Electronic Colloquium on Computational Complexity, Report No. 130 (2020)

Optimal Inapproximability of Satisfiable k-LIN over Non-Abelian Groups

Amey Bhangale* Subhash Khot†

Abstract

A seminal result of Håstad [Hås01] shows that it is NP-hard to find an assignment that satisfies $\frac{1}{|G|}+\varepsilon$ fraction of the constraints of a given k-LIN instance over an abelian group, even if there is an assignment that satisfies $(1-\varepsilon)$ fraction of the constraints, for any constant $\varepsilon > 0$. Engebretsen *et al.* [EHR04] later showed that the same hardness result holds for k-LIN instances over any finite non-abelian group.

Unlike the abelian case, where we can efficiently find a solution if the instance is satisfiable, in the non-abelian case, it is NP-complete to decide if a given system of linear equations is satisfiable or not, as shown by Russell and Goldmann [GR02].

Surprisingly, for certain non-abelian groups G, given a satisfiable k-LIN instance over G, one can in fact do better than just outputting a random assignment using a simple but clever algorithm. The approximation factor achieved by this algorithm varies with the underlying group. In this paper, we show that this algorithm is *optimal* by proving a tight hardness of approximation of satisfiable k-LIN instance over *any* non-abelian G, assuming $\mathbf{P} \neq \mathbf{NP}$.

As a corollary, we also get 3-query probabilistically checkable proofs with perfect completeness over large alphabets with improved soundness.

1 Introduction

Constraint satisfaction problems (CSPs) are the most fundamental problems in computer science. A simplest such CSP which we know how to solve is a system of k-LIN equations over some abelian group. More generally, an instance of Max-k-LIN over a group $G = (G, \bullet)$, not necessarily abelian, consists of a set of variables $x_1, x_2, \ldots, x_n$ and a set of constraints $C_1, C_2, \ldots, C_m$. Each C_i is a linear equation involving k variables, for example $a_1 \bullet x_{i_1} \bullet a_2 \bullet x_{i_2} \bullet \ldots a_k \bullet x_{i_k} = b$, for some group elements $a_1, a_2, \ldots, a_k, b \in G$. The task is to find an assignment to the variables that satisfies as many constraints as possible.

For *any* abelian group G, if there is a perfect solution to the given Max-k-LIN instance over G, then it can be found efficiently in polynomial time using Gaussian elimination. A given instance is *almost* satisfiable if there exists an assignment that satisfies $(1-\varepsilon)$-fraction of the constraints

*Department of Computer Science and Engineering, University of California, Riverside, CA, USA. Email: ameyrb@ucr.edu

†Department of Computer Science, Courant Institute of Mathematical Sciences, New York University, NY, USA. Email: khot@cs.nyu.edu

ISSN 1433-8092

for small constant $\varepsilon > 0$. If the given instance of Max-k-LIN over an abelian group G is almost satisfiable, then Håstad [Hås01] showed that it is **NP**-hard to even find an assignment that satisfies $\frac{1}{|G|} + \varepsilon$ of the constraints for every constant $\varepsilon > 0$. In other words, one cannot do significantly better than just outputting a random assignment.

The situation changes completely if the instance is a set of linear equations over a non-abelian group. In this case, Russell and Goldmann [GR02] showed that the problem of deciding if a given instance is satisfiable or not is **NP**-complete, for *every* non-abelian group.

An algorithm (folklore): It turns out, one can do much better than outputting a random assignment for some groups G, when the instance is satisfiable. Given an instance ϕ over G, consider an instance ϕ' over $H = G/[G,G]$ where $[G,G]$ is a commutator subgroup of G, i.e., the subgroup generated by the elements $\{g^{-1}h^{-1}gh \mid g,h \in G\}$. The instance ϕ' is same as ϕ except that all the group constants are replaced by their equivalence class in $G/[G,G]$. The important property of this quotient group H is that it is an abelian group. Since ϕ has a satisfying assignment over G, ϕ' has a satisfying assignment over H. Hence, we can find the satisfying assignment σ of ϕ' in polynomial time. The solution σ is an assignment of cosets of $[G,G]$ to the variables. We construct a random assignment to ϕ such that for every variable x, we select a random group element from $\sigma(x)$ and assign it to x. It is easy to see that each constraint is satisfied with probability equal to the $\frac{1}{|[G,G]|}$. Thus, this gives an assignment that satisfies $\frac{1}{|[G,G]|}$ fraction of the constraints in expectation. Therefore, if there exists a non-trivial commutator subgroup of G, then we get an algorithm which does better than the random assignment threshold.

If the instance is almost satisfiable, then it is not clear how to modify the above algorithm for almost satisfiable instances to get better than $\frac{1}{|G|}$ approximation. In fact, for almost satisfiable instances over any non-abelian group, [EHR04] showed that it is **NP**-hard to do better than outputting a random assignment.

This leaves an intriguing question of finding the correct approximation threshold for satisfiable instances over non-abelian groups. In this paper, we show that the above described algorithm for satisfiable instance over non-abelian groups is the best one can hope for. More specifically, we show

Theorem 1.1. *For any constant $\varepsilon > 0$, given a* satisfiable *instance of a Max-3-LIN over a finite non-abelian group G, it is* **NP***-hard to find an assignment that satisfies $\frac{1}{|[G,G]|} + \varepsilon$ fraction of the constraints.*

The theorem can be extended to Max-k-LIN for any $k \geqslant 3$ to imply a similar hardness result for all Max-k-LIN problems over G.

If G is a *simple group*, i.e. $|[G,G]| = |G|$, then Theorem 1.1 implies an NP-hardness of approximating satisfiable Max-3-CSP instances over an alphabet of size q to within a factor of $\frac{1}{q} + \varepsilon$, for every constant $\varepsilon > 0$. As a direct consequence, we get improved soundness of 3-query probabilistically checkable proofs (PCPs) with perfect completeness over large alphabets. Since PCPs are not the main focus of this paper, we refer interested readers to the book by Arora-Barak [AB09, Chapter 18] to see the relation between PCPs and CSPs.

Corollary 1.2. *For infinitely many $q \in \mathbb{Z}^+$, any language in* **NP** *is decided by a nonadaptive PCP with answers from a domain of size q that queries three positions in the proof, has perfect completeness and soundness $\frac{1}{q} + \varepsilon$ for any constant $\varepsilon > 0$.*

This improves a result of Engebretsen-Holmerin [EH05] where they constructed PCPs with soundness $\frac{1}{q} + \frac{1}{q^2} + \varepsilon$, and also a result of [Tan09] in which they showed a conditional result with soundness $\frac{1}{q} + \frac{1}{q^2} - \frac{1}{q^3} + \varepsilon$, for any constant $\varepsilon > 0$.[1]

1.1 Techniques

We assume some familiarity with the Fourier analysis of functions over abelian groups (for instance, Chapter 8 of Ryan O'Donnell's book [O'D14]). Throughout the section, $\varepsilon > 0$ is an arbitrarily small constant and $\delta(\varepsilon) > 0$ decays with ε. We only discuss 3-LIN here, however the argument is similar for k-LIN in general.

Given a well established field of hardness of approximation where the starting point is the Label Cover problem (see Definition 2.1), at the heart of these reductions are the *dictatorship tests*. A function $f : G^n \to G$ is a dictator function if it depends only on one variable, i.e., $f(x_1, x_2, \ldots, x_n) = x_i$ for some $i \in [n]$. On the other hand, we have functions which are *far* from dictator functions.

To understand a notion of distance from a dictator function, which is useful for a reduction to work, define the influence of the i^{th} coordinate on the function to be 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:

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

Thus, the i^{th} dictator function has $\mathbf{Inf}_i(f) = 1$. At the first attempt, it might make sense to define functions which are close to dictators are the functions with a coordinate with large influence. However, note that there are linear functions $\ell_S = \sum_{i \in S} \beta_i x_i$ where $S \subseteq [n]$ such that ℓ_S has all the variables $i \in S$ with influences 1. We would like to isolate these functions with large $|S|$ from the dictator functions. This motivates to define a more refined notion of low degree influence of a variable i as follows:

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

where $|\alpha|$ is the number of non-zero coordinates of α. Thus, for the i^{th} dictator function, its low degree ($d = 1$) influence of the coordinate i is 1 (and rest of the influences are 0). A function is far from any dictator function if all the low degree influences, for some $d = O(1)$ which is independent of n, of the function are small, say at most ε.

Although the above definition is the correct definition for most reductions, in case of linear equations, we work with an even weaker notion of the distance. We consider the following definition. A function is far from dictator functions if for every α such that $|\alpha| = O(1)$ which is

[1] The theorem of [EH05] holds for every $q \geqslant 3$, and the theorem of [Tan09] holds for every $q \geqslant 4$. Our theorem holds for q such that there are simple groups of cardinality q.

independent of n, $|\hat{f}(\alpha)|^2 \leqslant \varepsilon$. In other words, all the *low degree* Fourier coefficients of f have small weights. Note that this notion still isolates ℓ_S with large $|S|$.

