Optimal Inapproximability of Generalized Linear Equations over a Finite Group

Amey Bhangale

Department of Computer Science and Engineering, University of California, Riverside, CA, USA

Yezhou Zhang[1]

Department of Computer Science and Engineering, University of California, Riverside, CA, USA

Abstract

Constraint satisfaction problems (CSPs) consist of a set of variables taking values from some finite domain and a set of local constraints on these variables. The objective is to find an assignment to the variables that maximizes the fraction of satisfied constraints.

In this work, we study the CSP where the constraints are *generalized linear equations* over a finite group G. More specifically, for a given $S \subseteq G$, the constraints in this CSP are of the form addition of the values to the variables (similarly, product for non-abelian groups) belongs to the set S. We give an approximation algorithm for this problem on satisfiable instances and show that it is optimal for certain S assuming $\mathbf{P} \neq \mathbf{NP}$.

This natural predicate is one of the very few known predicates that are approximation resistant on almost satisfiable instances, assuming $\mathbf{P} \neq \mathbf{NP}$, but admits a non-trivial approximation algorithm on satisfiable instances.

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness

Keywords and phrases Constraint satisfaction problems, inapproximability, approximation algorithms, non-abelian groups, Fourier analysis

Digital Object Identifier 10.4230/LIPIcs.ICALP.2026.30

Category Track A: Algorithms, Complexity and Games

Related Version *Full Version*: `https://arxiv.org/abs/2605.10010`

Funding *Amey Bhangale*: Supported by the Hellman Fellowship award and NSF CAREER award 2440882.

Yezhou Zhang: Supported by the Hellman Fellowship award and NSF CAREER award 2440882.

1 Introduction

Constraint satisfaction problems (CSPs) are one of the most fundamental problems in theoretical computer science. A Max-P-CSP instance φ for a given predicate $P : \Sigma^k \to \{0, 1\}$ consists of a set of variables $x_1, x_2, \ldots, x_n$ taking values from the domain Σ and a collection of constraints $C_1, C_2, \ldots, C_m$ where each C_i consists of a constraint of the form $P(x_{i_1}, x_{i_2}, \ldots, x_{i_k})$. The constraints might involve *literals* instead of just the variables. The objective is to assign values to the variables that maximize the fraction of satisfied constraints from φ. An α-approximation for Max-P-CSP is an algorithm that, given an instance φ of Max-P-CSP, outputs an assignment that satisfies at least $\alpha \cdot \mathrm{OPT}$, where OPT is the optimal value of the instance φ.

A systematic study of the complexity of CSPs was started by Schaefer in 1978 [36], who showed that for every predicate P over a 2-element set, the problem P-CSP is either solvable in polynomial time or is **NP**-complete. A famous dichotomy conjecture of Feder and Vardi [17, 18], which was resolved recently in huge breakthroughs by Bulatov [11] and

[1] Corresponding author

Zhuk [38] independently, states that for every predicate P, checking the satisfiability of a P-CSP is either in $\mathbf{P}$ or is $\mathbf{NP}$-complete. However, when it comes to approximation algorithms for Max-P-CSP, the question is wide open.

The optimal approximability results for Max-P-CSP for various predicates P are known starting with the seminal work of [24]. Håstad showed that Max-3SAT, where the predicate is over a Boolean alphabet and is an OR of three literals, is $\mathbf{NP}$-hard to approximate within a factor of $\frac{7}{8}+\varepsilon$ for every $\varepsilon > 0$. The above hardness result also holds even if the given instance is guaranteed to have an assignment with value 1. It is also easy to get a $\frac{7}{8}$-approximation algorithm for Max-3SAT – for each variable, pick a random value from $\{0,1\}$. This algorithm satisfies $\frac{7}{8}$-fraction of the constraints in expectation. It is also easy to derandomize this algorithm by the method of conditional expectation. Thus, the hardness result is optimal. In the same work on the optimal hardness of Max-3SAT, Håstad [24] also studied the inapproximability of Max-3LIN over an abelian group G. In Max-3LIN over a group G, the variables take values from the group $(G,+)$, and constraints are of the type $x_{i_1} + x_{i_2} + x_{i_3} = c$ for some $c \in G$. If the given instance of Max-3LIN is fully satisfiable, then it is known [21] that the satisfying assignment can be found in polynomial time using techniques similar to Gaussian elimination. Furthermore, a random assignment to the variables satisfies $\frac{1}{|G|}$ fraction of the constraints in expectation. Håstad [24] showed that this algorithm is optimal for general instances.

The Max-3LIN over a non-abelian group $(G, \bullet)$ is another interesting problem. Unlike the abelian group case, here, Goldmann and Russell showed [21] that it is $\mathbf{NP}$-complete to check the satisfiability of a given instance for every non-abelian group G. Engebretsen, Holmerin, and Russell gave [16] similar inapproximability results as in the case of abelian groups. More specifically, they showed that for every $\varepsilon > 0$, if an instance of Max-3LIN over a non-abelian group $(G, \bullet)$ is given with value at least $1-\varepsilon$, then it is $\mathbf{NP}$-hard to find an assignment that satisfies at least $\frac{1}{|G|}+\varepsilon$ fraction of the constraints. There is, however, a better than $\frac{1}{|G|}$ approximation algorithm for certain groups on satisfiable instances. It is folklore to get a $\frac{1}{|[G,G]|}$-approximation algorithm for Max-3LIN over a non-abelian group $(G, \bullet)$ on satisfiable instances, where $[G,G]$ is the commutator subgroup of G. A commutator of two group elements g and h is a group element $g^{-1} \bullet h^{-1} \bullet g \bullet h$, and a commutator subgroup $[G,G]$ is the subgroup generated by all the commutators of the group G. Recently, Bhangale and Khot [4] showed that this algorithm is optimal on satisfiable instances assuming $\mathbf{P} \neq \mathbf{NP}$.

Given these approximation algorithms and inapproximability results concerning Max-3LIN, it is natural to ask what role the abelian nature of a group G plays in such results. In this work, we study a generalized version of the Max-3LIN problem, denoted by Max-Ek-LIN$_S(G)$, as defined below.

Fix a group $(G, \bullet)$ and a subset $S \subseteq G$. In the instance of Max-Ek-LIN$_S(G)$, the variables take values from G. A *literal* of a variable x is given by $(g \bullet x)$ for some $g \in G$.

▶ **Definition 1** (Max-Ek-LIN$_S(G)$). *Consider a multiset of linear equations $(C_1, C_2, \ldots, C_m)$ over variables $\{x_1, x_2, \ldots, x_n | x_i \in G\}$ for a group G. Each constraint consists of a tuple of exactly k literals and will be considered satisfied when their sum/product over G is in $S \subseteq G$. The Max-Ek-LIN$_S(G)$ problem is to find an assignment to $\{x_1, x_2, \ldots, x_n\}$ that maximizes the number of satisfied constraints.*

Observe that the problem Max-3LIN over a group G that was discussed earlier is Max-Ek-LIN$_{\{1_G\}}(G)$ where 1_G is the identity element of the group G. In Max-Ek-LIN$_S(G)$, we allow more satisfying assignments in the predicate, i.e, ignoring the literals for simplicity, the condition that a tuple (a,b,c) satisfies a given constraint depends on the value of $a \bullet b \bullet c$, and hence we call the predicate a *generalized linearity predicate* over G.[2]

[2] This nomenclature was also used by Chattopadhyay and Wigderson [13] to describe similar predicates.

Approximation Algorithm for Max-E3-LIN$_S(G)$

To keep things simple, we restrict to the setting when G is an abelian group for this discussion. We start with a simple approximation algorithm for Max-E3-LIN$_S(G)$. Let Φ be an instance of Max-E3-LIN$_S(G)$ with constraints set $(C_1, C_2, \ldots, C_m)$ over the variables $X = \{x_1, x_2, \ldots, x_n\}$. Here, the constraint C_i is of the form

$$(a_{i_1} + x_{i_1}) + (a_{i_2} + x_{i_2}) + (a_{i_3} + x_{i_3}) \in S.$$

One can try to replace checking this condition by a linear equation over a certain abelian group, hoping to solve this system using Gaussian elimination, and from this solution get a non-trivial solution to the original instance. The basic algorithmic idea is to relax the constraint by passing to a quotient group. Suppose $H \leqslant G$ is a subgroup such that $S \subseteq g + H$ for some $g \in G$, and let $Q := G/H$. Then the above constraint implies

$$([a_{i_1}]_Q + [x_{i_1}]_Q) + ([a_{i_2}]_Q + [x_{i_2}]_Q) + ([a_{i_3}]_Q + [x_{i_3}]_Q) = 0_Q, \tag{1}$$

Thus, we convert the constraints in Φ to a system of equations $\widetilde{\Phi}$ over Q. Since Φ is satisfiable, the system $\widetilde{\Phi}$ is satisfiable as well. Hence, when Q is abelian, we can find a satisfying assignment to $\widetilde{\Phi}$ in polynomial time using Gaussian elimination. Given such an assignment, we set each x_i to be a uniformly random element from the corresponding coset in G. For every constraint, the resulting sum is uniformly distributed over the appropriate coset of H, and hence the constraint is satisfied with probability $|S|/|H|$. One natural first choice is to take $H = \mathrm{cl}(S)$, the subgroup generated by S, but this choice is not always optimal. For instance, when $G = \mathbb{Z}_4 \times \mathbb{Z}_4$ and $S = \{(0,1), (1,0)\}$, the subgroup $\mathrm{cl}(S)$ is all of G, whereas S is contained in a coset of a smaller subgroup, giving a better approximation ratio. Therefore, the right choice is the smallest subgroup H such that S is contained in a coset of H. For non-abelian groups, we additionally require H to be normal and to contain the commutator subgroup $[G, G]$, so that the quotient G/H is abelian. This leads to the following theorem.

▶ **Theorem 2.** *Fix any finite group G and $S \subseteq G$. Let H_S is the smallest normal subgroup such that*

1. $[G, G] \subseteq H_S$.
2. *$S \subseteq gH_S$ for some $g \in G$, i.e., S is a subset of some coset of H_S.*

then the problem Max-Ek-LIN$_S(G)$ is approximable within a factor of $\frac{|S|}{|H_S|}$ on satisfiable instances. Furthermore, if $S^{-1}S$ generates H_S, assuming $\mathbf{P} \neq \mathbf{NP}$*, for every $\varepsilon > 0$, it is* $\mathbf{NP}$*-hard to approximate Max-Ek-LIN$_S(G)$ within a factor of $\frac{|S|}{|H_S|} + \varepsilon$.*

Note that for any abelian group G, the subgroup $[G, G]$ consists only of the identity element of G. For non-abelian groups, the condition $[G, G] \subseteq H_S$ guarantees that the quotient group $Q := G/H_S$ is abelian, which was crucial for the above-discussed approximation algorithm. See Theorem 13 for a straightforward generalization of the above approximation algorithm for non-abelian groups. The above theorem shows that even for abelian groups G, the problem Max-Ek-LIN$_S(G)$ is $\mathbf{NP}$-hard on satisfiable instances, in general.

A predicate P is called approximation resistant if it is $\mathbf{NP}$-hard to do better than the random assignment algorithm. We are aware of two results, one by Håstad [25] on satisfying degree-d equations over $\mathrm{GF}[2]^n$, and another by Bhangale-Khot [4] on Max-3LIN over non-abelian groups G where G is not a *simple* group. Both these predicates are approximation resistant in general (i.e., on almost satisfiable instances), but have non-trivial approximation algorithms on satisfiable instances. Thus, our result adds the generalized linear equation predicate to the small class of predicates that are known to be approximation resistant in general but have non-trivial approximation algorithms on satisfiable instances. We hope that our result gives another piece of information to help understand the approximability of Max-CSPs on satisfiable instances.

In Section 3, we give a simple $\frac{|S|}{|H_S|}$-approximation algorithm for Max-Ek-LIN$_S(G)$ for every finite group G. The algorithm implicitly uses the abelian embedding of the generalized linearity predicate as defined in the series of work [6, 7, 8, 9] towards understanding the approximability of satisfiable CSPs. To show this optimal hardness result, we design a novel decoding procedure in the soundness analysis of the reduction, which may be of independent interest towards understanding the approximability of satisfiable CSPs.

