Electronic Colloquium on Computational Complexity, Report No. 62 (2015)

$2^{(\log N)^{1/4-o(1)}}$ Hardness for Hypergraph Coloring

Sangxia Huang [*]

April 15, 2015

Abstract

We show that it is quasi-NP-hard to color 2-colorable 8-uniform hypergraphs with $2^{(\log N)^{1/4-o(1)}}$ colors, where N is the number of vertices. There has been much focus on hardness of hypergraph coloring recently. In [15], Guruswami, Håstad, Harsha, Srinivasan and Varma 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 standard Label-Cover with an inner-verifier based on Low-Degree-Long-Code, using Reed-Muller code testing results by Dinur and Guruswami [10]. Using a different approach in [27], Khot and Saket constructed a new variant of Label-Cover, and composed it with Quadratic-Code to show quasi-NP-hardness of coloring 2-colorable 12-uniform hypergraphs with $2^{(\log N)^c}$ colors, for some c around $1/20$. Their construction of Label-Cover is based on a new notion of *superposition complexity* for CSP instances. The composition with inner-verifier was subsequently improved by Varma, giving the same hardness result for 8-uniform hypergraphs [33].

Our construction uses both Quadratic-Code and Low-Degree-Long-Code, and builds upon the work by Khot and Saket. We present a different approach to construct CSP instances with superposition hardness by observing that when the number of assignments is odd, satisfying a constraint in superposition is the same as *odd-covering* a constraint. We employ Low-Degree-Long-Code in order to keep the construction efficient. In the analysis, we also adapt and generalize one of the key theorems by Dinur and Guruswami [10] in the context of analyzing probabilistically checkable proof systems.

1 Introduction

For an integer $k \geq 2$, a k-uniform hypergraph $H = (V, F)$ consists of vertex set V and edge set $F \subseteq \binom{V}{k}$. A set of vertices $S \subseteq V$ is an *independent set* if for all $f \in F$, $f \nsubseteq S$, i.e., no edge is completely inside S. A hypergraph is q-colorable if its vertices can be partitioned into q disjoint independent sets.

Coloring a graph or a hypergraph using few colors is a classical combinatorial optimization problem, and is one of the most well-studied problems in theoretical computer science. It is also closely related to other problems such as finding maximum independent sets, PCPs with certain special properties, and also inapproximability of constraint satisfaction problems. In addition to being an important theoretical challenge, graph coloring also has a number of applications such as scheduling and register allocation.

[*] `huang.sangxia@gmail.com` KTH Royal Institute of Technology, Stockholm, Sweden.

ISSN 1433-8092

We use $\alpha(H)$ to denote the fractional size of the maximum cardinality independent set of H, also known as the *fractional independence number*, and we use $\chi(H)$ to denote the minimum q such that H is q-colorable. It is easy to verify that we have $\chi(H)\alpha(H) \geq 1$ for any H.

In the ordinary graph case, corresponding to $k = 2$, deciding whether a graph G has a 2-coloring is the same as deciding whether it is a bipartite graph, and can be easily solved in polynomial time. In general, however, determining the chromatic number of a graph exactly is NP-hard [14]. In fact, even coloring 3-colorable graphs with 4 colors is NP-hard. For general q-colorable graphs, it is NP-hard to color with $q + 2\lfloor \frac{q}{3} \rfloor - 1$ colors [24, 17]. For sufficiently large q, it was shown that it is NP-hard to color a q-colorable graph with $2^{\Omega(q^{1/3})}$ colors [20], improving on an earlier lower-bound of $q^{\frac{1}{25}\log q}$ by Khot [25]. Assuming a variant of Khot's 2-to-1 Conjecture, Dinur, Mossel and Regev [11] proved that it is NP-hard to q'-color a q-colorable graph for any $3 \leq q < q'$. The dependency between the hardness of graph coloring and the parameters of 2-to-1 Label-Cover was made explicit and improved by Dinur and Shinkar [13], who showed that it is NP-hard to $(\log n)^c$-color a 4-colorable graph for some constant $c > 0$ assuming the 2-to-1 Conjecture. As for algorithms, there have been many results as well [34, 6, 21, 7]. For 3-colorable graphs, the best algorithm is by Kawarabayashi and Thorup [23] which uses $O(n^{0.19996})$ colors, based on results by Arora and Chlamtac [2] Chlamtac [9] and the earlier work of Kawarabayashi and Thorup [22]. As we can see, there is still a huge gap between the best approximation guarantee and the best hardness result.

For $k \geq 3$, even determining whether a k-uniform hypergraph has a 2-coloring is NP-hard. In terms of approximation algorithms, the best algorithm for 2-colorable 3-uniform hypergraphs still requires $n^{\Omega(1)}$ colors [29, 1, 8].

From the hardness side, the first super-constant hardness result was proved in [16]. The main result there is that for 2-colorable 4-uniform hypergraphs, finding a coloring with any constant number of colors is NP-hard, and finding a coloring with $O(\log\log n/\log\log\log n)$ colors is quasi-NP-hard. For 2-colorable 3-uniform hypergraphs, a similar hardness result was proved in [12]. Khot [26] proved that coloring 3-colorable 3-uniform hypergraphs with any constant number of colors is hard, and for q-colorable 4-uniform hypergraphs, coloring with $(\log n)^{\Omega(q)}$ colors is quasi-NP-hard for $q \geq 7$. The analysis in [16] was improved by Holmerin, who proved that even finding an independent set of fractional size $\Omega(\log\log\log n/\log\log n)$ is quasi-NP-hard [19]. The construction was further improved recently by Saket [32], who proved that it is quasi-NP-hard to find independent set of size $n/(\log n)^{\Omega(1)}$ in 2-colorable 4-uniform hypergraphs [32]. There has also been work on the hardness of finding independent sets in almost 2-colorable hypergraphs — hypergraphs that becomes 2-colorable after removing a small fraction of vertices. Much stronger result is known, albeit at the cost of imperfect completeness. We refer to [28] for more details.

Recently, in [15], Guruswami, Harsha, Håstad, Srinivasan and Varma proved the first super-polylogarithmic hardness result for hypergraph coloring, showing hardness for coloring 2-colorable 8-uniform hypergraphs with $2^{2^{\Omega(\sqrt{\log\log n})}}$ colors. Their reduction uses the Low-Degree-Long-Code proposed in [5], based on techniques for testing Reed-Muller codes developed in [10].

Using a very different approach, Khot and Saket gave another exponential improvement in [27], showing a quasi-NP-hardness for coloring 2-colorable 12-uniform hypergraphs with $\exp((\log n)^{\Omega(1)})$ colors, where the constant in $\Omega(1)$ is around $1/20$, although it might be improved with a more careful analysis of their reduction. Part of their analysis was subsequently simplified by Varma in [33] using ideas from [15].

In this work, we give another improvement for hardness of hypergraph coloring. Our main result is as follows.

Theorem 1.1. *It is quasi-*NP*-hard to color a* 2*-colorable* 8*-uniform hypergraph of size* N *with* $2^{(\log N)^{1/4-o(1)}}$ *colors.*

1.1 Proof Overview

We start by describing the PCP reduction of proving hypergraph coloring hardness used in many previous works. Most of these results show hardness of finding an independent set of large fractional size. We can view the output of these reductions as $\textsc{NotAllEqual}_k$ CSP instances. The variables correspond to the vertices of a hypergraph, and the $\textsc{NotAllEqual}_k$ constraints correspond to the hyperedges. Note that for hypergraph coloring results, all variables appear positively in such instances and no negations are allowed. An assignment that satisfies all the $\textsc{NotAllEqual}_k$ constraints thus gives a perfect 2-coloring for the hypergraph. In the other direction, a set of vertices in the hypergraph naturally corresponds to a $\{0, 1\}$ assignment to the variables in the $\textsc{NotAllEqual}_k$ instance, and the vertices form an independent set if for all constraints in the $\textsc{NotAllEqual}_k$ instances, there is at least 1 variable that is assigned 0.

The starting point of the reduction is usually some Label-Cover hardness. We then encode the supposed labeling for the Label-Cover instance with some coding scheme, and design a PCP to test the consistency of the labeling.

One classical choice of encoding is the Long-Code, which encodes m bits of information with 2^{2^m} bits. This huge blowup makes it impossible to prove hardness results better than $\operatorname{polylog} n$ via the Label-Cover plus Long-Code approach.

A much more efficient encoding is the Hadamard code, which only uses 2^m bits to encode m bits of information. However, the disadvantage of the Hadamard code is that one can only enforce linear constraints on the codewords, which means that we can only start from hard problems involving only linear constraints, and as a result, we lose perfect completeness and can only prove results about almost coloring.

The Low-Degree-Long-Code proposed in [5] lies somewhere between Long-Code and Hadamard code. We can view Hadamard code as encoding m bits by writing down the evaluation of all m-variable functions of degree at most 1 on these m bits, and Long-Code as writing down the evaluation of all possible m-variable functions — that is, degree up to m — on these m bits. Low-Degree-Long-Code has a parameter d, the degree, and the encoding writes down the evaluation of all polynomials of degree at most d. Dinur and Guruswami [10] obtained hardness result for a variant of hypergraph coloring based on Low-Degree-Long-Code, and the techniques were soon adapted in [15] to get a hardness result of $2^{2^{\Omega(\sqrt{\log\log n})}}$.

The aforementioned result by Khot and Saket [27] uses Quadratic Code, which is the same as Low-Degree-Long-Code with $d = 2$. Their construction, however, is completely different from that in [15].

One can view the Quadratic Code used in [27] as the Hadamard encoding of matrix M that is symmetric and has rank 1, that is, there exists some $u \in \mathbb{F}_2^m$ such that $M = u \otimes u$. Khot and Saket described a 6-query test such that if some encoding function $f : \mathbb{F}_2^{m\times m} \to \mathbb{F}_2$ passes the test with non-trivial probability, then we can decode it into a low rank matrix.

In order to use this encoding, it seems natural that one would like to construct some variant of LABEL-COVER where the labels are now matrices, with some linear constraints on the entries of the matrices (since as discussed above we are using Hadamard code to encode the matrices). In the completeness case, we would like to have some matrix labelings of rank 1 that satisfies all linear constraints on the vertices as well as projection constraints on the edges, and in the soundness case, not even labelings with low rank matrices can satisfy more than a small fraction of them.