A (non-adaptive) dictatorship test queries the function f at a few locations and based on the values it sees, decides if the function is a dictator function or far from it. This choice of predicate is tightly connected to the specific constraint satisfaction problem (CSP) for which we want to show a NP-hardness result. Furthermore, the gap between the test passing probability in the *completeness case* (when f is a dictator function) and the *soundness case* (when f is far from any dictator function) translates into the inapproximability factor of NP-hardness of the CSP.[2]

1.1.1 Abelian Groups

Let us look at a candidate dictatorship test where the predicate is a linear equation in 3 variables over an abelian group $G \cong \mathbb{Z}_q$. Here, 0 is the identity element of G.

- Select $x, y \sim G^n$ uniformly at random.
- Set $z = x + y$.
- Check if $f(x) + f(y) = f(z)$.

It is clear that if f is an i^{th} dictator then the test passes with probability 1. The non-trivial thing is to analyze the test passing probability if f is far from any dictator. It is easy to see that there are functions which are far from dictators and still the test passes with probability 1 on them. A family of such functions are the linear functions of the form $\ell_S = \sum_{i \in S} \beta_i x_i$ where $S \subseteq [n]$ and $\beta_i \in G \setminus \{0\}$, with large $|S|$. It is not hard to see that these functions pass the test with probability 1. In fact, Blum-Luby-Rubinfeld [BLR93] showed that this is a good test for the linear functions (instead of the dictator functions).

One must be able to design a test such that ℓ_S for large S passes with small probability. To design such a test, Håstad [Hås01] introduced the so called *noise* to each coordinate. The modified test is as follows:

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

This noise takes care of the earlier mentioned counterexamples, i.e., functions ℓ_S for large S now pass the test with probability roughly $\frac{1}{|G|}$. In general, Håstad [Hås01] showed that if f is far

[2] This is not totally correct as one has to overcome other important issues when such a test is used in the actual reduction, starting with the Label Cover instance.

from dictator functions then the test passes with probability at most $\frac{1}{|G|} + \delta$ for a small constant $\delta > 0$. The proof of this statement uses Fourier analysis over abelian groups. Note that this bound is optimal as even a random function passes the test with probability $\frac{1}{|G|}$.

However, now the guarantee in the completeness case is no longer the same. We only get that dictator functions pass this test with probability $1 - \varepsilon$ (instead of 1). This gap in the test passing probability is translated into the NP-hardness of 3-LIN over abelian group and coincidentally, in this abelian case, the NP-hardness result is *optimal*. More precisely, given a system of linear equations over an abelian group G, where each equation involves 3 variables, it is NP-hard to distinguish between the cases when there exists an assignment that satisfies at least $(1 - \varepsilon)$-fraction of the constraints vs. no assignment can satisfy more than $\frac{1}{|G|} + \delta$ fraction of the constraints.

1.1.2 Non-Abelian Groups

We now look into the non-abelian case. Since we would like to design a test which passes with probability 1 (or $(1 - \varepsilon)$) in the completeness case, there is a natural generalization of the above mentioned tests to a non-abelian group G. Here we denote the group operation by the symbol $\cdot$ and the identity element of G by 1_G.

We first describe the test with completeness $(1 - \varepsilon)$, which is similar to the test over abelian group with noise we described earlier.

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

The analysis of this test is implicit in the work of Engebretsen *et al.* [EHR04]. Firstly, it is easy to see that the dictator functions pass this test with probability $(1 - \varepsilon)$. The soundness of this test is analyzed by [EHR04] where they show that in the soundness case, the test passes with probability at most $\frac{1}{|G|} + \delta$ for small constant $\delta > 0$. Their proof goes via Fourier analysis over non-abelian groups. As in the abelian case, in this case also, it can be shown that the noise takes care of *high degree* Fourier terms.[3] This implies a NP-hardness result of approximating 3-LIN instances over non-abelian group, which is similar to the abelian case.

Although the proof of the soundness of the test in [EHR04] uses representation theory and Fourier analysis of functions on non-abelian groups, the proof now follows from the more general statement called the *invariance principle* of Mossell [Mos10]. The distribution on the tuple (x, y, z) is a product distribution $\mu^{\otimes n}$ where μ is a distribution on (x_i, y_i, z_i) (note that for all i it is the same distribution). Since we add noise to each coordinate with some non-zero probability, the distri-

[3] Although we have not formally defined a degree of a Fourier coefficient of a function in this non-abelian setting, think of it as the number of non-trivial irreducible representations in α (See Proposition 2.28).

bution μ is *connected* and hence we can easily take care of high degree functions in the analysis. Furthermore, the distribution μ is pairwise independent. These two conditions imply that in the soundness, the test passes with probability at most $\frac{1}{|G|} + \delta$. This statement is implicit in [AM09].

We now state the (obvious) dictatorship test with perfect completeness.

- Select $x, y \sim G^n$ uniformly at random.
- For each $i \in [n]$, set $z_i = y_i^{-1} \cdot x_i^{-1}$.
- Check if $f(x) \cdot f(y) \cdot f(z) = 1_G$.

Our main contribution is the soundness analysis of the above dictatorship test over non-abelian groups *without noise*. Note that we cannot use the invariance principle based techniques in this case, as the distribution does not satisfy the condition of *connectedness*.

Proof overview: Our proof of the soundness analysis is inspired by the magic that was discovered by Gowers [Gow08] to show that there are non-abelian groups where the size of any product free set is sublinear in $|G|$.[4] Gowers' trick worked only for quasi-random groups[5] as he was interested in $o(|G|)$ bound on the product free sets, whereas we are able to carry out our reduction for every non-abelian group.

The trick is elegantly captured by the following inequality by Babai, Nikolov, and Pyber [BNP08]. For any functions $f, g : G \to \mathbb{C}$ with at least one of f, g having mean zero:

$$\|f * g\|_{L^2(G)} \leqslant \frac{1}{\sqrt{D}} \|f\|_{L^2(G)} \|g\|_{L^2(G)}, \tag{1}$$

where D is the smallest dimension of a non-trivial representation of G.[6] In comparison, a trivial application of Cauchy-Schwartz inequality gives an upper bound of $\|f\|_{L^2(G)} \|g\|_{L^2(G)}$. Thus, Equation (1) has a multiplicative improvement of a factor $\frac{1}{\sqrt{D}}$ over a trivial upper bound.

Coming back to analyzing the soundness of our test, its analysis boils down to analyzing the following expression:

$$\begin{aligned} \mathbf{E}[g_1(x) g_2(y) g_3(z)] &= \mathbf{E}[(g_1 * g_2 * g_3)(1_{G^n})] \\ &\leqslant \sum_{\alpha \in \mathsf{Irrep}(G^n)} \dim(\alpha) \cdot \|\hat{g_1}(\alpha)\|_{\mathsf{HS}} \cdot \|\hat{g_2}(\alpha)\|_{\mathsf{HS}} \cdot \|\hat{g_3}(\alpha)\|_{\mathsf{HS}}, \end{aligned} \tag{2}$$

where g_i are bounded functions, derived from f, from $G^n \to \mathbb{C}$, i.e., $\|g_i\|_2 \leqslant 1$. $\hat{g_i}(\alpha)$ is the "fourier coefficient" of g_i corresponding to the irreducible representation α of G^n and $\|\cdot\|_{\mathsf{HS}}$ is the Hilbert–Schmidt norm of a matrix. Here, the inequality follows by using the Fourier expansion of $(g_1 * g_2 * g_3)$ and a triangle inequality.

[4] unlike the abelian case where one can always find, in this case, a 'sum-free' set of size $\Omega(|G|)$.

[5] A group G is called *quasi-random* if the smallest dimension of any non-trivial representation of G is large.

[6] see Definition 2.11 and Definition 2.16 for the definitions of $\|\cdot\|_{L^2(G)}$ and the convolution operator $*$, and Section 2.2.1 for representation theory.

Once we have this expression, similar to Equation (1), it is easy to bound the terms with large $\dim(\alpha)$: By applying Cauchy-Schwartz inequality and using Parseval's identity, we can show the following:

$$\sum_{\dim(\alpha)\geqslant D} \dim(\alpha) \cdot \|\hat{g_1}(\alpha)\|_{\mathsf{HS}} \cdot \|\hat{g_2}(\alpha)\|_{\mathsf{HS}} \cdot \|\hat{g_3}(\alpha)\|_{\mathsf{HS}} \leqslant \frac{1}{\sqrt{D}} \|g_1\|_2 \cdot \|g_2\|_2 \cdot \|g_3\|_2 \leqslant \frac{1}{\sqrt{D}}.$$

Thus, we can effectively bound the *higher dimension* terms in the expression. Therefore, if the original expectation is δ, then we essentially get

$$\sum_{\dim(\alpha)\leqslant D} \dim(\alpha) \cdot \|\hat{g_1}(\alpha)\|_{\mathsf{HS}} \cdot \|\hat{g_2}(\alpha)\|_{\mathsf{HS}} \cdot \|\hat{g_3}(\alpha)\|_{\mathsf{HS}} \approx \delta,$$

for a small D. Taking the maximum $\|\hat{g_1}(\alpha)\|_{\mathsf{HS}}$ out in the summation, the remaining sum can be upper bounded by 1. Therefore, we get that there exists an α such that the dimension of α is at most D and $\|\hat{g_1}(\alpha)\|_{\mathsf{HS}} \approx \delta - \frac{1}{\sqrt{D}}$.