1.1 Related Work

In this section, we go over relevant work on the inapproximability of Max-CSPs. The PCP Theorem [1, 2, 19] shows that Max-P-CSPs are **NP**-hard to approximate within a factor of $(1-\delta)$ for some constant $\delta > 0$ if checking satisfiability of P-CSP is **NP**-complete. Håstad in his seminal work [24] greatly improved the hardness of approximation results for a few CSPs. Notable examples of the CSPs from his work include Max-3SAT and Max-3LIN. For Max-CUT, which is a 2-ary CSP, Håstad showed that it is **NP**-hard to approximate within a factor of $\frac{16}{17}$. Khot [29] formulated the *Unique Games Conjecture* (UGC), which is a conjecture on the hardness of the Label Cover instances (see Definition 4) restricted to the constraints being 1-to-1. Samorodnitsky and Trevisan [35] showed that Boolean Max-kCSP (with arbitrary k-ary Boolean predicates) is **NP**-hard to approximate beyond $O(\frac{k}{2^k})$ assuming the UGC, matching the best algorithm up to a constant factor. Khot, Kindler, Mossel, and O'Donnell [30] showed that Max-CUT is **NP**-hard to approximate within a factor of 0.878 assuming the UGC, which matches the approximation guarantee of the Goemans-Williamson [20] algorithm for Max-CUT.

Raghavendra [33] presented an elegant result that generalizes the above Max-CUT reduction and establishes, for any Max-P-CSP instance, the (c, s)-integrality gap of the basic Semidefinite programming relaxation implies finding an $s + \varepsilon$ satisfying assignment on $c - \varepsilon$ satisfiable instances is **NP**-hard assuming the UGC. This result fully characterizes the approximability of Max-P-CSPs assuming the UGC on almost satisfiable instances. Furthermore, given a fixed predicate P, Raghavendra's result does not explicitly give the optimal hardness factor for Max-P-CSP. Austrin and Mossel [3] gave the right threshold for predicates that support a uniform and pairwise independent distribution. A distribution μ on $P^{-1}(1) \subseteq \Sigma^k$ is said to be pairwise independent if for every distinct pair $i, j \in [k]$, the marginal of μ restricted to the coordinates (i, j) is uniform over Σ^2. Austrin and Mossel showed that such predicates are approximation resistant on almost satisfiable instances, assuming the UGC.

Chan [12] established a general criterion for approximation resistance, resolving the NP-hardness of Max-k-CSP up to a constant factor and assuming $\mathbf{P} \neq \mathbf{NP}$. Specifically, he proved the hardness for Max-CSPs where the domain is an abelian group G and the predicate $P^{-1}(1) \subseteq G^k$ is a subgroup that satisfies a condition analogous to that identified by Austrin and Mossel.

The question of finding the optimal approximation algorithm (even assuming certain conjectures, like d-to-1 conjecture [29] or Rich 2-to-1 conjecture [10]) for satisfiable instances of Max-P-CSPs is wide open. In a recent series of work, Bhangale, Khot, and Minzer [6, 7] defined a property of abelian embeddability of the predicate towards understanding the approximability of satisfiable CSPs. A predicate $P : \Sigma^k \to \{0, 1\}$ is said to have an abelian embedding in an abelian group G, if there are maps $\alpha_i : \Sigma \to G$, not all constant, such that $\sum_i \alpha_i(a_i) = 0_G$ for every $(a_1, a_2, \dots, a_k) \in P^{-1}(1)$. They gave an optimal dictatorship test for 3-ary predicates that have no abelian embedding. Very recently, for certain 3-ary predicates that have an abelian embedding, they gave an approximation algorithm [8, 9] for satisfiable instances that uses a combination of Gaussian elimination as well as the SDP rounding algorithm. They also showed the "optimality" of this algorithm by giving a dictatorship test with matching parameters.

1.2 Techniques

In this section, we give an overview of the techniques in the inapproximability results of our main theorem.

1.2.1 Abelian Groups

We begin with the case of abelian groups to highlight one of the main differences in the analysis of the reduction compared to the seminal work of Håstad on Max-3LIN over an abelian group. We assume some familiarity with the Fourier analysis of functions over abelian groups (for instance, Chapter 8 of Ryan O'Donnell's book [32]). Throughout the section, $\varepsilon > 0$ is an arbitrarily small constant.

Starting with a work of Håstad [24], a typical way to prove the hardness of approximation is to start with an **NP**-hard problem called the *Label Cover* [22, 27, 28, 14, 31, 26, 23]. A gadget is built on top of the Label Cover instance to create an instance of a given CSP. For simplicity of the presentation, we focus here is on the most important and technical component of the reduction. This component is the construction of *dictatorship tests* and analyzing the tests.

A function $f : [q]^n \to [q]$ is called a dictator function if $f(x_1, x_2, \dots, x_n) = x_j$ for some $j \in [n]$. A dictatorship test for a predicate $P : [q]^k \to \{0, 1\}$ consists of a distribution μ that is (almost) supported on the set of satisfying assignments of P. The test samples k inputs $z_1, z_2, \dots, z_k$ as follows: For each coordinate $i \in [n]$, the tuple $((z_1)_i, (z_2)_i, \dots, (z_k)_i)$ is sampled independently from μ. The test accepts f if $P(f(z_1), f(z_2), \dots, f(z_k))$ evaluates to 1, i.e., $f(z_1), f(z_2), \dots, f(z_k)$ forms a satisfying assignment for P. It is clear that if f is a dictator function, then the test passes with probability (almost) 1. This is because, in this case, we are checking if the i^{th} coordinate of the inputs is from $P^{-1}(1)$, which is always true by construction.

Once the distribution is fixed, the next step is to analyze the soundness of the dictatorship test, i.e., the probability with which the test passes if f is "far from the dictator functions". The notion of far from dictator functions changes based on the hardness reduction. A typical notion that is used is that the function has all the variables with degree $d = O(1)$ influences low. The influence of the i^{th} coordinate on the function is the probability that, on a random input, changing the i^{th} coordinate changes the values of the function. In terms of the Fourier coefficients of f, this is equal to the following quantity:

$$\mathrm{Inf}_i(f) := \sum_{\alpha:\alpha_i \neq 0} |\hat{f}(\alpha)|^2.$$

Thus, the i^{th} dictator function has $\mathrm{Inf}_i(f) = 1$. A degree-d influence of the i^{th} variable is given by the following expression,

$$\mathrm{Inf}_i^{\leqslant d}(f) := \sum_{\alpha:\alpha_i \neq 0 \wedge |\alpha| \leqslant d} |\hat{f}(\alpha)|^2,$$

where $|\alpha|$ is the number of non-zero coordinates of α. With this notion of far from dictator functions in mind, the dictatorship test used in the hardness reduction [24] of Max-3LIN works as follows.

1. Select $\boldsymbol{x}, \boldsymbol{y} \sim G^n$ uniformly at random.
2. Set $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$.
3. For each $i \in [n]$, resample (x_i, y_i, z_i) from G^3 uniformly at random, with probability ε.
4. Check if $f(\boldsymbol{x}) + f(\boldsymbol{y}) = f(\boldsymbol{z})$.

It is clear that any dictator function passes the above test with probability at least $1 - \varepsilon$. To analyze the soundness of the test, we can express the test passing probability as follows:

$$\Pr[\text{Test passes}] = \frac{1}{|G|} \mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} \left[\sum_{\chi_\rho \in \hat{G}} \chi_\rho(f(\boldsymbol{x}) + f(\boldsymbol{y}) + f(\boldsymbol{z})) \right],$$

where the summation is over all the characters of the group G. The term with χ_ρ being the trivial character gives $\frac{1}{|G|}$. For the remaining terms with χ_ρ being a non-trivial character, we are left with analyzing the following expectation

$$\mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} [\chi_\rho(f(\boldsymbol{x}) + f(\boldsymbol{y}) + f(\boldsymbol{z}))] = \mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} [\chi_\rho(f(\boldsymbol{x})) \cdot \chi_\rho(f(\boldsymbol{y})) \cdot \chi_\rho(f(\boldsymbol{z}))] .$$

By letting $F(\boldsymbol{w}) = \chi_\rho(f(\boldsymbol{w}))$ and expanding the function F with the Fourier basis over G^n, and doing some simplifications, we get that the above expectation is upper bounded as follows.

$$\begin{aligned} &\mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} [\chi_\rho(f(\boldsymbol{x})) \cdot \chi_\rho(f(\boldsymbol{y})) \cdot \chi_\rho(f(\boldsymbol{z}))] \leqslant \sum_{\alpha} |\hat{F}(\alpha)|^3 \cdot (1-\varepsilon)^{|\alpha|} \\ &= \sum_{\alpha \wedge |\alpha| \leqslant d} |\hat{F}(\alpha)|^3 \cdot (1-\varepsilon)^{|\alpha|} + \sum_{\alpha \wedge |\alpha| > d} |\hat{F}(\alpha)|^3 \cdot (1-\varepsilon)^{|\alpha|}. \end{aligned}$$

Decoding Strategy

Now, because of the '*noise*' (i.e., step 3 in the test), the second term can be shown to be negligible for some large d as a function of ε, by a simple application of Cauchy-Schwarz inequality and using Parseval's identity. As for the first term, if it is non-negligible, then the structure of the function can be used in the hardness reduction starting from the Label Cover instance.

More precisely, the starting point of the reduction is a Label Cover instance (see Definition 4). It is **NP**-hard to distinguish between the Label Cover instances with value 1 from instances with value at most δ for small $\delta > 0$. In the reduction, each vertex v of the Label Cover instance is replaced with a cloud of vertices $C[v]$ where $|C[v]|$ is G^M where M is the label size of the vertex. These constitute the variables/literals of the reduced instance of Max-3LIN. The distribution on the constraints is specified by the above dictatorship test. If the value of the Label Cover instance is 1, then the value of the reduced Max-3LIN instance is at least $1 - \varepsilon$. Now, similar to the above analysis, if we fix an assignment f to the Max-3LIN instance with value $\frac{1}{|G|} + \varepsilon$ for some $\varepsilon > 0$, then the assignment restricted to most of $C[v]$, call it f_v, has high degree-d influential variables. From this, one can come up with a labeling to the Label Cover instance with value greater than $\delta = \delta(\varepsilon)$, thereby showing the soundness of the reduction. The strategy simply picks a random non-trivial character ρ, a Fourier coefficient $\widehat{\chi_\rho(f_v)}(\alpha)$ of f_v with probability $|\widehat{\chi_\rho(f_v)}(\alpha)|^2$ and assign a label i such that α_i is not a trivial character.

We now give the natural extension of the above dictatorship test to the generalized linear equation predicate that we consider in this paper. Towards this, fix an abelian group $(G, +)$ and a subset $S \subseteq G$. Consider the following dictatorship test.

1. Select $\boldsymbol{x}, \boldsymbol{y} \sim G^n$ uniformly at random.
2. Select $\boldsymbol{w} \sim S^n$ uniformly at random.
3. Set $\boldsymbol{z} = -\boldsymbol{x} - \boldsymbol{y} + \boldsymbol{w}$.
4. Check if $f(\boldsymbol{x}) + f(\boldsymbol{y}) + f(\boldsymbol{z}) \in S$.

It is clear that every dictator function passes the test with probability 1. Similar to the above analysis, the soundness of the test can be expressed as follows:

$$\Pr[\text{Test passes}] = \frac{1}{|G|} \mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} \left[\sum_{s \in S} \sum_{\chi_\rho \in \hat{G}} \chi_\rho(f(\boldsymbol{x}) + f(\boldsymbol{y}) + f(\boldsymbol{z}) - s) \right],$$

Once again, the terms that correspond to the trivial character give $\frac{|S|}{|G|}$ (which corresponds to the approximation ratio of the algorithm that picks a random assignment). Note that our goal is to show that the test passes with probability almost $\frac{|S|}{|H_S|} + \varepsilon$ in the soundness towards proving Theorem 2. Once again, the term that corresponds to a given $s \in S$ and a character χ_ρ gives,

$$\mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} [\chi_\rho(f(\boldsymbol{x}) + f(\boldsymbol{y}) + f(\boldsymbol{z}) - s)] = \mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} [\chi_\rho(f(\boldsymbol{x})) \cdot \chi_\rho(f(\boldsymbol{y})) \cdot \chi_\rho(f(\boldsymbol{z})) \cdot \chi_\rho(s)] .$$

Ignoring the constant shift $\chi_\rho(s)$, we are again left with analyzing the expectation

$$\mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} [\chi_\rho(f(\boldsymbol{x})) \cdot \chi_\rho(f(\boldsymbol{y})) \cdot \chi_\rho(f(\boldsymbol{z}))] .$$

Recall that the subgroup H_S is the smallest subgroup such that $S \subseteq g + H_S$ for some $g \in G$. Now, unlike the Max-3LIN, we cannot expect this expectation to be small when the functions F have negligible degree-d influences for some $d = O(1)$. To see this, consider a character χ_ρ that is constant on the subgroup H_S. If we let $f(\boldsymbol{x}) = \sum_i x_i$ where the sum is the group operation, then the derived function F has all degree-d influences 0. However, the expectation becomes $\mathbb{E}[\chi_\rho(\sum_i w_i)]$, since $S \subseteq g + H_S$ for some g, we get that the $\sum_i w_i \in ng + H_S$, i.e, it always belongs to a fixed coset of H_S. Since the character χ_ρ is constant on the subgroup H_S (and hence constant on every coset of H_S), we get that the expectation is 1 in absolute value. The number of such characters (including the trivial character) for which we cannot bound the expectation is precisely $|G|/|H_S|$. This gives the right factor $\frac{|S|}{|H_S|}$ in the test passing probability that we need for our Theorem 2.

