} x_i x_j = 0$ in instance $\mathcal{A}$ is now written as $c + \sum_{1 \le i \le j \le m} c_{ij} x_i x_j = 0$ in instance $\mathcal{C}$ where $c_{ii} = c_i$. The claim in the YES case follows from the analogous claim in Theorem 3.2, so we focus on the NO case.

We show that if there are t assignments over $\mathbb{F}[q]$ that satisfy all equations in the instance $\mathcal{C}$ in super-position, then there are $t \cdot s$ assignments over $\mathbb{F}[2]$ that satisfy all equations in the instance $\mathcal{A}$ in super-position and $s \le 2r$. Let a typical equation in the instance $\mathcal{C}$ be $c + \sum_{1 \le i \le j \le m} c_{ij} x_i x_j = 0$, where $c, c_{ij} \in \mathbb{F}[2]$. Suppose there are $\mathbb{F}[q]$-valued assignments $a^{(1)}, \ldots, a^{(t)} \in \mathbb{F}[q]^m$ that satisfy the equation in super-position, i.e.

$$c + \sum_{1 \le i \le j \le m} c_{ij} \left(\sum_{\ell=1}^{t} a_i^{(\ell)} a_j^{(\ell)} \right) = 0.$$

The computations above are over $\mathbb{F}[q]$. Fixing an arbitrary representation of $\mathbb{F}[q]$ as a r-dimensional vector space over $\mathbb{F}[2]$, the above equation must hold in the last bit of the vector representation, i.e. in the notation of Lemma 3.5,

$$c + \sum_{1 \le i \le j \le m} c_{ij} \left(\sum_{\ell=1}^{t} (a_i^{(\ell)} a_j^{(\ell)})_{\mathsf{last}} \right) = 0.$$

However by Lemma 3.5, there are vectors $\lambda_1, \ldots, \lambda_s \in \mathbb{F}[2]^r$, that *capture* the computation of the last bit of a product of two elements in $\mathbb{F}[q]$. Hence, the above equation can be written as

$$c + \sum_{1 \le i \le j \le m} c_{ij} \left(\sum_{\ell=1}^{t} \sum_{p=1}^{s} \langle a_i^{(\ell)}, \lambda_p \rangle \cdot \langle a_j^{(\ell)}, \lambda_p \rangle \right) = 0.$$

Since all values now are in $\mathbb{F}[2]$, we can separate the diagonal terms and re-write them as linear terms (note $c_i = c_{ii}$), i.e.

$$c + \sum_{1 \le i \le m} c_i \left(\sum_{\ell=1}^{t} \sum_{p=1}^{s} \langle a_i^{(\ell)}, \lambda_p \rangle \right) + \sum_{1 \le i < j \le m} c_{ij} \left(\sum_{\ell=1}^{t} \sum_{p=1}^{s} \langle a_i^{(\ell)}, \lambda_p \rangle \cdot \langle a_j^{(\ell)}, \lambda_p \rangle \right) = 0.$$

This is same as saying that the $t \cdot s$ many $\mathbb{F}[2]$-valued assignments given by $\langle a^{(\ell)}, \lambda_p \rangle$ for $\ell \in [t], p \in [s]$ satisfy the corresponding equation in the instance $\mathcal{A}$ in super-position. Noting that the choice of the equation is arbitrary, the theorem follows by the guarantee on the NO case in Theorem 3.2. □

Lemma 3.5. *Let $\mathbb{F}[q]$ be an extension field of $\mathbb{F}[2]$ with $q = 2^r$. Any $x \in \mathbb{F}[q]$ can be thought of as a (row) vector in $\mathbb{F}[2]^r$ in some fixed representation of $\mathbb{F}[q]$ as a r-dimensional vector space over $\mathbb{F}[2]$. For $x \in \mathbb{F}[q]$, let $(x)_{\mathsf{last}}$ denote the last bit of the corresponding vector. Then there exist vectors $\lambda_1, \ldots, \lambda_s \in \mathbb{F}[2]^r$, $s \le 2r$ such that*

$$\forall x, y \in \mathbb{F}[q] \qquad (xy)_{\mathsf{last}} = \sum_{i=1}^{s} \langle x, \lambda_i \rangle \cdot \langle y, \lambda_i \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product over $\mathbb{F}[2]^r$ and while computing the expression $\langle x, \lambda_i \rangle$, x is being thought of as a vector in $\mathbb{F}[2]^r$.

Proof. The map $(x, y) \mapsto (xy)_{\mathsf{last}}$ can be thought of as a symmetric bilinear map $\mathbb{F}[2]^r \times \mathbb{F}[2]^r \mapsto \mathbb{F}[2]$. Hence there is a $r \times r$ symmetric matrix Λ over $\mathbb{F}[2]$ such that

$$\forall x, y \in \mathbb{F}[q] \qquad (xy)_{\mathsf{last}} = x \cdot \Lambda \cdot y^T.$$

The matrix Λ can be written as $\sum_{i=1}^{s} \lambda_i \otimes \lambda_i$ for some $s \le 2r$ and $\lambda_i \in \mathbb{F}[2]^r$ by Lemma 2.1. The same s and λ_i satisfy the conclusion of the lemma. □

The conclusion in the NO case (i.e. *soundness*) of Theorem 3.4 can be boosted via a standard trick, so that a constant fraction of equations must fail instead of at least one equation failing. Suppose the instance $\mathcal{C}$ in Theorem 3.4 has L equations written as $E_1 = 0, \ldots, E_L = 0$. One can take a $M \times L$ matrix Γ over $\mathbb{F}[2]$, $M = O(L)$, that is a generator matrix of a linear code of constant relative distance, say 0.10, and construct a new system $\mathcal{C}'$ of equations

$$\sum_{j=1}^{L} \Gamma_{ij} E_j = 0 \qquad i = 1, \ldots, M.$$

Clearly, a satisfying assignment to $\mathcal{C}$ is also a satisfying assignment to $\mathcal{C}'$. On other other hand, if no t assignments satisfy *all equations* in $\mathcal{C}$ in super-position, then no t assignments satisfy even 0.90 fraction of the equations in $\mathcal{C}'$ in super-position. With this observation, we re-state Theorem 3.4 as:

Theorem 3.6. *Let $\mathbb{F}[q]$ be an extension field of $\mathbb{F}[2]$ with $q = 2^r$. There is a reduction from* 3SAT *to an instance $\mathcal{C}$ of quadratic equations over $\mathbb{F}[q]$ such that,*

- *The equations have $\mathbb{F}[2]$ coefficients and no linear terms.*
- YES Case. *If the* 3SAT *instance is satisfiable then there is an assignment to $\mathcal{C}$ that satisfies all the equations. In fact there is such an assignment that is $\mathbb{F}[2]$ valued.*

- NO Case. *If the* 3SAT *instance is unsatisfiable then there are no t assignments to $\mathcal{C}$ that are $\mathbb{F}[q]$ valued and satisfy 0.90 fraction of the equations in super-position for any $1 \leq t \leq k$. Here k is a parameter and the reduction runs in time $N^{O(r \log k)}$ where N is the size of the* 3SAT *instance.*

Note however that the equations in the instance above have unbounded arity, i.e. a typical equation may depend on almost all the variables. In the next section, we show how to reduce arity while preserving (in the NO case) the high super-position complexity and *soundness* that is appreciably bounded away from 1.

4 A Low Arity Quadratic CSP with Superposition and Approximation Gap

In this section, we prove a theorem that is analogous to Theorem 3.6, but the equations depend only on polylog variables. The soundness suffers a little and is $1 - \frac{1}{\mathsf{polylog}}$ instead of 0.90, but this presents no problem in further reductions. The idea is to start with the instance $\mathcal{C}$ of quadratic equations in Theorem 3.6 and produce a new system of quadratic equations by running the (initial part of) the algebraic proof of PCP Theorem and in particular the Sum-Check Protocol.

Theorem 4.1. *Let $\mathbb{F}[q]$ be an extension field of $\mathbb{F}[2]$. There is a reduction from* 3SAT *to an instance $\mathcal{D}$ of quadratic equations over $\mathbb{F}[q]$, with N variables and $\mathsf{poly}(N)$ equations, such that*

- *Every equation depends on $O(\frac{\log^2 N}{\log \log N})$ variables and $q = O(\log^{O(1)} N)$.*
- *The set of variables is partitioned into sets U, V, W. The equations are either linear that involve variables only from $U \cup V$ or quadratic of the form $y = xx'$ where $y \in V$ and $x, x' \in W$.*
- YES Case. *If the* 3SAT *instance is satisfiable then there is an $\mathbb{F}[q]$ valued assignment to $\mathcal{D}$ that satisfies all the equations.*
- NO Case. *If the* 3SAT *instance is unsatisfiable then there are no t assignments to $\mathcal{D}$ that are $\mathbb{F}[q]$ valued and satisfy $1 - \frac{1}{1000k}$ fraction of the equations in super-position for any $1 \leq t \leq k$. Here the parameter k can be taken as $k = (\log N)^c$ where $c > 10$ is a constant that can be chosen as desired.*
- *The reduction runs in time $R^{O((\log \log R)^{O(1)})}$ if R is the size of the 3SAT instance.*

Remark: *Assignments $\rho_1, \ldots, \rho_t : U \cup V \cup W \mapsto \mathbb{F}[q]$ are said to satisfy an equation of instance $\mathcal{D}$ in super-position if either (1) the equation is linear that involves variables only from $U \cup V$ and the assignment $\rho = \sum_{i=1}^{t} \rho_i$ satisfies the equation or (2) the equation is of the form $y = xx'$ where $y \in V$ and $x, x' \in W$ and $\rho(y) = \sum_{i=1}^{t} \rho_i(x)\rho_i(x')$.*

The PCP Construction

We start with an instance $\mathcal{C}$ given by Theorem 3.6 consisting of n variables $X_1, \ldots, X_n$ and n' equations $E_1, \ldots, E_{n'}$. We will construct a PCP over the alphabet $\mathbb{F}[q]$ such that the PCP proof is partitioned into three parts U, V, W. The proof locations then correspond to the sets of variables U, V, W for the instance $\mathcal{D}$ and the PCP tests correspond to linear or quadratic equations in these variables.

Let $h = \lceil \log n \rceil$, $m = \lceil \log n / \log \log n \rceil$, so that $h^m \geq n$. Let $d := 2m(h-1)$. The size of the field $\mathbb{F}[q]$ is chosen to be $(10d)^3 m$. The parameter k in Theorem 3.6 is chosen to be $(\log n)^{c+4}$. It can be verified that

if R is the size of the 3SAT instance which reduces to Theorem 3.6, then $q^m = \exp(\log R(\log\log R)^{O(1)})$. The number of variables produced by the present reduction will be $\text{poly}(q^m) = \text{poly}(n)$.

Let S be any subset of $\mathbb{F}[q]$ of size h. We identify the indices $\{1, 2, \ldots, n\}$ with points in S^m. Let $\sigma : \{X_1, \ldots, X_n\} \mapsto \mathbb{F}[q]$ be a supposed satisfying assignment to the instance $\mathcal{C}$. It is now thought of as an assignment $\sigma : S^m \mapsto \mathbb{F}[q]$. Given such an assignment, there is a unique m-variate polynomial g with degree $h-1$ in each co-ordinate such that g agrees with the assignment σ on S^m. In literature, g is known as the *low degree extension* of the supposed satisfying assignment σ. Let the r^{th} equation in the instance $\mathcal{C}$ be

$$E_r : \sum_{1 \le i \le j \le n} c_{r,i,j} X_i X_j = c_r.$$

Let $u_r(\overline{z}, \overline{w})$ be the unique $2m$-variate polynomial with degree $h-1$ in each co-ordinate, such that $u_r(\overline{a}, \overline{b}) = c_{r,i,j}$ where the index i is identified with $\overline{a} \in S^m$ and j with $\overline{b} \in S^m$. Note that u_r can be computed from the coefficients $c_{r,i,j}$. The equation can now be written as

$$E_r : \sum_{\overline{a} \in S^m, \overline{b} \in S^m} u_r(\overline{a}, \overline{b}) g(\overline{a}) g(\overline{b}) = c_r.$$

The goal of the verifier is to check whether g is indeed a low degree extension of a satisfying assignment σ to the instance $\mathcal{C}$. It will be convenient to let f be the $2m$-variate polynomial defined as $f(\overline{z}, \overline{w}) = g(\overline{z}) g(\overline{w})$. Thus f is a polynomial with degree $h-1$ in each co-ordinate and hence of total degree $d = 2m(h-1)$. The equation E_r can be re-written as

$$E_r : \sum_{\overline{a} \in S^m, \overline{b} \in S^m} u_r(\overline{a}, \overline{b}) f(\overline{a}, \overline{b}) = c_r.$$

We are now ready to describe the PCP. The PCP verifier expects the following *proof*.