Our analysis shows that if the test passes with probability greater than $\frac{1}{|[G,G]|} + \delta$, then there exists an α such that $\dim(\alpha) \leqslant O_\delta(1)$ and $\|\hat{\tilde{f}}(\alpha)\|_{\mathsf{HS}} \approx \delta$, for some function $\tilde{f}$ derived from f. Note that this conclusion is different than what we had aimed for, i.e., concluding that there exists a *low degree* Fourier coefficient with large magnitude. However, we show that such a bound is enough to carry out the actual soundness analysis of the reduction.

Since we do not introduce noise to each coordinate, there are functions with the Fourier mass concentrated on large dimensional α which pass this test with probability $\frac{1}{|[G,G]|}$. Thus, our analysis of the test is also optimal.

Although the dictatorship test works, there are many complications that arise when we compose this test with the Label Cover instance. We briefly discuss three issues here:

1. As observed before, in the soundness analysis we conclude that there is a low dimension Fourier coefficient whose norm is large, if the test passes with non-trivial probability. However, there are terms with dimension 1 but with high degree, which are problematic for the final *decoding strategy* in our reduction. In [EHR04], these problematic terms were handled by adding noise, a technique which is similar to the abelian case. In our case, we do not have the noise. However, we observe a stronger property of the folded functions.[7] Namely, if f is folded, then the function $\rho(f(x))_{ij}$, where ρ is any irreducible representation of G of dimension at least 2, has all the Fourier coefficients with dimension 1 *zero*. Thus, we can just focus on terms with dimensions at least 2.

2. Our decoding strategy is different from the one in [EHR04]. The decoding strategy in [EHR04] is based on non-empty low degree Fourier coefficients, which was similar to Håstad's decoding strategy. In our reduction, we slightly changed the decoding strategy — it is based on the Fourier coefficients whose dimension is at least 2, but can be of high degree. This condition is forced on us by the way we are handling the higher dimensions terms. Fortunately, the decoding strategy works without much trouble.

[7] A function f is called folded if $f(c \bullet x) = c \bullet f(x)$ for all $x \in G^n$ and $c \in G$

3. Because of the d-to-1 nature of the projection constraints, we have to take care of many potential scenarios in the actual reduction when the error term can be large. We handle this collectively by using a careful choice permuting the columns of the matrices (i.e., the Fourier coefficients and the representation matrices) involved in the soundness analysis.

2 Preliminaries

2.1 Label Cover

We start by defining the LABEL-COVER problem which we use as a starting point for our reduction.

Definition 2.1 (LABEL-COVER). *An instance $\mathcal{H} = (\mathcal{U}, \mathcal{V}, E, [L], [R], \{\pi_e\}_{e \in E})$ of the* LABEL-COVER *constraint satisfaction problem consists of a bi-regular bipartite graph $(\mathcal{U}, \mathcal{V}, E)$, two sets of alphabets $[L]$ and $[R]$ and a surjective projection map $\pi_e : [R] \to [L]$ for every edge $e \in E$. Given a labeling $\ell : \mathcal{U} \to [L], \ell : \mathcal{V} \to [R]$, an edge $e = (u, v)$ is said to be satisfied by ℓ if $\pi_e(\ell(v)) = \ell(u)$.*

$\mathcal{H}$ is said to be satisfiable *if there exists a labeling that satisfies all the edges. $\mathcal{H}$ 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 [AS98, ALM+98, FGL+96] and Raz's Parallel Repetition Theorem [Raz98]. The additional structural property on the hard instances (item 2 below) is proved by Håstad [Hås01, Lemma 6.9].

Theorem 2.2 (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 $\mathcal{H} = (\mathcal{U}, \mathcal{V}, E, [L], [R], \{\pi_e\}_{e \in E})$ of* LABEL-COVER *with the following properties:*

1. *$|\mathcal{U}|, |\mathcal{V}| \leqslant n^{O(r)}$; $L, R \leqslant 2^{O(r)}$; $\mathcal{H}$ is bi-regular with degrees bounded by $2^{O(r)}$.*
2. *(Smoothness) There exists a constant $d_0 \in (0, 1/3)$ such that for any $v \in \mathcal{V}$ and $\alpha \subseteq [R]$, for a random neighbor u,*
$$\mathop{\mathbf{E}}_{u} \left[|\pi_{uv}(\alpha)|^{-1} \right] \leqslant |\alpha|^{-2d_0},$$
where $\pi_{uv}(\alpha) := \{i \in [L] \mid \exists j \in \alpha \text{ s.t. } \pi_{uv}(j) = i\}$. This implies that
$$\forall v, \alpha, \qquad \mathbf{Pr}_u \left[|\pi_{uv}(\alpha)| < |\alpha|^{d_0} \right] \leqslant \frac{1}{|\alpha|^{d_0}}.$$
3. *There is a constant $s_0 \in (0, 1)$ such that,*
   - *YES Case : If the 3-SAT instance is satisfiable, then $\mathcal{H}$ is satisfiable.*
   - *NO Case : If the 3-SAT instance is unsatisfiable, then $\mathcal{H}$ is at most $2^{-s_0 r}$-satisfiable.*

2.2 Fourier analysis

In this section, we give a brief overview of the representation theory of non-abelian group and Fourier analysis over non-abelian groups. For more comprehensive understanding, we refer the reader to the book of Terras [Ter99]. We state many propositions in the following subsection, and the proofs of these propositions can be found in the same book [Ter99].

2.2.1 Representation Theory

In this paper, we only consider non-abelian groups which are *finite*. Let $G = (G, \bullet)$ be a finite non-abelian group. The identity element of a group is denoted by 1_G.

Definition 2.3. *A representation (V, ρ) of G is a vector space V together with a group homomorphism $\rho : G \to \mathsf{GL}(V)$ from G to the group $\mathsf{GL}(V)$ of invertible $\mathbb{C}$-linear transformations from V to V. The dimension of the vector space V is denoted by $\dim(\rho)$.*

For convenience, we just use the letter ρ to denote a representation of G and use ρ_V to denote the underlying vector space. We view a representation $\rho(\cdot)$ as its corresponding matrix of the linear transformation. Thus $\rho(\cdot)_{ij}$ is used to denote the $(i,j)^{th}$ entry of that matrix. We always work with representations which are unitary. There is one representation which is obvious – just map everything to $1 \in \mathbb{C}$. This representation is called the *trivial representation* which has dimension 1. We will denote the trivial representation by $\{\mathbf{1}\}$.

Definition 2.4. *Let ρ and τ be representations of G. An isomorphism from ρ_V to τ_V is an invertible linear transformation $\phi : \rho_V \to \tau_V$ such that*

$$\phi \circ \rho(g) = \tau(g) \circ \phi,$$

for all $g \in G$. We say that ρ_V and τ_V are isomorphic and write $\rho_V \cong \tau_V$ if there exists an isomorphism from ρ_V to τ_V.

Definition 2.5. *Let ρ be a representation of G. A vector subspace $W \subset \rho_V$ is G-invariant if $\rho(g)w \in W$ for all $g \in G$ and $w \in W$.*

If a representation (V, ρ) has a G-invariant subspace W other than $\{0\}$ and V itself, then the action on W itself is a representation of G. This leads to the following important definition of irreducible representations.

Definition 2.6. *A representation ρ of G is irreducible if $\rho_V \neq \emptyset$ and ρ_V has no G-invariant subspaces other than $\{0\}$ and ρ_V.*

We will denote the set of all irreducible representations of G up to isomorphism by $\mathsf{Irrep}(G)$.

Fact 2.7. *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.*

Definition 2.8. *The tensor product of two representations ρ and τ of a group G is the representation $\rho \otimes \tau$ on $\rho_V \otimes \tau_V$ defined by the condition*

$$(\rho \otimes \tau)(g)(v \otimes w) = \rho(g)(v) \otimes \tau(g)(w),$$

and extended to all vectors in $\rho_V \otimes \tau_V$ by linearity.

Definition 2.9. *The direct sum of two representations ρ and τ is the space $\rho_V \oplus \tau_V$ with the block-diagonal action $\rho \oplus \tau$ of G.*

If the representation in not irreducible, then by an appropriate change of basis ρ can be converted into a block diagonal matrix with blocks corresponding to the invariant subspaces. Thus, any representation can be completely decomposed into a direct sum of irreducible representations of G, by applying an appropriate unitary transformation. Note that this decomposition is *unique*. We use the following notation to denote the decomposition of a reducible representation: If ρ is a reducible representation of G then $\rho \cong \oplus_i n_i \rho_i$, where each i we have *distinct* $\rho_i \in \mathsf{Irrep}(G)$ and n_i denotes the multiplicity of ρ_i in the decomposition. It will be convenient to think of this representation as a block diagonal matrices with ρ_i as the blocks along the diagonal with multiplicity n_i.

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

$$\langle (\rho)_{ij} \mid (\tau)_{k\ell} \rangle_G = 0, \tag{3}$$

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,

$$\langle (\rho)_{ij} \mid (\rho)_{k\ell} \rangle_G = \frac{\delta_{i\ell}\delta_{jk}}{\dim(\rho)}, \tag{4}$$

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

2.2.2 Fourier analysis on non-abelian group

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 $\mathbb{C}$.

Definition 2.11. *Define the inner product $\langle \cdot, \cdot \rangle_{L^2(G)}$ on $L^2(G)$ by*

$$\langle f, g \rangle_{L^2(G)} = \mathop{\mathbf{E}}_{x \in G} [f(x)\overline{g(x)}].$$

We can define a character for every representation of a group.