Now, consider the character χ_ρ, which is not constant on H_S. Similar to Håstad's analysis, we can upper bound the corresponding expectation as follows

$$\mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} [\chi_\rho(f(\boldsymbol{x})) \cdot \chi_\rho(f(\boldsymbol{y})) \cdot \chi_\rho(f(\boldsymbol{z}))] \leqslant \sum_{\alpha} |\hat{F}(\alpha)|^3 \cdot (1 - \eta)^{|\alpha|_S},$$

where $\eta > 0$ is a non-zero constant that only depends on $|G|$ and $|\alpha|_S$ is the number of coordinates of α where the character α_i is *not constant* on the subgroup H_S. In order to work with this expression, we modify the notion of far from dictator functions that will be useful for our hardness reduction as follows.

Modified low-degree influences

A modified degree-d influence of the i^{th} variable is expressed as the following expression,

$$\mathrm{Inf}_i^{\leqslant d}(f) := \sum_{\alpha:\alpha_i\neq 0\wedge|\alpha|_S\leqslant d} |\hat{f}(\alpha)|^2,$$

where $|\alpha|_S$ is the number of coordinates of α where the character α_i is not constant on the subgroup H_S. With this change, we split the summation as follows:

$$\sum_{\alpha} |\hat{F}(\alpha)|^3 \cdot (1-\eta)^{|\alpha|_S} = \sum_{\alpha\wedge|\alpha|_S\leqslant d} |\hat{F}(\alpha)|^3 \cdot (1-\eta)^{|\alpha|_S} + \sum_{\alpha\wedge|\alpha|_S>d} |\hat{F}(\alpha)|^3 \cdot (1-\eta)^{|\alpha|_S}.$$

Again, the second term is negligible for some large $d = O_{|G|}(1)$.

Modified Decoding Strategy

In our analysis of the reduction, we need to show that the terms that are similar to the first term are not negligible, then there is a decoding strategy (similar to the one described in the reduction to Max-3LIN above) in the hardness reduction starting from the Label Cover instance. The following strategy works. The modified strategy picks a random non-trivial irreducible representation ρ that is not constant on H_S, a Fourier coefficient $\widehat{\chi_\rho(f_v)}(\alpha)$ of f_v with probability $|\widehat{\chi_\rho(f_v)}(\alpha)|^2$ and assign a label i such that the character corresponding to α_i is not constant on H_S.

1.2.2 Non-abelian Groups

The above dictatorship test can be easily modified for the non-abelian case in a natural way

1. Select $\boldsymbol{x}, \boldsymbol{y} \sim G^n$ uniformly at random.
2. Select $\boldsymbol{s} \in S^n$ uniformly at random.
3. For each $i \in [n]$, set $z_i = y_i^{-1} \bullet x_i^{-1} \bullet s_i$.
4. Check if $f(\boldsymbol{x}) \bullet f(\boldsymbol{y}) \bullet f(\boldsymbol{z}) \in S$.

The completeness case is trivial. In the soundness case, we again express the test passing probability as

$$\Pr[\text{Test passes}] = \frac{1}{|G|} \mathop{\mathbb{E}}_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} \left[\sum_{s\in S} \left[\sum_{\rho\in\mathsf{Irrep}(G)} \dim(\rho) \cdot \chi_\rho(f(\boldsymbol{x}) \bullet f(\boldsymbol{y}) \bullet f(\boldsymbol{z}) \bullet s^{-1}) \right] \right],$$

where the inner sum is over all irreducible representations of $(G, \bullet)$ and $\dim(\rho)$ is the dimension of the representation ρ. Similarly to the abelian case, for certain representations of dimension 1 (that are constant on the subgroup H_S defined in Theorem 2), we bound the expectation by 1 in absolute value. This gives the factor $\frac{|S|}{|H_S|}$. The analysis for the remaining dimension 1 representations, Lemma 16, is analogous to the one described in the abelian case but focuses on the natural structure of non-abelian groups and generalized linear equation predicates. For some technical reasons, for non-abelian groups, we could analyze this under the assumption that $S^{-1}S$ generates the subgroup H_S.

Regarding representations with dimension $\geqslant 2$, we treat the analysis of Bhangale and Khot [4] as a black box (Lemma 19) to conclude that the associated expectations are small unless they yield a decoding strategy for the Label Cover instance. For some technical reasons, the analysis of this part of the reduction requires the use of Layered Label Cover instead of the bipartite Label Cover. Therefore, in our main reduction, we also use a Layered

Label Cover instance as a starting point, but we only use the layered version in the proof of Lemma 19 to adopt results from [5]. For the primary part of this paper, we only need a bipartite Label Cover instance.

▶ Remark 3. The conference version of the paper [4] had an error, which the authors fixed with the use of Layered Label Cover as a starting point.[3] This fix is reflected in the Lemma 19 [5, Claim 4.5] that we use as black-box.

1.3 Organization

We begin Section 2 by defining the Label Cover instance and the hardness of approximation of Label Cover, which is the starting point of our reduction. In Section 2.2, we go over the basics of Fourier analysis over general finite groups. We formally give the approximation algorithm described in the introduction in Section 3. Finally, in Section 4 we give our hardness reduction for Max-E3-LIN$_S(G)$ and analyze the reduction. The hardness for Max-Ek-LIN$_S(G)$ for $k \geqslant 4$ follows easily from a similar reduction, thereby proving the main Theorem 2.

2 Preliminaries

2.1 Label Cover

We start by defining the Label Cover and Layered Label Cover problem, which we use as a starting point for our reduction.

▶ **Definition 4** (Label Cover). *An instance $\Psi = (U, V, E, [L], [R], \{\pi_e\}_{e\in E})$ of the Label Cover constraint satisfaction problem consists of a bi-regular bipartite graph (U, V, E), alphabets $[L]$ and $[R]$ and a surjective projection map $\pi_e : [L] \to [R]$ for every edge $e \in E$. Given a labeling $\ell : U \to [L], \ell : V \to [R]$, an edge $e = (u, v)$ is said to be satisfied by ℓ if $\pi_e(\ell(u)) = \ell(v)$.*

Ψ is said to be satisfiable *if there exists a labeling that satisfies all the edges. Ψ is said to be* at most δ-satisfiable *if every labeling satisfies at most a δ fraction of the edges.*

The hardness of Label Cover stated below follows from the PCP Theorem [2, 1, 19] and Raz's Parallel Repetition Theorem [34].

▶ **Theorem 5** (Hardness of Label Cover). *For every $r \in \mathbb{N}$, there is a deterministic $n^{O(r)}$-time reduction from a 3-SAT instance of size n to an instance $\Psi = (U, V, E, [L], [R], \{\pi_e\}_{e\in E})$ of Label Cover with the following properties: $|U|, |V| \leqslant n^{O(r)}$; $L, R \leqslant 2^{O(r)}$; Ψ is bi-regular with degrees bounded by $2^{O(r)}$.*

YES Case: If the 3-SAT instance is satisfiable, then Ψ is satisfiable.

NO Case: If the 3-SAT instance is unsatisfiable, then Ψ is at most 2^{-r} satisfiable

For our hardness result, we need the following variant of the Label Cover problem.

▶ **Definition 6** (Layered Label Cover). *An T-Layered Label Cover instance, given by $\mathcal{H} = (\mathcal{V} = \{V_1, \dots, V_T\}, \{\Pi_{i,j}\}_{1\leqslant i<j\leqslant T}\}, \{[R_i]_{i\in[T]}\}$ consist of T sets of vertices $\mathcal{V} = \{V_1, \dots, V_T\}$. The label set of vertices in layer i is denoted by $[R_i]$. Every pair of layers $1 \leq i < j \leq l$ has a set of constraints $\Pi_{i,j}$ between the vertices in V_i and V_j. The constraint between $v \in V_i$ and $u \in V_j$ (if it exists in $\Pi_{i,j}$) is denoted by π_{vu}. Moreover, every constraint between a pair of vertices is a projection constraint: for every assignment $k \in [R_i]$ to $v \in V_i$, there is a unique assignment to $u \in V_j$ that satisfies the constraint π_{vu}.*

[3] personal communication [5]

▶ **Theorem 7** ([15], Hardness of Layered Label Cover). *For any constant parameters $T \geq 2$, $r \in \mathbb{Z}$, the following problem is NP-hard. Given an T-Layered Label Cover instance $\mathcal{H} = (\mathcal{V} = \{V_1, ..., V_T\}, \{\Pi_{i,j}\}_{1 \leq i < j \leq T}, \{[R_t]\}_{t \in [T]})$ where all variable ranges R_t are of size $2^{O(Tr)}$, distinguish between the following two cases:*

Completeness. *There is an assignment satisfying all the constraints of the Label Cover instance. In this case, we say that $\mathcal{H}$ is fully satisfiable.*

Soundness. *For every $1 \leq i < j \leq T$, no assignment satisfies more than a 2^{-r} fraction of the set of constraints $\Pi_{i,j}$ between layers i and j. In this case, we say that $\mathcal{H}$ is at most 2^{-r}-satisfiable.*

2.2 Fourier analysis over non-abelian group

We assume familiarity with the representation theory and Fourier analysis of functions over finite non-abelian groups. We refer the reader to the book by Terras [37] and the work of Bhangale and Khot [4] for standard definitions (such as irreducible representations, characters, Plancherel's, and Parseval's identity) which we omit here for brevity.

In this paper, we only consider non-abelian groups that are finite. Let $G = (G, \bullet)$ be a finite non-abelian group. The identity element of a group is denoted by 1_G. We assume familiarity with basic definitions of representations, dimensions, invariant subspaces, and tensor products. We will denote the set of all irreducible representations of G up to isomorphism by $\mathsf{Irrep}(G)$.

▶ **Fact 8.** *Let G be a group and H be any subgroup of G. If $\rho \in \mathsf{Irrep}(G)$ then ρ restricted to H is also a (not necessarily irreducible) representation of H.*

The following proposition shows that matrix entries of irreducible representations are "orthogonal" with respect to a symmetric bilinear form, unless they are conjugates of each other, in which case the corresponding product is the inverse of the dimension of the representation.

▶ **Proposition 9.** *If ρ and τ are two non-isomorphic irreducible representations of G then for any i, j, k, l we have*

$$\langle (\rho)_{ij} \mid (\tau)_{kl} \rangle_G = 0, \tag{2}$$

where $\langle f_1 \mid f_2 \rangle_G := \frac{1}{|G|} \sum_{g \in G} f_1(g) f_2(g^{-1})$ (called a "symmetric bilinear form"). Also,

$$\langle (\rho)_{ij} \mid (\rho)_{kl} \rangle_G = \frac{\delta_{il}\delta_{jk}}{\dim(\rho)}, \tag{3}$$

where δ_{ij} is the delta-function which is 1 if $i = j$ and 0 otherwise.

We use Proposition 9 many times in the proof. For convenience, we note an important identity that follows from Proposition 9 (by setting τ to be the trivial map $\{\mathbf{1}\}$).