- The table of values of the supposed polynomial $f(\overline{z}, \overline{w})$ at all points $(\overline{z}, \overline{w}) \in \mathbb{F}[q]^{2m}$. It is ensured that f is symmetric, i.e $f(\overline{z}, \overline{w}) = f(\overline{w}, \overline{z})$ for all $\overline{z}, \overline{w} \in \mathbb{F}[q]^m$, by identifying the corresponding proof locations.
- For each line $\ell = \overline{\alpha} + t\overline{\beta}$ in $\mathbb{F}[q]^{2m}$, the coefficients of a univariate polynomial $\phi_\ell(t)$ of degree at most d which is supposed to be the restriction of the polynomial f to the line ℓ.
- For the r^{th} equation E_r, $m' \in \{0, 1, \ldots, 2m-1\}$, and $(\theta_1, \ldots, \theta_{m'}) \in \mathbb{F}[q]^{m'}$, the coefficients of a (supposed) univariate polynomial of degree $2(h-1)$, denoted as $p_{r,\theta_1,\ldots,\theta_{m'}}(y_{m'+1})$. This polynomial is supposed to be

$$\sum_{y_{m'+2}, \ldots, y_{2m} \in S} u_r(\theta_1, \ldots, \theta_{m'}, y_{m'+1}, y_{m'+2}, \ldots, y_{2m}) f(\theta_1, \ldots, \theta_{m'}, y_{m'+1}, y_{m'+2}, \ldots, y_{2m}). \quad (9)$$

  Note that in the above sum, $y_{m'+1}$ is the only formal variable, so the sum is a polynomial in that variable. The polynomial for $m' = 0$ is denoted by $p_{r,\emptyset}(y_1)$. These polynomials are referred to as the *partial sum polynomials*.
- The table of values of the supposed polynomial $g(\overline{z})$ at all points $\overline{z} \in \mathbb{F}[q]^m$.

The proof locations are partitioned into sets U, V, W as follows. V consists of the table of values of f, W consists of the table of values of g and U consists of the rest (i.e. the line polynomials and the partial sum polynomials).

Test of the Verifier.

Perform one of the three tests below with equal probability. In the second test, perform one of the $2m+1$ sub-tests with equal probability.

1. (Low Degree Test) Pick a line $\ell = \overline{\alpha} + t\overline{\beta}$ in $\mathbb{F}[q]^{2m}$ and $t \in \mathbb{F}[q]$ uniformly at random. Check that $\phi_\ell(t) = f(\ell(t))$.
2. (Sum Check Protocol) Pick an equation E_r uniformly at random to verify. Pick $\theta = (\theta_1, \ldots, \theta_{2m}) \in \mathbb{F}[q]^{2m}$ uniformly at random.
   a. Check that $\sum_{y_1 \in S} p_{r,\emptyset}(y_1) = c_r$.
   b. Check that for $j \in \{1, \ldots, 2m-1\}$,
   $$\sum_{y_{j+1} \in S} p_{r,\theta_1,\ldots,\theta_j}(y_{j+1}) = p_{r,\theta_1,\ldots,\theta_{j-1}}(\theta_j).$$
   c. Check that $p_{r,\theta_1,\ldots,\theta_{2m-1}}(\theta_{2m}) = u_r(\theta_1, \ldots, \theta_{2m}) f(\theta_1, \ldots, \theta_{2m})$.
3. (Consistency) For a uniformly random pair $\overline{z}, \overline{w} \in \mathbb{F}[q]^m$, check that $f(\overline{z}, \overline{w}) = g(\overline{z})g(\overline{w})$.

Note that the low degree test produces a linear equation. The Sum-Check protocol produces a set of $2m+1$ linear equations. Finally, the consistency test produces a quadratic equation. Note that the linear equations have variables only from the set $U \cup V$ whereas the quadratic equation is of the type $y = xx'$ with $y \in V$ and $x, x' \in W$. Every equation depends on at most $O(d)$ variables. This completes the description of the instance $\mathcal{D}$ produced by our reduction.

YES Case

In the YES case, we show that there is an assignment $U \cup V \cup W \mapsto \mathbb{F}[q]$ that satisfies all equations in $\mathcal{D}$. Indeed, there is an assignment σ to the variables $X_1, \ldots, X_n$ that satisfies the equations of $\mathcal{C}$. Let g be its low degree extension, let $f(\overline{z}, \overline{w}) = g(\overline{z})g(\overline{w})$, and let the line polynomials ϕ_ℓ and the partial sum polynomials be as they ought to be. It is clear that this yields an assignment $U \cup V \cup W \mapsto \mathbb{F}[q]$ that satisfies all equations in $\mathcal{D}$.

NO Case

We wish to show that there are no k assignments $\rho_1, \ldots, \rho_k : U \cup V \cup W \mapsto \mathbb{F}[q]$ that satisfy $1 - \delta$ fraction of the equations in super-position where $\delta = \left(\frac{1}{1000k}\right)$. Assume on the contrary that this is the case. Let $\rho = \sum_{i=1}^{k} \rho_i$. Note that if a linear equation (over variables in $U \cup V$) is satisfied in super-position, this amounts to saying that ρ satisfies it. On the other hand if a quadratic equation $y = xx'$ is satisfied in super-position for $y \in V$, $x, x' \in W$, this amounts to saying that $\rho(y) = \sum_{i=1}^{k} \rho_i(x)\rho_i(x')$.

Let $f(\cdot, \cdot)$ be the f-table (i.e. assignment to V) according to the assignment ρ. Let $\phi_\ell(\cdot)$, $p_{r,\theta_1,\ldots,\theta_j}(\cdot)$ be polynomials (given as their coefficients, i.e. assignment to U) according to the assignment ρ. Let

$g_1(\cdot), \ldots, g_k(\cdot)$ be the g-tables (i.e. assignment to W) according to assignments $\rho_1, \ldots, \rho_k$ respectively. Let M denote the $\mathbb{F}[q]^m \times \mathbb{F}[q]^m$ matrix whose entries are $M(\overline{z}, \overline{w}) = \sum_{i=1}^{k} g_i(\overline{z}) g_i(\overline{w})$. Note that $f, g_1, \ldots, g_k$ are supposed to be low degree polynomials, but this might not be the case in cheating proof(s).

Since $1 - \delta$ fraction of the equations overall are satisfied (in super-position), the Low Degree Test on function f and the line polynomials $\{\phi_\ell\}$ passes with probability $1 - 3\delta$, and by Theorem 5.1 of [Aro94], there is a polynomial F of total degree d which agrees with f at $1 - 6\delta$ fraction of points. Since f is symmetric, i.e. $f(\overline{z}, \overline{w}) = f(\overline{w}, \overline{z})$, so is F. Otherwise $F(\overline{z}, \overline{w})$ and $F(\overline{w}, \overline{z})$ would be different polynomials, disagreeing almost everywhere, and F would not be close to f.

By averaging argument, for 99% of the equations E_r, with probability at least $1 - O(\delta m)$, all the tests in the Sum-Check Protocol succeed. We choose parameters so that this probability is at least $\frac{1}{2}$ and larger than the *error probability* $O(md/q)$ of the Sum-Check Protocol. Fix any such "good' equation E_r. We claim that

$$\sum_{\overline{a} \in S^m, \overline{b} \in S^m} u_r(\overline{a}, \overline{b}) F(\overline{a}, \overline{b}) = c_r. \tag{10}$$

This is because otherwise, by the standard analysis of the Sum-Check Protocol, the polynomials $p_{r,\theta_1,\ldots,\theta_j}$ for $j = 0, 1, \ldots, 2m-1$ given in the proof will be different from the (correct) partial sums of the polynomial $u_r(\cdot,\cdot)F(\cdot,\cdot)$ (as in Equation (9)) and then for $j = 2m - 1$, Test 2.c would fail with high probability given that f and F are $(1 - O(\delta))$-close.

Finally, the consistency test passes with probability $1 - 3\delta$ in super-position. This amounts to saying that f, and hence F, agrees with matrix M on $1 - 9\delta$ fraction of the entries. The rank of M is at most k and by Lemma 4.2, and the setting of $\delta = (1/1000k)$, the rank of F is also at most k. Using Lemma 2.1, one can write $F = \sum_{i=1}^{s} f_i \otimes f_i$ with $s \leq 2k$. We have shown that for 99% of the equations E_r,

$$\sum_{\overline{a} \in S^m, \overline{b} \in S^m} u_r(\overline{a}, \overline{b}) \sum_{i=1}^{s} f_i(\overline{a}) f_i(\overline{b}) = c_r. \tag{11}$$

This is same as saying that $f_i : S^m \mapsto \mathbb{F}[q]$ thought of as assignments $f_i : \{X_1, \ldots, X_n\} \mapsto \mathbb{F}[q]$ satisfy 99% of the equations E_r in super-position. This is a contradiction to the guarantee in the NO case of Theorem 3.6.

Lemma 4.2. *Let $F : \mathbb{F}[q]^m \times \mathbb{F}[q]^m \mapsto \mathbb{F}[q]$ be a degree d polynomial in $2m$ variables. Written as a matrix of its values, assume that F agrees with a matrix M of rank at most k in $1 - \delta$ fraction of the entries and $\delta < \frac{1}{16k}, \frac{d}{q} < \frac{1}{2}$. Then as a matrix F also has rank at most k.*

Proof. F agrees with M on at least $1 - \delta$ fraction of the entries. By averaging, there are at least $\frac{3}{4}$ fraction of the columns – call them as "good" columns – such that in each of these columns F and M agree on $1 - 4\delta$ fraction of the entries. Since M is of rank at most k, its sub-matrix given by the good columns is also of rank at most k. Thus there are k good columns – say $M(\cdot, \overline{b}_1), \ldots, M(\cdot, \overline{b}_k)$ – whose span contains all the good columns of M. For any good column, say $M(\cdot, \overline{b})$, we have $M(\cdot, \overline{b}) = \sum_{i=1}^{k} \zeta_i \cdot M(\cdot, \overline{b}_i)$ for some $\zeta_i \in \mathbb{F}[q]$. Consider the combination $\sum_{i=1}^{k} \zeta_i \cdot F(\cdot, \overline{b}_i)$. This combination agrees with $M(\cdot, \overline{b})$ in $1 - 4k\delta$ fraction of the entries and thus agrees with $F(\cdot, \overline{b})$ at $1 - 4(k+1)\delta$ fraction of the entries. By Schwartz-Zippel Lemma, since each column of F is a degree d polynomial, we obtain that $\sum_{i=1}^{k} \zeta_i \cdot F(\cdot, \overline{b}_i) = F(\cdot, \overline{b})$ provided $1 - 4(k+1)\delta > \frac{d}{q}$. Thus the sub-matrix of F given by its good columns has column rank at most k. Since column rank is the same as row rank, there is a row basis of size at most k for this sub-matrix of F. Let $F(\overline{a}_1, \cdot), \ldots, F(\overline{a}_k, \cdot)$ be the rows that span the row space of this sub-matrix. Any other row $F(\overline{a}, \cdot)$ is a linear combination $F(\overline{a}, \cdot) = \sum_{i=1}^{k} \zeta_i \cdot F(\overline{a}_i, \cdot)$ in the entries corresponding to the good columns. However

since all the rows are degree d polynomials, this linear combination is an identity of polynomials and hence holds in *all* the columns of F, completing the proof. □

5 Gap Amplification

We will need a theorem that is similar to Theorem 4.1 in spirit, but rather different in form.