Definition 2.12. *The character of a representation ρ is the function $\chi_\rho : G \to \mathbb{C}$ defined by $\chi_\rho(g) = \mathsf{tr}(\rho(g))$.*

The following proposition shows that the characters corresponding to the irreducible representations of a group are orthogonal to each other.

Proposition 2.13 (Orthogonality of characters)**.** *For $\rho, \tau \in \mathsf{Irrep}(G)$, we have*

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

We use Proposition 2.10 many times in the proof. For convenience, we note an important identity that follows from Proposition 2.10 (by setting τ to be the trivial map $\{\mathbf{1}\}$).

Proposition 2.14. *If $\rho \in \mathsf{Irrep}(G)\backslash\{\mathbf{1}\}$, $\sum_{g\in G} \rho(g) = 0$.*

We have a following proposition. It also shows that the maximum dimension of any irreducible representation of G is at most $\sqrt{G}$.

Proposition 2.15.

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

This implies the following:,

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

Definition 2.16. *For two functions $f, g \in L^2(G)$ their* convolution *$f * g \in L^2(G)$ is defined as*

$$(f * g)(x) := \mathop{\mathbf{E}}_{y\in G}[f(y)g(y^{-1}x)].$$

For an abelian group, any function $f : G \to \mathbb{C}$ can be written as linear combinations of characters, i.e., the characters span the whole space $L^2(G)$. However, for non-abelian groups, characters form an orthonormal basis only for the set of *class functions* – maps which are constant on *conjugacy classes*. A conjugacy class in G is a nonempty subset H of G such that the following two conditions hold: Given any $x, y \in H$, there exists $g \in G$ such that $gxg^{-1} = y$, and if $x \in H$ and $g \in G$ then $gxg^{-1} \in H$. Since this is an equivalence class, any group is a collection of disjoint conjugacy classes.

As in the abelian case, we can understand operations like inner product, convolution etc., using the Fourier transform which is defined as follows:

Definition 2.17. *For a function $f \in L^2(G)$, define the Fourier transform of f to be the element $\hat{f} \in \prod_{\rho\in\mathsf{Irrep}(G)} \mathsf{End}\,\rho_V$ given by*

$$\hat{f}(\rho) = \mathop{\mathbf{E}}_{x\in G}[f(x)\rho(x)] \in \mathsf{End}\,\rho_V.$$

Definition 2.18. *Let V be a finite-dimensional complex inner product space. Define an inner product $\langle\cdot,\cdot\rangle_{\mathsf{End}\,V}$ on $\mathsf{End}\,V$ by*

$$\langle A, B\rangle_{\mathsf{End}\,V} = \mathsf{tr}(AB^\star).$$

We can now state the Fourier inversion theorem.

Proposition 2.19 (Fourier inversion theorem)**.** *For $f \in L^2(G)$ we have*

$$f(x) = \sum_{\rho\in\mathsf{Irrep}(G)} \dim(\rho) \cdot \langle\hat{f}(\rho), \rho(x)\rangle_{\mathsf{End}\,\rho_V}.$$

We have the following simple identities (See [Ter99] for the proofs).

Proposition 2.20 (Plancherel's identity)**.**

$$\langle f, g\rangle_{L^2(G)} = \sum_{\rho\in \mathsf{Irrep}(G)} \dim(\rho)\cdot \langle \hat{f}(\rho), \hat{g}(\rho)\rangle_{\mathsf{End}\,\rho_V}.$$

Proposition 2.21 (Parseval's identity)**.**

$$\mathop{\mathbf{E}}_{x\in G}[|f(x)|^2] = \sum_{\rho\in \mathsf{Irrep}(G)} \dim(\rho)\cdot \|\hat{f}(\rho)\|_{\mathsf{HS}}^2,$$

where $\|A\|_{\mathsf{HS}} := \sqrt{\langle A, A\rangle_{\mathsf{End}\,V}} = \sqrt{\mathsf{tr}(AA^\star)} = \sqrt{\sum_{ij}|A_{ij}|^2}$.

Note that the norm $\|\cdot\|_{\mathsf{HS}}$ satisfies a triangle inequality.

Claim 2.22. $\|AB\|_{\mathsf{HS}} \leqslant \|A\|_{\mathsf{HS}}\cdot\|B\|_{\mathsf{HS}}$.

Proof. $\|AB\|_{\mathsf{HS}}^2 = \sum_{ij}|(AB)_{ij}|^2 \leqslant \sum_{ij}\left(\sum_k |A_{ik}B_{kj}|\right)^2$. Using the Cauchy-Schwartz inequality on the inner sum,

$$\|AB\|_{\mathsf{HS}}^2 \leqslant \sum_{ij}\left(\sum_k |A_{ik}|^2\right)\left(\sum_\ell |B_{\ell j}|^2\right) = \sum_{ijk\ell}|A_{ik}|^2|B_{\ell j}|^2 = \left(\sum_{ik}|A_{ik}|^2\right)\left(\sum_{\ell j}|B_{\ell j}|^2\right) = \|A\|_{\mathsf{HS}}^2\cdot\|B\|_{\mathsf{HS}}^2.$$

□

Claim 2.23. *Let A be any matrix and U be any unitary matrix, then* $\|UA\|_{\mathsf{HS}} = \|A\|_{\mathsf{HS}}$.

Proof. Let V be an unitary matrix which converts U to the identity matrix, i.e., $VUV^\star = I$. Since the change of basis does not change the $\|\cdot\|_{\mathsf{HS}}$, we have

$$\|UA\|_{\mathsf{HS}} = \|VUAV^\star\|_{\mathsf{HS}} = \|VUV^\star VAV^\star\|_{\mathsf{HS}} = \|IVAV^\star\|_{\mathsf{HS}} = \|A\|_{\mathsf{HS}}.$$

□

Proposition 2.24 (Convolution theorem)**.** *For $f, g \in L^2(G)$ we have*

$$\hat{f \ast g}(\rho) = \hat{f}(\rho)\hat{g}(\rho).$$

2.3 Important claims

In this section, we prove a few statements that will be used in the soundness analysis. The following claim shows that the character functions always come in 'pairs' with respect to the complex conjugation.

Claim 2.25. *Let G be any non abelian group. For every $\rho \in \mathsf{Irrep}(G)$, such that* $\dim(\rho) = 1$*, there exists $\widetilde{\rho} \in \mathsf{Irrep}(G)$ with* $\dim(\widetilde{\rho}) = 1$ *such that*

$$\chi_\rho(g) = \overline{\chi_{\widetilde{\rho}}(g)}, \qquad \forall g \in G.$$

Proof. We claim that the set of characters corresponding to dimension 1 irreducible representations of G forms a group under point-wise multiplication. This will be enough to show the claim.

Let $G' = G/[G,G]$ be the abelian quotient group. Assume ρ is a degree 1 representation of G. Then it satisfies $\rho(a)\rho(b) = \rho(ab)$ for all $a, b \in G$. Define a map $\Gamma_\rho : G' \to \mathbb{C}$ as $\Gamma_\rho(g') = \rho(g)$ where $g' = g[G,G]$. This is a well defined map as

$$\rho(aba^{-1}b^{-1}) = \rho(a)\rho(b)\rho(a^{-1})\rho(b^{-1}) = \rho(a)\rho(a^{-1})\rho(b)\rho(b^{-1}) = 1.$$

Thus, the map ρ is constant on every coset of $[G,G]$ and hence Γ_ρ is well defined. The set of all $\{\Gamma_\rho \mid \rho \in \mathsf{Irrep}(G), \dim(\rho) = 1\}$ is the set of all the multiplicative characters of the abelian group G' and hence form a group under coordinate-wise multiplication. There is a one-to-one correspondence between the coordinate wise multiplicative action of Γ_ρ's and ρ's. Thus, $\{\chi_\rho \mid \rho \in \mathsf{Irrep}(G), \dim(\rho) = 1\}$ form a group under point-wise multiplication. □

The following lemma shows that the direct sum decomposition of tensors of large dimension irreducible representations cannot contain overwhelming copies of a single dimension 1 representation.

Lemma 2.26. *Let $\rho = \otimes_{k=1}^{t} \rho_{i_k}$ be a representation of G where each $\rho_{i_k} \in \mathsf{Irrep}(G)$ and $\dim(\rho_{i_k}) \geqslant 2$ for all $k \in [t]$. Suppose following is the decomposition of ρ into its irreducible components*

$$\otimes_{k=1}^{t} \rho_{i_k} \cong \oplus_{\ell=1}^{r} n_{j_\ell} \rho_{j_\ell},$$

where ρ_{j_ℓ} and $\rho_{j_{\ell'}}$ are distinct for every $\ell \neq \ell'$. Then for all $\ell \in [r]$, $n_{j_\ell} \leqslant \left(1 - \frac{1}{|G|}\right) \dim(\rho)$.