Such LABEL-COVER hardness result is not readily available. Khot and Saket proposed the notion of *superposition complexity* for quadratic equations. Briefly speaking, let $q(x) = c + \sum_{i=1}^{m} c_i x_i + \sum_{1 \leq i < j \leq m} c_{ij} x_i x_j = 0$ be a quadratic equation on m $\mathbb{F}_2$-variables. We say that t assignments $a^{(1)}, \cdots, a^{(t)} \in \mathbb{F}_2^m$ satisfy the equation $q(x) = 0$ *in superposition* if

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

If we have a system of quadratic equations, then we say that t assignments satisfy the system of quadratic equations in superposition if each quadratic equation is satisfied in superposition. Having a small number of assignments satisfying quadratic constraints in superposition is exactly the same as having a symmetric low-rank matrix satisfying the linearized version of the constraints, as we discuss in more detail in Section 2.

Through a remarkable chain of reductions, Khot and Saket established the inapproximability of quadratic equations with superposition complexity, as well as the actual construction of the LABEL-COVER with matrix labels. They started with superposition hardness for E3-SAT with gap of $1/n$, and use low-degree testing and sum-check protocol like in the original proof of the PCP theorem [3, 4] to achieve a superposition hardness result for systems of quadratic equations with good soundness and moderate blowup. This is then followed by a Point versus Surface test which produces the actual LABEL-COVER instance.

The focus of this work is also the construction of such LABEL-COVER instances. Let t be some odd natural number. A set of t assignments odd-covers an equation (or more generally, a constraint) if the number of assignments that satisfy the equation is odd. We show in Section 2 that the notion of odd-covering is equivalent to satisfaction in superposition when the number of assignments is odd. This viewpoint enables us to construct the kind of LABEL-COVER instance used in [27] very easily. In fact, the reduction in Section 3 looks very much like a classical CSP inapproximability proof.

Although simpler, the above observation alone is not sufficient to give us a hardness result better than [27]. The issue here is that for the reduction in Section 3 to work for our choice of parameters, the soundness of the LABEL-COVER that we start with needs to be sub-constant, and a typical LONG-CODE reduction will again blow up the size of the instance by too much. Hence, for this step, we employ LOW-DEGREE-LONG-CODE. Our technical contribution here is Theorem 2.27, a generalization of the Reed-Muller code testing result of [10].

2 Preliminaries

Before we discuss the relation between superposition, odd-covering and low rank matrices, we define an operation on vectors and matrices that we will use frequently.

Definition 2.1. *Define $D_1 : \mathbb{F}_2^{m+1} \to \mathbb{F}_2^m$ as the operator that removes the first coordinate of a vector. Define D_1 similarly for matrices as the operator that removes the first row and column of a given matrix.*

2.1 Superposition and Odd-Covering

Khot and Saket [27] defined the notion of satisfying in superposition as follows.

Definition 2.2 (Superposition)**.** *Let $a^{(1)}, \cdots, a^{(t)} \in \mathbb{F}_2^m$ be t assignments and $q(x) = 0$ be a quadratic equation in m $\mathbb{F}_2$-variables with*

$$q(x) = c + \sum_{i=1}^{m} c_i x_i + \sum_{1 \le i < j \le m} c_{ij} x_i x_j \,.$$

We say that the t assignments satisfy the equation $q(x) = 0$ in superposition *if*

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

Definition 2.3. *Given a system of quadratic equations $\{q_i(x) = 0\}_{i=1}^{L}$ on variables $x_1, \cdots, x_m$, its* superposition complexity *is the minimum number t, if it exists, such that there are t assignments $a^{(1)}, \cdots, a^{(t)} \in \mathbb{F}_2^m$ that satisfy each equation $q_i(x) = 0$ in superposition.*

We define the odd superposition complexity *(or* even superposition complexity*) to be the minimum odd integer t (or even integer t, respectively) such that there are t assignments that satisfy all equations in superposition.*

Note that by simply adding all 0 assignments, we can argue that the above three notions of superposition complexity differ by at most 1.

We now explain the relation between superposition complexity of quadratic equations and low rank matrices. Assume for simplicity of exposition that the quadratic equation $q(x) = 0$ as defined above is homogeneous, that is, the constant term c and the linear coefficients c_i are all 0.

We can express a homogeneous quadratic equation $q(x) = 0$ with a matrix by defining $C \in \mathbb{F}_2^{m \times m}$, where $C_{ij} = c_{ij}$ for $1 \le i < j \le m$, and $C_{ij} = 0$ otherwise. Let $x = (x_1 \; x_2 \; \cdots \; x_m)$. Then $q(x) = 0$ is the same as $\langle C, x \otimes x \rangle = x^T C x = 0$, where $\langle \cdot, \cdot \rangle$ denotes the entry-wise dot product of two matrices. Note that $x \otimes x$ is a symmetric rank-1 matrix.

Suppose now that we have a symmetric matrix A such that $\langle C, A \rangle = 0$. For a fixed C, this is a linear constraint on the entries of A. If in addition A has rank 1, then there exists x_a, such that $A = x_a \otimes x_a$, and by the above, we have that x_a satisfies $q(x_a) = 0$. Therefore, if A is a symmetric rank 1 matrix and $\langle C, A \rangle = 0$, then A encodes an assignment that satisfies the quadratic equation $q(x) = 0$.

The following decomposition lemma from [27] illustrates the situation when A has low rank.

Lemma 2.4. *Let $A \in \mathbb{F}_2^{m \times m}$ be a symmetric matrix of rank k over $\mathbb{F}_2$. Then there exists $l \le 3k/2$ and vectors $v_1, \cdots, v_l$ in the column space of A, such that $A = \sum_{i=1}^{l} v_i \otimes v_i$.*

Let A be a low rank matrix and $v_1, \cdots, v_l$ be $l \leq 3k/2$ assignments given by Lemma 2.4. Then

$$\begin{aligned} 0 = \langle C, A \rangle &= \sum_{t=1}^{l} \langle C, v_t \otimes v_t \rangle \\ &= \sum_{t=1}^{l} \sum_{1 \leq i < j \leq m} c_{ij} v_{ti} v_{tj} \\ &= \sum_{1 \leq i < j \leq m} c_{ij} \sum_{t=1}^{l} v_{ti} v_{tj} \,. \end{aligned}$$

Therefore we have that $v_1, \cdots, v_l$ satisfy $q(x) = 0$ in superposition.

The notion we will now consider is the following, which we call *odd-covering.*

Definition 2.5 (Odd-covering)**.** *Let $a^{(1)}, \cdots, a^{(t)} \in \mathbb{F}_2^m$ be t assignments and $q(x) = 0$ be a quadratic equation in m $\mathbb{F}_2$-variables as defined above. We say that the t assignments* odd-cover *the equation $q(x) = 0$ if the number of assignments $a^{(l)}$ that satisfies $q(a^{(l)}) = 0$ is odd.*

The key observation is that odd-covering and satisfying in superposition are equivalent when the number of assignments involved is odd.

Lemma 2.6. *Let t be an odd integer and $a^{(1)}, \cdots, a^{(t)} \in \mathbb{F}_2^m$ be t assignments, and $q(x) = 0$ be a quadratic equation in m $\mathbb{F}_2$-variables as defined above. Then the t assignments satisfy $q(x) = 0$ in superposition if and only if the t assignments odd-cover $q(x) = 0$.*

Proof. Using the fact that t is odd, we have the following

$$\begin{aligned} \sum_{l=1}^{t} q(a^{(l)}) &= \sum_{l=1}^{t} \left(c + \sum_{i=1}^{m} c_i a_i^{(l)} + \sum_{1 \leq i < j \leq m} c_{ij} a_i^{(l)} a_j^{(l)} \right) \\ &= t \cdot c + \sum_{l=1}^{t} \sum_{i=1}^{m} c_i a_i^{(l)} + \sum_{l=1}^{t} \sum_{1 \leq i < j \leq m} c_{ij} a_i^{(l)} a_j^{(l)} \\ &= c + \sum_{i=1}^{m} c_i \left(\sum_{l=1}^{t} a_i^{(l)} \right) + \sum_{1 \leq i < j \leq m} c_{ij} \left(\sum_{l=1}^{t} a_i^{(l)} a_j^{(l)} \right) \,. \end{aligned}$$

Now observe that the t assignments odd-cover $q(x) = 0$ if and only if the number of assignments that does not satisfy $q(x) = 0$ is even, which is equivalent to saying that the left hand side of the above equation is 0, and that by definition means that the t assignments satisfy $q(x) = 0$ in superposition. □

In the description above, we assumed that the quadratic equation $q(x) = 0$ is homogeneous, which allows us to encode it with a matrix $C \in \mathbb{F}_2^{m \times m}$ and express the whole equation as $\langle C, A \rangle = 0$, where $A = x \otimes x$. For quadratic equations that are not homogeneous, we encode them with a $(m + 1) \times (m + 1)$ matrix. In particular, for $q(x) = c + \sum c_i x_i + \sum c_{ij} x_i x_j = 0$, we have matrix C, where $C_{11} = c$, $C_{1i} = c_{i-1}$ for $i = 2, \cdots, m+1$, and $C_{ij} = c_{i-1,j-1}$ for $2 \leq i < j \leq m+1$. As for the variable vector, we add to x an entry that is always 1.

Definition 2.7. *Given a matrix $A \in \mathbb{F}_2^{(m+1)\times(m+1)}$. We say that A is* pseudo-quadratic *if the following holds:*

- *A is symmetric.*
- *$A_{1,1} = 1$.*
- *For all $i = 2, \cdots, m+1$, $A_{1,i} = A_{i,1} = A_{i,i}$.*

Note that for vector $v \in \mathbb{F}_2^{m+1}$ such that $v_1 = 1$, $v \otimes v$ is a pseudo-quadratic rank-1 matrix.

We prove a stronger form of Lemma 2.4 for pseudo-quadratic matrices where we decode a low rank pseudo-quadratic matrix into an odd number of assignments.