▶ **Proposition 10.** *If $\rho \in \mathsf{Irrep}(G) \backslash \{\mathbf{1}\}$, $\sum_{g \in G} \rho(g) = 0$.*

In this paper, we will be interested in studying $L^2(G)$, the space of functions from a finite group G to the complex numbers $\mathbb{C}$. Define the inner product $\langle \cdot, \cdot \rangle_{L^2(G)}$ on $L^2(G)$ by $\langle f, g \rangle_{L^2(G)} = \mathbb{E}_{x \in G}[f(x)\overline{g(x)}]$. The character of a representation ρ is the function $\chi_\rho : G \to \mathbb{C}$ defined by $\chi_\rho(g) = \mathsf{tr}(\rho(g))$.

We will use the following propositions.

▶ **Proposition 11** (Orthogonality of characters). *For $\rho, \tau \in \mathsf{Irrep}(G)$, we have*

$$\frac{1}{|G|}\sum_{g\in G}\chi_\rho(g)\overline{\chi_\tau(g)} = \begin{cases}1 & \rho_V \cong \tau_V, \\ 0 & \textit{otherwise}.\end{cases}$$

This proposition also shows that the maximum dimension of any irreducible representation of G is at most $\sqrt{G}$.

▶ **Proposition 12.**

$$\sum_{\rho\in\mathsf{Irrep}(G)} \dim(\rho)\chi_\rho(g) = \begin{cases}|G| & g = 1_G, \\ 0 & \textit{otherwise}.\end{cases}$$

This implies $\sum_{\rho\in\mathsf{Irrep}(G)} \dim(\rho)^2 = |G|$.

3 An Approximation Algorithm

In this section, we give an approximation algorithm for Max-E3-LIN$_S(G)$ for any group $(G, \bullet)$ and $S \subseteq G$. This algorithm is a straightforward generalization of the algorithm for abelian groups G discussed in the introduction. We repeat it here for completeness.

▶ **Theorem 13.** *There exists a $\frac{|S|}{|H_S|}$-approximation algorithm for Max-E3-LIN$_S(G)$, where H_S is the smallest normal subgroup such that (i) $[G, G] \subseteq H_S$, and (ii) S is a subset of some coset of H_S.*

Proof. Let Φ be an instance of Max-E3-LIN$_S(G)$ with constraints $(C_1, C_2, \dots, C_m)$ over the variables $X = \{x_1, x_2, \dots, x_n\}$. We first convert the set of constraints to a system of equations, denoted by $\tilde{\Phi}$, over the group $(Q, +) := G/H_S$ with variables $Y = \{y_1, y_2, \dots, y_n\}$. Note that H_S is a normal subgroup of G containing the commutator subgroup such that $S \subseteq gH_S$ for some $g \in G$. Thus, Q is an abelian group.

Consider a constraint C_i which is of the form $(a_{i_1} \bullet x_{i_1}) \bullet (a_{i_2} \bullet x_{i_2}) \bullet (a_{i_3} \bullet x_{i_3}) \in S$. We convert this to the equation over Q as

$$[a_{i_1}]_Q + y_{i_1} + [a_{i_2}]_Q + y_{i_2} + [a_{i_3}]_Q + y_{i_3} = [S]_Q,$$

where $[S]_Q$ is an element of Q that corresponds to the coset of H_S containing S, and the element $[g]_Q$ corresponds to the coset of H_S containing g.

As Φ is satisfiable, consider the satisfying assignment $\boldsymbol{\alpha} : X \to G$ to Φ. Consider the assignment $\tilde{\boldsymbol{\alpha}} : Y \to Q$ given by the natural map $\tilde{\boldsymbol{\alpha}}(y_i) = [\boldsymbol{\alpha}(x_i)]_Q$. It is easy to see that $\tilde{\boldsymbol{\alpha}}$ satisfies all the equations from the instance $\tilde{\Phi}$, and hence, $\tilde{\Phi}$ is satisfiable.

Since $\tilde{\Phi}$ is a system of equations over an abelian group $(Q, +)$, we can find a satisfying assignment to $\tilde{\Phi}$ in polynomial time using Gaussian elimination [21]. Let $\tilde{\boldsymbol{\beta}}$ be the assignment returned by this procedure. To construct the final assignment to the X variables, we simply set x_i to be a random element from the coset $\tilde{\boldsymbol{\beta}}(y_i)$. Let $\boldsymbol{\beta} : X \to G$ be the random assignment given by the above procedure. It can be easily observed that $\boldsymbol{\beta}$ satisfies a given constraint C_i in Φ with probability $\frac{|S|}{|H_S|}$ and hence $\boldsymbol{\beta}$ satisfies $\frac{|S|}{|H_S|}$ fraction of the constraints in expectation. The randomized algorithm can be easily derandomized using the method of conditional expectations. ◀

4 Hardness of Max-E3-LIN$_S(G)$

We start with some basic facts. For a nonabelian group G, the quotient group $G/[G,G]$ is an abelian group. The dual of $G/[G,G]$ is isomorphic to

$$\{\chi_\rho \mid \rho \in \mathsf{Irrep}(G), \dim(\rho) = 1\}$$

We denote this subgroup of $\hat{G}$ as $\widehat{G/[G,G]}$.

Similarly, consider any normal subgroup $H \trianglelefteq G$ such that $[G,G] \subseteq H$, the quotient group G/H is an abelian group that is isomorphic to a subgroup of $G/[G,G]$. Furthermore, the dual of G/H is isomorphic to

$$\{\chi_\rho \mid \rho \in \mathsf{Irrep}(G), \dim(\rho) = 1, \chi_\rho(h) = 1, \forall h \in H\}.$$

Main Reduction

We now give a reduction from a Layered Label Cover instance, denoted by, $\mathcal{H} = (\mathcal{V} = \{V_1, \ldots, V_T\}, \{\Pi_{i,j}\}_{1 \leqslant t < t' \leqslant T}, \{[R_t]_{t \in [T]}\}$ to a Max-E3-LIN$_S(G)$ instance Φ over a nonabelian group G. For $\delta > 0$, we will use the following setting of T and r in Theorem 7

$$2^{-r} \leqslant \min\left\{\frac{\delta^{10}}{(2|G|)^{20}}, \frac{\delta^2}{10|G|^{10K}}\right\}, \quad T \geqslant \left(\frac{8|G|^3}{\delta}\right)^4,$$

where $K := \frac{8|G|^6}{\delta^2}$.

Consider the Layered Label Cover instance $\mathcal{H}$. For all $t \in [T]$ and for each $v \in V_t$, we create a cluster $C[v]$ of literals of size $|G|^{R_t}$. In each cluster $C[v]$, each literal is indexed by a string of length R_t. For any string $(1_G, \boldsymbol{y}) \in G^{R_t}$, its corresponding literals are $g \bullet (1_G, \boldsymbol{y})$ for $g \in G$, where the string $g \bullet (1_G, \boldsymbol{y})$ is $(g, g \bullet y_1, \ldots, g \bullet y_{R_{t-1}})$.

An assignment to the instance that we are going to create is given by the maps $A_v : G^{R_t} \to G$ for all $v \in V_t$ and all $t \in [T]$. Note that any such assignment is assumed to be folded, i.e., $A_v(g \bullet (1_G, \boldsymbol{y})) = g \bullet A_v((1_G, \boldsymbol{y}))$.

The distribution on the constraint of the reduced instance Φ of Max-E3-LIN$_S(G)$ is given by the following PCP verifier.

1. Pick a uniformly random pair (t, t') satisfying $1 \leqslant t < t' \leqslant T$.
2. Choose an edge constraint $\pi_{uv} : [R_t] \to [R_{t'}]$ from $\mathcal{H}$ uniformly at random.
3. Sample a string $\boldsymbol{x} \sim G^{R_{t'}}$ and $\boldsymbol{y} \sim G^{R_t}$ independently and uniformly at random.
4. Sample an element $\boldsymbol{s} \in S^{R_t}$ uniformly at random.
5. For each $j \in [R_t]$, set $z_j = y_j^{-1} \bullet x_{\pi_{uv}(j)}^{-1} \bullet s_j$.
6. Accept if and only if $A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \in S$.

4.1 Completeness

If $\mathcal{H}$ is fully satisfiable, then there exists a corresponding assignment σ such that all the constraints are satisfiable. Let $A_v(\boldsymbol{x}) = x_{\sigma(v)}$ and $A_u(\boldsymbol{y}) = y_{\sigma(u)}$, i.e., the dictator functions. Then, the test passes as,

$$\begin{aligned}
&A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \\
&= x_{\sigma(v)} \bullet y_{\sigma(u)} \bullet z_{\sigma(u)} \\
&= x_{\sigma(v)} \bullet y_{\sigma(u)} \bullet (y_{\sigma(u)})^{-1} \bullet (x_{\pi_{u,v}(\sigma(u))})^{-1} \bullet s_{\sigma(u)} \\
&= x_{\sigma(v)} \bullet y_{\sigma(u)} \bullet (y_{\sigma(u)})^{-1} \bullet (x_{\sigma(v)})^{-1} \bullet s_{\sigma(u)} && \text{(Using } \pi_{u,v}(\sigma(u)) = \sigma(v)\text{)} \\
&= s_{\sigma(u)} \in S.
\end{aligned}$$

Hence, the test always passes. Thus, the value of the instance Φ is 1.

4.2 Soundness

In this section, we prove the soundness of the analysis.

▶ **Lemma 14.** *For every $\delta > 0$, if the Layered Label Cover instance $\mathcal{H}$ is at most 2^{-r}, then the Max-E3-LIN$_S(G)$ instance Φ is at most $\frac{|S|}{|H_S|} + \delta$ satisfiable.*

Proof. Fix the assignment $\{A_v\}_{v \in \mathcal{V}}$ to the instance Φ. We define the value of an assignment A, value(A), as the probability that the above test passes. The following expression gives the value of this assignment A,

$$\mathop{\mathbb{E}}_{1 \leqslant t < t' \leqslant T} \left[\mathop{\mathbb{E}}_{\substack{\pi_{uv} \in \Pi_{t,t'} \\ (\boldsymbol{x},\boldsymbol{y},\boldsymbol{s})}} \left[\sum_{s \in S} \left[\frac{1}{|G|} \sum_{\rho \in \mathsf{Irrep}(G)} \dim(\rho) \cdot \chi_\rho(A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \bullet s^{-1}) \right] \right] \right].$$

By Proposition 12, this expression equals 1 if and only if $A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \in S$, and 0 otherwise. We can rewrite this expression according to the representations $\rho \in \mathsf{Irrep}(G)$,

$$\begin{aligned}
&\text{value}(A) \\
&= \frac{1}{|G|} \mathop{\mathbb{E}}_{1 \leqslant t < t' \leqslant T} \left[\sum_{\rho \in \mathsf{Irrep}(G)} \sum_{s \in S} \mathop{\mathbb{E}}_{\substack{\pi_{uv} \in \Pi_{t,t'} \\ \boldsymbol{x},\boldsymbol{y},\boldsymbol{s}}} \left[\dim(\rho) \cdot \chi_\rho(A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \bullet s^{-1})\right] \right] \\
&= \frac{1}{|G|} \mathop{\mathbb{E}}_{1 \leqslant t < t' \leqslant T} \left[\sum_{\substack{\rho \in \widehat{G/H_S} \\ \dim(\rho)=1}} \sum_{s \in S} \mathop{\mathbb{E}}_{\substack{\pi_{uv} \in \Pi_{t,t'} \\ \boldsymbol{x},\boldsymbol{y},\boldsymbol{s}}} \left[\dim(\rho)\chi_\rho(A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \bullet s^{-1})\right] \right] &(4)\\
&+ \frac{1}{|G|} \mathop{\mathbb{E}}_{1 \leqslant t < t' \leqslant T} \left[\sum_{\substack{\rho \notin \widehat{G/H_S} \\ \dim(\rho)=1}} \sum_{s \in S} \mathop{\mathbb{E}}_{\substack{\pi_{uv} \in \Pi_{t,t'} \\ \boldsymbol{x},\boldsymbol{y},\boldsymbol{s}}} \left[\dim(\rho)\chi_\rho(A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \bullet s^{-1})\right] \right] &(5)\\
&+ \frac{1}{|G|} \mathop{\mathbb{E}}_{1 \leqslant t < t' \leqslant T} \left[\sum_{\dim(\rho) \geqslant 2} \sum_{s \in S} \mathop{\mathbb{E}}_{\substack{\pi_{uv} \in \Pi_{t,t'} \\ \boldsymbol{x},\boldsymbol{y},\boldsymbol{s}}} \left[\dim(\rho) \cdot \chi_\rho(A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \bullet s^{-1})\right] \right]. &(6)
\end{aligned}$$

Term (4) is a constant between any two layers. As for term (5), we prove that they can be used to decode a valid assignment to any pair of layers unless they are negligible along a random path p. Finally, for term (6), we use [5] as a black box and show they are negligible along a random path p.