Proof. As $\sum_{\ell=1}^{r} n_{j_\ell} \dim(\rho_{j_\ell}) = \dim(\rho)$, the claim is trivially true for ℓ such that $\dim(\rho_{j_\ell}) \geqslant 2$. Thus, we will show the conclusion for ℓ such that $\dim(\rho_{j_\ell}) = 1$. We first prove the lemma when $t = 2$ and then prove it for arbitrary t. Let $\rho = \rho_1 \otimes \rho_2$. The only way the conclusion cannot be true for this ρ is when $\rho \cong \tau \cdot I$ where I is a $\dim(\rho)$ sized identity matrix and $\dim(\tau) = 1$ (i.e, all the irreducible components are the same and are of dimension 1). This is because, $\dim(\rho_i)$ is always upper bounded by $\sqrt{G} - 1$ (Proposition 2.15). Thus, $\dim(\rho) < |G|$ and hence if the conclusion is not true for τ then $\lceil \left(1 - \frac{1}{|G|}\right) \dim(\rho) \rceil = \dim(\rho)$. We now show that $\rho \cong \tau \cdot I$ cannot happen. Since τ is a scalar,

$$\rho \cong \tau \cdot I \implies (\rho_1 \otimes (\tau\rho_2)) \cong I.$$

Now, both ρ_1 and $(\tau\rho_2)$ are irreducible representations of G. Since, the eigenvalues of a tensor are the pairwise product of eigenvalues of individual matrices, only way $(\rho_1 \otimes (\tau\rho_2)) \cong I$ can happen is if there exists ω, with $|\omega| = 1$, such that all the eigenvalues of $\rho_1(g)$ are ω for all $g \in G$ as well as that of $(\tau\rho_2)(g)$ are $\overline{\omega}$ for all $g \in G$. This means $\chi_{\rho_1}(g) = \dim(\rho_1) \cdot \omega$ for all $g \in G$ as the trace of a matrix is equal to sum of the eigenvalues of the matrix. This contradicts Proposition 2.13, i.e., $\sum_{g \in G} \chi_{\rho_1}(g) = |G| \dim(\rho_1) \cdot \omega \neq 0$.

Now consider $\rho = \otimes_{k=1}^{m+1} \rho_{i_k} = \otimes_{k=1}^{m} \rho_{i_k} \otimes \rho_{i_{m+1}}$, where $m \geqslant 2$. We have,

$$\begin{aligned} \rho &= \otimes_{k=1}^{m+1} \rho_{i_k} \\ &= \otimes_{k=1}^{m} \rho_{i_k} \otimes \rho_{i_{m+1}} \end{aligned}$$

$$\begin{aligned}
&\cong (\oplus_{\ell=1}^{r'} n_{j_\ell}\rho_{j_\ell}) \otimes \rho_{i_{m+1}} \\
&= \oplus_{\ell=1}^{r'} n_{j_\ell}(\rho_{j_\ell} \otimes \rho_{i_{m+1}}) \\
&\cong \oplus_{\ell=1}^{r'} n_{j_\ell} \left(\oplus_{\ell'=1}^{r''} n_{\ell'}^{\ell}\rho_{j_{\ell'}^{\ell}}\right).
\end{aligned}$$

Using the $t = 2$ case, we have $n_{\ell'}^{\ell} \leqslant \left(1 - \frac{1}{|G|}\right) \dim(\rho_{j_\ell}) \dim(\rho_{i_{m+1}})$. We also know that for two different indices $\ell_1' \neq \ell_2'$, $\rho_{j_{\ell_1'}^{\ell}} \neq \rho_{j_{\ell_2'}^{\ell}}$ by definition. Consider any representation τ of dimension 1. Let $(\ell, \ell') = (\ell, \ell_\tau')$ be the unique index in the inner direct sum where it appears (it might not appear at all in which case we treat $n_{\ell_\tau'}^{\ell} = 0$). The total count of the occurrences of τ in the direct sum is upper bounded by

$$\begin{aligned}
\sum_{\ell=1}^{r} n_{j_\ell} \cdot n_{\ell_\tau'}^{\ell} &\leqslant \sum_{\ell=1}^{r} n_{j_\ell} \cdot \left(1 - \frac{1}{|G|}\right) \dim(\rho_{j_\ell}) \dim(\rho_{i_{m+1}}) \\
&= \left(1 - \frac{1}{|G|}\right) \sum_{\ell=1}^{r} n_{j_\ell} \cdot \dim(\rho_{j_\ell}) \dim(\rho_{i_{m+1}}) \\
&= \left(1 - \frac{1}{|G|}\right) \dim(\rho).
\end{aligned}$$

□

We have a following corollary that follows from the previous lemma.

Corollary 2.27. *Let $\rho = \otimes_{k=1}^{t} \rho_{i_k}$ be a representation of G where each $\rho_{i_k} \in \mathsf{Irrep}(G)$ for all $k \in [t]$, and $\dim(\rho) \geqslant 2$. Suppose following is the decomposition of ρ into its irreducible components*

$$\otimes_{k=1}^{t} \rho_{i_k} \cong \oplus_{\ell=1}^{r} n_{j_\ell} \rho_{j_\ell},$$

where ρ_{j_ℓ} and $\rho_{j_{\ell'}}$ are distinct for every $\ell \neq \ell'$. Then for all $\ell \in [r]$, $n_{j_\ell} \leqslant \left(1 - \frac{1}{|G|}\right) \dim(\rho)$.

Proof. Assume without loss of generality that the first t' terms are all the dimension 1 representations in the tensor product ρ. Now, the (tensor) product of dimension 1 representations is also a dimension 1 representation of G. Suppose $\tau = (\otimes_{k=1}^{t'} \rho_{i_k})$ where $\dim(\tau) = 1$. We can write ρ as:

$$\rho = \otimes_{k=1}^{t} \rho_{i_k} = (\otimes_{k=1}^{t'} \rho_{i_k}) \otimes \rho_{i_{t'+1}} \otimes (\otimes_{k=t'+2}^{t'} \rho_{i_k}) = (\tau \rho_{i_{t'+1}}) \otimes (\otimes_{k=t'+2}^{t'} \rho_{i_k}).$$

Now, $\tau\rho_{i_{t'+1}}$ itself is a irreducible representation of G of dimension at least 2. Therefore, the conclusion follows from Lemma 2.26. □

2.4 Functions on G^n

For any non-abelian group G and $n \geqslant 1$, we have a group G^n where the the group operation is defined coordinate wise. The irreducible representations of G^n are precisely those representations obtained by taking tensor products of n irreducible representations of G.

Proposition 2.28 ([Ter99]). *The set of irreducible representations of G^n is given by*

$$\mathsf{Irrep}(G^n) = \{\alpha \mid \alpha = \otimes_{i\in[n]}\rho_i \text{ where } \rho_i \in \mathsf{Irrep}(G)\}.$$

We denote α by the corresponding tuple $(\rho_1, \rho_2, \ldots, \rho_n)$. We define the weight of a representation $\alpha = (\rho_1, \rho_2, \ldots, \rho_n)$ (denoted by $|\alpha|$) to be the number of non-trivial representations in $(\rho_1, \rho_2, \ldots, \rho_n)$.

We will be working with functions $f : G^n \to G$ which are *folded*. f is said to be folded if $f(c\boldsymbol{x}) = cf(\boldsymbol{x})$ for all $c \in G$ and $\boldsymbol{x} \in G^n$. The following claim shows that for all functions $g(\boldsymbol{x}) := \rho(f(\boldsymbol{x}))_{ij}$ where $\dim(\rho) \geqslant 2$ and $1 \leqslant i, j \leqslant \dim(\rho)$, all the Fourier coefficients corresponding to representations of dimension 1 are zero, if f is folded.

Lemma 2.29. *Let $f : G^n \to G$ be any folded function and $g(\boldsymbol{x}) := \rho(f(\boldsymbol{x}))_{ij}$ where $\rho \in \mathsf{Irrep}(G), \dim(\rho) \geqslant 2$ and $1 \leqslant i, j \leqslant \dim(\rho)$. Let α be* any *representation of G^n such that $\dim(\alpha) = 1$, then $\hat{g}(\alpha) = 0$.*

Proof. Recall, for any $\boldsymbol{x} \in G^n$, $\alpha(\boldsymbol{x})$ is a scalar as $\dim(\alpha) = 1$. f is folded which means that $f(c\boldsymbol{x}) = cf(\boldsymbol{x})$ for all $c \in G$ and $\boldsymbol{x} \in G^n$. Since $\rho(.)$ has dimension at least 2, in the following analysis, we use $[\rho(.)]$ to denote that matrix of linear transformation for clarity.

$$\begin{aligned}
\hat{g}(\alpha) = \mathop{\mathbf{E}}_{\boldsymbol{x}}[g(\boldsymbol{x})\alpha(\boldsymbol{x})] &= \mathop{\mathbf{E}}_{\boldsymbol{x}}[[\rho(f(\boldsymbol{x}))]_{ij} \cdot \alpha(\boldsymbol{x})] \\
&= \frac{1}{|G|}\mathop{\mathbf{E}}_{\boldsymbol{x}}\left[\sum_{c\in G}[\rho(f(c\boldsymbol{x}))]_{ij} \cdot \alpha(c\boldsymbol{x})\right] \\
&= \frac{1}{|G|}\mathop{\mathbf{E}}_{\boldsymbol{x}}\left[\sum_{c\in G}[\rho(cf(\boldsymbol{x}))]_{ij} \cdot \alpha(c)\alpha(\boldsymbol{x})\right] \\
&= \frac{1}{|G|}\mathop{\mathbf{E}}_{\boldsymbol{x}}\left[\sum_{c\in G}(\alpha(c)[\rho(c)] \cdot [\rho(f(\boldsymbol{x}))])_{ij}\,\alpha(\boldsymbol{x})\right] \\
&= \frac{1}{|G|}\mathop{\mathbf{E}}_{\boldsymbol{x}}\left[\left(\left(\sum_{c\in G}\alpha(c)[\rho(c)]\right) \cdot [\rho(f(\boldsymbol{x}))]\right)_{ij}\alpha(\boldsymbol{x})\right].
\end{aligned}$$