Definition 5.1. *Let C be a constraint that is a conjunction of quadratic equations $\{q_i(x) = 0\}_{i=1}^{s}$ over $\mathbb{F}[2]$ and over boolean variables $x = (x_1, \ldots, x_n)$. A set of t global assignments $\sigma_1, \ldots, \sigma_t : \{x_1, \ldots, x_n\} \mapsto \mathbb{F}[2]$ is said to satisfy the constraint C "k-locally" if there is a subset $T \subseteq \{\sigma_1, \ldots, \sigma_t\}$ such that $|T| \leq k$ and the assignments in T satisfy each quadratic equation $q_i(x) = 0$ of C in super-position.*

Theorem 5.2. *For any large enough constant $b > 0$, there is a reduction from 3SAT to a system of constraints $\mathcal{A}$ over boolean variables $x_1, \ldots, x_n$ such that:*

- *Every constraint C is a conjunction of $(\log n)^{8b+2}$ quadratic equations, each of which depends on at most $O(\log^2 n)$ variables.*
- *YES Case: There is a boolean assignment that satisfies all the constraints.*
- *NO Case: Given any $t = 2^{(\log n)^{2b}}$ global assignments $\sigma_1, \ldots, \sigma_t$, the fraction of the constraints that are satisfied k-locally is at most $2^{-(\log n)^{5b}}$ and $k = (\log n)^{3b}$.*
- *The reduction runs in time $\exp\left((\log R)^{8b+5}\right)$ where R is the size of the 3SAT instance. The number of constraints produced is $n^{(\log n)^{8b+2}+O(1)}$.*

Proof. We first observe that the instance $\mathcal{D}$ given by Theorem 4.1 can be turned into a boolean instance $\mathcal{D}'$ where all variables and constraints are over $\mathbb{F}[2]$. This is simply by writing every $\mathbb{F}[q]$ valued variable, $q = 2^r$, by a string of r boolean variables that represent it when $\mathbb{F}[q]$ is considered as a r-dimensional vector space over $\mathbb{F}[2]$. It is easily observed that every linear or quadratic equation in $\mathcal{D}$ leads to an equivalent system of r linear or quadratic equations over the new boolean variables, one for each of the r "co-ordinates". Each equation now involves $O(\frac{\log^2 N}{\log\log N}) \cdot \log q = O(\log^2 N)$ variables. In the YES case, $\mathcal{D}$ has a satisfying assignment and the same assignment, viewed as boolean assignment to the new variables, satisfies the instance $\mathcal{D}'$. In the NO case, we are guaranteed that no $k = (\log N)^{3b}$ assignments ($c = 3b$ in the notation of Theorem 4.1) that are $\mathbb{F}[q]$ valued satisfy $\left(1 - \frac{1}{1000(\log N)^{3b}}\right)$ fraction of the equations in $\mathcal{D}$ in super-position. This statement is almost preserved: when an equation over $\mathbb{F}[q]$ fails, at least one from the system of r equations over $\mathbb{F}[2]$ that is equivalent to it, must fail. Thus we can conclude that in the NO case, no k assignments that are $\mathbb{F}[2]$ valued satisfy $\left(1 - \frac{1}{1000r\cdot(\log N)^{3b}}\right)$ fraction of the equations in $\mathcal{D}'$ in super-position. Let's be generous and upper bound this fraction by $\left(1 - \frac{1}{(\log N)^{3b+1}}\right)$. Let $n = rN$ denote the number of variables in $\mathcal{D}'$ so that $\log n \approx \log N$. Thus, we can rewrite the statement of Theorem 4.1 over n $\mathbb{F}[2]$-valued variables and poly(n) quadratic equations – each involving $O((\log n)^2)$ variables – with $k = (\log n)^{3b}$ and the soundness in the NO case $\left(1 - \frac{1}{(\log n)^{3b+1}}\right)$. This holds for every constant $b > 4$.

Now we construct the desired instance $\mathcal{A}$ whose variables are same as in the instance $\mathcal{D}'$ and whose constraints are all s-wise conjunctions of the equations in $\mathcal{D}'$ where $s = (\log n)^{8b+2}$. The satisfying assignment in the YES case is preserved. In the NO case, let $\sigma_1, \ldots, \sigma_t$ be any $t = 2^{(\log n)^{2b}}$ global assignments. We

know that any $k = (\log n)^{3b}$ of these assignments satisfy at most $\left(1 - \frac{1}{(\log n)^{3b+1}}\right)$ fraction of the equations in $\mathcal{D}'$ in super-position. Thus, the fraction of the new constraints satisfied in super-position is at most $\left(1 - \frac{1}{(\log n)^{3b+1}}\right)^s$ which is at most $2^{-(\log n)^{5b+1}}$. One can now take a union bound over all subsets of the t global assignments of size at most k and the theorem follows. Note that in this reduction, the number of constraints produced is $n^{(\log n)^{8b+2}+O(1)}$. □

We will need analogue of Theorem 5.2 over a large field $\mathbb{F}[q]$ of characteristic 2. Applying the same trick as in Theorem 3.4 and Lemma 3.5, we easily obtain the following theorem from Theorem 5.2. The parameter k denoting the super-position complexity suffers a loss of $2 \log q$. Setting $q = 2^{(\log n)^{2b}}$, the parameter k is now reduced to $\frac{1}{2}(\log n)^b$.

Theorem 5.3. *For any large enough constant $b > 0$, there is a reduction from 3SAT to a system of constraints $\mathcal{A}$ over $\mathbb{F}[q]$-valued variables $x_1, \ldots, x_n$, where $q = 2^{(\log n)^{2b}}$, such that:*

- *Every constraint C is a conjunction of $(\log n)^{8b+2}$ quadratic equations, each of which depends on at most $O(\log^2 n)$ variables. The quadratic equations have no linear terms.*
- *YES Case: There is a boolean assignment that satisfies all the constraints.*
- *NO Case: Given any $t = 2^{(\log n)^{2b}}$ global assignments $\sigma_1, \ldots, \sigma_t$, the fraction of the constraints that are satisfied k-locally is at most $2^{-(\log n)^{5b}}$ where $k = \frac{1}{2}(\log n)^b$.*
- *The reduction runs in time $\exp\left((\log R)^{8b+5}\right)$ where R is the size of the 3SAT instance. The number of constraints produced is $n^{(\log n)^{8b+2}+O(1)}$.*

6 The Outer PCP

We are now ready to present our construction of the Outer PCP, a.k.a. the Label Cover problem. It is constructed algebraically via a "point versus ruled-surface" low degree test analogous to the point versus line test in [AS03]. The analysis is rather straightforward using [AS03] result as a black-box. We start with the instance $\mathcal{A}$ produced by the reduction in Theorem 5.3.

Let $m = \lceil \log n / \log \log n \rceil$ and $h = \lceil \log n \rceil$ so that $h^m \geq n$ and let $d := m(h-1)$. We identify the variables of $\mathcal{A}$ with S^m where $S \subseteq \mathbb{F}[q]$ is of size h. Any $\mathbb{F}[q]$ valued assignment ρ to the variables of $\mathcal{A}$ is interpreted as an assignment $\rho : S^m \mapsto \mathbb{F}[q]$ and can be extended to a corresponding polynomial $g : \mathbb{F}[q]^m \mapsto \mathbb{F}[q]$ of degree at most $(h-1)$ in each of the m coordinates, i.e. of total degree d. Let $\mathcal{C}$ denote the set of constraints of $\mathcal{A}$, each over at most $l = O((\log n)^{8b+4})$ variables. For convenience, we shall denote every constraint $C \in \mathcal{C}$ as $C[\{\overline{x_i}\}_{i=1}^l]$, where $\{\overline{x_i}\}_{i=1}^l$ is the set of points in S^m corresponding to the l variables of $\mathcal{A}$ on which the constraint is defined.

Definition 6.1. *A curve $\overline{\omega} : \mathbb{F}[q] \mapsto \mathbb{F}[q]^m$ of degree r is a mapping $\overline{\omega}(t) := (\omega_1(t), \ldots, \omega_m(t))$ where each ω_j is a degree r univariate polynomial in t.*

Fix, for the rest of this section, distinct values $t_0^*, t_1^*, \ldots, t_l^* \in \mathbb{F}[q]$. It is easy to see that for any set of points $\overline{x}, \overline{x_1}, \ldots, \overline{x_l} \in \mathbb{F}[q]^m$, there is a degree l curve $\overline{\omega}$ such that $\overline{\omega}(t_0^*) = \overline{x}$ and $\overline{\omega}(t_i^*) = \overline{x_i}$ for $i = 1, \ldots, l$. A curve $\overline{\omega}$ is said to correspond to a constraint $C = C[\{\overline{x_i}\}_{i=1}^l]$ and an additional point $\overline{x}$ if the said condition holds (the points $t_0^*, t_1^*, \ldots, t_l^* \in \mathbb{F}[q]$ are understood implicitly). We have the following observation.

Observation 6.2. *Given a constraint C, consider a random curve $\overline{\omega}$ corresponding to C and a uniformly random point $\overline{x}$. Then, for any $t \in \mathbb{F}[q] \setminus \{t_1^*, \ldots, t_l^*\}$, the point $\overline{\omega}(t)$ is uniformly distributed in $\mathbb{F}[q]^m$.*

Definition 6.3. *A ruled surface $R = R[\overline{\omega}, \overline{y}]$ where $\overline{\omega}(t)$ is a curve and $\overline{y} \in \mathbb{F}[q]^m$ is a direction, is a surface parametrized by two parameters $t, s \in \mathbb{F}[q]$ where,*

$$R[\overline{\omega}, \overline{y}](t, s) = \overline{\omega}(t) + s\overline{y}.$$

For a constraint C, a point $\overline{x}$ and a direction $\overline{y}$, let $R[\overline{\omega}, \overline{y}]$ be a ruled surface where $\overline{\omega}$ is the curve of degree l corresponding to C and $\overline{x}$. Let $\mathcal{R}_C$ be the class of all such ruled surfaces corresponding to a constraint $C \in \mathcal{C}$ and let $\mathcal{R} := \cup_{C \in \mathcal{C}} \mathcal{R}_C$. Suppose $g : \mathbb{F}[q]^m \mapsto \mathbb{F}[q]$ is a (global) polynomial of degree d. The restriction of g to a ruled surface $R \in \mathcal{R}_C$, for any constraint C, is a bivariate polynomial – in t and s – of degree at most $d^* := ld \leq (\log n)^{8b+6}$ in variable t and at most d in variable s. The total number coefficients of such a polynomial is at most $d^* d = (\log n)^{8b+8}$. We are now ready to describe the Label Cover instance $\mathcal{L}$.

Left vertex set: This consists of all points in $\mathbb{F}[q]^m$. The label set – same for each vertex – is the set of values $\mathbb{F}[q]$ that can be assigned to the points.

Right vertex set: The set of right vertices is $\mathcal{R}$, namely the class of all ruled surfaces over all constraints $C \in \mathcal{C}$. The label set for a ruled surface $R \in \mathcal{R}$ is the set of all bivariate polynomials in t and s of degree at most d^* in the variable t and at most d in the variable s. Such a polynomial is represented by a vector of its coefficients, at most $(\log n)^{8b+8}$ in number as mentioned before. For a ruled surface R corresponding to a constraint C, there is a conjunction of $(\log n)^{8b+2}$ quadratic equations on these coefficients that determines whether the values given by the polynomial at points $\{(t_i^*, s = 0)\}_{i=1}^{l}$ satisfy C.

Edges: For every ruled surface $R \in \mathcal{R}$ and every point $\overline{v} \in R$, there is an edge between $\overline{v}$ and R. The edge is satisfied by a labeling $\overline{g}$ to the surface R and a label p to the point $\overline{v}$ if $\overline{g}(\overline{v}) = p$, i.e. if the surface polynomial and the value at the point are consistent. Note that the computation of $\overline{g}(\overline{v})$ is linear in the coefficients of $\overline{g}$. In the rest of this section we shall prove the correctness of our reduction.

Theorem 6.4. *Let $k = \frac{1}{2}(\log n)^b$, $q = 2^{(\log n)^{2b}}$ and $\delta = 2^{-(\log n)^{2b-1}}$. Let $\mathcal{L}$ be the Label Cover instance described above.*

1. *The labels are elements of $\mathbb{F}[q]$ for the points in $\mathbb{F}[q]^m$ and coefficient vectors of length $(\log n)^{8b+8}$ over $\mathbb{F}[q]$ for the surfaces. The projection maps, mapping a coefficient vector of a surface polynomial to its value at a point on the surface, are homogeneous and $\mathbb{F}[q]$-linear. The coefficient vector is supposed to satisfy a constraint C that is a conjunction of quadratic equations over $\mathbb{F}[q]$. The size of the instance $\mathcal{L}$ is at most $2^{(\log n)^{8b+4}}$ (taking into account the choice of the constraint C and the corresponding ruled surface).*

2. YES Case. *If the instance $\mathcal{A}$ given by Theorem 5.3 is a YES instance then there is a labeling to the vertices of $\mathcal{L}$ that satisfies all the edges. Further, the label (coefficient vector) for any ruled surface R satisfies the associated constraint $C \in \mathcal{C}$.*

3. NO Case. *((k, δ)-Soundness) If the instance $\mathcal{A}$ given by Theorem 5.3 is a NO instance then the following <u>cannot hold</u> for the instance $\mathcal{L}$:*

   - *For every $\overline{v} \in \mathbb{F}[q]^m$ there are k labels $p_1^{\overline{v}}, \ldots, p_k^{\overline{v}} \in \mathbb{F}[q]$.*

- *For every ruled surface $R \in \mathcal{R}$, there are k labels (polynomials given as coefficient vectors) $\overline{g}_1^R, \ldots, \overline{g}_k^R$, such that the constraint C corresponding to the ruled surface is satisfied in super-position by the k labels.*
- *For δ fraction of the edges of $\mathcal{L}$, between a point $\overline{v}$ and a ruled surface R:*

$$\overline{g}_j^R(\overline{v}) = p_j^{\overline{v}} \quad \forall j \in \{1, \ldots, k\}.$$

4. δ-Smoothness*: For any surface R, let $\overline{g}$ be a non-zero label. Then*

$$\Pr_{\overline{v} \in R}\left[\overline{g}(\overline{v}) = 0\right] \leq \delta. \tag{12}$$

We note that in the statement above, $\mathsf{poly}(\delta) \gg \frac{\mathsf{poly}(lk)}{q}$, which shall be useful for our analysis. The properties listed in the first item are clear from the construction. The YES and NO cases of the above theorem are proved as follows.