For term 4 and term 5, since they have $\dim(\rho) = 1$, the character χ_ρ is a homomorphism. Thus, we have

$$\dim(\rho)\chi_\rho(A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \bullet s^{-1}) = \chi_\rho(A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z})) \cdot \chi_\rho(s^{-1}).$$

For term (4), as χ_ρ is a 1-bounded function, it's upper bounded by

$$\frac{1}{|G|} \sum_{\substack{\rho \in \widehat{G/H_S} \\ \dim(\rho)=1}} \sum_{s \in S} 1 = \frac{1}{|G|} |S| \left| \frac{G}{H_S} \right| = \frac{|S|}{|H_S|}.$$

4.2.1 Bounding expressions in (5)

We now bound the term (5). Using $|\chi_\rho(s^{-1})| \leqslant 1$, we have (5) is at most

$$\frac{|S|}{|G|} \sum_{\substack{\rho \notin \widehat{G/H_S} \\ \dim(\rho)=1}} \mathop{\mathbb{E}}_{\substack{(u,v) \\ \boldsymbol{x},\boldsymbol{y},\boldsymbol{s}}} [\chi_\rho(A_v(\boldsymbol{x})) \cdot \chi_\rho(A_u(\boldsymbol{y})) \cdot \chi_\rho(A_u(\boldsymbol{z}))],$$

We argue that if this expression is large, then a decoding strategy exists for the Label Cover instance. We need the following simple lemma.

▶ **Lemma 15.** *Let $h : G^n \to G$ be any folded function, β is a representation of G^n. Define $s(\beta) := \{i \mid \beta_i \notin \widehat{G/H_S}\}$. Let $g(\boldsymbol{x}) = \chi_\rho(h(\boldsymbol{x}))$ for $\chi_\rho \notin \widehat{G/H_S}$ and* $\dim(\rho) = 1$*, then $\hat{g}(\beta) = 0$ for all β with $s(\beta) = \emptyset$.*

Proof. By definition, we have

$$\begin{aligned}
\hat{g}(\beta) &= \mathop{\mathbb{E}}_{\boldsymbol{x}\in G^n} [\chi_\rho(h(\boldsymbol{x}))\beta(\boldsymbol{x})] \\
&= \mathop{\mathbb{E}}_{\substack{\boldsymbol{y}\in G^{n-1} \\ x_1 = 1_G}} \left[\mathop{\mathbb{E}}_{c\in G} [\chi_\rho(h(c \bullet (1_G, \boldsymbol{y})))\beta(\boldsymbol{c} \bullet (1_G, \boldsymbol{y}))] \right] \\
&= \mathop{\mathbb{E}}_{\boldsymbol{y}\in G^{n-1}} \left[\mathop{\mathbb{E}}_{c\in G} [\chi_\rho(c) \cdot \chi_\rho(h((1_G, \boldsymbol{y}))) \cdot \beta(\boldsymbol{c}) \cdot \beta((1_G, \boldsymbol{y}))] \right] \\
&= \mathop{\mathbb{E}}_{\boldsymbol{y}\in G^{n-1}} \chi_\rho(h((1_G, \boldsymbol{y})))\beta((1_G, \boldsymbol{y})) \cdot \mathop{\mathbb{E}}_{c\in G} [\chi_\rho(c)\beta(\boldsymbol{c})].
\end{aligned}$$

If β satisfies $|s(\beta)| = 0$, then $\beta(\boldsymbol{c}) = \bigotimes_{i=1}^{L} \beta_i(c)$ is a complex number. Therefore, there always exists a ρ' such that $\dim(\rho') = 1$ and $\chi_\rho(c)\beta(\boldsymbol{c}) = \chi_{\rho'}(c)$. Furthermore, such $\rho' \notin \widehat{G/H_S}$ since for any element $q \in G/H_S$, $\beta(q) = 1$ and $\chi_\rho(q) \neq 1$, indicating that $\chi_{\rho'}(q) \neq 1$. Hence, by Proposition 10,

$$\hat{g}(\beta) = \mathop{\mathbb{E}}_{\boldsymbol{y}\in G^{n-1}} \chi_\rho(h((1_G, \boldsymbol{y})))\beta((1_G, \boldsymbol{y})) \cdot \mathop{\mathbb{E}}_{c\in G} [\chi_{\rho'}(c)] = 0. \qquad ◀$$

We now prove the following main lemma from this section.

▶ **Lemma 16.** *If the Layered Label Cover instance $\mathcal{H}$ is at most 2^{-r} satisfiable, then for any* $\dim(\rho) = 1$ *such that $\rho \notin \widehat{G/H_S}$,*

$$\left| \mathop{\mathbb{E}}_{1\leqslant t<t'\leqslant T} \left[\mathop{\mathbb{E}}_{\pi_{uv}\in\Pi_{t,t'}} \left[\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}} [\chi_\rho(A_v(\boldsymbol{x})) \cdot \chi_\rho(A_u(\boldsymbol{y})) \cdot \chi_\rho(A_u(\boldsymbol{z}))] \right] \right] \right| \leqslant \frac{\delta}{2|G|}, \tag{7}$$

Proof. Consider two layers U and V whose alphabets are $[L]$ and $[R]$ and an edge constraint $e = (u, v)$ such that $u \in U$ and $v \in V$. Let π denote the projection constraint on e. Let $f_v(\boldsymbol{x}) = \chi_\rho(A_v(\boldsymbol{x}))$, $g_u(\boldsymbol{x}) = \chi_\rho(A_u(\boldsymbol{x}))$ and $h_u^{\boldsymbol{s}}(\boldsymbol{x}) = \chi_\rho(A_u(\boldsymbol{x} \bullet \boldsymbol{s}))$. With these notations, we have

$$\chi_\rho(A_u(\boldsymbol{y})) \cdot \chi_\rho(A_u(\boldsymbol{z})) = g_u(\boldsymbol{y}) \cdot h_u^{\boldsymbol{s}}(\boldsymbol{y}^{-1} \bullet (\boldsymbol{x} \circ \pi)^{-1}) = (g_u * h_u^{\boldsymbol{s}})((\boldsymbol{x} \circ \pi)^{-1}),$$

where $(\boldsymbol{x} \circ \pi)_j = x_{\pi(j)}$. Thus, the inner expectation can be written as

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{s}} [f_v(\boldsymbol{x}) \cdot (g_u * h_u^{\boldsymbol{s}})((\boldsymbol{x} \circ \pi)^{-1})] \\
&= \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{s}} \left[\sum_{\alpha \in \mathsf{Irrep}(G^R)} \dim(\alpha) \mathsf{tr}(\hat{f}_v(\alpha) \cdot \alpha(\boldsymbol{x}^{-1})) \sum_{\beta \in \mathsf{Irrep}(G^L)} \dim(\beta) \mathsf{tr}(\hat{g_u}(\beta) \hat{h_u^{\boldsymbol{s}}}(\beta) \cdot \beta(\boldsymbol{x} \circ \pi)) \right] \\
&= \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{s}} \left[\sum_{\alpha,\beta} \dim(\alpha) \dim(\beta) \mathsf{tr}(\hat{f}_v(\alpha) \cdot \alpha(\boldsymbol{x}^{-1})) \mathsf{tr}(\hat{g_u}(\beta) \hat{h_u^{\boldsymbol{s}}}(\beta) \cdot \beta(\boldsymbol{x} \circ \pi)) \right] \\
&= \sum_{\alpha,\beta} \dim(\alpha) \dim(\beta) \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{s}} \left[\mathsf{tr}(\hat{f}_v(\alpha) \cdot \alpha(\boldsymbol{x}^{-1})) \mathsf{tr}(\hat{g_u}(\beta) \hat{h_u^{\boldsymbol{s}}}(\beta) \cdot \beta(\boldsymbol{x} \circ \pi)) \right].
\end{aligned}$$

Denote

$$\mathbf{Term}^e(\alpha, \beta) := \dim(\alpha) \dim(\beta) \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{s}} [\mathsf{tr}(\hat{f}_v(\alpha) \cdot \alpha(\boldsymbol{x}^{-1})) \mathsf{tr}(\hat{g_u}(\beta) \hat{h_u^{\boldsymbol{s}}}(\beta) \cdot \beta(\boldsymbol{x} \circ \pi))],$$

we have

$$\begin{aligned}
&\mathbf{Term}^e(\alpha, \beta) \\
&= d_\alpha \dim(\beta) \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{s}} \left[\sum_{1 \leqslant p,q \leqslant d_\alpha} \hat{f}_v(\alpha)_{pq} \cdot \alpha(\boldsymbol{x}^{-1})_{qp} \sum_{1 \leqslant i,k \leqslant \dim(\beta)} \hat{g_u}(\beta) \hat{h_u^{\boldsymbol{s}}}(\beta)_{ik} \cdot \beta(\boldsymbol{x} \circ \pi)_{ki} \right] \\
&= d_\alpha \dim(\beta) \mathop{\mathbb{E}}_{\boldsymbol{s}} \left[\sum_{p,q,i,k} \hat{f}_v(\alpha)_{pq} \cdot (\hat{g}_u(\beta) \hat{h_u^{\boldsymbol{s}}}(\beta))_{ik} \cdot \mathop{\mathbb{E}}_{\boldsymbol{x}} \left[\alpha(\boldsymbol{x}^{-1})_{qp} \cdot \beta(\boldsymbol{x} \circ \pi)_{ki} \right] \right],
\end{aligned}$$

where (i, k) are the tuples $i = (i_1, i_2, \ldots, i_L)$ and $k = (k_1, k_2, \ldots, k_L)$. Similarly, (p, q) are tuples $p = (p_1, p_2, \ldots, p_R)$ and $q = (q_1, q_2, \ldots, q_R)$. Then,

$$\begin{aligned}
\mathop{\mathbb{E}}_{\boldsymbol{x}} \left[\alpha(\boldsymbol{x}^{-1})_{qp} \cdot \beta(\boldsymbol{x} \circ \pi)_{ki} \right] &= \mathop{\mathbb{E}}_{\boldsymbol{x}} \left[\prod_{l=1}^{R} \alpha_l(x_l^{-1})_{q_l p_l} \prod_{l'=1}^{L} \beta_{l'}(x_{l'})_{k_{l'} i_{l'}} \right] && (8) \\
&= \prod_{l=1}^{R} \mathop{\mathbb{E}}_{\boldsymbol{x}} \left[\alpha_l(x_l^{-1})_{q_l p_l} \prod_{l' \in \pi^{-1}(l)} \beta_{l'}(x_{l'})_{k_{l'} i_{l'}} \right]. && (9)
\end{aligned}$$

We argue that $s(\alpha) \subseteq \pi(s(\beta))$ to make this expression non-zero. Suppose that there exists a ℓ such that for all $\ell' \in \pi^{-1}(\ell)$, $\dim(\beta_{\ell'}) = 1$, then the product of all such $\beta_{\ell'}$ must be 1 dimensional. According to Proposition 9, the expectation is 0 unless α_ℓ is also 1 dimensional and isomorphic to the product of $\beta_{\ell'}$. However, for an $\alpha_l \notin \widehat{G/H_S}$, if $\beta_{\ell'} \in \widehat{G/H_S}$ for all ℓ', then the product of $\beta_{\ell'}$ belongs to $\widehat{G/H_S}$ and is not isomorphic to α_ℓ. Therefore, for the expectation to be non-zero, there exists some ℓ' such that $\beta_{\ell'} \notin \widehat{G/H_S}$, indicating $s(\alpha) \subseteq s(\beta)$.