Now, for $\alpha \in \mathsf{Irrep}(G^n)$, let $\widetilde{\alpha} \in \mathsf{Irrep}(G^n)$ be the dimension 1 representation satisfying the condition in Claim 2.25. We have:

$$\begin{aligned}
\sum_{c\in G}\alpha(c)[\rho(c)] &= \sum_{c\in G}\overline{\widetilde{\alpha}(c)} \cdot [\rho(c)] \\
&= \sum_{c\in G}\widetilde{\alpha}(c^{-1}) \cdot [\rho(c)] \\
&= \sum_{c\in G}(\otimes_{i=1}^{n}\widetilde{\alpha}_i(c^{-1})) \cdot [\rho(c)] \\
&= \sum_{c\in G}\tau(c^{-1}) \cdot [\rho(c)] && \dim(\tau) = 1 \\
&= 0, && \text{(Using Proposition 2.14)}
\end{aligned}$$

where in the second last step, we used the fact that the product of dimension 1 representations $(\otimes_{i=1}^{n}\widetilde{\alpha}_i)$ og G is itself a dimension 1 representation (τ) of G. Therefore, $\hat{g}(\alpha) = 0$.

□

Fix any surjective projection map $\pi : [R] \to [L]$ for some $R \geqslant L$. Consider the following subgroup of G^R given by the elements

$$\{(x \circ \pi) \in G^R \mid x \in G^L\},$$

where $(x \circ \pi)_i = x_{\pi(i)}$. Let us denote this group by $\pi(G^R)$. Note that this group is isomorphic to G^L. Thus, any representation $\alpha \in \mathsf{Irrep}(G^R)$ (which is a representation of G^L using Fact 2.7), can be decomposed into irreducible representations of G^L.

The following lemma says that if α satisfies certain property, then for each irreducible representation occurring in the decomposition, either its dimension is large or its multiplicity is small.

Lemma 2.30. *Let $\pi : [R] \to [L]$ be any surjective projection map. Let $\varepsilon_0 \in (0, \frac{1}{2}]$ and $c \geqslant 10|G| \log(\frac{1}{\varepsilon_0})$. Suppose $\alpha \in \mathsf{Irrep}(G^R)$,*

$$\alpha = \bigotimes_{i=1}^{R} \rho_i = \bigotimes_{\ell=1}^{L} \underbrace{\left(\bigotimes_{j \in \pi_{uv}^{-1}(\ell)} \rho_j \right)}_{=:B_\ell}$$

such that number of ℓ with $\dim(B_\ell) \geqslant 2$ is at least c. If $\alpha \cong \oplus_m n_m \beta_m$ be the decomposition of α into irreducible representations of $\pi(G^L) \cong G^L$, then for every m either $\dim(\beta_m) \geqslant c$ or $n_m \leqslant \varepsilon_0^2 \cdot \dim(\alpha)$.

Proof. We can decompose α as follows:

$$\alpha = \bigotimes_{i=1}^{R} \rho_i = \bigotimes_{\ell=1}^{L} \underbrace{\left(\bigotimes_{j \in \pi_{uv}^{-1}(\ell)} \rho_j \right)}_{B_\ell} \cong \bigotimes_{\ell=1}^{L} \left(\oplus_{k=1}^{t_\ell} n_k^\ell \rho_k^\ell \right) = \oplus_m n_m \beta_m,$$

where for every ℓ and k, $\rho_k^\ell \in \mathsf{Irrep}(G)$. Let $d_\ell = \dim(B_\ell)$. By assumption, there are at least c coordinates ℓ such that $d_\ell \geqslant 2$. Let us denote this subset by $S \subseteq [L]$. Fix any $\beta_m = (\rho_{k_1}^1, \rho_{k_2}^2, \ldots, \rho_{k_L}^L)$ in the direct sum, such that $\dim(\beta_m) \leqslant c$. Then we have,

$$n_m = \prod_{\ell=1}^{L} n_{k_\ell}^\ell.$$

As the dimension of β_m is at most c, it must be the case that for at least $c - \log c$ many $\ell \in S$, $\dim(\rho_{k_\ell}^\ell) = 1$. Let us denote these coordinates by $S' \subseteq S$. Therefore, using Corollary 2.27,

$$n_m = \prod_{\ell \in S'} n_{k_\ell}^\ell \prod_{\ell \notin S'} n_{k_\ell}^\ell \leqslant \prod_{\ell \in S'} \left(1 - \frac{1}{|G|}\right) d_\ell \prod_{\ell \notin S'} d_\ell \leqslant \left(1 - \frac{1}{|G|}\right)^{c - \log c} \prod_{\ell=1}^{L} d_\ell.$$

Since $\prod_{\ell=1}^{L} d_\ell = \dim(\alpha)$, we have

$$\frac{n_m}{\dim(\alpha)} \leqslant \left(1 - \frac{1}{|G|}\right)^{c - \log c} \leqslant e^{-\frac{c - \log c}{|G|}} \leqslant e^{-\frac{c}{2|G|}} \leqslant \varepsilon_0^2,$$

where we used the fact that $\frac{c}{2} \geqslant \log c$. □

2.5 Notations

Whenever possible, we use the notation α, β to denote the representations of group G^n and ρ, τ for group G. Also, we use bold letters $\boldsymbol{x}, \boldsymbol{c}$ to denote the elements of G^n.

For a representation $\alpha \in \mathsf{Irrep}(G^n)$ where $\alpha = \otimes_{i=1}^{n} \rho_i$, we use the notation $\dim_{\geqslant k}(\alpha)$ to denote the number of $i \in [n]$ such that $\dim(\rho_i) \geqslant k$.

3 Warm-up: Dictatorship Test

In this section, we analyze the dictatorship test where the test involves checking some linear equation over a non-abelian group. The analysis will highlight a few important differences between our test and the linearity test over abelian groups.

Fix a non-abelian group G. Let $f : G^n \to G$ be a function. A function is called a dictator function if it is for the form $f(\boldsymbol{x}) = x_i$ for some $i \in [n]$. We use $\bullet$ to denote the group operation. Consider the following 3-query dictatorship test for f:

1. Sample $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ from G^n uniformly at random.
2. Sample $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ from G^n uniformly at random.
3. Calculate $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)$ such that $c_i = b_i^{-1} a_i^{-1}$.
4. Check if $f(\boldsymbol{a}) \bullet f(\boldsymbol{b}) \bullet f(\boldsymbol{c}) = 1_G$.

Completeness is trivial: If f is an i^{th} dictator function, i.e., $f(x_1, x_2, \ldots, x_n) = x_i$, then the test passes with probability 1. This is because we are essentially checking if $a_i \bullet b_i \bullet c_i = 1_G$, which is always true by the definition of c_i.

We analyze the soundness of the test. The following lemma says that if the test passes with some non-trivial probability then it must be the case that f (or a minor variation of f) has a low dimension Fourier coefficient whose Hilbert-Schmidt norm is large. The actual conclusion is somewhat stronger than this. In the next section, we will show that such a conclusion can be used to analyze the soundness of the final reduction (which is also presented in next section).

Lemma 3.1. *Assume f is folded. For all $\varepsilon > 0$ and $\delta > 0$, if f passes the test with probability $\frac{1}{|[G,G]|} + \varepsilon$, then there exist $\rho \in \mathsf{Irrep}(G)$ and $1 \leqslant i, j \leqslant \dim(\rho)$ such that for $h(\boldsymbol{x}) := \rho(f(\boldsymbol{x}))_{ij}$,*

$$\max_{\substack{\alpha, \\ \dim(\alpha) \geqslant 2, \\ \dim_{\geqslant 2}(\alpha) < \frac{1}{2\delta^2}.}} \|\hat{h}(\alpha)\|_{\mathsf{HS}} \geqslant \frac{\varepsilon}{|G|} - \delta.$$

Proof. Using Proposition 2.15, the probability that the test passes can be expressed as follows:

$$\mathbf{Pr}[\text{Test passes}] = \frac{1}{|G|} \sum_{\rho \in \mathsf{Irrep}(G)} \dim(\rho) \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f(\boldsymbol{a}) \bullet f(\boldsymbol{b}) \bullet f(\boldsymbol{c}))]$$

$$= \frac{1}{|G|} \sum_{\substack{\rho \in \mathsf{Irrep}(G), \\ \dim(\rho)=1}} \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f(\boldsymbol{a}) \bullet f(\boldsymbol{b}) \bullet f(\boldsymbol{c}))]$$
$$+ \frac{1}{|G|} \sum_{\substack{\rho \in \mathsf{Irrep}(G), \\ \dim(\rho)\geqslant 2}} \dim(\rho) \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f(\boldsymbol{a}) \bullet f(\boldsymbol{b}) \bullet f(\boldsymbol{c}))].$$