Lemma 2.8. *Let $A \in \mathbb{F}_2^{(m+1)\times(m+1)}$ be a pseudo-quadratic matrix of rank k over $\mathbb{F}_2$. Then there exists an odd integer $k_0 < 3k/2+1$, and vectors $v_1, \cdots, v_{k_0} \in \mathbb{F}_2^{m+1}$, such that for all $i \in [k_0]$, $v_{i,1} = 1$, and $A = \sum_{i=1}^{k_0} v_i \otimes v_i$. Moreover, for all $i \in [k_0]$, $D_1(v_i)$ is in the column space of $D_1(A)$.*

Proof. Let $A' = D_1(A)$. Note that A' is symmetric and has rank at most k. Therefore by Lemma 2.4, there exists $l < 3k/2$ vectors $u_1, \cdots, u_l \in \mathbb{F}_2^m$, such that $A' = \sum_{i=1}^{l} u_i \otimes u_i$. Now consider vectors $v_1, \cdots, v_l \in \mathbb{F}_2^{m+1}$, where for each i, $v_{i,1} = 1$ and $v_{i,j} = u_{i,j-1}$ for $j = 2, \cdots, m+1$. Let $A'' = \sum_{i=1}^{l} v_i \otimes v_i$, and $B = A - A''$. For $j, j' \in \{2, \cdots, m+1\}$, we have

$$A''_{j,j'} = \sum_{i=1}^{l} v_{i,j} v_{i,j'} = \sum_{i=1}^{l} u_{i,j-1} u_{i,j'-1} = A'_{j-1,j'-1} = A_{j,j'} .$$

Moreover, we have

$$A''_{1,j} = \sum_{i=1}^{l} v_{i,1} v_{i,j} = \sum_{i=1}^{l} v_{i,j} v_{i,j} = A''_{j,j} = A_{j,j} = A_{1,j} .$$

We conclude that for all $(i,j) \neq (1,1)$, $A_{i,j} = A''_{i,j}$. Note that $A''_{1,1} = (l \bmod 2)$. Therefore if $A''_{1,1} = 1 = A_{1,1}$, then we have l is odd and $A = \sum_{i=1}^{l} v_i \otimes v_i$ as promised. Otherwise l is even. Let $e = (1\ 0\ \cdots\ 0) \in \mathbb{F}_2^{m+1}$. Then $A = \sum_{i=1}^{l} v_i \otimes v_i + e \otimes e$ gives the desired decomposition. □

The following lemma summarizes the discussion at the beginning of this section and relates odd superposition complexity with low-rank pseudo-quadratic matrices.

Lemma 2.9. *Let $q_1(x) = 0, \cdots, q_s(x) = 0$ be a set of s quadratic equations on variable $x_1, \cdots, x_m$, and let $Q_1, \cdots, Q_s \in \mathbb{F}_2^{(m+1)\times(m+1)}$ be their corresponding matrix forms. Suppose there is a pseudo-quadratic matrix $A \in \mathbb{F}_2^{(m+1)\times(m+1)}$ such that $\operatorname{rank}(A) \leq k$ and for all $i \in [s]$, $\langle Q_i, A\rangle = 0$, then there exists $l < 3k/2+1$ vectors $a^{(1)}, \cdots, a^{(l)} \in \mathbb{F}_2^{m+1}$ in the column space of A, for some odd integer l, such that $A = \sum_{i=1}^{l} a^{(i)} \otimes a^{(i)}$. This implies that the assignments $D_1(a^{(1)}), \cdots, D_1(a^{(l)})$ satisfy all equations $q_1(x) = 0, \cdots, q_s(x) = 0$ in superposition.*

Proof. Apply Lemma 2.8 to A, and let $v_1, \cdots, v_l$ be the vectors we get, with $v_{i1} = 1$ for $i \in [l]$, and $A = \sum_{i \in [l]} v_i \otimes v_i$. We now verify that $D_1(v_1), \cdots, D_1(v_l)$ satisfy all equations in superposition.

Consider equation i for $i \in [s]$. We have

$$\begin{aligned} 0 = \langle Q_i, A \rangle &= \sum_{i=1}^{l} \langle Q_i, v_i \otimes v_i \rangle \\ &= \sum_{i=1}^{l} q_i(v_i) \,. \end{aligned}$$

By definition, we have that $v_1, \cdots, v_l$ satisfy q_i in superposition. □

2.2 Label-Cover

The starting point of our reduction is the Label-Cover hardness obtained from E3-Sat instances. We use Label-Cover instances obtained by applying the PCP Theorem [3, 4] and the Parallel Repetition Theorem [31]. The exact formulation below is from [15].

Definition 2.10. *Let ϕ be a* E3-Sat *instance with X as the set of variables and $\mathcal{C}$ the set of clauses. The r-repeated* Label-Cover *instance $\mathcal{L}(r, \phi)$ is specified by:*

- *A bipartite graph $G = (U, V, E)$, where $V := \mathcal{C}^r$ and $U := X^r$.*
- *Label set for U, denote by $L := \{0,1\}^r$, and label set for V, denote by $R := \{0,1\}^{3r}$.*
- *There is an edge $\{u, v\} \in E$ if for each $i \in [r]$, u_i is a variable appearing in clause v_i.*
- *For edge $\{u, v\}$, the constraint $\pi_{uv} : \{0,1\}^{3r} \to \{0,1\}^r$ is the projection of the assignment of the $3r$ clause variables in v to the assignment of the r variables in u.*
- *For each $v \in V$, there is a set of r functions $\{f_i^v : \{0,1\}^{3r} \to \{0,1\}^r\}_{i \in [r]}$, such that $f_i^v(a) = 0$ if and only if the assignment a satisfies the clause v_i. Note that each f_i^v depends only on 3 entries of a.*

A labeling $\sigma : U \to L, V \to R$ satisfies an edge $\{u, v\}$ iff $\pi_{uv}(\sigma(V)) = \sigma(U)$, and $\sigma(V)$ satisfies all clauses in v. The value of $\mathcal{L}(r, \phi)$ is the maximum fraction of edges that can be simultaneously satisfied by any labeling.

We have the following hardness result for Label-Cover.

Theorem 2.11. *Given a* E3-Sat *instance ϕ on n variables and $r \in \mathbb{N}$, there is an algorithm that constructs $\mathcal{L}(r, \phi)$ in time $n^{O(r)}$, and that the output* Label-Cover *instance has the following properties:*

- *If ϕ is satisfiable, then the value of $\mathcal{L}(r, \phi)$ is 1.*
- *If ϕ is unsatisfiable, then the value of $\mathcal{L}(r, \phi)$ is at most $2^{-\varepsilon_0 r}$, for some universal constant $\varepsilon_0 \in (0, 1)$.*

In our construction of Label-Cover instance with matrix labels, we need to use the following Parallel Repetition theorem from Rao [30], which applies to projection games (Label-Cover), with the advantage that the rate at which the soundness decreases is independent of the label size of the original instance.

Theorem 2.12 (Parallel Repetition [30]). *There is a universal constant $\alpha > 0$, such that for a* Label-Cover *instance Ψ, if $\mathrm{Opt}(\Psi) \le 1-\varepsilon$, then $\mathrm{Opt}(\Psi^n) \le (1-\varepsilon/2)^{\alpha\varepsilon n}$.*

2.3 Low-Degree-Long-Code

In this section, we review the basics of Low-Degree-Long-Code. The formulation here is from [10] and [15]. Towards the end of this section, we prove a key lemma that we will use for proving our superposition hardness results.

For a positive integer m, denote by P_m the vector space of m-variable functions $\mathbb{F}_2^m \to \mathbb{F}_2$. For $f, g \in \mathsf{P}_m$, let $\Delta(f,g)$ be the Hamming distance between f and g. For a subset of functions $\mathcal{F} \subseteq \mathsf{P}_m$, the distance between g and $\mathcal{F}$ is defined as $\Delta(g,\mathcal{F}) = \min_{f\in\mathcal{F}} \Delta(f,g)$.

We define the following dot product on P_m.

Definition 2.13 (Dot Product). *For $f, g \in \mathsf{P}_m$, the dot product is defined as $\langle f, g\rangle = \sum_{x\in\mathbb{F}_2^m} f(x)g(x)$.*

Denote by $\mathsf{P}_{m,d}$ be the space of functions with degree at most d. For a subspace $\mathcal{A} \subseteq \mathsf{P}_{m,d}$, denote its dual by $\mathcal{A}^\perp = \{g \in \mathsf{P}_m \mid \forall f \in \mathcal{A}, \langle f, g\rangle = 0\}$. It is well known that $\mathsf{P}_{m,d}^\perp = \mathsf{P}_{m,m-d-1}$.

For $\beta \in \mathsf{P}_m$, denote by $\mathrm{supp}(\beta)$ the support of β, that is $\mathrm{supp}(\beta) = \{x \mid \beta(x) = 1\}$. Define $\mathrm{wt}(\beta) = |\mathrm{supp}(\beta)|$.

Definition 2.14 (Low-Degree-Long-Code). *The* Low-Degree-Long-Code *encoding for an m-bit string $a \in \mathbb{F}_2^m$ is a function $A_a : \mathsf{P}_{m,d} \to \mathbb{F}_2$, defined as $A_a(g) = g(a)$, for all $g \in \mathsf{P}_{m,d}$.*

Definition 2.15 (Character Set). *For $\beta \in \mathsf{P}_m$, define the corresponding* character function *$\chi_\beta : \mathsf{P}_{m,d} \to \mathbb{R}$ as $\chi_\beta(f) = (-1)^{\langle\beta,f\rangle}$.*

Define the character set *$\Lambda_{m,d}$ to be the set of functions $\beta \in \mathsf{P}_m$ which are minimum weight functions in the cosets of $\mathsf{P}_m/\mathsf{P}_{m,d}^\perp$, where ties are broken arbitrarily.*

We have the following result about the character set and the "Fourier decomposition" for functions $\mathsf{P}_{m,d} \to \mathbb{R}$ from [10].

Lemma 2.16.

- *For any $\beta, \beta' \in \mathsf{P}_m$, $\chi_\beta = \chi_{\beta'}$ if and only if $\beta - \beta' \in \mathsf{P}_{m,d}^\perp$.*
- *For $\beta \in \mathsf{P}_{m,d}^\perp$, χ_β is the constant 1 function.*
- *For any β, there exists β', such that $\beta - \beta' \in \mathsf{P}_{m,d}^\perp$, and $|\mathrm{supp}(\beta')| = \Delta(\beta, \mathsf{P}_{m,d}^\perp)$. We call such β' the minimum support function for the coset $\beta + \mathsf{P}_{m,d}^\perp$.*
- *The characters in the character set $\Lambda_{m,d}$ form an orthonormal basis under the inner product $\langle A, B\rangle = \mathbf{E}_{f\in\mathsf{P}_{m,d}}[A(f)B(f)]$.*
- *Any function $A : \mathsf{P}_{m,d} \to \mathbb{R}$ can be uniquely decomposed as*

$$A(g) = \sum_{\beta\in\Lambda_{m,d}} \widehat{A}_\beta \chi_\beta(g)\,.$$

- *Parseval's identity: For any $A : \mathsf{P}_{m,d} \to \mathbb{R}$, $\sum_{\beta \in \Lambda_{m,d}} \widehat{A}_\beta^2 = \mathbf{E}_{f \sim \mathsf{P}_{m,d}}[A(f)^2]$.*

The following lemma relates characters from different domains related by coordinate projections and is from [10].

Lemma 2.17. *Let $n \le m$, and $S \subseteq [m]$ with $|S| = n$, and let $\pi : \mathbb{F}_2^m \to \mathbb{F}_2^n$ be a projection, mapping $x \in \mathbb{F}_2^m$ to $x|_S \in \mathbb{F}_2^n$. Then for $f \in \mathsf{P}_{n,d}$ and $\beta \in \mathsf{P}_m$, we have*

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

where $\pi_2(\beta)(y) = \sum_{x \in \pi^{-1}(y)} \beta(x)$.