Define a function $F_\beta^{ki}(\boldsymbol{x}^{-1}) := \beta(\boldsymbol{x} \circ \pi)_{ki}$, and note that

$$\sum_i \|F_\beta^{ki}\|_2^2 = \sum_i \mathop{\mathbb{E}}_{\boldsymbol{x}} [|\beta(\boldsymbol{x}^{-1} \circ \pi)_{ki}|^2] = \mathop{\mathbb{E}}_{\boldsymbol{x}} \left[\sum_i |\beta(\boldsymbol{b}^{-1} \circ \pi)_{ki}|^2 \right] = 1, \tag{10}$$

where the last equality follows from the fact that the sum expression is exactly the norm of the k-th row of the representation β. Since $\beta(.)$ is unitary, the norm of its row is always 1. Using the function F_β^{ki}, we further simplify the expectation $\mathbb{E}_{\boldsymbol{x}} \left[\alpha(\boldsymbol{x}^{-1})_{qp} \cdot \beta(\boldsymbol{x} \circ \pi)_{ki} \right]$ as follows.

$$
\begin{aligned}
\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\alpha(\boldsymbol{x}^{-1})_{qp}\cdot\beta(\boldsymbol{x}\circ\pi)_{ki}\right] &= \mathop{\mathbb{E}}_{\boldsymbol{x}}[\alpha(\boldsymbol{x}^{-1})_{qp}F_{\beta}^{ki}(\boldsymbol{x}^{-1})]\\
&= \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\alpha(\boldsymbol{x}^{-1})_{qp}\sum_{\gamma}\dim(\gamma)\mathsf{tr}(\hat{F_{\beta}^{ki}}(\gamma)\gamma(\boldsymbol{x}))\right]\\
&= \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\sum_{\gamma}\dim(\gamma)\sum_{p',q'}\hat{F_{\beta}^{ki}}(\gamma)_{p'q'}\gamma(\boldsymbol{x})_{q'p'}\alpha(\boldsymbol{x}^{-1})_{qp}\right]\\
&= \sum_{\gamma}\dim(\gamma)\sum_{p',q'}\hat{F_{\beta}^{ki}}(\gamma)_{p'q'}\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\gamma(\boldsymbol{x})_{q'p'}\alpha(\boldsymbol{x}^{-1})_{qp}\right].
\end{aligned}
$$

By Proposition 9, the expectation is 0 unless $\alpha=\gamma$, $p'=q$ and $q'=p$. Hence, we have

$$
\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\alpha(\boldsymbol{x}^{-1})_{qp}\cdot\beta(\boldsymbol{x}\circ\pi)_{ki}\right] = \hat{F_{\beta}^{ki}}(\alpha)_{pq}.
$$

Thus, we can express $\mathbf{Term}^{e}(\alpha,\beta)$ as,

$$
\mathbf{Term}^{e}(\alpha,\beta) = d_{\alpha}\dim(\beta)\sum_{p,q,i,k}\hat{f}_v(\alpha)_{pq}\cdot\mathop{\mathbb{E}}_{\boldsymbol{s}}\left[(\hat{g}_u(\beta)\hat{h_u^{\boldsymbol{s}}}(\beta))_{ik}\right]\cdot\hat{F_{\beta}^{ki}}(\alpha)_{pq}.
$$

In addition, Lemma 15 indicates that $s(\alpha)$ and $s(\beta)$ are non-empty. Therefore,

$$
\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{s}}[f_v(\boldsymbol{x})\cdot(g*h^{\boldsymbol{s}})((\boldsymbol{x}\circ\pi)^{-1})] = \underbrace{\sum_{\substack{\alpha,\beta\\ |s(\alpha)|,|s(\beta)|\neq 0\\ |s(\beta)|<C\\ s(\alpha)\subseteq\pi(s(\beta))}}\mathbf{Term}^{e}(\alpha,\beta)}_{\Theta^{e(u,v)}_{\text{Low}}} + \underbrace{\sum_{\substack{\alpha,\beta\\ |s(\alpha)|,|s(\beta)|\neq 0\\ |s(\beta)|\geqslant C\\ s(\alpha)\subseteq\pi(s(\beta))}}\mathbf{Term}^{e}(\alpha,\beta)}_{\Theta^{e(u,v)}_{\text{High}}},
$$

where in the last expression we break the summation based on $|s(\beta)|$.

If (7) is not true, then we have

$$
\left|\mathop{\mathbb{E}}_{e(u,v)}\left[\Theta^{e(u,v)}_{\text{Low}}+\Theta^{e(u,v)}_{\text{High}}\right]\right|\geqslant\delta'.
$$

where $\delta'=\frac{\delta}{2|G|}$. We will later show that that $\left|\mathbb{E}_{e(u,v)}\left[\Theta^{e(u,v)}_{\text{High}}\right]\right|\leqslant\delta'/2$. Assuming this, we have

$$
\left|\mathop{\mathbb{E}}_{e(u,v)}\left[\Theta^{e(u,v)}_{\text{Low}}\right]\right|\geqslant\delta'/2.
$$

We now show how to come up with a decoding strategy based on the above lower bound.

Bounding the $\Theta_{\text{Low}}^{e(u,v)}$ term. Next, we argue that if $|\mathbb{E}_e[\Theta_{\text{Low}}^{e(u,v)}]|$ is large, then we can decode an assignment to the Label Cover instance $\mathcal{H}$. We first simplify the expression,

$$\begin{aligned}
&|\Theta_{\text{Low}}^{e(u,v)}|^2 \\
&= \left|\mathop{\mathbb{E}}_{\boldsymbol{s}}\left[\sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\sum_{p,q,i,k}\hat{f}_v(\alpha)_{pq}\cdot(\hat{g}_u(\beta)\hat{h}_u^{\boldsymbol{s}}(\beta))_{ik}\cdot\hat{F_\beta^{ki}}(\alpha)_{pq}\right]\right|^2 \\
&= \left|\mathop{\mathbb{E}}_{\boldsymbol{s}}\left[\sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\sum_{p,q,i,j,k}\hat{f}_v(\alpha)_{pq}\cdot\hat{g}_u(\beta)_{ij}\cdot\hat{h}_u^{\boldsymbol{s}}(\beta)_{jk}\cdot\hat{F_\beta^{ki}}(\alpha)_{pq}\right]\right|^2 \\
&\leqslant \left(\sum_{\alpha,\beta}d_\alpha d_\beta\sum_{\substack{p,q\\ i,j,k}}|\hat{f}_v(\alpha)_{pq}|^2|\hat{g}_u(\beta)_{ij}|^2\right)\left(\sum_{\alpha,\beta}d_\alpha d_\beta\sum_{\substack{p,q\\ i,j,k}}|\hat{F_\beta^{ki}}(\alpha)_{pq}|^2\mathop{\mathbb{E}}_{\boldsymbol{s}}\left[\hat{h}_u^{\boldsymbol{s}}(\beta)_{jk}\right]^2\right)
\end{aligned}$$

The second term is bounded by 1 as

$$\begin{aligned}
&\sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\sum_{\substack{p,q\\ i,j,k}}|\hat{F_\beta^{ki}}(\alpha)_{pq}|^2\mathop{\mathbb{E}}_{\boldsymbol{s}}\left[\hat{h}_u^{\boldsymbol{s}}(\beta)_{jk}\right]^2 \\
&\leqslant \mathop{\mathbb{E}}_{\boldsymbol{s}}\left[\sum_{\beta}\dim(\beta)\sum_{j,k}|\hat{h}_u^{\boldsymbol{s}}(\beta)_{jk}|^2\sum_i\sum_\alpha\dim(\alpha)\sum_{p,q}|\hat{F_\beta^{ki}}(\alpha)_{pq}|^2\right] \\
&= \mathop{\mathbb{E}}_{\boldsymbol{s}}\left[\sum_{\beta}\dim(\beta)\sum_{j,k}|\hat{h}_u^{\boldsymbol{s}}(\beta)_{jk}|^2\sum_i\|F_\beta^{ki}\|^2\right] \\
&= \mathop{\mathbb{E}}_{\boldsymbol{s}}\left[\sum_{\beta}\dim(\beta)\sum_{j,k}|\hat{h}_u^{\boldsymbol{s}}(\beta)_{jk}|^2\right] && \text{(Using Equation (10))} \\
&\leqslant \mathop{\mathbb{E}}_{\boldsymbol{s}}\left[\|h_u^{\boldsymbol{s}}\|_2^2\right] = 1,
\end{aligned}$$

Based on the above bound, the term $|\Theta_{\text{Low}}^{e(u,v)}|^2$ is upper bounded by

$$|\Theta_{\text{Low}}^{e(u,v)}|^2 \leqslant \sum_{\substack{\alpha,\beta\\ |s(\alpha)|,|s(\beta)|\neq 0\\ |s(\beta)|<C\\ s(\alpha)\subseteq\pi(s(\beta))}}\dim(\alpha)\dim(\beta)\sum_{\substack{p,q\\ i,j,k}}|\hat{f}_v(\alpha)_{pq}|^2|\hat{g}_u(\beta)_{ij}|^2$$

Since for β such that $|s(\beta)|\leqslant C$, $\dim(\beta)=\prod_{i=1}^L\dim(\beta_i)=\prod_{i,\dim(\beta_i)\geqslant 2}\dim(\beta_i)\leqslant(\sqrt{|G|})^C$ and the index i varies over the dimension of β,

$$\begin{aligned}
|\Theta_{\text{Low}}^{e(u,v)}|^2 &\leqslant |G|^{\frac{C}{2}}\sum_{\substack{\alpha,\beta\\ |s(\alpha)|,|s(\beta)|\neq 0\\ |s(\beta)|<C\\ s(\alpha)\subseteq\pi(s(\beta))}}\dim(\alpha)\dim(\beta)\sum_{\substack{p,q\\ i,j}}|\hat{f}_v(\alpha)_{pq}|^2|\hat{g}_u(\beta)_{ij}|^2 \\
&= |G|^{\frac{C}{2}}\sum_{\substack{\alpha,\beta\\ |s(\alpha)|,|s(\beta)|\neq 0\\ |s(\beta)|<C\\ s(\alpha)\subseteq\pi(s(\beta))}}\dim(\alpha)\dim(\beta)\left\|\hat{f}_v(\alpha)\right\|_{\mathsf{HS}}^2\|\hat{g}_u(\beta)\|_{\mathsf{HS}}^2.
\end{aligned}$$

Now, we can present the decoding strategy for a typical edge $e=(u,v)$:

Decoding strategy.

1. For each $u \in U$, consider a function $g_u(\boldsymbol{x}) = \chi_\rho(A_u(\boldsymbol{x}))$, sample a β with probability $\dim(\alpha)\,\|\hat{g}_u(\beta)\|_{\mathsf{HS}}^2$ and select a random coordinate j s.t. $\chi_{\beta_j} \notin \widehat{G/H_S}$. If there is no such j, then return $\perp$.
2. For each $v \in V$, consider a function $f_v(\boldsymbol{x}) = \chi_\rho(A_v(\boldsymbol{x}))$, sample an α with probability $\dim(\alpha)\,\|f_v(\alpha)\|_{\mathsf{HS}}^2$ and select a random coordinate i s.t. $\chi_{\alpha_i} \notin \widehat{G/H_S}$. If there is no such i, then return $\perp$.