In the first summation, for any $\rho \in \mathsf{Irrep}(G)$ such that $\dim(\rho) = 1$, using the multiplicativity of the characters, we have

$$\begin{aligned}
\chi_\rho(f(\boldsymbol{a}) \bullet f(\boldsymbol{b}) \bullet f(\boldsymbol{c})) &= \chi_\rho(f(\boldsymbol{a}))\chi_\rho(f(\boldsymbol{b}))\chi_\rho(f(\boldsymbol{c})) \\
&\leqslant \left|\chi_\rho(f(\boldsymbol{a}))\right| \cdot \left|\chi_\rho(f(\boldsymbol{b}))\right| \cdot \left|\chi_\rho(f(\boldsymbol{c}))\right| \\
&= 1. && \text{(unitary representations)}
\end{aligned}$$

As the number of dimension 1 representations of a group G is equal to the size of the quotient $G/[G,G]$, we get

$$\mathbf{Pr}[\text{Test passes}] = \frac{1}{|[G,G]|} + \frac{1}{|G|} \sum_{\substack{\rho \in \mathsf{Irrep}(G), \\ \dim(\rho)\geqslant 2}} \dim(\rho) \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f(\boldsymbol{a}) \bullet f(\boldsymbol{b}) \bullet f(\boldsymbol{c}))]. \tag{5}$$

Now, fix any $\rho \in \mathsf{Irrep}(G)$ such that $\dim(\rho) \geqslant 2$. For $1 \leqslant i, j \in \dim(\rho)$, let $g_{ij} : G^n \to \mathbb{C}$ be defined as $g_{ij}(\boldsymbol{x}) := \rho(f(\boldsymbol{x}))_{ij}$. Using the definition of characters, we have

$$\begin{aligned}
\mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f(\boldsymbol{a}) \bullet f(\boldsymbol{b}) \bullet f(\boldsymbol{c}))] &= \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\mathsf{tr}(\rho(f(\boldsymbol{a}) \cdot f(\boldsymbol{b}) \cdot f(\boldsymbol{c})))] \\
(\rho \text{ is a homomorphism}) \qquad &= \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\mathsf{tr}(\rho(f(\boldsymbol{a})) \cdot \rho(f(\boldsymbol{b}) \cdot \rho(f(\boldsymbol{c}))] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \left[\sum_{1\leqslant i,j,k\leqslant \dim(\rho)} \rho(f(\boldsymbol{a}))_{ij} \cdot \rho(f(\boldsymbol{b})_{jk} \cdot \rho(f(\boldsymbol{c}))_{ki} \right] \\
&= \sum_{1\leqslant i,j,k\leqslant \dim(\rho)} \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \left[g_{ij}(\boldsymbol{a}) g_{jk}(\boldsymbol{b}) g_{ki}(\boldsymbol{c}) \right] \\
&= \sum_{1\leqslant i,j,k\leqslant \dim(\rho)} (g_{ij} * g_{jk} * g_{ki})(1_{G^n}). && (6)
\end{aligned}$$

Since we assume that the test passes with probability $\frac{1}{|[G,G]|} + \varepsilon$, from Equation (5) and Equation (6) (and using $\dim(\rho) \leqslant \sqrt{|G|}$), we conclude that there exists ρ and $1 \leqslant i, j, k \leqslant \dim(\rho)$ such that

$$|(g_{ij} * g_{jk} * g_{ki})(1_{G^n})| \geqslant \frac{\varepsilon}{|G|}.$$

We now analyze the term $(g_{ij} * g_{jk} * g_{ki})(1_{G^n})$ for a fixed (i, j, k). For the ease of notation, we write $h_1 := g_{ij}, h_2 := g_{jk}$ and $h_3 := g_{ki}$.

$$\frac{\varepsilon}{|G|} \leqslant |(g_{ij} * g_{jk} * g_{ki})(1_{G^n})| = |(h_1 * h_2 * h_3)(1_{G^n})|$$

$$
\begin{aligned}
&= \left| \sum_{\alpha \in \mathsf{Irrep}(G^n)} \dim(\alpha) \cdot \mathsf{tr}(h_1 * \hat{h_2} * h_3(\alpha)) \right| \\
&\leqslant \sum_{\alpha \in \mathsf{Irrep}(G^n)} \dim(\alpha) \cdot |\mathsf{tr}(\hat{h}_1(\alpha)\hat{h}_2(\alpha)\hat{h}_3(\alpha))| \\
&= \sum_{\alpha \in \mathsf{Irrep}(G^n)} \dim(\alpha) \cdot \left| \langle \hat{h}_1(\alpha)\hat{h}_2(\alpha), \hat{h}_3(\alpha)^\star \rangle_{\mathsf{End}\, \alpha_V} \right| \\
&\leqslant \sum_{\alpha \in \mathsf{Irrep}(G^n)} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)^\star\|_{\mathsf{HS}}.
\end{aligned}
$$

We now use Lemma 2.29 to conclude that for all $1 \leqslant i \leqslant 3$, $\hat{h}_i(\alpha) = 0$ if $\dim(\alpha) = 1$. Using this, we continue as follows:

$$
\begin{aligned}
|(h_1 * h_2 * h_3)(1_G)| &\leqslant \sum_{\substack{\alpha \in \mathsf{Irrep}(G^n), \\ \dim(\alpha) \geqslant 2}} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)^\star\|_{\mathsf{HS}} \\
&= \sum_{\substack{\alpha, \\ \dim(\alpha) \geqslant 2}} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}}.
\end{aligned}
$$

Let $D := \frac{1}{2\delta^2}$. Now, we split the sum into two parts $|(h_1 * h_2 * h_3)(1_{G^n})| \leqslant \Theta_{\mathsf{low}} + \Theta_{\mathsf{high}}$ where

$$
\Theta_{\mathsf{low}} = \sum_{\substack{\alpha, \\ \dim(\alpha) \geqslant 2, \\ \dim_{\geqslant 2}(\alpha) < D}} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}},
$$

and

$$
\Theta_{\mathsf{high}} = \sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}}.
$$

3.1 Bounding higher order terms

In this section, we show that the high degree terms can be upper bounded by a small constant, even though the *three* queries are perfectly correlated.

We bound Θ_{high} as follows:

$$
\begin{aligned}
\Theta_{\mathsf{high}} &= \sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}} \\
&\leqslant \sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\|_{\mathsf{HS}} \|\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}} && \text{(Claim 2.22)} \\
&\leqslant \frac{1}{\sqrt{2D}} \sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha)^{3/2} \cdot \|\hat{h}_1(\alpha)\|_{\mathsf{HS}} \|\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}}.
\end{aligned}
$$

Here, we used that fact that all the representations α of G with $\dim_{\geqslant 2}(\alpha) \geqslant D$ have dimensions at least $2D$. At this point, we would like to point out the main source of effectively bounding the higher order terms. It is the size of $\dim(\alpha)$ in the summation. In Gowers' [Gow08] proof, a similar expression appears in the analysis, with the same condition that all the representations in the summation have large dimension. It is in some sense the main difference between the abelian and the non-abelian setting (both in this work and Gowers'), similar to the Equation (1) mentioned in the introduction.

Now, using Cauchy-Schwartz inequality,

$$\begin{aligned}
\Theta_{\mathsf{high}} &\leqslant \frac{1}{\sqrt{2D}} \sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha)^{3/2} \cdot \|\hat{h}_1(\alpha)\|_{\mathsf{HS}} \|\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}} \\
&\leqslant \frac{1}{\sqrt{2D}} \left(\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\|_{\mathsf{HS}}^2 \right)^{1/2} \cdot \left(\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha)^2 \cdot \|\hat{h}_2(\alpha)\|_{\mathsf{HS}}^2 \|\hat{h}_3(\alpha)\|_{\mathsf{HS}}^2 \right)^{1/2} \\
&\leqslant \frac{1}{\sqrt{2D}} \left(\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\|_{\mathsf{HS}}^2 \right)^{1/2} \cdot \\
&\qquad \left(\left(\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha) \cdot \|\hat{h}_2(\alpha)\|_{\mathsf{HS}}^2 \right) \cdot \left(\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) \geqslant D}} \dim(\alpha) \cdot \|\hat{h}_3(\alpha)\|_{\mathsf{HS}}^2 \right) \right)^{1/2} \\
&\leqslant \frac{1}{\sqrt{2D}} \cdot \|h_1\|_{L^2(G^n)} \cdot \|h_2\|_{L^2(G^n)} \cdot \|h_3\|_{L^2(G^n)}. \qquad \text{(Using Proposition 2.21)}
\end{aligned}$$

Si nce h_1 was defined as $h_1(\boldsymbol{x}) = \rho(f(\boldsymbol{x}))_{ij}$ where $\rho \in \mathsf{Irrep}(G)$, $|h_1(\boldsymbol{x})| \leqslant 1$. Same is true for h_2 and h_3, and hence all the norms are bounded by 1. Therefore,

$$\Theta_{\mathsf{high}} \leqslant \frac{1}{\sqrt{2D}} = \delta.$$

3.2 Bounding lower order terms

It remain to show that Θ_{low} is related to the Fourier mass of h_3 on the low dimension representations.

$$\begin{aligned}
\Theta_{\mathsf{low}} &= \sum_{\substack{\alpha, \\ \dim(\alpha) \geqslant 2, \\ \dim_{\geqslant 2}(\alpha) < D}} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\hat{h}_2(\alpha)\|_{\mathsf{HS}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}} \\
&\leqslant \max_{\substack{\alpha, \\ \dim(\alpha) \geqslant 2, \\ \dim_{\geqslant 2}(\alpha) < D}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}} \cdot \left(\sum_{\alpha} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\hat{h}_2(\alpha)\|_{\mathsf{HS}} \right)
\end{aligned}$$