Like in the classical Long-Code reductions, we enforce special structures on the tables. This is a technique known as *folding.* The following properties of the Fourier coefficients of folded functions were also studied in [10].

Definition 2.18. *A table $A : \mathsf{P}_{m,d} \to \mathbb{R}$ is folded over constant if for any $f \in \mathsf{P}_{m,d}$, we have $A(f+1) = -A(f)$.*

Lemma 2.19. *If $A : \mathsf{P}_{m,d} \to \mathbb{R}$ is folded over constant, then for any α such that $\widehat{A}_\alpha \neq 0$, we have $\sum_{x \in \mathbb{F}_2^m} \alpha(x) = 1$. In particular, we have $\mathrm{supp}(\alpha) \neq \emptyset$.*

Definition 2.20. *Let $q_1, \cdots, q_k \in \mathsf{P}_{m,3}$, and let*

$$J(q_1, \cdots, q_k) := \left\{ \sum_i r_i q_i \mid r_i \in \mathsf{P}_{m,d-3} \right\}.$$

We say that a function $A : \mathsf{P}_{m,d} \to \mathbb{R}$ is folded over J *if A is constant over cosets of J in $\mathsf{P}_{m,d}$.*

The following lemma shows that a function folded over J does not have weight on small support characters that are non-zero on J.

Lemma 2.21. *Let $\beta \in \mathsf{P}_m$ be such that $\mathrm{wt}(\beta) < 2^{d-3}$, and there exists some $i \in [k]$ and $x \in \mathrm{supp}(\beta)$ with $q_i(x) \neq 0$. Then if $A : \mathsf{P}_{m,d} \to \mathbb{R}$ is folded over J, then $\widehat{A}_\beta = 0$.*

In the actual reduction, $q_1, \cdots, q_k$ will be the set of functions associated with vertices in the Label-Cover instance, as described in Definition 2.10.

In [10], Dinur and Guruswami proved the following theorem about Reed-Muller codes over $\mathbb{F}_2$.

Theorem 2.22. *Let d be a multiple of 4. If $\beta \in \mathsf{P}_m$ is such that $\Delta(\beta, \mathsf{P}_{m,d}) \ge 2^{d/2}$, then*

$$\mathop{\mathbf{E}}_{g \sim \mathsf{P}_{m,d/4}} \left[\left| \mathop{\mathbf{E}}_{h \sim \mathsf{P}_{m,3d/4}} [\chi_\beta(gh)] \right| \right] \le 2^{-4 \cdot 2^{d/4}}\,.$$

Note that $\chi_\beta(gh) = (-1)^{\langle \beta g, \beta h \rangle}$. The key lemma we will now prove is a generalization of the above theorem. The setting is that we have an additional t functions $A_1, \cdots, A_t : \mathsf{P}_{m,d} \to \mathbb{F}_2$. We show that as long as t is small compared to $2^{d/2}$, the expectation $\mathbf{E}_{g,h}[(-1)^{\langle \beta g, \beta h \rangle + \sum_{i=1}^t A_i(g) A_i(h)}]$ is still close to 0 for arbitrary $A_1, \cdots, A_t$.

We use some of the key steps in [10].

Definition 2.23. *For β and $k \leq d$, define*

$$B_{d,k}^{(m)} := \{g \in \mathsf{P}_{m,k} \mid \beta g \in \mathsf{P}_{m,m-d-1+k}\} \ .$$

Note that $B_{d,k}^{(m)}$ is a subspace of $\mathsf{P}_{m,k}$.

For positive integers d, k, define $\Phi_{d,k} : \mathbb{N} \to \mathbb{N}$ as follows: if $d < k$, then $\Phi_{d,k}$ is identically 0, otherwise

$$\Phi_{d,k}(D) = \min_{\substack{m>d \\ \beta \in \mathsf{P}_m : \Delta(\beta, P(m,m-d-1)) \geq D}} \left\{ \dim(P(m,k)) - \dim(B_{d,k}^{(m)}(\beta)) \right\} \ .$$

The following two claims are from [10], which serve as the basis step and induction step for their lower-bound for $\Phi_{d,k}(D)$.

Claim 2.24. *For $d \geq k$ and $D \geq 1$, $\Phi_{d,k}(D) \geq 1$.*

Claim 2.25. *For all $d \geq k$ and $40 < D < 2^d$, $\Phi_{d,k}(D) \geq \Phi_{d-1,k}(D/4) + \phi_{d-1,k-1}(D/4)$.*

For $D = 2^{d-4} = 4^{d/2-2}$ and $k = d/2$, applying the above for a depth of $d/2 - 4$, reducing D from $4^{d/2-2}$ to 16, we have $\Phi_{d,d/2}(2^{d-4}) \geq 2^{d/2-4}$. This gives the following theorem.

Theorem 2.26. *For all integers m, d such that $m > d > 0$ and $4|d$, if $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$ has distance more than 2^{d-4} from $\mathsf{P}_{m,m-d-1}$, then the subspace $B_{d,d/2}^{(m)}(\beta)$ (as a subspace of $\mathsf{P}_{m,d/2}$) has codimension at least $2^{d/2-4}$.*

We remark that Dinur and Guruswami used different degree parameters in [10] for their application. Otherwise, the above theorem is the same as in [10].

We are now ready to prove the main theorem of this section.

Theorem 2.27. *Let $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$ be a polynomial with distance more than 2^{d-4} from $\mathsf{P}_{m,m-d-1}$. Let $t \in \mathbb{N}$ and $A_1, \cdots, A_t : \mathsf{P}_{m,d/2} \to \mathbb{F}_2$ be some arbitrary t functions. Let μ be the uniform distribution on $\mathsf{P}_{m,d/2}$. Then*

$$\begin{aligned} &\mathop{\mathbf{E}}_{g,h\sim\mu}\left[\chi_\beta(gh) \cdot (-1)^{\sum_{i=1}^t A_i(g)A_i(h)}\right] \\ = &\mathop{\mathbf{E}}_{g,h\sim\mu}\left[(-1)^{\langle \beta g, \beta h\rangle + \sum_{i=1}^t A_i(g)A_i(h)}\right] \leq 2^{-(2^{d/2-4}-t)/2} \ . \end{aligned}$$

Proof. Denote by $\mathcal{W}$ the quotient space $\mathsf{P}_{m,d/2}/B_{d,d/2}^{(m)}(\beta)$. By Theorem 2.26, we have $w := \dim(\mathcal{W}) = \operatorname{codim}(B_{d,d/2}^{(m)}(\beta)) \geq 2^{d/2-4}$.

The expectation we are considering can be written as

$$\mathop{\mathbf{E}}_{g_0,h_0\sim\mathcal{W}} \mathop{\mathbf{E}}_{\substack{g:g-g_0\in B_{d,d/2}^{(m)}(\beta) \\ h:h-h_0\in B_{d,d/2}^{(m)}(\beta)}} \left[(-1)^{\langle \beta g, \beta h\rangle + \sum_{i=1}^t A_i(g)A_i(h)}\right] \ . \qquad (1)$$

Consider $f \in \mathsf{P}_{m,d/2}$ and $g \in B_{d,d/2}^{(m)}(\beta)$. We have $\langle \beta f, \beta g\rangle = \langle \beta g, f\rangle = 0$, because $f \in \mathsf{P}_{m,d/2}$ and $\beta g \in \mathsf{P}_{m,m-d/2-1} = \mathsf{P}_{m,d/2}^{\perp}$. This allows us to define "dot product" between elements in $\mathcal{W}$. In particular, for any $f, f', g, g' \in \mathsf{P}_{m,d/2}$ such that $f - f', g - g' \in B^{(m)}_{d,d/2}(\beta)$, we have

$$\begin{aligned}&\langle \beta f', \beta g' \rangle \\ &= \langle \beta f', \beta g' \rangle + \langle \beta(f - f'), \beta g' \rangle + \langle \beta f', \beta(g - g') \rangle + \langle \beta(f - f'), \beta(g - g') \rangle \\ &= \langle \beta f, \beta g \rangle\,.\end{aligned}$$

This means that taking any representative from $\mathcal{W}$ will give the same result for this "dot product".

We can thus further rewrite the expectation as

$$(1) = \mathop{\mathbf{E}}_{g_0,h_0 \sim \mathcal{W}} \left[(-1)^{\langle \beta g_0, \beta h_0 \rangle} \mathop{\mathbf{E}}_{\substack{g: g - g_0 \in B^{(m)}_{d,d/2}(\beta) \\ h: h - h_0 \in B^{(m)}_{d,d/2}(\beta)}} \left[(-1)^{\sum_{i=1}^t A_i(g) A_i(h)} \right] \right]. \tag{2}$$

Consider the matrix $M \in \mathbb{R}^{2^{w+t} \times 2^{w+t}}$, where the rows and columns are indexed by a pair (f_0, a) where $f_0 \in \mathcal{W}$ and $a \in \mathbb{F}_2^t$, and the entries are

$$M_{(f_0,a),(g_0,b)} = (-1)^{\langle \beta f_0, \beta g_0 \rangle + \sum_{i=1}^t a_i b_i}\,.$$

Define vector $u \in \mathbb{R}^{2^{w+t}}$ as

$$u_{f_0,a} = \Pr_{g \sim \mathsf{P}_{m,d/2}} \left[g - f_0 \in B^{(m)}_{d,d/2}(\beta) \wedge \forall i \in [t], A_i((g)) = a_i \right]\,.$$

Since in (2), g and h are sampled independently, we can verify that the expectation in (2) is exactly $u^T M u$. Moreover, since g is chosen uniformly random from $\mathsf{P}_{m,d/2}$, the probability that $g - f_0 \in B^{(m)}_{d,d/2}(\beta)$ is exactly 2^{-w}, thus all entries in u have value at most 2^{-w}, and therefore $\|u\|_2 \le 2^{-w/2}$.

We finish the proof by studying the spectrum of M. Observe that M can be written as the tensor product of a $2^w \times 2^w$ matrix and a $2^t \times 2^t$ matrix as follows. Define $W \in \mathbb{R}^{2^w \times 2^w}$ as

$$W_{f_0,g_0} = (-1)^{\langle \beta f_0, \beta g_0 \rangle}\,,$$

for $f_0, g_0 \in \mathcal{W}$. Define $H \in \mathbb{R}^{2^t \times 2^t}$ as

$$H_{a,b} = (-1)^{\sum_{i=1}^t a_i b_i}\,.$$

We can easily verify that $M = W \otimes H$.