6.1 YES Case

In the YES case, there is an assignment to each variable of $\mathcal{A}$ given by Theorem 5.3 that satisfies all the constraints. Therefore, there is a degree d polynomial $f : \mathbb{F}[q]^m \mapsto \mathbb{F}[q]$ that gives this assignment to the corresponding points in S^m. The left vertices of $\mathcal{L}$ are labeled using the assignment given by f. Each ruled surface R is labeled by the polynomial given by the restriction of f to R. This assignment satisfies the constraint C associated with R and is consistent with the assignment to all the points $\overline{v} \in R$.

6.2 NO Case

For the sake of contradiction assume that there exists a labeling as in the NO case of Theorem 6.4. An averaging argument shows that for $\frac{\delta}{2}$ fraction of the constraints C, we have

$$\Pr_{\substack{R \in \mathcal{R}_c \\ \overline{v} \in R}}\left[\bigwedge_{j=1}^{k}\left(\overline{g}_j^R(\overline{v}) = p_j^{\overline{v}}\right)\right] \geq \frac{\delta}{2}. \tag{13}$$

Call such a constraint "good". For the rest of the analysis we fix a good constraint C. Our analysis shall show that there exists a set $\mathcal{F}$ of $\mathsf{poly}(k/\delta)$ global assignments to the variables of $\mathcal{A}$ such that for every "good" constraint C, there is a subset $\beta_C \subseteq \mathcal{F}$ such that $|\beta_C| \leq k$, and the assignments in β_C satisfy the constraint C in super-position. This implies that $\mathcal{F}$ k-locally satisfies $\frac{\delta}{2}$ fraction of the constraints of $\mathcal{A}$. This yields a contradiction to the NO Case of Theorem 5.3 by our choice of parameters.

We say that a line $\ell(s)$ is contained in a ruled surface $R(t, s)$ if it is obtained by fixing a value of t in $R(t, s) = \overline{\omega}(t) + s\overline{y}$. Since choosing a random point on a ruled surface R is equivalent to choosing a random line ℓ contained in R and then choosing a random point on ℓ, the Equation (13) can be rewritten as:

$$\Pr_{\substack{R \in \mathcal{R}_c \\ \ell \in R,\, \overline{v} \in \ell}}\left[\bigwedge_{j=1}^{k}\left(\overline{g}_j^R(\overline{v}) = p_j^{\overline{v}}\right)\right] \geq \frac{\delta}{2}. \tag{14}$$

From Observation 6.2, it is easy to see that the above probability is essentially equal to the probability obtained by first picking a random line ℓ and then a random $R \in \mathcal{R}_C$ containing the line. Let $\mathcal{R}_C^\ell$ be the set

of ruled surfaces $R \in \mathcal{R}_C$ that contain ℓ. Thus we have,

$$\Pr_{\substack{\ell,\ \overline{v}\in\ell \\ R\in\mathcal{R}_C^{\ell}}}\left[\bigwedge_{j=1}^{k}\left(\overline{g}_j^R(\overline{v}) = p_j^{\overline{v}}\right)\right] \geq \frac{\delta}{3}. \tag{15}$$

Let us define k *points tables*, $f^1, \ldots, f^k : \mathbb{F}[q]^m \mapsto \mathbb{F}[q]$ as $f^j(\overline{v}) := p_j^{\overline{v}}$ for all $\overline{v} \in \mathbb{F}[q]^m$ and $j \in \{1, \ldots, k\}$. We also have a *randomized* k-tuple of *lines tables* $H = (h^1, \ldots, h^k)$. For each line ℓ, $H(\ell) = (h^1(\ell), \ldots, h^k(\ell))$ where each $h^j(\ell)$ is a degree d univariate polynomial. The randomized tuple H is constructed as follows: for each line ℓ, choose a random $R \in \mathcal{R}_C^{\ell}$ and let $h^j(\ell)$ be the univariate restriction of the bivariate polynomial $\overline{g}_j^R$ to the line ℓ. Let E be the following event (over the choice of ℓ, $\overline{v} \in \ell$ and H):

$$E \equiv \bigwedge_{j=1}^{k}\left(f^j(\overline{v}) = h^j(\ell)(\overline{v})\right). \tag{16}$$

Equation (15) can be re-stated as $\Pr_{\ell,\overline{v}\in\ell,H}[E] \geq \frac{\delta}{3}$. Applying the result of Arora and Sudan [AS03], stated as Theorem 2.6, the following holds. For each $j \in \{1, \ldots, k\}$, let $\{P_1^j, \ldots, P_u^j\}$ be the *list* of degree d polynomials over $\mathbb{F}[q]^m$ which agree with f^j on at least $\mathsf{poly}(\delta/k)$ fraction of points. Here the size of the list u is upper bounded by $\mathsf{poly}(k/\delta)$. Then, for each $j \in \{1, \ldots, k\}$,

$$\Pr_{\ell,\overline{v}\in\ell,H}\left[f^j(\overline{v}) \notin \{P_1^j(\overline{v}), \ldots, P_u^j(\overline{v})\}\ ,\ f^j(\overline{v}) = h^j(\ell)(\overline{v})\right] \leq \frac{\delta}{6k}. \tag{17}$$

A point to note here is that the statement above holds if $h^j(\cdot)$ is a *deterministic* lines table and hence also if it is a *randomized* table. Taking a union bound over $j \in \{1, \ldots, k\}$, we obtain,

$$\Pr_{\ell,\overline{v}\in\ell,H}\left[\bigvee_{j=1}^{k}\left(f^j(\overline{v}) \notin \{P_1^j(\overline{v}), \ldots, P_u^j(\overline{v})\}\ ,\ f^j(\overline{v}) = h^j(\ell)(\overline{v})\right)\right] \leq \frac{\delta}{6}. \tag{18}$$

If E' denotes the event in the above equation, then it follows that

$$\Pr_{\ell,\overline{v}\in\ell,H}\left[\bigwedge_{j=1}^{k}\left(h^j(\overline{v}) \in \{P_1^j(\overline{v}), \ldots, P_u^j(\overline{v})\}\right)\right] \geq \Pr[E] - \Pr[E'] \geq \frac{\delta}{3} - \frac{\delta}{6} = \frac{\delta}{6}. \tag{19}$$

Incurring a negligible loss, we switch the order of random choices made to the original order, i.e. choosing a random $R \in \mathcal{R}_C$, then a random line $\ell \in R$ and then a point $\overline{v} \in \ell$. The value $h^j(\overline{v})$ is then same as $\overline{g}_j^R(\overline{v})$ and we obtain:

$$\Pr_{\substack{R\in\mathcal{R}_C \\ \overline{v}\in R}}\left[\bigwedge_{j=1}^{k}\left(\overline{g}_j^R(\overline{v}) \in \{P_1^j(\overline{v}), \ldots, P_u^j(\overline{v})\}\right)\right] \geq \frac{\delta}{12}. \tag{20}$$

In particular, there is one $R \in \mathcal{R}_C$ such that,

$$\Pr_{\overline{v}\in R}\left[\bigwedge_{j=1}^{k}\left(\overline{g}_j^R(\overline{v}) \in \{P_1^j(\overline{v}), \ldots, P_u^j(\overline{v})\}\right)\right] \geq \frac{\delta}{12}.$$

By the setting of our parameters and an application of the Schwartz-Zippel Lemma, the above cannot occur unless $\overline{g}_j^R \in \{P_1^j(R), \ldots, P_u^j(R)\}$ for each $j \in \{1, \ldots, k\}$, where $P_i^j(R)$ is the restriction of the polynomial P_i^j to surface R. From our assumption, the labels $\{\overline{g}_j^R\}_{j=1}^k$ satisfy the constraint C associated with the surface R in super-position. Let $\mathcal{P}$ be the set of at most ku assignments to $\mathbb{F}[q]^m$ given by the polynomials $\{P_i^j \mid j = 1, \ldots, k;\ i = 1, \ldots, u\}$. Thus, there is a subset β_C of k global assignments from $\mathcal{P}$ such that they satisfy C in super-position. This holds for every good constraint C and at least $\frac{\delta}{2}$ fraction of constraints are good. Noting that from our setting of parameters and analysis above, $ku \leq 2^{(\log n)^{2b}}$ and $\frac{\delta}{2} \geq 2^{-(\log n)^{2b}}$. This yields a contradiction to the NO case of Theorem 5.3 and thus completes the analysis of the NO case of Theorem 6.4.

6.3 Smoothness Property

For the Label Cover constructed in this section the following smoothness property holds. For any ruled surface R, let $\overline{g}$ be a non-zero label i.e. a vector over $\mathbb{F}[q]$ representing coefficients of a non-zero degree d^* polynomial on R. The Schwartz-Zippel Lemma implies,

$$\Pr_{\overline{v} \in R}[\overline{g}(\overline{v}) = 0] \leq \frac{d^*}{q} \leq \delta, \tag{21}$$

by the setting of parameters. We end this section by noting that Observation 6.2 implies that the Label Cover instance constructed is essentially bi-regular.

7 Abstracting out the Outer PCP

The Label Cover instance constructed in the previous section is defined over algebraic and geometric objects which makes it rather cumbersome to use directly in a hardness reduction. So we provide an abstraction of this Label Cover instance in two steps. The first abstraction, *Label-Cover-Intermediate* is a clean restatement of the Label Cover instance (along with a useful trick) where the labels are boolean vectors. The second abstraction, *Label-Cover-Final* is a more compact representation of the former, where the labels are boolean matrices instead of boolean vectors. The matrix representation allows us to re-state the super-position complexity succinctly in terms of the rank of matrices.

7.1 The First Abstraction

Consider the Label Cover instance from Theorem 6.4. We first replace the instance over $\mathbb{F}[q]$ by an equivalent instance over $\mathbb{F}[2]$. Every element of $\mathbb{F}[q]$ is now written as a $\mathbb{F}[2]$-vector of length $\log q$. An $\mathbb{F}[q]$ valued assignment to a set of s variables is now thought of as a $\mathbb{F}[2]$ valued assignment to a corresponding set of $s \cdot \log q$ many $\mathbb{F}[2]$ valued variables. A quadratic equation in the $\mathbb{F}[q]$-variables is replaced by an equivalent system of $\log q$ quadratic equations in the $\mathbb{F}[2]$-variables, one for each of the $\log q$ "co-ordinates". If a set of $\mathbb{F}[q]$ valued assignments satisfies a quadratic equation in super-position, then the set of corresponding $\mathbb{F}[2]$ valued assignments satisfies the equivalent system of quadratic equations in super-position, and vice-versa. The maps that are $\mathbb{F}[q]$-linear are now $\mathbb{F}[2]$-linear. Clearly, re-writing everything over $\mathbb{F}[2]$ increases the number of variables and constraints by a factor of $\log q$.

Thus the Label Cover instance now consists of a set of vertices U on the left hand side, a set of vertices V on the right hand side, and a set of edges $E \subseteq U \times V$. A label to a vertex $u \in U$ is a vector from $\mathbb{F}[2]^{r-1}$ and a label to a vertex $v \in V$ is a vector from $\mathbb{F}[2]^{m-1}$ (the reason why the label lengths are denoted $r - 1$

and $m-1$ instead of r and m respectively will be clear shortly). For every edge $e=(u,v)$, there is a homogeneous linear map $\tilde{\rho}^e : \mathbb{F}[2]^{m-1} \mapsto \mathbb{F}[2]^{r-1}$. The edge is satisfied by a labeling $\overline{w}_u$ and $\overline{z}_v$ to vertices u and v respectively if and only if $\tilde{\rho}^e(\overline{z}_v) = (\overline{w}_u)$. Also, the label $\overline{z}_v$ is supposed to satisfy a constraint $\tilde{C}_v$ that is a conjunction of quadratic equations (we do not care at this point as to the number of constraints involved in the conjunction).