For α, β such that $s(\alpha)$, $s(\beta)$ are nonempty and $s(\alpha) \subseteq \pi(s(\beta))$, the strategy will succeed with probability at least $1/|s(\beta)|$. This is because for any label ℓ returned by player v, the condition $s(\alpha) \subseteq \pi(s(\beta))$ guarantees that there exists a $\ell' \in \pi_e^{-1}(\ell)$ such that $\beta_{\ell'} \notin \widehat{G/H_S}$, and the player u returns this ℓ' with probability $1/|s(\beta)|$. Therefore, the expected value of the labeling returned by the strategy is given by

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{e(u,v)}\left[\sum_{\substack{\alpha,\beta\\ |s(\alpha)|,|s(\beta)|\neq 0\\ |s(\beta)|<C\\ s(\alpha)\subseteq\pi(s(\beta))}} d_\alpha \dim(\beta)\left\|\hat{f}_v(\alpha)\right\|_{\mathsf{HS}}^2 \|\hat{g}_u(\beta)\|_{\mathsf{HS}}^2 \frac{1}{|s(\beta)|}\right]\\
&\geqslant \frac{1}{C}\mathop{\mathbb{E}}_{e(u,v)}\left[\sum_{\substack{\alpha,\beta\\ |s(\alpha)|,|s(\beta)|\neq 0\\ |s(\beta)|<C\\ s(\alpha)\subseteq\pi(s(\beta))}} \dim(\alpha)\dim(\beta)\left\|\hat{f}_v(\alpha)\right\|_{\mathsf{HS}}^2 \|\hat{g}_u(\beta)\|_{\mathsf{HS}}^2\right]\\
&\geqslant \frac{1}{C\cdot|G|^{\frac{C}{2}}}\mathop{\mathbb{E}}_{e(u,v)}\left[|\Theta_{\mathrm{Low}}^{e(u,v)}|^2\right]\\
&\geqslant \frac{1}{C\cdot|G|^{\frac{C}{2}}}\left|\mathop{\mathbb{E}}_{e(u,v)}\left[\Theta_{\mathrm{Low}}^{e(u,v)}\right]\right|^2 \geqslant \frac{\delta'^2}{4C|G|^{C/2}}. \qquad \text{(Cauchy-Schwarz inequality)}
\end{aligned}
$$

We set $C = \Omega_G(r - 2\log\delta')$ so that $\frac{\delta'^2}{4C|G|^{C/2}} > O(2^{-r})$, which contradicts that hardness of Label Cover. Therefore, $|\mathbb{E}_{e(u,v)}[\Theta_{\mathrm{Low}}^{e(u,v)}]| \leqslant \delta'/4$ for any two layers in $\mathcal{H}$, and hence

$$
\left|\mathop{\mathbb{E}}_{1\leqslant t<t'\leqslant T}\left[\mathop{\mathbb{E}}_{\pi_{u,v}\in\Pi_{t,t'}}\left[\Theta_{\mathrm{Low}}^{e(u,v)}\right]\right]\right| \leqslant \frac{\delta'}{4}.
$$

Bounding the $\Theta_{\mathrm{High}}^{e(u,v)}$ term. It remains to bound $|\mathbb{E}_e[\Theta_{\mathrm{High}}^e]|$. We divide the term into two parts for D such that $C \geqslant \Omega_G(\log D - \log\delta')$.

$$
\Theta_{\mathrm{High}}^{e(u,v)} = \sum_{\substack{\alpha,\beta\\ |s(\alpha)|,|s(\beta)|\neq 0\\ |s(\beta)|\geqslant C\\ s(\alpha)\subseteq\pi(s(\beta))\\ \dim(\beta)\leqslant D}} \mathbf{Term}^e(\alpha,\beta) + \sum_{\substack{\alpha,\beta\\ |s(\alpha)|,|s(\beta)|\neq 0\\ |s(\beta)|\geqslant C\\ s(\alpha)\subseteq\pi(s(\beta))\\ \dim(\beta)> D}} \mathbf{Term}^e(\alpha,\beta),
$$

We denote the first term by $\Theta_{\mathrm{High},\leqslant D}^{e(u,v)}$ and the second term by $\Theta_{\mathrm{High},>D}^{e(u,v)}$. We bound these terms separately.

Bounding $\Theta^{e(u,v)}_{\text{High},\leqslant D}$. Starting with the simplified expression for $\mathbf{Term}^e(\alpha,\beta)$, we have

$$\Theta^{e(u,v)}_{\text{High},\leqslant D} = \sum_{\alpha,\beta} \dim(\alpha)\dim(\beta) \sum_{p,q,i,k} \hat{f}_v(\alpha)_{pq} \cdot \mathop{\mathbb{E}}_{\boldsymbol{s}} \left[(\hat{g}_u(\beta) \hat{h^{\boldsymbol{s}}_u}(\beta))_{ik} \right] \cdot \hat{F^{ki}_{\beta}}(\alpha)_{pq}.$$

We now simplify the expectation over $\boldsymbol{s}$. Recall that the function $h^{\boldsymbol{s}}_u(\boldsymbol{x}) := \chi_\rho(A_u(\boldsymbol{x} \bullet \boldsymbol{s})) = g_u(\boldsymbol{x} \bullet \boldsymbol{s})$. We now express the Fourier coefficient of $h^{\boldsymbol{s}}_u$ in terms of the Fourier coefficient of $g^{\boldsymbol{s}}_u$. By the Fourier inversion formula,

$$\begin{aligned} h^{\boldsymbol{s}}_u(\boldsymbol{x}) = g_u(\boldsymbol{x} \bullet \boldsymbol{s}) &= \sum_{\beta} \dim(\beta) \mathsf{tr}(\hat{g}_u(\beta)\beta(\boldsymbol{x} \bullet \boldsymbol{s})^\star) \\ &= \sum_{\beta} \dim(\beta) \mathsf{tr}(\hat{g}_u(\beta)(\beta(\boldsymbol{x})\beta(\boldsymbol{s}))^\star) && \text{(Using homomorphism of } \beta) \\ &= \sum_{\beta} \dim(\beta) \mathsf{tr}(\hat{g}_u(\beta)\beta(\boldsymbol{s})^\star \beta(\boldsymbol{x})^\star) \end{aligned}$$

As the Fourier expansion is unique, we have $\hat{h}^{\boldsymbol{s}}_u(\beta) = \hat{g}_u(\beta)\beta(\boldsymbol{s})^\star$. Using this, we have

$$\begin{aligned} \mathop{\mathbb{E}}_{\boldsymbol{s}} \left[(\hat{g}_u(\beta) \hat{h^{\boldsymbol{s}}_u}(\beta))_{ik} \right] &= \mathop{\mathbb{E}}_{\boldsymbol{s}} \left[(\hat{g}_u(\beta)\hat{g}_u(\beta)\beta(\boldsymbol{s}^{-1}))_{ik} \right] \\ &= \mathop{\mathbb{E}}_{\boldsymbol{s}} \left[\sum_{j,j'} \hat{g}_u(\beta)_{ij} \hat{g}_u(\beta)_{jj'} \beta(\boldsymbol{s}^{-1})_{j'k} \right] \\ &= \sum_{j,j'} \hat{g}_u(\beta)_{ij} \hat{g}_u(\beta)_{jj'} \mathop{\mathbb{E}}_{\boldsymbol{s}} \left[\beta(\boldsymbol{s}^{-1})_{j'k} \right] \\ &= \sum_{j,j'} \hat{g}_u(\beta)_{ij} \hat{g}_u(\beta)_{jj'} \mathop{\mathbb{E}}_{\boldsymbol{s}} \left[\prod_{\ell=1}^{L} \beta_\ell(s_\ell^{-1})_{j'_\ell,k_\ell} \right] \\ &= \sum_{j,j'} \hat{g}_u(\beta)_{ij} \hat{g}_u(\beta)_{jj'} \prod_{\ell=1}^{L} \mathop{\mathbb{E}}_{\boldsymbol{s}_\ell} \left[\beta_\ell(s_\ell^{-1})_{j'_\ell,k_\ell} \right] \end{aligned}$$

Now, we are in the setting when $\dim(\beta) \leqslant D$ but $s(\beta) \geqslant C$. Hence, the number of dimension $\geqslant 2$ representations in $(\beta_1, \beta_2, \ldots, \beta_L)$ is upper bounded by $\log_2 D$. Thus, there are at least $C - \log_2 D$ coordinates $\ell \in [L]$ such that $\beta_\ell \notin \widehat{G/H_S}$ and $\dim(\beta_\ell) = 1$. For each such β_ℓ we can apply the following claim.

▷ Claim 17. For i such that $i \in s(\beta)$, we have $\|\mathbb{E}_{s\sim S}[\beta_i(s)^\star]\|^2_{\mathsf{op}} \leqslant 1 - \delta_G$ for some $\delta_G > 0$.

Proof. We refer the reader to the full version of the paper for the proof. ◁

When $\dim(\beta_\ell) = 1$ and $\beta_\ell \notin \widehat{G/H_S}$, the above claim gives

$$\exists\, \varepsilon_G > 0, \quad |\mathop{\mathbb{E}}_{s\in S}[\beta(s^{-1})]| \leqslant 1 - \varepsilon_G.$$

Correspondingly, we have

$$\begin{aligned} \mathop{\mathbb{E}}_{\boldsymbol{s}} \left[(\hat{g}_u(\beta) \hat{h^{\boldsymbol{s}}_u}(\beta))_{ik} \right] &= \sum_{j,j'} \hat{g}_u(\beta)_{ij} \hat{g}_u(\beta)_{jj'} \prod_{\ell=1}^{L} \mathop{\mathbb{E}}_{\boldsymbol{s}_\ell} \left[\beta_\ell(s_\ell^{-1})_{j'_\ell,k_\ell} \right] \\ &\leqslant (1 - \varepsilon_G)^{(C - \log_2 D)} \sum_{j,j'} |\hat{g}_u(\beta)_{ij}| \cdot |\hat{g}_u(\beta)_{jj'}|. \end{aligned}$$

Plugging this upper bound, we get

$$\begin{aligned}\Theta^{e(u,v)}_{\mathrm{High},\leqslant D} &= \sum_{\alpha,\beta} \dim(\alpha)\dim(\beta) \sum_{p,q,i,k} \hat{f}_v(\alpha)_{pq} \cdot \mathop{\mathbb{E}}_{\boldsymbol{s}}\left[(\hat{g}_u(\beta)\hat{h^{\boldsymbol{s}}_u}(\beta))_{ik}\right] \cdot \hat{F^{ki}_\beta}(\alpha)_{pq} \\ &\leqslant (1-\varepsilon_G)^{(C-\log_2 D)} \sum_{\alpha,\beta} d_\alpha d_\beta \sum_{\substack{p,q,i,k\\ j,j'}} |\hat{f}_v(\alpha)_{pq}| \cdot |\hat{g}_u(\beta)_{ij}| \cdot |\hat{g}_u(\beta)_{jj'}| \cdot |\hat{F^{ki}_\beta}(\alpha)_{pq}|\end{aligned}$$

Applying the Cauchy-Schwarz inequality,

$$\begin{aligned}|\Theta^{e(u,v)}_{\mathrm{High},\leqslant D}|^2 \leqslant (1-\varepsilon_G)^{2(C-\log_2 D)} &\left(\sum_{\alpha,\beta} \dim(\alpha)\dim(\beta) \sum_{\substack{p,q,i,k\\ j,j'}} |\hat{f}_v(\alpha)_{pq}|^2 |\hat{g}_u(\beta)_{ij}|^2\right) \\ &\left(\sum_{\alpha,\beta} \dim(\alpha)\dim(\beta) \sum_{\substack{p,q,i,k\\ j,j'}} |\hat{F}^{ki}_\beta(\alpha)_{pq}|^2 |\hat{g}_u(\beta)_{jj'}|^2\right).\end{aligned}$$

Using the fact that i, j, j' and k vary over the dimension of β, which is at most D, the first term is at most,

$$\begin{aligned}&\left(\sum_{\alpha,\beta} \dim(\alpha)\dim(\beta) \sum_{\substack{p,q,i,k\\ j,j'}} |\hat{f}_v(\alpha)_{pq}|^2 |\hat{g}_u(\beta)_{ij}|^2\right) \\ &\leqslant D^2 \left(\sum_{\alpha} \dim(\alpha) \sum_{p,q} |\hat{f}_v(\alpha)_{pq}|^2\right)\left(\sum_{\beta} \dim(\beta) \sum_{i,j} |\hat{g}_u(\beta)_{ij}|^2\right) \leqslant D^2 \|f_v\|_2^2 \|g_u\|_2^2 \leqslant D^2.\end{aligned}$$

Similarly, the second term is

$$\begin{aligned}&\sum_{\alpha,\beta} \dim(\alpha)\dim(\beta) \sum_{\substack{p,q,i,k\\ j,j'}} |\hat{F}^{ki}_\beta(\alpha)_{pq}|^2 |\hat{g}_u(\beta)_{jj'}|^2 \\ &\qquad\leqslant D \sum_{\beta} \dim(\beta) \sum_{j,j'} |\hat{h}^{\boldsymbol{s}}_u(\beta)_{jj'}|^2 \sum_{i} \sum_{\alpha} \dim(\alpha) \sum_{p,q} |\hat{F}^{ki}_\beta(\alpha)_{pq}|^2 \\ &\qquad= D \sum_{\beta} \dim(\beta) \sum_{jj'} |\hat{g}_u(\beta)_{jj'}|^2 \sum_{i} \|F^{ki}_\beta\|^2 \\ &\qquad= D \sum_{\beta} \dim(\beta) \sum_{j,j'} |\hat{g}_u(\beta)_{jj'}|^2 \qquad \text{(Using Equation (10))} \\ &\qquad\leqslant D \cdot \|g_u\|_2^2 = D,\end{aligned}$$

Therefore, we have,

$$\left|\mathop{\mathbb{E}}_{e(u,v)}\left[\Theta^{e(u,v)}_{\mathrm{High},\leqslant D}\right]\right|^2 \leqslant \mathop{\mathbb{E}}_{e(u,v)}\left[|\Theta^{e(u,v)}_{\mathrm{High},\leqslant D}|^2\right] \leqslant (1-\varepsilon_G)^{2(C-\log_2 D)} D^3.$$

We verify that a setting of D satisfies $(1-\varepsilon_G)^{2(C-\log_2 D)} D^3 \leqslant \frac{\delta'^2}{16}$:

$$(1-\varepsilon_G)^{2(C-\log_2 D)} D^3 \leqslant \frac{\delta'^2}{16} \implies C \geqslant \Omega_G\left(\log D - \log \delta'\right).$$

Bounding $\Theta^{e(u,v)}_{\text{High},>D}$. We use the same high-degree argument as in the proof of the [4, Claim 4.5].