The matrix H satisfies $HH^T = 2^t \cdot I$, where I is the identity matrix, therefore we have that the eigenvalues of H all have absolute value exactly $2^{t/2}$. For the spectrum of W, let $f_0, g_0 \in \mathcal{W}$ be two rows of W. Consider the dot product of row f_0 and g_0 of matrix W

$$W_{f_0}^T W_{g_0} = \sum_{h_0 \in \mathcal{W}} (-1)^{\langle \beta(f_0 + g_0), \beta h_0 \rangle} = \sum_{h_0 \in \mathcal{W}} (-1)^{\langle \beta(f_0 + g_0), h_0 \rangle}\,.$$

The above sum is 2^w if $\beta(f_0 + g_0) \in \mathsf{P}_{m,m-d/2-1}$, or in other words f_0 and g_0 belong to the same coset in $\mathcal{W}$, and otherwise the sum is 0. Hence we have $WW^T = 2^w \cdot I$, and thus the eigenvalues of W all have absolute value $2^{w/2}$. We conclude that the tensor product matrix $M = W \otimes H$ has eigenvalues with absolute value $2^{(w+t)/2}$.

We can now upper-bound the absolute value of the expectation by $|u^T M u| \le 2^{(w+t)/2} \cdot \|u\|_2^2 = 2^{-(w-t)/2}$. □

3 Superposition Hardness for Gap TSA

Let b be some large integer parameter. The Tri-Sum-And (TSA) predicate is a predicate on 5 $\mathbb{F}_2$-variables defined as follows

$$\mathrm{TSA}(x_1, \cdots, x_5) = 1 + x_1 + x_2 + x_3 + x_4 x_5\,.$$

From the definition, we can see that TSA instances are systems of quadratic equations, each involving exactly 5 $\mathbb{F}_2$-variables.

The predicate was studied in [18] as a starting point of an efficient PCP construction. For the predicate itself, Håstad and Khot proved that it is approximation resistant on satisfiable instances.

In this section, we prove a superposition hardness result for TSA.

Theorem 3.1. *There is a reduction that takes as input a* E3-Sat *instance of size n, and outputs a* TSA *instance of size $n^{O(b\log\log n)}$ with the following properties:*

- *If the* E3-Sat *instance is satisfiable, then there is an assignment that satisfies all TSA constraints.*
- *If the* E3-Sat *instance is unsatisfiable, then for any odd integer $t < (\log n)^b$, and any t assignments, at most a $15/16$ fraction of the* TSA *constraints are satisfied in superposition.*

The reduction runs in time $n^{O(b\log\log n)}$.

Proof. The reduction follows a similar approach as a typical inapproximability hardness reduction.

Given a E3-Sat instance, we apply Theorem 2.11 with soundness $1/(1000(\log n)^{2b})$ to get a Label-Cover instance. This gives the parameter $r = (2b\log\log n + O(1))/\varepsilon_0$, where ε_0 is some universal constant. The vertex set of the bipartite graph has size $n^{O(b\log\log n)}$, and the label sets are $L = \{0,1\}^r$ and $R = \{0,1\}^{3r}$. Let $d = \Theta(b\log\log n)$ be such that $2^{d/2-4} \approx (\log n)^b + 3$. This implies also that $2^d \approx 256(\log n)^{2b}$.

For each $u \in U$ and $v \in V$, we expect functions $f_u : \mathsf{P}_{r,d} \to \{-1,1\}$ and $g_v : \mathsf{P}_{3r,d} \to \{-1,1\}$. We assume that all functions are folded over constant. The entries of the functions correspond to variables of some TSA instance. Therefore the number of variables in the output instance is $n^{O(b\log\log n)} \cdot (3r)^{(1+o(1))d} = n^{O(b\log\log n)}$, and the number of constraints is polynomial in the number of variables.

Consider the following test:

1. Sample random edge $e = \{u_1, u_2\} \sim E$. Let π be the projection on the edge, and let f and g be the functions associated with u_1 and u_2.
2. Sample uniformly random query $x \sim \mathsf{P}_{r,d}$, $y \sim \mathsf{P}_{3r,d}$, and $v, w \sim \mathsf{P}_{3r,d/2}$.
3. Construct query $z := x \circ \pi + y + vw \in \mathsf{P}_{3r,d}$.
4. Accept iff $f(x)g(y)g(z)(g(v) \wedge g(w)) = 1$, where $\wedge$ here denotes the binary operator that evaluates to -1 when both operands are -1, and 1 otherwise.

The completeness is straightforward. In this case, the Label-Cover instance has a perfect labeling. Setting the functions to be the Low-Degree-Long-Code encoding of the labels gives an assignment that satisfies all TSA constraints.

In the soundness case, there exists some $t < (\log n)^b$ assignments that satisfy in superposition a $15/16$ fraction of the constraints. That is, for each $u_1 \in U$ and $u_2 \in V$, there are t functions that are folded over constant, $f^{(1)}, \cdots, f^{(t)} : \mathsf{P}_{r,d} \to \{-1, 1\}$ and $g^{(1)}, \cdots, g^{(t)} : \mathsf{P}_{3r,d} \to \{-1, 1\}$ such that over random sample of edges $\{u_1, u_2\}$ and queries x, y, z, v, w, with probability at least $15/16$, the number of $i \in [t]$ such that $f^{(i)}(x)g^{(i)}(y)g^{(i)}(z)(g^{(i)}(v) \wedge g^{(i)}(w)) = 1$ is odd. By an averaging argument, we have that for at least $3/4$ of the edges, over random sample of queries, the above holds with probability at least $3/4$. Call such an edge *good*.

We assume that the functions are folded in the same way. Recall that when applying folding, we partition the domain of the functions into equivalence classes, define the function value in one of the equivalence classes, and then extend to the full domain by adding appropriate constants. For our reduction, we identify one equivalence class for each vertex, and the t functions associated with it supply value only for that equivalence class. This is to make sure $f^{(1)}, \cdots, f^{(t)}$ and $g^{(1)}, \cdots, g^{(t)}$ corresponds exactly to t assignments in superposition.

Fix a good edge for now, and we drop the subscripts u_1 and u_2. Then we have the following

$$\frac{1}{2} + \frac{1}{2} \mathop{\mathbf{E}}_{x,y,z,v,w} \left[\prod_{i=1}^{t} \Big(f^{(i)}(x)g^{(i)}(y)g^{(i)}(z)(g^{(i)}(v) \wedge g^{(i)}(w)) \Big) \right] \geq \frac{3}{4} \,,$$

or

$$\mathop{\mathbf{E}}_{x,y,z,v,w} \left[\prod_{i=1}^{t} \Big(f^{(i)}(x)g^{(i)}(y)g^{(i)}(z)(g^{(i)}(v) \wedge g^{(i)}(w)) \Big) \right] \geq \frac{1}{2} \,.$$

Let $f' = \prod_{i=1}^{t} f^{(i)}$, and $g' = \prod_{i=1}^{t} g^{(i)}$. Since t is odd, we have that f' and g' are both folded over constant. Taking the Fourier expansion of f' and g', we have the following

$$\begin{aligned}
\frac{1}{2} \;&\leq\; \mathop{\mathbf{E}}_{x,y,z,v,w} \left[\prod_{i=1}^{t} \Big(f^{(i)}(x)g^{(i)}(y)g^{(i)}(z)(g^{(i)}(v) \wedge g^{(i)}(w)) \Big) \right] \\
&= \mathbf{E} \left[f'(x)g'(y)g'(z) \prod_{i=1}^{t} (g^{(i)}(v) \wedge g^{(i)}(w)) \right] \\
&= \sum_{\substack{\alpha \in \Lambda_{r,d} \\ \beta_1, \beta_2 \in \Lambda_{3r,d}}} \widehat{f'}_{\alpha} \widehat{g'}_{\beta_1} \widehat{g'}_{\beta_2} \\
&\qquad \mathop{\mathbf{E}}_{x,y,z,v,w} \left[\chi_{\alpha}(x)\chi_{\beta_1}(y)\chi_{\beta_2}(x \circ \pi + y + vw) \prod_{i=1}^{t} (g^{(i)}(v) \wedge g^{(i)}(w)) \right] \\
&= \sum_{\beta \in \Lambda_{3r,d}} \widehat{f'}_{\pi_2(\beta)} \widehat{g'}^{2}_{\beta} \mathop{\mathbf{E}}_{vw} \left[\chi_{\beta}(vw) \prod_{i=1}^{t} (g^{(i)}(v) \wedge g^{(i)}(w)) \right] \,.
\end{aligned}$$

Applying Cauchy-Schwarz and using Parseval, we have

$$\begin{aligned}
\frac{1}{4} &\le \left(\sum_{\beta\in\Lambda_{3r,d}} \widehat{g'}^{2}_{\beta}\right)\left(\sum_{\beta\in\Lambda_{3r,d}} \widehat{f'}^{2}_{\pi_2(\beta)}\widehat{g'}^{2}_{\beta} \mathop{\mathbf{E}}_{vw}\left[\chi_\beta(vw)\prod_{i=1}^{t}(g^{(i)}(v)\wedge g^{(i)}(w))\right]^2\right) \\
&= \sum_{\beta\in\Lambda_{3r,d}:\mathrm{wt}(\beta)\le 2^{d-4}} \widehat{f'}^{2}_{\pi_2(\beta)}\widehat{g'}^{2}_{\beta} \mathop{\mathbf{E}}_{vw}\left[\chi_\beta(vw)\prod_{i=1}^{t}(g^{(i)}(v)\wedge g^{(i)}(w))\right]^2 + \\
&\quad \sum_{\beta\in\Lambda_{3r,d}:\mathrm{wt}(\beta)> 2^{d-4}} \widehat{f'}^{2}_{\pi_2(\beta)}\widehat{g'}^{2}_{\beta} \mathop{\mathbf{E}}_{vw}\left[\chi_\beta(vw)\prod_{i=1}^{t}(g^{(i)}(v)\wedge g^{(i)}(w))\right]^2 .
\end{aligned}$$

For the terms where $\mathrm{wt}(\beta) > 2^{d-4}$, we apply Theorem 2.27 to get

$$\left|\mathop{\mathbf{E}}_{vw}\left[\chi_\beta(vw)\prod_{i=1}^{t}(g^{(i)}(v)\wedge g^{(i)}(w))\right]\right| \le 2^{-(2^{d/2-4}-t)/2} ,$$

and therefore

$$\begin{aligned}
&\sum_{\beta\in\Lambda_{3r,d}:\mathrm{wt}(\beta)> 2^{d-4}} \widehat{f'}^{2}_{\pi_2(\beta)}\widehat{g'}^{2}_{\beta} \\
&\mathop{\mathbf{E}}_{vw}\left[\chi_\beta(vw)\prod_{i=1}^{t}(g^{(i)}(v)\wedge g^{(i)}(w))\right]^2 \le 2^{-(2^{d/2-4}-t)} < \frac{1}{8} .
\end{aligned}$$