It will be convenient in future to extend the labels by an additional co-ordinate that is supposed to be constant 1. Thus the label sets are now $\mathbb{F}[2]^r$ and $\mathbb{F}[2]^m$ respectively and the new labels $\overline{x}_u$ and $\overline{y}_v$ are now supposed to be $\overline{x}_u = (\overline{w}_u, 1)$ and $\overline{y}_v = (\overline{z}_v, 1)$ respectively. A new homogeneous linear map $\rho_e : \mathbb{F}[2]^m \mapsto \mathbb{F}[2]^r$ is defined as $(\tilde{\rho}^e, \mathsf{id})$, i.e. it maps the first $m-1$ co-ordinates according to $\tilde{\rho}^e$ and retains the last co-ordinate. The convenience offered by this extension is that the constraint $\tilde{C}_v$ can now be replaced by a constraint C_v that is a conjunction of *homogeneous* quadratic equations. This is by simply replacing a quadratic equation

$$c + \sum_{1 \le i \le m-1} c_i \, \overline{z}_{v,i} + \sum_{1 \le i < j \le m-1} c_{ij} \, \overline{z}_{v,i} \, \overline{z}_{v,j} = 0, \tag{22}$$

by a homogeneous[7] quadratic equation

$$c \cdot \overline{y}_{v,m} + \sum_{1 \le i \le m-1} c_i \, \overline{y}_{v,i} + \sum_{1 \le i < j \le m-1} c_{ij} \, \overline{y}_{v,i} \, \overline{y}_{v,j} = 0. \tag{23}$$

Since the last co-ordinate of the label $\overline{y}_v$ is supposed to be $\overline{y}_{v,m} = 1$, the equation is equivalent to the original equation provided $\overline{y}_{v,m}$ indeed equals 1. On the other hand, if k assignments $\overline{y}_v^1 = (\overline{z}_v^1, b_1), \ldots, \overline{y}_v^k = (\overline{z}_v^k, b_k)$ satisfy Equation (23) in super-position and moreover that $\sum_{\ell=1}^k b_\ell = 1$, then the k assignments $\overline{z}_v^1, \ldots, \overline{z}_v^k$ satisfy Equation (22) in super-position. In other words, while considering equations in k-wise super-position, Equation (23) is equivalent to (22) provided that the sum of the last bits of the k (new) assignments equals 1.

After the minor transformations above, we note the quantitative parameters of the Label Cover instance in Theorem 6.4. We have $|V| = 2^{(\log n)^{8b+4}}$ (the "larger" side of Label Cover), $m = (\log n)^{10b+8}$ (length of labels to the "larger" side), $\delta = 2^{-(\log n)^{2b-1}}$ and $k = \frac{1}{2}(\log n)^b$. Let $|V| = N$ denote the instance size and b be a large enough constant. Note that the reductions in Theorem 5.3 and Theorem 6.4 are both quasi-polynomial in the size of the original 3SAT instance. We now state our first abstraction:

Theorem 7.1. *For any small enough constant $\varepsilon > 0$, there is a quasi-polynomial time reduction from an instance of* 3SAT *to a bi-regular instance $\mathcal{A}$ of* Label-Cover-Intermediate *such that*

- *Vertex sets U and V are bounded in size by N.*
- *Lengths of labels r and m are bounded in size by $(\log N)^{5/4+\varepsilon}$.*
- *For each edge $e=(u,v)$, the map $\rho^e : \mathbb{F}[2]^m \mapsto \mathbb{F}[2]^r$ is homogeneous linear.*
- *For each vertex $v \in V$, there is a constraint C_v that is a conjunction of homogenous quadratic equations over the label set $\mathbb{F}[2]^m$.*
- *$\delta = 2^{-(\log N)^{1/4-\varepsilon}}$ and $k = (\log N)^{1/8-\varepsilon}$.*

The reduction satisfies:

[7] By *homogeneous* we mean here that there is no constant term. This usage of the term is somewhat abused. If one wishes, one can replace the linear terms by squared terms which is equivalent over $\mathbb{F}[2]$ and then the term homogeneous will be correct.

1. YES Case. *If the* 3SAT *instance is satisfiable then there is a labeling* $\overline{x}_u \in \mathbb{F}[2]^r$ *and* $\overline{y}_v \in \mathbb{F}[2]^m$ *for vertices* $u \in U$ *and* $v \in V$ *such that for all vertices* u, v *and all edges* $e = (u, v)$*, we have (i) the last coordinates of* $\overline{x}_u$ *and* $\overline{y}_v$ *are* 1*, i.e.* $\overline{x}_{u,r} = \overline{y}_{v,m} = 1$ *(ii) the labeling satisfies the edge, i.e* $\rho^e(\overline{y}_v) = (\overline{x}_u)$ *and (iii) the label* $\overline{y}_v$ *satisfies the constraint* C_v*.*

2. NO Case. *(*(k, δ) *Soundness) If the* 3SAT *instance is not satisfiable then the following cannot hold:*

   - *There are* k *labels* $\overline{x}_u^1, \ldots, \overline{x}_u^k$ *for each* $u \in U$*.*
   - *There are* k *labels* $\overline{y}_v^1, \ldots, \overline{y}_v^k$ *for each* $v \in V$ *that satisfy* C_v *in super-position and* $\sum_{j=1}^k \overline{y}_{v,m}^j = 1$*, i.e. the sum of their last co-ordinates is* 1*.*
   - *For* δ *fraction of the edges* $e = (v, u)$*, we have* $\rho^e(\overline{y}_v^j) = \overline{x}_u^j, \ \ \forall j \in \{1, \ldots, k\}$*.*

3. δ-Smoothness. *For any* $v \in V$ *and* $\overline{y}_v \in \mathbb{F}[2]^m$*,* $\overline{y}_v \neq 0$*, over the choice of a random a edge* $e = (u, v)$ *incident on* v*,*

$$\Pr_{e=(u,v)} [\rho^e(\overline{y}_v) = 0] \leq \delta.$$

Remark: To see the *smoothness* property above, consider any $v \in V$ and $\overline{y}_v = (\overline{z}_v, b) \neq 0$. For any $e = (u, v)$, $\rho^e(\overline{y}_v) = (\tilde{\rho}^e(\overline{z}_v), b)$. Clearly, if $b \neq 0$, then the property holds. Otherwise, if $\overline{z}_v \neq 0$, then it follows from the smoothness property given in Equation (12) and Section 6.3.

7.2 The Second Abstraction: Label Cover with Matrix Labels

This section provides a further abstraction of Label-Cover-Intermediate instance given in Theorem 7.1. As mentioned earlier, in this abstraction, the labels are matrices instead of vectors. Specifically, if $\overline{y}_v \in \mathbb{F}[2]^m$ is a label before, then the new label is the rank one symmetric matrix $M_v = \overline{y}_v \otimes \overline{y}_v$. This is more convenient in two respects: firstly, the homogeneous quadratic constraints on (the entries of) the vector $\overline{y}_v$ can now be viewed as homogeneous linear constraints on (the entries of) the matrix M_v. Secondly, in the NO Case, an upper bound on the number of assignments used in a super-position can be stated more succinctly as an upper bound on the rank of the matrix. This transformation from a Label-Cover-Intermediate instance leads to a Label-Cover-Final instance summarized below. The proof of its correctness is a bit delicate.

Theorem 7.2. *For any small enough constant* $\varepsilon > 0$*, there is a quasi-polynomial time reduction from an instance of* 3SAT *to a bi-regular instance* $\mathcal{B}$ *of* Label-Cover-Final *such that*

- *Vertex sets* U *and* V *are bounded in size by* N*.*
- *Lengths of labels are* r^2 *and* m^2 *where* r *and* m *are bounded by* $(\log N)^{5/4+\varepsilon}$*. The labels are viewed as* $r \times r$ *and* $m \times m$ *matrices respectively.*
- *For each edge* $e = (u, v)$*, the map* $\pi^e : \mathbb{F}[2]^{m \times m} \mapsto \mathbb{F}[2]^{r \times r}$ *is homogeneous linear. Moreover, every* $r \times r$ *matrix* A *can be "lifted" to a* $m \times m$ *matrix* $A \circ \pi^e$ *such that for all* $m \times m$ *matrices* M*, we have* $\langle A \circ \pi^e, M \rangle = \langle A, \pi^e(M) \rangle$*.*
- *For each vertex* $v \in V$*, there is a constraint* C_v *that is a conjunction of homogenous linear equations over* $\mathbb{F}[2]^{m \times m}$ *(i.e. in the entries of an* $m \times m$ *matrix).*
- $\delta = 2^{-(\log N)^{1/4-2\varepsilon}}$ *and* $k = (\log N)^{1/8-2\varepsilon}$*.*

The reduction satisfies:

1. YES Case. *If the* 3*SAT instance is satisfiable then*

   - *For each $u \in U$, there is a labeling $\overline{x}_u \otimes \overline{x}_u$ for some $\overline{x}_u \in \mathbb{F}[2]^r$.*
   - *For each $v \in V$, there is a labeling $\overline{y}_v \otimes \overline{y}_v$ for some $\overline{y}_v \in \mathbb{F}[2]^m$ such that the m^{th} coordinate of $\overline{y}_v$ is 1.*
   - *Each edge $e = (u, v)$ is satisfied by the above labeling, i.e. $\pi^e(\overline{y}_v \otimes \overline{y}_v) = \overline{x}_u \otimes \overline{x}_u$. Further, for every $v \in V$, the matrix $\overline{y}_v \otimes \overline{y}_v$ satisfies the constraint C_v.*

2. NO Case. *((k, δ) Soundness) If the* 3*SAT instance is not satisfiable then the following cannot hold:*

   - *There is a symmetric matrix $M_u \in \mathbb{F}[2]^{r \times r}$ for each $u \in U$.*
   - *There is a symmetric matrix $M_v \in \mathbb{F}[2]^{m \times m}$ for each $v \in V$ such that $\mathsf{rank}(M_v) \leq k$, the (m, m) entry of M_v is 1 and it satisfies the constraint C_v.*
   - *For δ fraction of the edges $e = (u, v)$, we have $\pi^e(M_v) = M_u$.*

3. Smoothness. *For any $v \in V$ and any symmetric non-zero matrix M_v with $\mathsf{rank}(M_v) \leq k$, over the choice of a random edge $e = (u, v)$ incident on v,*

$$\Pr_{e=(v,u)}[\pi^e(M_v) = 0] \leq \frac{\delta}{2}. \tag{24}$$

The rest of this section is devoted to proving the above theorem. We appropriately transform the Label-Cover-Intermediate instance $\mathcal{A}$ given by Theorem 7.1. The vertex sets U and V as well as the edge set remain unchanged. The parameters r, m remain unchanged whereas the parameter k, δ change slightly. As mentioned before, the labels are now $r \times r$ and $m \times m$ matrices respectively. The construction of new projection maps π^e from the previous ones ρ^e is somewhat delicate.

Projections: Consider an edge $e = (u, v)$. In $\mathcal{A}$, there is a homogeneous linear map $\rho = \rho^e : \mathbb{F}[2]^m \mapsto \mathbb{F}[2]^r$. We can write,

$$\rho(\overline{x}) := \Gamma \overline{x},$$

for some $r \times m$ matrix Γ. Define the homogeneous linear mapping $\pi = \pi^e : \mathbb{F}[2]^{m \times m} \mapsto \mathbb{F}[2]^{r \times r}$ as:

$$\pi(M) = \Gamma M \Gamma^{\mathsf{T}}.$$

Note that, if M is symmetric, then $\pi(M)$ is symmetric as well. For any $\overline{x}, \overline{y} \in \mathbb{F}[2]^m$, it is easily seen that

$$\pi(\overline{x} \otimes \overline{y}) = \rho(\overline{x}) \otimes \rho(\overline{y}). \tag{25}$$

Defining for an $r \times r$ matrix A, a "lifted" matrix $A \circ \pi \in \mathbb{F}[2]^{m \times m}$ as $A \circ \pi = \Gamma^{\mathsf{T}} A \Gamma$, it is easily seen that for any $M \in \mathbb{F}[2]^{m \times m}$, we have $\langle A \circ \pi, M \rangle = \langle A, \pi(M) \rangle$.