▷ Claim 18. For every edge $e = (u, v)$, and $C = \Omega_G(\log(1/\delta'))$, we have $\left|\mathbb{E}_{\boldsymbol{s}}\left[\Theta^{e(u,v)}_{\text{High},>D}\right]\right| \leqslant \frac{\delta'}{4}$.

Proof. We refer the reader to the full version of the paper for the proof of this claim. ◁

Finishing the proof. Using these bounds, we have,

$$\begin{aligned}\left|\mathop{\mathbb{E}}_{e(u,v)}\left[\Theta^{e(u,v)}_{\text{High}}\right]\right| &= \left|\mathop{\mathbb{E}}_{e(u,v)}\left[\Theta^{e(u,v)}_{\text{High},\leqslant D} + \Theta^{e(u,v)}_{\text{High},>D}\right]\right| \\ &\leqslant \left|\mathop{\mathbb{E}}_{e(u,v)}\left[\Theta^{e(u,v)}_{\text{High},\leqslant D}\right]\right| + \left|\mathop{\mathbb{E}}_{e(u,v)}\left[\Theta^{e(u,v)}_{\text{High},>D}\right]\right| \leqslant \frac{\delta'}{4} + \frac{\delta'}{4} = \frac{\delta'}{2} \leqslant \frac{\delta}{4|G|},\end{aligned}$$

as required. ◀

Therefore, the term (5) collectively can be upper bounded by $\frac{\delta}{2}$.

4.2.2 Bounding expressions in (6)

The analysis for representations with $\dim(\rho) \geqslant 2$ shares structural similarities with the proof of Max-3LIN over non-abelian groups in [4]. However, our setting introduces additional noise terms $\boldsymbol{s} \in S^{R_t}$ and $s \in S$ that must be carefully handled. We provide a high-level proof outline here and defer the exhaustive algebraic expansion to the full version of the paper. The core of the bound relies on the following lemma:

▶ **Lemma 19.** *Let $K := \frac{8|G|^{10}}{\delta^2}$. If the Layered Label Cover instance $\mathcal{H}$ is at most $\min\left(\frac{\delta^2}{10|G|^{10K}}, \frac{\delta^2}{2|G|^{20}}\right)$-satisfiable, then*

$$\left|\mathop{\mathbb{E}}_{1\leqslant t<t'\leqslant T}\left[\mathop{\mathbb{E}}_{\pi_{uv}\in\Pi_{t,t'}}\left[\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}}[\chi_\rho(A_v(\boldsymbol{x}) \bullet A_u(\boldsymbol{y}) \bullet A_u(\boldsymbol{z}) \bullet s^{-1})]\right]\right]\right| \leqslant \frac{\delta}{|G|^3}.$$

Proof Sketch. The analysis for representations with $\dim(\rho) \geqslant 2$ is structurally similar to the analysis in [5]. By fixing a noise vector $\boldsymbol{s}$, we can follow the approach in [5]. we introduce a threshold parameter $K := \frac{8|G|^{10}}{\delta^2}$ to divide the remaining summation into a low-degree component (Γ_{low}) and a high-degree component (Γ_{high}), based on whether the number of dimensions $\geqslant 2$ in the representation β is strictly less than K or at least K. See the full version of the paper for the actual proof. ◀

By summing the result of Lemma 19 over all representations ρ with $\dim(\rho) \geqslant 2$ and all $s \in S$, the total contribution to the expectation is at most $\frac{\delta}{2}$.

Combining this with the bound of $\frac{\delta}{2}$ on the terms for 1-dimensional representations (Equation 5), we conclude that if the Layered Label Cover instance is at most 2^{-r}-satisfiable, then value$(A) \leqslant \frac{|S|}{|H_S|} + \delta$. This completes the proof of Lemma 14. ◀

References

1 Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM (JACM)*, 45(3):501–555, 1998. doi:10.1145/278298.278306.

2 Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM (JACM)*, 45(1):70–122, 1998. doi:10.1145/273865.273901.

3 Per Austrin and Elchanan Mossel. Approximation resistant predicates from pairwise independence. *Computational Complexity*, 18(2):249–271, 2009. doi:10.1007/s00037-009-0272-6.

4 Amey Bhangale and Subhash Khot. Optimal inapproximability of satisfiable k-lin over non-abelian groups. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 1615–1628. ACM, 2021. doi:10.1145/3406325.3451003.

5 Amey Bhangale and Subhash Khot. Optimal inapproximability of satisfiable k-LIN over non-abelian groups. *(updated paper)*, 2026.

6 Amey Bhangale, Subhash Khot, and Dor Minzer. On Approximability of Satisfiable k-CSPs: I. In *54th Annual Symposium on Theory of Computing (STOC'22)*, pages 976–988, New York, NY, USA, 2022. ACM. doi:10.1145/3519935.3520028.

7 Amey Bhangale, Subhash Khot, and Dor Minzer. On Approximability of Satisfiable K-CSPs: II. In *55th Annual Symposium on Theory of Computing (STOC'23)*, pages 632–642, New York, NY, USA, 2023. ACM. doi:10.1145/3564246.3585120.

8 Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: Iv. In *56th Annual Symposium on Theory of Computing (STOC'24)*, pages 1423–1434, New York, NY, USA, 2024. ACM. doi:10.1145/3618260.3649610.

9 Amey Bhangale, Subhash Khot, and Dor Minzer. On approximability of satisfiable k-csps: V, 2025. doi:10.1145/3717823.3718127.

10 Mark Braverman, Subhash Khot, and Dor Minzer. On rich 2-to-1 games. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference, ITCS 2021, Virtual Conference, January 6-8, 2021*, LIPIcs, pages 27:1–27:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.ITCS.2021.27.

11 Andrei A. Bulatov. A dichotomy theorem for nonuniform csps. In *58th Annual Symposium on Foundations of Computer Science (FOCS'17)*, pages 319–330, Piscataway, NJ, USA, 2017. IEEE. doi:10.1109/FOCS.2017.37.

12 Siu On Chan. Approximation resistance from pairwise-independent subgroups. *Journal of the ACM (JACM)*, 63(3):1–32, 2016. doi:10.1145/2873054.

13 Arkadev Chattopadhyay and Avi Wigderson. Linear Systems over Composite Moduli . In *50th Annual Symposium on Foundations of Computer Science (FOCS'09)*, Los Alamitos, CA, USA, October 2009. IEEE. doi:10.1109/FOCS.2009.17.

14 Irit Dinur and Venkatesan Guruswami. PCPs via low-degree long code and hardness for constrained hypergraph coloring. In *54th Annual Symposium on Foundations of Computer Science (FOCS'13)*, pages 340–349, Piscataway, NJ, USA, 2013. IEEE. doi:10.1109/FOCS.2013.44.

15 Irit Dinur, Venkatesan Guruswami, Subhash Khot, and Oded Regev. A new multilayered pcp and the hardness of hypergraph vertex cover. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 595–601, New York, NY, USA, 2003. Association for Computing Machinery. doi:10.1145/780542.780629.

16 Lars Engebretsen, Jonas Holmerin, and Alexander Russell. Inapproximability results for equations over finite groups. *Theoretical Computer Science*, 312(1):17–45, 2004. doi:10.1016/S0304-3975(03)00401-8.

17 Tomás Feder and Moshe Y. Vardi. Monotone monadic SNP and constraint satisfaction. In S. Rao Kosaraju, David S. Johnson, and Alok Aggarwal, editors, *Proceedings of the Twenty-Fifth Annual ACM Symposium on Theory of Computing, May 16-18, 1993, San Diego, CA, USA*, pages 612–622. ACM, 1993. doi:10.1145/167088.167245.

18 Tomás Feder and Moshe Y Vardi. The computational structure of monotone monadic snp and constraint satisfaction: A study through datalog and group theory. *SIAM Journal on Computing*, 28(1):57–104, 1998. doi:10.1137/S0097539794266766.

19 Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM (JACM)*, 43(2):268–292, 1996. doi:10.1145/226643.226652.

20 Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995. doi:10.1145/227683.227684.

21 Mikael Goldmann and Alexander Russell. The complexity of solving equations over finite groups. *Information and Computation*, 178(1):253–262, 2002. doi:10.1006/inco.2002.3173.

22 Venkatesan Guruswami, Johan Hastad, and Madhu Sudan. Hardness of approximate hypergraph coloring. *SIAM Journal on Computing*, 31(6):1663–1686, 2002. doi:10.1137/S0097539700377165.

23 Venkatesan Guruswami, Johan Håstad, Rajsekar Manokaran, Prasad Raghavendra, and Moses Charikar. Beating the random ordering is hard: Every ordering csp is approximation resistant. *SIAM Journal on Computing*, 40(3):878–914, 2011. doi:10.1137/090756144.

24 Johan Håstad. Some optimal inapproximability results. *Journal of the ACM (JACM)*, 48(4):798–859, 2001. doi:10.1145/502090.502098.

25 Johan Håstad. Satisfying degree-d equations over $gf[2]^n$. *Theory of Computing*, 9(27):845–862, 2013. doi:10.4086/toc.2013.v009a027.

26 Johan Håstad. On the np-hardness of max-not-2. *SIAM Journal on Computing*, 43(1):179–193, 2014. doi:10.1137/120882718.

27 Subhash Khot. Hardness results for approximate hypergraph coloring. In *34th Annual Symposium on Theory of Computing (STOC'02)*, pages 351–359, New York, NY, USA, 2002. ACM. doi:10.1145/509907.509962.

28 Subhash Khot. Hardness results for coloring 3 -colorable 3 -uniform hypergraphs. In *43rd Annual Symposium on Foundations of Computer Science (FOCS'02)*, pages 23–32, Piscataway, NJ, USA, 2002. IEEE. doi:10.1109/SFCS.2002.1181879.

29 Subhash Khot. On the power of unique 2-prover 1-round games. In *34th Annual Symposium on Theory of Computing (STOC'02)*, pages 767–775, New York, NY, USA, 2002. ACM, ACM. doi:10.1145/509907.510017.

30 Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM Journal on Computing*, 37(1):319–357, 2007. doi:10.1137/S0097539705447372.

31 Subhash Khot and Muli Safra. A two-prover one-round game with strong soundness. *Theory of Computing*, 9:863–887, 2013. doi:10.4086/toc.2013.v009a028.

32 Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, USA, 2014. doi:10.1017/CBO9781139814782.

33 Prasad Raghavendra. *Approximating np-hard problems efficient algorithms and their limits*. PhD thesis, University of Washington, Seattle, WA, USA, 2009. AAI3377316. doi:10.5555/1751149.

34 Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, 1998. doi:10.1137/S0097539795280895.

35 Alex Samorodnitsky and Luca Trevisan. A pcp characterization of np with optimal amortized query complexity. In *32nd Annual Symposium on Theory of Computing (STOC'00)*, pages 191–199, New York, NY, USA, 2000. ACM. doi:10.1145/335305.335329.

36 Thomas J. Schaefer. The complexity of satisfiability problems. In *10th Annual Symposium on Theory of Computing (STOC'78)*, pages 216–226, New York, NY, USA, 1978. ACM. doi:10.1145/800133.804350.

37 Audrey Terras. *Fourier analysis on finite groups and applications*. Cambridge University Press, USA, 1999. doi:10.1017/CBO9780511626265.

38 Dmitriy Zhuk. A proof of the csp dichotomy conjecture. *Journal of the ACM (JACM)*, 67(5), August 2020. doi:10.1145/3402029.