This gives us

$$\begin{aligned}
&\sum_{\beta\in\Lambda_{3r,d}:\mathrm{wt}(\beta)\le 2^{d-4}} \widehat{f'}^{2}_{\pi_2(\beta)}\widehat{g'}^{2}_{\beta} \\
\ge &\sum_{\beta\in\Lambda_{3r,d}:\mathrm{wt}(\beta)\le 2^{d-4}} \widehat{f'}^{2}_{\pi_2(\beta)}\widehat{g'}^{2}_{\beta} \mathop{\mathbf{E}}_{vw}\left[\chi_\beta(vw)\prod_{i=1}^{t}(g^{(i)}(v)\wedge g^{(i)}(w))\right]^2 \ge \frac{1}{8} .
\end{aligned}$$

Let $\{u_1, u_2\}$ be a good edge. Consider the following labeling strategy: for u_1, pick α with probability $\widehat{f'}^{2}_{\alpha}$ and pick a random label from $\mathrm{supp}(\alpha)$, and for u_2, pick β with probability $\widehat{g'}^{2}_{\beta}$ and pick a random label from $\mathrm{supp}(\beta)$. The procedure is well defined because f' and g' are all folded, and thus by Lemma 2.19, $\mathrm{supp}(\alpha)$ and $\mathrm{supp}(\beta)$ are nonempty. Also, for β such that $\mathrm{wt}(\beta) \le 2^{d-4} < 2^{d-3}$, by Lemma 2.21, the assignments in $\mathrm{supp}(\beta)$ all satisfy the clauses in u_2. Then the probability that the labeling of u_1 and u_2 satisfies the projection constraint on a good edge $\{u_1, u_2\}$ is at least $\frac{1}{2^{d-4}}\sum_{\beta:wt(\beta)\le 2^{d-4}} \widehat{f'}^{2}_{\pi_2(\beta)}\widehat{g'}^{2}_{\beta} \ge 1/(8\cdot 2^{d-4}) > 1/(100(\log n)^{2b})$. Since there are at least a 3/4 fraction of good edges, overall the labeling satisfies more than $(3/4)\cdot(1/(100(\log n)^{2b})) > 1/(1000(\log n)^{2b})$, contradicting the fact that in the soundness case the Label-Cover instance does not have labeling with value larger than $1/(1000(\log n)^{2b})$. This completes the proof. □

4 Label Cover with Matrix Labels

We now use Theorem 3.1 to construct a Label-Cover instance with properties similar to that in [27].

Let b be some large integer parameter, and $t \approx (\log n)^b$ be an odd integer. Given a TSA instance with t-superposition hardness gap of 15/16 from Theorem 3.1, consider the following 2-Prover-1-Round projection game:

1. The referee picks a TSA constraint, which we denote as $\mathcal{C}(x_1, x_2, x_3, x_4, x_5)$, and then picks randomly $i \in [5]$.
2. The referee sends x_i to Alice and $\mathcal{C}$ to Bob.
3. Alice replies with $a \in \mathbb{F}_2^t$, and Bob replies with $b \in (\mathbb{F}_2^t)^5$.
4. The referee accepts iff b, interpreted as t $\mathbb{F}_2$ assignments, satisfies $\mathcal{C}$ in superposition, and $b_i = a$.

This is a projection game with perfect completeness and soundness 79/80.

Using Theorem 2.12, we get the following Label-Cover construction. Note that it is important that we use Theorem 2.12 by [30] instead of [31], because the answer size is non-constant and it is important that the rate at which soundness decreases is independent of that.

Theorem 4.1. *There exists a reduction that takes a* E3-Sat *instance of size n, and outputs a* Label-Cover *instance $(U, V, E, L, R, \Pi, \Gamma)$ with the following property:*

- *The bipartite graph (U, V, E) has size $\exp((\log n)^{(2+o(1))b})$, and the reduction runs in time $\exp((\log n)^{(2+o(1))b})$.*
- *The label set $R = \mathbb{F}_2^{m_r}$, $L = \mathbb{F}_2^{m_l}$, where $m_l, m_r = (\log n)^{(2+o(1))b}$.*
- *For each $v \in V$, there is a set of quadratic $\mathbb{F}_2$ equations, each involving 5 of the m_r coordinates of the labeling of v. The set of valid labelings $\Gamma(v)$ are those that satisfy all quadratic equations.*
- *For each edge $e \in E$, there is a set $S_e \subseteq [m_r]$, such that $\pi_e : \mathbb{F}_2^{m_r} \to \mathbb{F}_2^{m_l}$ is defined as $\pi_e(r) = r_{S_e}$.*
- *If the* E3-Sat *instance is satisfiable, then there is a labeling that satisfies all quadratic equation constraints for all vertices $v \in V$, and all projection constraints for all edges.*
- *If the* E3-Sat *instance is unsatisfiable, then for any odd integer $l < (\log n)^b$, any labeling $\sigma^{(1)}, \cdots, \sigma^{(l)}$ for the vertices in U and V, the following does not hold simultaneously:*
  - *For each $v \in V$, and for each equation q associated with v, the assignment given by $\sigma^{(1)}(v), \cdots, \sigma^{(l)}(v)$ satisfy q in superposition.*
  - *For at least $2^{-(\log n)^{(2+o(1))b}}$ fraction of the edges $e = \{u, v\}$, we have $\pi_e(\sigma^{(j)}(v)) = \sigma^{(j)}(u)$, $\forall j \in [l]$.*

We now convert the above into a Label-Cover instance with matrix label and rank soundness constraint.

Theorem 4.2. *There exists a reduction that takes a* E3-Sat *instance of size n, and outputs a* Label-Cover *instance $(U, V, E, L, R, \Pi, \Gamma)$ with the following property:*

- *The bipartite graph (U, V, E) has size $\exp((\log n)^{(2+o(1))b})$, and the reduction runs in time $\exp((\log n)^{(2+o(1))b})$.*

- *The label sets are matrices* $R = \mathbb{F}_2^{(m_r+1)\times(m_r+1)}$, $L = \mathbb{F}_2^{(m_l+1)\times(m_l+1)}$, *where* $m_l, m_r = (\log n)^{(2+o(1))b}$.
- *For each* $v \in V$, *there is a set of homogeneous linear* $\mathbb{F}_2$ *equations involving entries of the labeling of* v. *The set of valid labelings* $\Gamma(v)$ *consists of matrices that satisfy all the associated linear equations.*
- *For each edge* $e \in E$, *there is a set* $S_e \subseteq [m_r+1]$, $1 \in S_e$, *such that* $\pi_e : \mathbb{F}_2^{(m_r+1)\times(m_r+1)} \to \mathbb{F}_2^{(m_l+1)\times(m_l+1)}$ *is defined as* $\pi_e(r) = r_{S_e}$.
- *If the* E3-Sat *instance is satisfiable, then for each* $u \in U$, *there is a labeling* $M_u = x_u \otimes x_u$ *where* $x_{u,1} = 1$, *and for each* $v \in V$, *there is a labeling* $M_v = x_v \otimes x_v$ *where* $x_{v,1} = 1$, *such that for all* $v \in V$, $M_v \in \Gamma(v)$, *and for all* $e \in E$, $\pi_e(M_v) = M_v|_{S_e} = M_u$.
- *If the* E3-Sat *instance is unsatisfiable, then for any labeling* σ *for the vertices in* U *and* V, *the following does not hold simultaneously:*
  - *For each* $v \in V$, *the matrix* $\sigma(v)$ *is pseudo-quadratic, has* $\operatorname{rank}(\sigma(v)) \leq (\log n)^b/2$, *and is valid* $\sigma(v) \in \Gamma(v)$.
  - *For at least* $2^{-(\log n)^b}$ *fraction of the edges* $e = \{u, v\}$, *we have* $\pi_e(\sigma(v)) = \sigma(u)$.

Proof. We start with the Label-Cover instance from the previous theorem.

The underlying bipartite graph of the new instance is exactly the same. The parameters m_r and m_l are the same as before. The labels for $u \in U$ in the new instance are now matrices from $\mathbb{F}_2^{(m_l+1)\times(m_l+1)}$, and the labels for $v \in V$ are from $\mathbb{F}_2^{(m_r+1)\times(m_r+1)}$. The constraints for labelings for vertices in $v \in V$ are the following:

1. The matrix label M is symmetric, and for $i = 2, \cdots, m_r+1$, we have $M_{i,i} = M_{1,i} = M_{i,1}$. These are all homogeneous linear constraints. Note that if in addition we have $M_{1,1} = 1$, then we get that M is pseudo-quadratic. Here, however, we do not include the latter constraint as it is not homogeneous. In fact, this will be handled by the inner verifier.
2. For each quadratic constraint in the previous instance, we include the linearized version of it in the new instance. That is, term $x_i x_j$ is replaced by entry $(i+1, j+1)$ of the matrix, term x_i is replaced by entry $(1, i+1)$, and constant 1 is replaced by entry $(1,1)$.

For edge e, let S_e be the set associated with its projection in the old instance, then in the new instance is defined by the set $S'_e = \{1\} \cup \{i+1 \mid i \in S_e\}$.

The completeness case is straightforward. For the soundness case, suppose that there are pseudo-quadratic matrices M_u and M_v for each $u \in U$ and $v \in V$, such that M_v satisfies homogeneous linear constraints associated with v, $\operatorname{rank}(M_v) \leq k$, and that for $2^{-(\log n)^b}$ fraction of the edges e, $(M_v)|_{S_e} = M_u$.

For such an edge $e = \{u, v\}$, by Lemma 2.9, there exists odd integer $l < 3/2 \cdot (\log n)^b/2 < (\log n)^b$ vectors $v_1, \cdots, v_l \in \mathbb{F}_q^{m_r+1}$, where $v_{i,1} = 1$ for $i \in [l]$, such that $M_v = \sum_{i=1}^{l} v_i \otimes v_i$, and the assignments $D_1(v_1), \cdots, D_1(v_l)$ satisfy in superposition the quadratic constraints of the old Label-Cover instance. For vertex u, we have that $\operatorname{rank}(M_u) = \operatorname{rank}((M_v)|_S) \leq \operatorname{rank}(M_v)$. Also, $M_u = \sum_{i=1}^{l} v_i|_S \otimes v_i|_S$, and that $D_1(v_i)|_{S-\{1\}}$ are in the column space of $D_1(M_u)$. Therefore, for any $i \in [l]$, if we take a uniformly random vector in the column space of $D_1(M_u)$, then with probability at least $2^{-(\log n)^b/2}$, it will be equal to $(v_i')|_S$. Repeat this for all $i \in [l]$, and we have that these labelings of u all satisfy the projection constraint with probability at least $2^{-(\log n)^{2b}}$.