Constraints: In the instance $\mathcal{A}$, for a vertex v, there is a conjunction of homogeneous quadratic equations C_v on the co-ordinates of $\mathbb{F}[2]^m$. Each of these equations is now thought of as a homogeneous linear equation in the co-ordinates of $\mathbb{F}[2]^{m \times m}$. In addition, homogeneous linear equations ensuring that the matrix is symmetric are added as well. By a slight abuse of notation, let C_v also denote the conjunction of homogeneous linear equations so defined on the co-ordinates of $\mathbb{F}[2]^{m \times m}$.

The YES and NO cases of Theorem 7.2 are proved as follows.

7.2.1 YES Case

There is a labeling $\overline{x}_u \in \mathbb{F}[2]^r$ and $\overline{y}_v \in \mathbb{F}[2]^m$ for vertices $u \in U$ and $V \in V$ given by the YES Case of Theorem 7.1. Define $M_u = \overline{x}_u \otimes \overline{x}_u$, and $M_v = \overline{y}_v \otimes \overline{y}_v$. From the reduction above, for any edge $e = (u, v)$,

$$\pi^e(\overline{y}_v \otimes \overline{y}_v) = \rho^e(\overline{y}_v) \otimes \rho^e(\overline{y}_v) = \overline{x}_u \otimes \overline{x}_u,$$

since $\rho^e(\overline{y}_v) = \overline{x}_u$ for a satisfied edge $e = (u, v)$ in $\mathcal{A}$. Thus, the labeling satisfies all the edges in $\mathcal{B}$. Furthermore, we have that $\overline{y}_v$ satisfies the conjunction C_v of homogeneous quadratic equations and the m^{th} coordinate of $\overline{y}_v$ is 1. Therefore, $\overline{y}_v \otimes \overline{y}_v$ satisfies the corresponding conjunction C_v of homogeneous linear equations. This completes the YES Case.

7.2.2 NO Case

Assume for the sake of contradiction that there are symmetric matrices M_u and M_v for each $u \in U$ and $v \in V$ satisfying the conditions in the NO Case of Theorem 7.2. Let δ, k be as in the statement of the theorem. These parameters are slightly different (and worse) from the corresponding parameters δ', k' in Theorem 7.1 and related as $k = \lfloor \frac{2k'}{3} \rfloor$ and $\delta = 2^{k'^2+1}\delta'$. We start by showing that the projection maps π^e are essentially rank-preserving (which implies in particular the smoothness property in Theorem 7.2).

Lemma 7.3. *Fix any $v \in V$, a rank parameter $\ell \leq k$ and a rank ℓ matrix $M = M_v \in \mathbb{F}[2]^{m \times m}$. Then over the choice of a random edge $e = (u, v)$ incident on v and $\pi = \pi^e$, we have $\mathsf{rank}(\pi(M)) = \ell$ except with probability $\frac{\delta}{2}$.*

Proof. Using Lemma 2.1, M can be decomposed into the canonical form:

$$M = \sum_{j=1}^{s} \overline{z}_j \otimes \overline{z}_j + \sum_{j=1}^{t} \overline{z}_{s+2j-1} \otimes \overline{z}_{s+2j} + \overline{z}_{s+2j} \otimes \overline{z}_{s+2j-1}, \tag{26}$$

where $\ell = s + 2t$ is the rank of M. From Equation (25), we get that

$$\pi(M) = \sum_{j=1}^{s} \rho(\overline{z}_j) \otimes \rho(\overline{z}_j) + \sum_{j=1}^{t} \rho(\overline{z}_{s+2j-1}) \otimes \rho(\overline{z}_{s+2j}) + \rho(\overline{z}_{s+2j}) \otimes \rho(\overline{z}_{s+2j-1}), \tag{27}$$

where $\rho = \rho^e$. Moreover, since the vectors $\{\overline{z}_j\}_{j=1}^{\ell}$ are linearly independent, except with probability $(2^\ell - 1)\delta'$ over the choice of the edge $e = (u, v)$, the vectors $\{\rho(\overline{z}_j)\}_{j=1}^{\ell}$ are also linearly independent. The reason is that for every non-zero linear combination z of the vectors $\{\overline{z}_j\}_{j=1}^{\ell}$, the vector $\rho(z)$ is the corresponding linear combination of the vectors $\{\rho(\overline{z}_j)\}_{j=1}^{\ell}$ and is non-zero except with probability δ' by the smoothness guarantee of Theorem 7.1. A union bound over all $2^\ell - 1$ choices for z proves the claim. Thus, except with probability $(2^\ell - 1)\delta' \leq \frac{\delta}{2}$, Equation (27) is a canonical decomposition of the matrix $\pi(M)$ and hence $\mathsf{rank}(\pi(M)) = \ell$ as well. □

By hypothesis, for δ fraction of the edges (u, v), we have $\pi(M_v) = M_u$ and by Lemma 7.3, for $\frac{\delta}{2}$ fraction of the edges (u, v), it holds in addition that $\mathsf{rank}(M_v) = \mathsf{rank}(M_u) \leq k$. Call such edges "good". Using the alternate decomposition in Lemma 2.1, noting that $\frac{3k}{2} \leq k'$ and adding dummy zero vectors as summands if necessary, we can write $M_v = \sum_{j=1}^{k'} \overline{y}_v^j \otimes \overline{y}_v^j$ and $M_u = \pi(M_v) = \sum_{j=1}^{k'} \rho(\overline{y}_v^j) \otimes \rho(\overline{y}_v^j)$. Consider the following labeling to the instance $\mathcal{A}$. For each v, we assign the labels $\overline{y}_v^1, \ldots, \overline{y}_v^{k'}$. For each u,

choose at random k' uniformly random vectors $\overline{x}_u^1, \ldots, \overline{x}_u^{k'}$ from the column space of M_u. For every good edge, with probability at least $2^{-k \cdot k'}$, we get $\rho(\overline{y}_v^j) = \overline{x}_u^j \ \ \forall j \in \{1, \ldots, k'\}$. Moreover since the (m, m) entry of the matrix M_v equals 1, so is the sum of the m^{th} co-ordinates of the vectors $\overline{y}_v^1, \ldots, \overline{y}_v^{k'}$. Finally, since M_v satisfies the constraint C_v, so do the vectors $\overline{y}_v^1, \ldots, \overline{y}_v^{k'}$ in super-position. Since $\frac{\delta}{2} \cdot 2^{-k \cdot k'} \geq \delta'$, this labeling contradicts the NO Case of Theorem 7.1 completing our analysis.

8 Inner Verifier and the Proof of Theorem 1.1

This section gives the final hardness reduction proving the following theorem which implies Theorem 1.1.

Theorem 8.1. *For every constant $\varepsilon > 0$, there is a quasi-polynomial time reduction from 3SAT to a 12-uniform hypergraph $\mathcal{G}$ on n vertices such that,*

YES Case. *If the 3SAT instance is satisfiable then $\mathcal{G}$ is 2-colorable.*

NO Case. *If the 3SAT instance is unsatisfiable then $\mathcal{G}$ does not contain an independent set of relative size $2^{(\log n)^{\frac{1}{20}-\varepsilon}}$.*

As is standard, our reduction amounts to constructing a PCP over the alphabet $\mathbb{F}[2]$. The proof locations correspond to vertices of a hypergraph and the tests correspond to the hyperedges. Our test queries 12 locations from the proof and hence the resulting hypergraph is 12-uniform. In the YES case, a correct proof (i.e. assignment of $\mathbb{F}[2]$ values to the proof locations) corresponds to a valid 2-coloring of the hypergraph. In the NO case, we show that there is no independent set of relative size s for an appropriate setting of the parameter s. In fact, our analysis (as is common) shows that every set of relative size s contains a fraction $(1 - o(1))s^{12}$ fraction of hyperedges completely inside it. The PCP is obtained by composing an Outer PCP, i.e. the Label-Cover-Final instance in Theorem 7.2, with an appropriate Inner PCP. The composed PCP is described below after defining the Hadamard Code and the notion of *folding* over the Hadamard Code.

Definition 8.2. *The Hadamard Code of a matrix $\alpha \in \mathbb{F}[2]^{m \times m}$ is indexed by all matrices $X \in \mathbb{F}[2]^{m \times m}$ and its value at the index X is $\langle \alpha, X \rangle \in \mathbb{F}[2]$. Let the $\{-1, 1\}$ valued function $\chi_\alpha : \mathbb{F}[2]^{m \times m} \mapsto \{-1, 1\}$ be defined as:*

$$\chi_\alpha(X) := (-1)^{\langle \alpha, X \rangle}.$$

It is well-known that the set $\{\chi_\alpha \mid \alpha \in \mathbb{F}[2]^{m \times m}\}$ forms an orthonormal basis for the space of real valued function over $\mathbb{F}[2]^{m \times m}$.

Let $\mathcal{A}$ be the instance of Label-Cover-Final given by Theorem 7.2 with parameters N and ε. The PCP proof consists of, for every vertex $v \in V$, the supposed Hadamard Code of the label/matrix $M_v \in \mathbb{F}[2]^{m \times m}$. Note that M_v is supposed to be a symmetric matrix of rank at most k and is supposed to satisfy a constraint C_v which is a conjunction of homogeneous linear equations. It is easily seen that there is a subspace $\mathcal{H}_v$ of $\mathbb{F}[2]^{m \times m}$ such that a matrix M_v is symmetric and satisfies C_v if and only if the Hadamard Code of M_v is constant over the cosets of $\mathcal{H}_v$. We therefore identify all the proof locations that correspond to the same coset of $\mathcal{H}_v$. This technique, known as "folding over constraint C_v", has the following consequence: let $A : \mathbb{F}[2]^{m \times m} \mapsto \{0, 1\}$ be an indicator function of a (supposed independent) set that is constant on the cosets of $\mathcal{H}_v$. Then for every non-zero Fourier coefficient $\hat{A}(\alpha)$, it must be the case that α is symmetric and satisfies the constraint C_v. We now describe the PCP.

Test of Verifier

1. Choose $u \in U$ uniformly at random and $v, w \in V$ uniformly and independently at random from the set of neighbors of u. Let π be the projection corresponding to the edge (u, v) and let π' be the projection corresponding to the edge (u, w). Uniformly and independently at random choose $X, Y, Z, X', Y', Z' \in \mathbb{F}[2]^{m \times m}$ and $\overline{x}, \overline{y}, \overline{x}', \overline{y}' \in \mathbb{F}[2]^m$ and $F \in \mathbb{F}[2]^{r \times r}$.

2. Let $K \in \mathbb{F}[2]^{m \times m}$ be the matrix with its (m, m) entry set to 1 and the rest of the entries set to 0. Let $\mathsf{Diag}(\overline{x}) \in \mathbb{F}[2]^{m \times m}$ be the diagonal matrix with $\overline{x}$ as the diagonal.

3. Let A_v and A_w be the supposed Hadamard Codes of the labels to v and w respectively. These are assumed to be folded over the constraints C_v and C_w respectively.

4. Accept if and only if the following 12 values are not all equal:

$$\begin{array}{ll} A_v(X) & A_v(X + \mathsf{Diag}(\overline{x})) \\ A_v(Y) & A_v(Y + \mathsf{Diag}(\overline{y})) \\ A_v(Z) & A_v(Z + (F \circ \pi) + \overline{x} \otimes \overline{y}) \\ A_w(X') & A_w(X' + \mathsf{Diag}(\overline{x}')) \\ A_w(Y') & A_w(Y' + \mathsf{Diag}(\overline{y}')) \\ A_w(Z') & A_w(Z' + (F \circ \pi') + \overline{x}' \otimes \overline{y}' + K), \end{array}$$

The rest of this section proves the YES and the NO Cases of Theorem 8.1 viewing the PCP construction as a hypergraph $\mathcal{G}$.

8.1 YES Case

Suppose $\mathcal{A}$ is a YES instance as given in Theorem 7.2. There are vectors $\overline{x}_u$ for each $u \in U$ and $\overline{y}_v$ for each $v \in V$ satisfying the conditions of the YES Case in Theorem 7.2. For each $v \in V$, let A_v be the Hadamard Code of the matrix $\overline{y}_v \otimes \overline{y}_v$. Since $\overline{y}_v \otimes \overline{y}_v$ is symmetric and satisfies the set of homogeneous linear constraints C_v, this Hadamard Code is folded as required. Also, note that the (m, m) entry of $\overline{y}_v \otimes \overline{y}_v$ is 1.