Overall, this labeling satisfies $2^{-(\log n)^b} 2^{-(\log n)^{2b}} = 2^{-(\log n)^{(2+o(1))b}}$ fraction of the edges in the old instance. □

5 Hypergraph Coloring Hardness

We now compose the Label-Cover from Theorem 4.2 with Quadratic Code inner-verifier to get inapproximability result for hypergraph coloring.

Theorem 5.1. *There is a reduction that takes as input a* E3-Sat *instance of size n, outputs a* 8*-uniform hypergraph H with the following properties:*

- *The size H and the running time of the reduction are both upper-bounded by* $\exp((\log n)^{(4+o(1))b})$.
- *If the* E3-Sat *instance is satisfiable, then H is* 2*-colorable.*
- *If the* E3-Sat *instance is unsatisfiable, then H does not have independent set of fractional size larger than* $2^{-O((\log n)^b)}$.

*In other words, it is quasi-*NP*-hard to color a* 2*-colorable* 8*-uniform hypergraph of size N with less than* $2^{(\log N)^{1/4-o(1)}}$ *colors.*

The following proof is based on a note by Girish Varma [33].

Given the Label-Cover instance from Theorem 4.2, we expect for each vertex $v \in V$ a function $f_v : \mathbb{F}_2^{(m_r+1)\times(m_r+1)} \to \mathbb{F}_2$. The expected encoding for matrix label $\sigma(v) = a_v \otimes a_v$ is $f_v(A) = \langle a_v \otimes a_v, A\rangle = a_v^T A a_v$. Let $\mathcal{H}_v \subseteq \mathbb{F}_2^{(m_r+1)\times(m_r+1)}$ be the dual of the subspace of the set of pseudo-quadratic matrices that satisfies the linear constraints associated with v. The function f_v is folded over $\mathbb{F}_2^{(m_r+1)\times(m_r+1)}/\mathcal{H}_v$.

Consider the following Boolean 8-uniform test:

- Choose $u \in U$ uniformly at random, and $v, w \in V$ uniformly and independently at random from the neighbors of u. Let $\pi, \sigma : \mathbb{F}_2^{(m_r+1)\times(m_r+1)} \to \mathbb{F}_2^{(m_l+1)\times(m_l+1)}$ be the projections corresponding to the edges (u, v) and (u, w) respectively, and let S_π and S_σ be the index set associated with them.
- Uniformly and independently sample $X_1, X_2, Y_1, Y_2 \in \mathbb{F}_2^{(m_r+1)\times(m_r+1)}$, $F \in \mathbb{F}_2^{(m_l+1)\times(m_l+1)}$, and $x, y, z, x', y', z' \in \mathbb{F}_2^{m_r+1}$. Let $e \in \mathbb{F}_2^{m_r+1}$ be the vector with only the 1-st entry 1 and the rest 0.
- Accept if and only if the following 8 values are not all equal:

$$\begin{array}{lll} f_v(X_1) & f_v(X_3) & \text{where } X_3 := X_1 + x \otimes y + F \circ \pi \\ f_v(X_2) & f_v(X_4) & \text{where } X_4 := X_2 + (x+e) \otimes z + F \circ \pi \\ f_w(Y_1) & f_w(Y_3) & \text{where } Y_3 := Y_1 + x' \otimes y' + F \circ \sigma + e \otimes e \\ f_w(Y_2) & f_w(Y_4) & \text{where } Y_4 := Y_2 + (x'+e) \otimes z' + F \circ \sigma + e \otimes e \end{array}$$

We denote by $\mathcal{T}$ the test distribution.

The vertex set of the hypergraph has size

$$\exp((\log n)^{(2+o(1))b}) \cdot 2^{(\log n)^{2(2+o(1))b}} = \exp((\log n)^{(4+o(1))b}) =: N\,.$$

5.1 Completeness

Let $y_v \otimes y_v$ for $v \in V$ and $x_u \otimes x_u$ for $u \in U$ be a perfect labeling for the Label Cover instance, with $y_{v,1} = x_{u,1} = 1$ and for each edge $e = \{u, v\} \in E$, we have $(y_v)|_{S_e} = x_u$. Consider the 2-coloring where for each $v \in V$, $f_v(X) = y_v^T X y_v = \langle X, y_v \otimes y_v\rangle$. Such a function is constant over cosets of $\mathcal{H}_v$. Let $x_1 := \langle X_1, y_v \otimes y_v\rangle$, $x_2 := \langle X_2, y_v \otimes y_v\rangle$, $y_1 := \langle Y_1, y_w \otimes y_w\rangle$, $y_2 := \langle Y_2, y_w \otimes y_w\rangle$, and $f := \langle F, x_u \otimes x_u\rangle$. Note that $\langle F, x_u \otimes x_u\rangle = \langle F, \pi_{u,v}(y_v \otimes y_v)\rangle = \langle F \circ \pi_{uv}, y_v \otimes y_v\rangle$. Also, $\langle e \otimes e, y_v \otimes y_v\rangle = \langle e, y_v\rangle = 1$. Therefore, the value of the 8 queries are

$$\begin{array}{ll}
x_1 & x_1 + \langle y_v, x\rangle\langle y_v, y\rangle + f \\
x_2 & x_2 + (\langle y_v, x\rangle + 1)\langle y_v, z\rangle + f \\
y_1 & y_1 + \langle y_w, x'\rangle\langle y_w, y'\rangle + f + 1 \\
y_2 & y_2 + (\langle y_w, x'\rangle + 1)\langle y_w, z'\rangle + f + 1
\end{array}$$

We finish the proof of the completeness case by a case analysis.

If $\langle y_v, y\rangle = \langle y_w, y'\rangle = 0$, then the sum of entries in the first and third row is 1, which means that there are different values. Similarly, we conclude that if $\langle y_v, z\rangle = \langle y_w, z'\rangle = 0$, then using similar argument as above, there are different values in the second and the fourth row. The same applies to the case when $\langle y_v, x\rangle = \langle y_2, x'\rangle = 1$, and the case when $\langle y_v, x\rangle = \langle y_w, x'\rangle = 0$.

Suppose now that $\langle y_v, x\rangle = 1$ and all entries are equal. Then from the second row, we have that $f = 0$, and from the first row, we get $\langle y_v, y\rangle = 0$. By the discussion above, we have that $\langle y_w, y'\rangle = 1$, and the third row gives us $\langle y_w, x'\rangle = 1$, but then the two entries on the last row are different.

Suppose otherwise that $\langle y_v, x\rangle = 0$ and all entries are equal. Then from the first row, we have $f = 0$, and the second row implies $\langle y_v, z\rangle = 0$. By the discussion above, we must have $\langle y_w, z'\rangle = 1$, and the last row gives $\langle y_w, x'\rangle = 0$, leaving two different entries in the third row.

Hence f_v gives a valid 2-coloring of $\mathcal{G}$.

5.2 Soundness

Let $\delta = 2^{-(\log n)^b}$ be the soundness parameter from Theorem 4.2 and $k = (\log n)^b/2$ be the rank upper-bound from Theorem 4.2.

Lemma 5.2. *If there is an independent set in $\mathcal{G}$ of relative size s, then*

$$s^8 \le \delta + \frac{1}{2^{k/2+1}}\,.$$

Proof. Consider any set $A \subseteq \mathcal{V}(\mathcal{G})$ of fractional size s. For every $v \in V$, let $f_v : \mathbb{F}_2^{(m_r+1)\times(m_r+1)} \to [0, 1]$ be the indicator function of A, extended such that it is constant over cosets of $\mathcal{H}_v$. The fractional size of A is given by

$$\mathop{\mathbf{E}}_{\substack{v\sim V\\ X\sim\mathbb{F}_2^{(m_r+1)\times(m_r+1)}}}[f_v(X)] = \mathop{\mathbf{E}}_{v\sim V}\left[\widehat{f}_{v,0}\right].$$

The set A is an independent set if and only if

$$\Theta := \mathop{\mathbf{E}}_{u,v,w}\mathop{\mathbf{E}}_{X_i,Y_i\sim\mathcal{T}}\prod_{i=1}^{4} f_v(X_i)f_w(Y_i) = 0\,. \tag{3}$$

Taking Fourier expansion and considering expectations over X_1, X_2, Y_1, Y_2, we get the following:

$$
\begin{aligned}
\Theta \; = \; & \mathop{\mathbf{E}}_{u,v,w} \sum_{\alpha_1,\alpha_2,\beta_1,\beta_2 \in \mathbb{F}_2^{(m_r+1)\times(m_r+1)}} \mathop{\mathbf{E}}_{F,x,x'} \Bigg[\\
& \widehat{f}^2_{v,\alpha_1} \mathop{\mathbf{E}}_{y}[\chi_{\alpha_1}(x \otimes y)]\chi_{\alpha_1}(F \circ \pi) \\
& \widehat{f}^2_{v,\alpha_2} \mathop{\mathbf{E}}_{z}[\chi_{\alpha_2}((x+e) \otimes z)]\chi_{\alpha_2}(F \circ \pi) \\
& \widehat{f}^2_{w,\beta_1} \mathop{\mathbf{E}}_{y'}[\chi_{\beta_1}(x' \otimes y')]\chi_{\beta_1}(F \circ \sigma)\chi_{\beta_1}(e \otimes e) \\
& \widehat{f}^2_{w,\beta_2} \mathop{\mathbf{E}}_{z'}[\chi_{\beta_2}((x'+e) \otimes z')]\chi_{\beta_2}(F \circ \sigma)\chi_{\beta_2}(e \otimes e) \Bigg] .
\end{aligned}
$$

Denote the term inside $\mathbf{E}_{F,x,x'}[\cdot]$ as $Term_{u,v,w}(\alpha_1, \alpha_2, \beta_1, \beta_2)$.

For the characters involving F, we have

$$
\begin{aligned}
& \mathop{\mathbf{E}}_{F}\left[\chi_{\alpha_1}(F \circ \pi)\chi_{\alpha_2}(F \circ \pi)\chi_{\beta_1}(F \circ \sigma)\chi_{\beta_2}(F \circ \sigma)\right] \\
= \; & \mathop{\mathbf{E}}_{F}\left[(-1)^{\langle \pi(\alpha_1+\alpha_2), F\rangle + \langle \sigma(\beta_1+\beta_2), F\rangle}\right] ,
\end{aligned}
$$

and since $F \in \mathbb{F}_2^{(m_l+1)\times(m_l+1)}$ is chosen uniformly at random, the above is 0 unless $\pi(\alpha_1+\alpha_2) = \sigma(\beta_1+\beta_2)$.

Let $\nu(\alpha) := \langle \alpha, e \otimes e\rangle$. Taking expectations over x, y, z, x', y', z', we have that when $\pi(\alpha_1+\alpha_2) \neq \sigma(\beta_1+\beta_2)$, $Term_{u,v,w}(\alpha_1, \alpha_2, \beta_1, \beta_2) = 0$, and otherwise

$$
\begin{aligned}
& Term_{u,v,w}(\alpha_1, \alpha_2, \beta_1, \beta_2) \\
= \; & (-1)^{\nu(\beta_1+\beta_2)} \widehat{f}^2_{v,\alpha_1} \widehat{f}^2_{v,\alpha_2} \widehat{f}^2_{w,\beta_1} \widehat{f}^2_{w,\beta_2} \\
& \Pr_{x}\left[\alpha_1 x = 0 \wedge \alpha_2 x = \alpha_2 e\right] \Pr_{x'}\left[\beta_1 x = 0 \wedge \beta_2 x' = \beta_2 e\right] .
\end{aligned}
$$

The terms that are potentially non-zero can now be partitioned into three parts:

$$
\begin{aligned}
\Theta_0 \; &= \; \mathop{\mathbf{E}}_{u,v,w} \sum_{\substack{\mathrm{rank}(\alpha_1+\alpha_2), \mathrm{rank}(\beta_1+\beta_2) \leq k \\ \pi(\alpha_1+\alpha_2) = \sigma(\beta_1+\beta_2) \\ \nu(\beta_1+\beta_2)=0}} Term_{u,v,w}(\alpha_1, \alpha_2, \beta_1, \beta_2) \\
\Theta_1 \; &= \; \mathop{\mathbf{E}}_{u,v,w} \sum_{\substack{\mathrm{rank}(\alpha_1+\alpha_2), \mathrm{rank}(\beta_1+\beta_2) \leq k \\ \pi(\alpha_1+\alpha_2) = \sigma(\beta_1+\beta_2) \\ \nu(\beta_1+\beta_2)=1}} Term_{u,v,w}(\alpha_1, \alpha_2, \beta_1, \beta_2) \\
\Theta_0 \; &= \; \mathop{\mathbf{E}}_{u,v,w} \sum_{\substack{\max\{\mathrm{rank}(\alpha_1+\alpha_2), \mathrm{rank}(\beta_1+\beta_2)\} > k \\ \pi(\alpha_1+\alpha_2) = \sigma(\beta_1+\beta_2)}} Term_{u,v,w}(\alpha_1, \alpha_2, \beta_1, \beta_2) .
\end{aligned}
$$

We first lower-bound Θ_0. Note that all terms in Θ_0 are positive. Consider the term corresponding to $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$. We have

$$
\mathop{\mathbf{E}}_{u,v,w} \widehat{f}^4_{v,0} \widehat{f}^4_{w,0} = \mathop{\mathbf{E}}_{u}\left(\mathop{\mathbf{E}}_{v} \widehat{f}^4_{v,0}\right)^2 \geq \left(\mathop{\mathbf{E}}_{u,v} \widehat{f}_{v,0}\right)^8 \geq s^8 .
$$

Therefore $\Theta_0 \ge s^8$.

For Θ_1, we have the following upper-bound

$$|\Theta_1| \le \mathop{\mathbf{E}}_{u,v,w} \sum_{\substack{\mathrm{rank}(\alpha_1+\alpha_2),\mathrm{rank}(\beta_1+\beta_2)\le k\\ \pi(\alpha_1+\alpha_2)=\sigma(\beta_1+\beta_2)\\ \nu(\beta_1+\beta_2)=1}} \widehat{f}^2_{v,\alpha_1}\widehat{f}^2_{v,\alpha_2}\widehat{f}^2_{w,\beta_1}\widehat{f}^2_{w,\beta_2}\,. \tag{4}$$

Consider the following randomized labeling strategy for vertices in $u \in U$ and $v \in V$: for $v \in V$, pick (β_1, β_2) with probability $\widehat{f}^2_{v,\beta_1}\widehat{f}^2_{v,\beta_2}$ and set its label to $\beta_1 + \beta_2$; for $u \in U$, pick a random neighbor v, and choose (α_1, α_2) with probability $\widehat{f}^2_{v,\alpha_1}\widehat{f}^2_{v,\alpha_2}$ and set its label to $\pi(\alpha_1 + \alpha_2)$. Due to folding, we have that β_1 and β_2 both satisfies the homogeneous linear constraints associated with v, and so does $\beta_1 + \beta_2$. Therefore the right hand side of (4) gives the probability that a random edge of the Label Cover is satisfied by this labeling. Thus $|\Theta_1| \le \delta$.

For Θ_2, note that if $\mathrm{rank}(\alpha) > k$, then for any fixed b, $\Pr_x[\alpha x = b] \le 1/2^{k+1}$. Therefore, for any fixed choice of u, v, w, all terms in Θ_2 have absolute value at most $1/2^{k/2+1}$. Combined with Parseval's identity, we conclude that $|\Theta_2| \le 1/2^{k/2+1}$. $\square$

We conclude that any independent set in $\mathcal{G}$ has fractional size at most $2^{-\log^b n/32}$, and therefore the chromatic number of $\mathcal{G}$ is at least $2^{\log^b n/32} = \exp((\log N)^{1/(4-o(1))})$.

Acknowledgments

The author would like to thank Johan Håstad for numerous inspiring discussions.

References

[1] Noga Alon, Pierre Kelsen, Sanjeev Mahajan, and Ramesh Hariharan. Approximate hypergraph coloring. *Nord. J. Comput.*, 3(4):425–439, 1996.

[2] Sanjeev Arora and Eden Chlamtac. New approximation guarantee for chromatic number. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, STOC '06, pages 215–224, New York, NY, USA, 2006. ACM.

[3] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998.

[4] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998.

[5] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, pages 370–379, 2012.

[6] Bonnie Berger and John Rompel. A better performance guarantee for approximate graph coloring. *Algorithmica*, 5(3):459–466, 1990.

[7] Avrim Blum and David R. Karger. An $\tilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs. *Inf. Process. Lett.*, 61(1):49–53, 1997.

[8] Hui Chen and Alan M. Frieze. Coloring bipartite hypergraphs. In *Integer Programming and Combinatorial Optimization, 5th International IPCO Conference, Vancouver, British Columbia, Canada, June 3-5, 1996, Proceedings*, pages 345–358, 1996.

[9] Eden Chlamtac. Approximation algorithms using hierarchies of semidefinite programming relaxations. In *FOCS*, pages 691–701, 2007.

[10] Irit Dinur and Venkatesan Guruswami. Pcps via low-degree long code and hardness for constrained hypergraph coloring. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 340–349, 2013.

[11] Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. *SIAM J. Comput.*, 39(3):843–873, 2009.

[12] Irit Dinur, Oded Regev, and Clifford D. Smyth. The hardness of 3-uniform hypergraph coloring. *Combinatorica*, 25(5):519–535, 2005.

[13] Irit Dinur and Igor Shinkar. On the conditional hardness of coloring a 4-colorable graph with super-constant number of colors. In *APPROX-RANDOM*, pages 138–151, 2010.

[14] M. R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[15] Venkatesan Guruswami, Prahladh Harsha, Johan Håstad, Srikanth Srinivasan, and Girish Varma. Super-polylogarithmic hypergraph coloring hardness via low-degree long codes. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 614–623, 2014.

[16] Venkatesan Guruswami, Johan Håstad, and Madhu Sudan. Hardness of approximate hypergraph coloring. *SIAM J. Comput.*, 31(6):1663–1686, 2002.

[17] Venkatesan Guruswami and Sanjeev Khanna. On the hardness of 4-coloring a 3-colorable graph. *SIAM J. Discrete Math.*, 18(1):30–40, 2004.

[18] Johan Håstad and Subhash Khot. Query efficient PCPs with perfect completeness. *Theory of Computing*, 1(7):119–148, 2005.

[19] Jonas Holmerin. Vertex cover on 4-regular hyper-graphs is hard to approximate within 2-epsilon. In *Proceedings on 34th Annual ACM Symposium on Theory of Computing, May 19-21, 2002, Montréal, Québec, Canada*, pages 544–552, 2002.

[20] Sangxia Huang. Improved hardness of approximating chromatic number. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 233–243. Springer, 2013.

[21] David R. Karger, Rajeev Motwani, and Madhu Sudan. Approximate graph coloring by semidefinite programming. *J. ACM*, 45(2):246–265, 1998.

[22] Ken-ichi Kawarabayashi and Mikkel Thorup. Combinatorial coloring of 3-colorable graphs. In *FOCS*, pages 68–75, 2012.

[23] Ken-ichi Kawarabayashi and Mikkel Thorup. Coloring 3-colorable graphs with o(nˆ{1/5}) colors. In *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), STACS 2014, March 5-8, 2014, Lyon, France*, pages 458–469, 2014.

[24] Sanjeev Khanna, Nathan Linial, and Shmuel Safra. On the hardness of approximating the chromatic number. *Combinatorica*, 20(3):393–415, 2000.

[25] Subhash Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring. In *FOCS*, pages 600–609, 2001.

[26] Subhash Khot. Hardness results for coloring 3-colorable 3-uniform hypergraphs. In *Proc. 43rd FOCS*, pages 23–32. IEEE Comp. Soc. Press, 2002.

[27] Subhash Khot and Rishi Saket. Hardness of coloring 2-colorable 12-uniform hypergraphs with $\exp(log^{\Omega(1)}n)$ colors. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, pages 206–215, 2014.

[28] Subhash Khot and Rishi Saket. Hardness of finding independent sets in 2-colorable and almost 2-colorable hypergraphs. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 1607–1625, 2014.

[29] Michael Krivelevich, Ram Nathaniel, and Benny Sudakov. Approximating coloring and maximum independent sets in 3-uniform hypergraphs. *Journal of Algorithms*, 41(1):99 – 113, 2001.

[30] Anup Rao. Parallel repetition in projection games and a concentration bound. *SIAM J. Comput.*, 40(6):1871–1891, 2011.

[31] Ran Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998.

[32] Rishi Saket. Hardness of finding independent sets in 2-colorable hypergraphs and of satisfiable csps. In *IEEE 29th Conference on Computational Complexity, CCC 2014, Vancouver, BC, Canada, June 11-13, 2014*, pages 78–89, 2014.

[33] Girish Varma. A note on reducing uniformity in khot-saket hypergraph coloring hardness reductions. *CoRR*, abs/1408.0262, 2014.

[34] Avi Wigderson. A new approximate graph coloring algorithm. In *STOC*, pages 325–329, 1982.