We show that test of the verifier accepts with probability 1. Observe first that for any edge $e = (u, v)$ and $F \in \mathbb{F}[2]^{r \times r}$,

$$\langle \overline{y}_v \otimes \overline{y}_v, F \circ \pi^e \rangle = \langle \pi^e(\overline{y}_v \otimes \overline{y}_v), F \rangle = \langle \overline{x}_u \otimes \overline{x}_u, F \rangle.$$

Further,

$$\langle \overline{y}_v \otimes \overline{y}_v, \mathsf{Diag}(\overline{x}) \rangle = \langle \overline{y}_v, \overline{x} \rangle,$$
$$\langle \overline{y}_v \otimes \overline{y}_v, \overline{x} \otimes \overline{y} \rangle = \langle \overline{y}_v, \overline{x} \rangle \langle \overline{y}_v, \overline{y} \rangle.$$

and,

$$\langle \overline{y}_v \otimes \overline{y}_v, K \rangle = 1.$$

Fix any choice made by the verifier in Step (1). Let

$$\begin{array}{lll} x = \langle X, \overline{y}_v \otimes \overline{y}_v \rangle, & y = \langle Y, \overline{y}_v \otimes \overline{y}_v \rangle, & z = \langle Z, \overline{y}_v \otimes \overline{y}_v \rangle \\ x' = \langle X', \overline{y}_w \otimes \overline{y}_w \rangle, & y' = \langle Y', \overline{y}_w \otimes \overline{y}_w \rangle, & z' = \langle Z', \overline{y}_w \otimes \overline{y}_w \rangle, \end{array}$$

and $f = \langle F, \overline{x}_u \otimes \overline{x}_u \rangle$. Using this, the assignment to the 12 query locations in the proof are:

$$\begin{array}{ll} x & x \ + \ \langle \overline{y}_v, \overline{x} \rangle \\ y & y \ + \ \langle \overline{y}_v, \overline{y} \rangle \\ z & z \ + \ \langle \overline{y}_v, \overline{x} \rangle \langle \overline{y}_v, \overline{y} \rangle + f \\ x' & x' \ + \ \langle \overline{y}_w, \overline{x}' \rangle \\ y' & y' \ + \ \langle \overline{y}_w, \overline{y}' \rangle \\ z' & z' \ + \ \langle \overline{y}_w, \overline{x}' \rangle \langle \overline{y}_w, \overline{y}' \rangle + f + 1 \end{array}$$

Clearly, if any of the four values $\langle \overline{y}_v, \overline{x} \rangle, \langle \overline{y}_v, \overline{y} \rangle, \langle \overline{y}_w, \overline{x}' \rangle, \langle \overline{y}_w, \overline{y}' \rangle$ equals 1 then the pair of values in the corresponding row are unequal. So assume that all of these four values are 0. Then, if $f = 1$ the pair of values in the third row are unequal and if $f = 0$, the pair of values in the last row are unequal. Therefore, in the YES Case, there is a proof that is accepted with probability 1. The assignment to the proof locations gives the two color classes of the hypergraph $\mathcal{G}$.

8.2 NO Case

Suppose $\mathcal{A}$ is a NO instance as given in Theorem 7.2 with parameters k and δ therein. We show that for an appropriate setting of the parameter s (chosen towards the end), any subset of vertices $\mathcal{I}$ of relative size s in the hypergraph $\mathcal{G}$ contains essentially s^{12} fraction of the hyperedges completely inside it and hence cannot be an independent set. Let $\mathcal{I}$ be any such subset. For any $v \in V$, let $A_v : \mathbb{F}[2]^{m \times m} \mapsto \{0, 1\}$ be the real-valued indicator function of the subset $\mathcal{I}$ for the locations in the supposed Hadamard Code of v. Thus,

$$\mathbb{E}_{v \in V, \ X \in \mathbb{F}[2]^{m \times m}} [A_v(X)] \geq s.$$

For convenience we write $A = A_v$ and $B = A_w$. The fraction of the hyperedges completely inside $\mathcal{I}$ is:

$$\begin{aligned} \Theta = \mathbb{E} \Big[& A(X) \, A(X + \mathsf{Diag}(\overline{x})) \, A(Y) \, A(Y + \mathsf{Diag}(\overline{y})) \, A(Z) A(Z + F \circ \pi + \overline{x} \otimes \overline{y}) \\ & B(X') \, B(X' + \mathsf{Diag}(\overline{x}')) \, B(Y') \, B(Y' + \mathsf{Diag}(\overline{y}')) \, B(Z') B(Z' + F \circ \pi' + \overline{x}' \otimes \overline{y}' + K) \Big], \end{aligned} \tag{28}$$

where the expectation is taken over the random choice of $u, v, w, X, Y, Z, X', Y', Z', \overline{x}, \overline{y}, \overline{x}', \overline{y}'$ and F. Expanding into the Fourier representation and using standard analysis we obtain,

$$\Theta = \sum_{\alpha, \beta, \gamma, \ \alpha', \beta', \gamma'} \eta(\alpha, \beta, \gamma, \alpha', \beta', \gamma'), \tag{29}$$

where we define,

$$\begin{aligned} \eta(\alpha, \beta, \gamma, \alpha', \beta', \gamma') := \mathbb{E} \Big[& \widehat{A}(\alpha)^2 \, \chi_\alpha(\mathsf{Diag}(\overline{x})) \, \widehat{A}(\beta)^2 \, \chi_\beta(\mathsf{Diag}(\overline{y})) \, \widehat{A}(\gamma)^2 \, \chi_\gamma(F \circ \pi) \\ & \chi_\gamma(\overline{x} \otimes \overline{y}) \, \widehat{B}(\alpha')^2 \, \chi_{\alpha'}(\mathsf{Diag}(\overline{x}')) \, \widehat{B}(\beta')^2 \, \chi_{\beta'}(\mathsf{Diag}(\overline{y}')) \\ & \widehat{B}(\gamma')^2 \, \chi_{\gamma'}(F \circ \pi') \, \chi_{\gamma'}(\overline{x}' \otimes \overline{y}') \, \chi_{\gamma'}(K) \Big]. \end{aligned} \tag{30}$$

Due to folding, the only terms that are possibly non-zero in Equation (29) are those where α, β, γ are symmetric and satisfy the homogeneous linear constraint C_v and similarly α', β', γ' are symmetric and satisfy the constraint C_w. Further, since $F \in \mathbb{F}[2]^{r \times r}$ is uniformly random,

$$\mathbb{E}_F\left[\chi_\gamma(F \circ \pi)\chi_{\gamma'}(F \circ \pi')\right] = \mathbb{E}_F\left[(-1)^{\langle \gamma, (F\circ\pi)\rangle + \langle \gamma', (F\circ\pi')\rangle}\right] = \mathbb{E}_F\left[(-1)^{\langle \pi(\gamma), F\rangle + \langle \pi'(\gamma'), F\rangle}\right],$$

which vanishes unless $\pi(\gamma) = \pi'(\gamma')$. Thus, the terms that are possibly non-zero in Equation (29) satisfy $\pi(\gamma) = \pi'(\gamma')$ and then we can drop the factor $\chi_\gamma(F \circ \pi)\chi_{\gamma'}(F \circ \pi')$ from further consideration. Hereafter, the analysis shall only consider these possibly non-zero terms and we omit explicitly stating these conditions. We split the expectation into three parts and analyze them separately. For an $m \times m$ matrix α, let $\nu(\alpha) \in \{0,1\}$ denote its (m,m) entry. Define:

$$\Theta_0 = \sum_{\substack{\mathsf{rank}(\gamma),\, \mathsf{rank}(\gamma') \le k \\ \nu(\gamma)=\nu(\gamma')=0}} \eta(\alpha,\beta,\gamma,\alpha',\beta',\gamma'), \tag{31}$$

$$\Theta_1 = \sum_{\substack{\mathsf{rank}(\gamma),\, \mathsf{rank}(\gamma') \le k \\ (\nu(\gamma)=1)\vee(\nu(\gamma')=1)}} \eta(\alpha,\beta,\gamma,\alpha',\beta',\gamma'), \tag{32}$$

$$\Theta_2 = \sum_{\max\{\mathsf{rank}(\gamma),\, \mathsf{rank}(\gamma')\} > k} \eta(\alpha,\beta,\gamma,\alpha',\beta',\gamma'). \tag{33}$$

Note that $\Theta = \Theta_0 + \Theta_1 + \Theta_2$. We upper bound $|\Theta_2|$ and $|\Theta_1|$ and lower bound Θ_0, yielding the desired lower bound $\Theta \ge \Theta_0 - |\Theta_1| - |\Theta_2|$.

8.2.1 Upper bound on $|\Theta_2|$

Lemma 8.3. *For any symmetric matrix γ such that $\mathsf{rank}(\gamma) > k$,*

$$\mathbb{E}_{\overline{x}}\left|\mathbb{E}_{\overline{y}}\left[\chi_\alpha(\mathsf{Diag}(\overline{x}))\chi_\beta(\mathsf{Diag}(\overline{y}))\chi_\gamma(\overline{x}\otimes\overline{y})\right]\right| \le 2^{-(k+1)}.$$

Proof. Let $\overline{\alpha}$ and $\overline{\beta}$ be the vectors given by the diagonals of α and β respectively. Using this,

$$\chi_\alpha(\mathsf{Diag}(\overline{x}))\chi_\beta(\mathsf{Diag}(\overline{y}))\chi_\gamma(\overline{x}\otimes\overline{y}) = (-1)^{\psi},$$

where $\psi := \psi(\overline{x},\overline{y}) = \langle\overline{\alpha},\overline{x}\rangle + \langle\overline{\beta},\overline{y}\rangle + \langle\gamma\overline{x},\overline{y}\rangle$. For any given $\overline{x}$, we have $\mathbb{E}_{\overline{y}}\left[(-1)^{\psi}\right] = 0$, whenever $\overline{\beta} \ne \gamma\overline{x}$. If $\overline{\beta}$ is not in the column space of γ then this is always true. On the other hand, if $\overline{\beta}$ is in the column space of γ, then over a random choice of $\overline{x}$, $\gamma\overline{x}$ is a uniformly random vector from the column space of γ, and $\overline{\beta} \ne \gamma\overline{x}$ with probability at least $1 - 2^{-(k+1)}$, since $\mathsf{rank}(\gamma) > k$. This completes the proof. □

Applying the above lemma to both γ and γ' and since $\overline{x}, \overline{y}, \overline{x}'$ and $\overline{y}'$ are independent, we obtain that for any fixed choice of u, v and w, the absolute value of the sum of terms on the RHS in (33) is at most,

$$2 \cdot 2^{-(k+1)} \left(\sum_\alpha \widehat{A}(\alpha)^2\right)^3 \left(\sum_\beta \widehat{B}(\beta)^2\right)^3 \le 2^{-k},$$

using the fact that A and B are $\{0,1\}$-valued functions and hence the sum of their squared coefficients is upper bounded by 1 (i.e. Parseval's inequality). Thus $|\Theta_2| \le 2^{-k}$.

8.2.2 Upper bound on $|\Theta_1|$

For a fixed choice of u, v and w, the absolute value of the RHS in Equation (32) can be bounded by

$$\sum_{\substack{\mathsf{rank}(\gamma),\mathsf{rank}(\gamma')\leq k,\ \pi(\gamma)=\pi'(\gamma'),\\ (\nu(\gamma)=1)\vee(\nu(\gamma')=1)}} \widehat{A}(\gamma)^2\widehat{B}(\gamma')^2 \tag{34}$$

where we again used Parseval. Under the random choice of v and w, the terms in the above sum satisfying $(\nu(\gamma)=1)\wedge(\nu(\gamma')=0)$ have the same expectation as those satisfying $(\nu(\gamma)=0)\wedge(\nu(\gamma')=1)$. Thus,

$$|\Theta_1| \leq 2\cdot \mathbb{E}_{u,v,w}\left[\sum_{\substack{\mathsf{rank}(\gamma),\mathsf{rank}(\gamma')\leq k,\\ \pi(\gamma)=\pi'(\gamma'),\ (\nu(\gamma)=1)}} \widehat{A}(\gamma)^2\widehat{B}(\gamma')^2\right]. \tag{35}$$

Consider the following strategy for labeling vertices $u\in U$ and $v\in V$. For $u\in U$, pick a random neighbor w and choose a matrix γ' with probability $\widehat{A}_w(\gamma')^2$ and assign to u the label $M_u=\pi'(\gamma')$ where $\pi'=\pi^{(u,w)}$. For $v\in V$, choose a matrix γ with probability $\widehat{A}_v(\gamma)^2$ and assign it $M_v=\gamma$. The above equation implies that for $|\Theta_1|/2$ fraction of edges $e=(u,v)$, M_v is of rank at most k, its (m,m) entry is 1, M_v satisfies the constraint C_v, and $\pi^e(M_v)=M_u$. The NO Case of Theorem 7.2 implies that $|\Theta_1|\leq 2\delta$ where δ is the soundness parameter therein.

8.2.3 Lower bound on Θ_0

For any vertex $v\in V$, define

$$\tau(v,\alpha,\beta,\gamma) := \mathbb{E}_{\overline{x},\overline{y}}\left[\widehat{A}(\alpha)^2\chi_\alpha(\mathsf{Diag}(\overline{x}))\widehat{A}(\beta)^2\chi_\beta(\mathsf{Diag}(\overline{y}))\widehat{A}(\gamma)^2\chi_\gamma(\overline{x}\otimes\overline{y})\right], \tag{36}$$

where $A=A_v$. Note that, for all terms in the RHS of Equation (31), $\chi_{\gamma'}(K)=1$. Thus, we can rewrite Θ_0 as (considering the possibilities for the common projection $\pi(\gamma)=\pi'(\gamma')=G$ separately),

$$\Theta_0 = \sum_{G\in\mathbb{F}[2]^{r\times r}} \mathbb{E}_{u,v,w}\left[\sum_{\substack{\mathsf{rank}(\gamma),\mathsf{rank}(\gamma')\leq k\\ \pi(\gamma)=\pi'(\gamma')=G,\ \nu(\gamma)=\nu(\gamma')=0}} \tau(v,\alpha,\beta,\gamma)\tau(w,\alpha',\beta',\gamma')\right]. \tag{37}$$

Fix a fixed vertex u, the choice of its neighbors v and w is independent and identical. Thus the expectation above (for a fixed G and u) is seen to be

$$\left(\mathbb{E}_v\left[\sum_{\substack{\mathsf{rank}(\gamma)\leq k\\ \pi(\gamma)=G,\ \nu(\gamma)=0}} \tau(v,\alpha,\beta,\gamma)\right]\right)^2 \tag{38}$$

which is always non-negative. For the purpose of lower bounding Θ_0, we may thus consider only the term $G=0$ and conclude

$$\Theta_0 \geq \mathbb{E}_{u,v,w}\left[\sum_{\substack{\mathsf{rank}(\gamma),\mathsf{rank}(\gamma')\leq k\\ \pi(\gamma)=\pi'(\gamma')=0,\ \nu(\gamma)=\nu(\gamma')=0}} \tau(v,\alpha,\beta,\gamma)\tau(w,\alpha',\beta',\gamma')\right] \tag{39}$$

We now observe that the contribution of terms with $\gamma \neq 0$ and $\gamma' \neq 0$ to the LHS in the equation above is negligible. Indeed, the expectation of terms involving $\gamma \neq 0$ is upper bounded by (and the same holds for terms involving $\gamma' \neq 0$),

$$\mathbb{E}_{u,v}\left[\sum_{\substack{\gamma\neq 0,\pi(\gamma)=0\\ \mathsf{rank}(\gamma)\leq k}} \widehat{A}(\gamma)^2\right]. \tag{40}$$

Using the smoothness property of Theorem 7.2 with smoothness parameter $\frac{\delta}{2}$, the expectation above is at most

$$\mathbb{E}_{v}\left[\sum_{\gamma\neq 0,\mathsf{rank}(\gamma)\leq k} \widehat{A}(\gamma)^2 \Pr_{u\sim v}\left[\pi(\gamma)=0\right]\right] \leq \frac{\delta}{2}. \tag{41}$$

Thus we conclude that the expectation on the LHS of Equation (39) is essentially due to terms with $\gamma = \gamma' = 0$ and

$$\Theta_0 \geq \mathbb{E}_{u,v,w}\left[\sum_{\alpha,\beta,\alpha',\beta'} \tau(v,\alpha,\beta,0)\tau(w,\alpha',\beta',0)\right] - \delta. \tag{42}$$

The following simple lemma allows us to complete the analysis.

Lemma 8.4. *For any α, β, $\tau(v,\alpha,\beta,0) \geq 0$. In particular, $\tau(v,0,0,0) = \widehat{A}(0)^6$.*

Proof. The second part of the lemma follows by the definition of $\tau(v,0,0,0)$. For the first part, let $\overline{\alpha}$ and $\overline{\beta}$ be the diagonals of α and β respectively. Observe that,

$$\begin{aligned}\mathbb{E}_{\overline{x},\overline{y}}\left[\chi_\alpha(\mathsf{Diag}(\overline{x}))\chi_\beta(\mathsf{Diag}(\overline{y}))\right] &= \mathbb{E}_{\overline{x},\overline{y}}\left[(-1)^{\langle\overline{\alpha},\overline{x}\rangle+\langle\overline{\beta},\overline{y}\rangle}\right]\\ &= \mathbb{1}_{\{(\overline{\alpha}=0)\wedge(\overline{\beta}=0)\}} \geq 0.\end{aligned}$$

□

Using the above lemma and the analogous property for $\tau(w,\alpha',\beta',0)$ in Equation (42), we obtain,

$$\begin{aligned}\Theta_0 &\geq \mathbb{E}_{u,v,w}\left[\widehat{A}(0)^6\widehat{B}(0)^6\right] - \delta\\ &= \mathbb{E}_{u}\left[\left(\mathbb{E}_{v}\left[\widehat{A}(0)^6\right]\right)^2\right] - \delta\\ &\geq \left(\mathbb{E}_{u,v}\left[\widehat{A}(0)\right]\right)^{12} - \delta\\ &\geq s^{12} - \delta. \end{aligned} \tag{43}$$

Combining the lower bound on Θ_0 with upper bounds for $|\Theta_1|$ and $|\Theta_2|$, we get

$$\Theta \geq \Theta_0 - |\Theta_1| - |\Theta_2| \geq s^{12} - \delta - 2\delta - 2^{-k}. \tag{44}$$

Since $k = (\log N)^{1/8-2\varepsilon}$ and $\delta = 2^{-(\log N)^{1/4-2\varepsilon}}$, we may choose $s = 2^{-(\log N)^{1/8-3\varepsilon}}$ and conclude that the fraction of hyperedges inside the subset $\mathcal{I}$ of vertices is essentially s^{12} as desired. The number of vertices in $\mathcal{G}$ is n which is at most $|V| \cdot 2^{m^2}$ which, in turn, is at most $N \cdot 2^{(\log N)^{10/4+2\varepsilon}}$. Thus, in terms of n, the hypergraph has no independent set of relative size $2^{-(\log n)^{1/20-4\varepsilon}}$, completing the proof of Theorem 8.1.

References

[ABSS97] S. Arora, L. Babai, J. Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes, and systems of linear equations. *Journal of Computer and System Sciences*, 54(2):317–331, 1997.

[ACC06] S. Arora, E. Chlamtac, and M. Charikar. New approximation guarantee for chromatic number. In *Proceedings of the ACM Symposium on the Theory of Computing*, pages 215–224, 2006.

[Aro94] S. Arora. *Probabilistic checking of proofs and the hardness of approximation problems*. PhD thesis, UC Berkeley, 1994.

[AS03] S. Arora and M. Sudan. Improved low-degree testing and its applications. *Combinatorica*, 23(3):365–426, 2003.

[BGH+12] B. Barak, P. Gopalan, J. Håstad, R. Meka, P. Raghavendra, and D. Steurer. Making the Long Code shorter. In *Proceedings of the Annual Symposium on Foundations of Computer Science*, pages 370–379, 2012.

[BGS98] M. Bellare, O. Goldreich, and M. Sudan. Free bits, PCPs, and nonapproximability-towards tight results. *SIAM Journal of Computing*, 27(3):804–915, 1998.

[BK97] A. Blum and D. R. Karger. An $\tilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs. *Information Processing Letters*, 61(1):49–53, 1997.

[BKS+10] A. Bhattacharyya, S. Kopparty, G. Schoenebeck, M. Sudan, and D. Zuckerman. Optimal testing of Reed-Muller codes. In *Proceedings of the Annual Symposium on Foundations of Computer Science*, pages 488–497, 2010.

[Blu94] A. Blum. New approximation algorithms for graph coloring. *Journal of the ACM*, 41(3):470–516, 1994.

[CF96] H. Chen and A. M. Frieze. Coloring bipartite hypergraphs. In *Proc. IPCO*, pages 345–358, 1996.

[Cha13] S. O. Chan. Approximation resistance from pairwise independent subgroups. In *Proceedings of the ACM Symposium on the Theory of Computing*, pages 447–456, 2013.

[DG13] I. Dinur and V. Guruswami. PCPs via low-degree long code and hardness for constrained hypergraph coloring. In *Proceedings of the Annual Symposium on Foundations of Computer Science*, 2013.

[DK13] I. Dinur and G. Kol. Covering CSPs. In *Proceedings of the Annual IEEE Conference on Computational Complexity*, pages 207–218, 2013.

[DKPS10] I. Dinur, S. Khot, W. Perkins, and M. Safra. Hardness of finding independent sets in almost 3-colorable graphs. In *Proceedings of the Annual Symposium on Foundations of Computer Science*, pages 212–221, 2010.

[DMR09] I. Dinur, E. Mossel, and O. Regev. Conditional hardness for approximate coloring. *SIAM Journal of Computing*, 39(3):843–873, 2009.

[DRS05] I. Dinur, O. Regev, and C. D. Smyth. The hardness of 3-uniform hypergraph coloring. *Combinatorica*, 25(5):519–535, 2005.

[GHH+14] V. Guruswami, J. Håstad, P. Harsha, S. Srinivasan, and G. Varma. Super-polylogarithmic hypergraph coloring hardness via low-degree long codes. *To Appear in STOC*, 2014.

[GHS02] V. Guruswami, J. Håstad, and M. Sudan. Hardness of approximate hypergraph coloring. *SIAM Journal of Computing*, 31(6):1663–1686, 2002.

[GK04] V. Guruswami and S. Khanna. On the hardness of 4-coloring a 3-colorable graph. *SIAM Journal of Discrete Mathematics*, 18(1):30–40, 2004.

[Hås01] J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.

[Hol02] J. Holmerin. Vertex cover on 4-regular hyper-graphs is hard to approximate within 2 - ε. In *Proceedings of the Annual IEEE Conference on Computational Complexity*, 2002.

[Hua13] S. Huang. Improved hardness of approximating chromatic number. In *APPROX-RANDOM*, pages 233–243, 2013.

[Kho01] S. Khot. Improved inaproximability results for maxclique, chromatic number and approximate graph coloring. In *Proceedings of the Annual Symposium on Foundations of Computer Science*, pages 600–609, 2001.

[Kho02a] S. Khot. Hardness results for approximate hypergraph coloring. In *Proceedings of the ACM Symposium on the Theory of Computing*, pages 351–359, 2002.

[Kho02b] S. Khot. Hardness results for coloring 3-colorable 3-uniform hypergraphs. In *Proceedings of the Annual Symposium on Foundations of Computer Science*, pages 23–32, 2002.

[KLS00] S. Khanna, N. Linial, and S. Safra. On the hardness of approximating the chromatic number. *Combinatorica*, 20(3):393–415, 2000.

[KMH96] P. Kelsen, S. Mahajan, and R. Hariharan. Approximate hypergraph coloring. In *Proc. SWAT*, pages 41–52, 1996.

[KMS98] D. R. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *Journal of the ACM*, 45(2):246–265, 1998.

[KNS01] M. Krivelevich, R. Nathaniel, and B. Sudakov. Approximating coloring and maximum independent sets in 3-uniform hypergraphs. *Journal of Algorithms*, 41(1):99–113, 2001.

[KS12] S. Khot and R. Saket. Hardness of finding independent sets in almost q-colorable graphs. In *Proceedings of the Annual Symposium on Foundations of Computer Science*, pages 380–389, 2012.

[KS14] S. Khot and R. Saket. Hardness of finding independent sets in 2-colorable and almost 2-colorable hypergraphs. *To Appear in SODA*, 2014.

[KT12] K. Kawarabayashi and M. Thorup. Combinatorial coloring of 3-colorable graphs. In *Proceedings of the Annual Symposium on Foundations of Computer Science*, pages 68–75, 2012.

[Sak14] R. Saket. Hardness of finding indpendent sets in 2-colorable hypergraphs and of satisfiable CSPs. *To Appear in CCC*, 2014.

[Wig83] A. Wigderson. Improving the performance guarantee for approximate graph coloring. *Journal of the ACM*, 30(4):729–735, 1983.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de