We can upper bound the summation by 1 using the Cauchy-Schwartz inequality as follows:

$$\begin{aligned}
\sum_{\alpha} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\hat{h}_2(\alpha)\|_{\mathsf{HS}} &\leqslant \sum_{\alpha} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\|_{\mathsf{HS}} \|\hat{h}_2(\alpha)\|_{\mathsf{HS}} \\
&\leqslant \left(\sum_{\alpha} \dim(\alpha) \cdot \|\hat{h}_1(\alpha)\|_{\mathsf{HS}}^2 \right)^{1/2} \cdot \left(\sum_{\alpha} \dim(\alpha) \cdot \|\hat{h}_2(\alpha)\|_{\mathsf{HS}}^2 \right)^{1/2}. \\
&= \|h_1\|_2 \cdot \|h_2\|_2 \qquad \text{(Proposition 2.21)} \\
&\leqslant 1.
\end{aligned}$$

where the last inequality uses the fact that $|h_1(\boldsymbol{x})|, |h_2(\boldsymbol{x})| \leqslant 1$ for all $\boldsymbol{x} \in G^n$. Using the upper bound on Θ_{high}, we have $\Theta_{\mathsf{low}} \geqslant \frac{\varepsilon}{|G|} - \delta$. Therefore, we get

$$\max_{\substack{\alpha, \\ \dim(\alpha) \geqslant 2, \\ \dim_{\geqslant 2}(\alpha) < D}} \|\hat{h}_3(\alpha)\|_{\mathsf{HS}} \geqslant \left(\frac{\varepsilon}{|G|} - \delta \right).$$

□

4 Main Reduction

In this section, we prove Theorem 1.1. We give a reduction from an instance of a LABEL-COVER, $\mathcal{H} = (\mathcal{U}, \mathcal{V}, E, [L], [R], \{\pi_e\}_{e \in E})$ as in Definition 2.1, to a 3-LIN instance $\mathcal{I}$ over a non-abelian group G.

The set of variables of $\mathcal{I}$ is $(\mathcal{U} \times G^L) \cup (\mathcal{V} \times G^R)$. Any assignment to the instance $\mathcal{I}$ is given by a set of functions $f_u : G^L \to G$ and $f_v : G^R \to G$ for each $u \in \mathcal{U}$ and $v \in \mathcal{V}$. We further assume that these functions are *folded*.

The distribution of the 3-LIN constraints in $\mathcal{I}$ is given by the following test:

1. Choose an edge $e(u, v) \in E$ of $\mathcal{H}$ uniformly at random.
2. Sample $\boldsymbol{a} = (a_1, a_2, \ldots, a_R)$ from G^R uniformly at random.
3. Sample $\boldsymbol{b} = (b_1, b_2, \ldots, b_L)$ from G^L uniformly at random.
4. Let $\boldsymbol{c} = (c_1, c_2, \ldots, c_R)$ be such that $c_i = (b \circ \pi_{uv})_i^{-1} \cdot a_i^{-1}$, here $\boldsymbol{x} \circ \pi \in G^R$ is the string defined as $(x \circ \pi)_i := x_{\pi(i)}$ for $i \in [R]$.
5. Test if $f_v(\boldsymbol{a}) \cdot f_u(\boldsymbol{b}) \cdot f_v(\boldsymbol{c}) = 1_G$.

The value of the instance $val(\mathcal{I})$ is the maximum probability that the above test is satisfied, where the maximum is over all folded functions $\{f_v\}_{v \in \mathcal{V}}, \{f_u\}_{u \in \mathcal{U}}$.

4.1 Analysis

Lemma 4.1 (Completeness). *If $\mathcal{H}$ is a satisfiable instance of* Label-Cover, *then $val(\mathcal{I}) = 1$.*

Proof. Fix a satisfying assignment $\ell : \mathcal{U} \to [L], \ell : \mathcal{V} \to [R]$ of $\mathcal{H}$. Consider the long code encoding of the labeling ℓ : $f_v(\boldsymbol{x}) = x_{\ell(v)}$ and $f_u(\boldsymbol{x}) = x_{\ell(u)}$, for every $v \in \mathcal{V}$ and $u \in \mathcal{U}$. We show that this assignment to $\mathcal{I}$ satisfies all the constraints.

$$\begin{aligned} f_v(\boldsymbol{a}) \bullet f_u(\boldsymbol{b}) \bullet f_v(\boldsymbol{c}) &= a_{\ell(v)} \bullet b_{\ell(u)} \bullet c_{\ell(v)} \\ &= a_{\ell(v)} \bullet b_{\ell(u)} \bullet (b \circ \pi_{uv})^{-1}_{\ell(v)} \bullet a^{-1}_{\ell(v)} \\ &= a_{\ell(v)} \bullet b_{\ell(u)} \bullet b^{-1}_{\pi_{uv}(\ell(v))} \bullet a^{-1}_{\ell(v)} \\ &= a_{\ell(v)} \bullet b_{\ell(u)} \bullet b^{-1}_{\ell(u)} \bullet a^{-1}_{\ell(v)} && (\pi_{uv}(\ell(v)) = \ell(u)) \\ &= 1_G. \end{aligned}$$

□

We now prove the main soundness lemma. Note that Lemma 4.1 and Lemma 4.2 along with the **NP**-hardness of Label-Cover from Theorem 2.2 for large enough r imply our main theorem Theorem 1.1 for any constant $\epsilon > 0$.

Lemma 4.2 (Soundness). *Let $\delta \in (0,1)$. Let C be a constant such that $C^{-d_0/2} \leqslant \frac{\delta^2}{12|G|^6}$, where d_0 is the constant from Theorem 2.2. If $\mathcal{H}$ is at most $\frac{\delta^2}{10|G|^{10C}}$-satisfiable, then $val(\mathcal{I}) \leqslant \frac{1}{|[G,G]|} + \delta$.*

Proof. Fix any assignment to the instance $\mathcal{I}$ given by a set of functions $f_u : G^L \to G$ and $f_v : G^R \to G$ for each $u \in \mathcal{U}$ and $v \in \mathcal{V}$. The value of the instance for this assignment is given by:

$$val(\mathcal{I}) = \mathop{\mathbf{E}}_{e(u,v)\in E} \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \left[\frac{1}{|G|} \sum_{\rho \in \mathsf{Irrep}(G)} \dim(\rho) \chi_\rho(f_v(\boldsymbol{a}) \bullet f_u(\boldsymbol{b}) \bullet f_v(\boldsymbol{c})) \right]$$

For any $\rho \in \mathsf{Irrep}(G)$ such that $\dim(\rho) = 1$, we have

$$\begin{aligned} \chi_\rho(f_v(\boldsymbol{a}) \bullet f_u(\boldsymbol{b}) \bullet f_v(\boldsymbol{c})) &= \chi_\rho(f_v(\boldsymbol{a})) \cdot \chi_\rho(f_u(\boldsymbol{b})) \cdot \chi_\rho(f_v(\boldsymbol{c})) \\ &\leqslant \left|\chi_\rho(f_v(\boldsymbol{a}))\right| \cdot \left|\chi_\rho(f_u(\boldsymbol{b}))\right| \cdot \left|\chi_\rho(f_v(\boldsymbol{c}))\right| \\ &= 1. && \text{(unitary representations)} \end{aligned}$$

As the number of dimension 1 representations of a group G is equal to the size of the quotient $G/[G,G]$, we get

$$val(\mathcal{I}) \leqslant \frac{1}{|[G,G]|} + \frac{1}{|G|} \mathop{\mathbf{E}}_{e(u,v)\in E} \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \left[\sum_{\substack{\rho \in \mathsf{Irrep}(G), \\ \dim(\rho) \geqslant 2}} \dim(\rho) \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f_v(\boldsymbol{a}) \bullet f_u(\boldsymbol{b}) \bullet f_v(\boldsymbol{c}))] \right]$$

$$\leqslant \frac{1}{|[G,G]|} + \sum_{\substack{\rho\in \mathsf{Irrep}(G),\\ \dim(\rho)\geqslant 2}} \left| \mathop{\mathbf{E}}_{e(u,v)\in E} \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f_v(\boldsymbol{a}) \bullet f_u(\boldsymbol{b}) \bullet f_v(\boldsymbol{c}))] \right|.$$

The lemma follows from the following Claim 4.3. □

Claim 4.3. *If $\mathcal{H}$ is at most $\frac{\delta^2}{10|G|^{10C}}$-satisfiable, then for every $\rho \in \mathsf{Irrep}(G)$ such that $\dim(\rho) \geqslant 2$,*

$$\left| \mathop{\mathbf{E}}_{e(u,v)\in E} \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f_v(\boldsymbol{a}) \bullet f_u(\boldsymbol{b}) \bullet f_v(\boldsymbol{c}))] \right| \leqslant \frac{\delta}{|G|}.$$

Proof. Fix any $\rho \in \mathsf{Irrep}(G)$ such that $\dim(\rho) \geqslant 2$. Let

$$\Theta := \mathop{\mathbf{E}}_{e(u,v)\in E} \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f_v(\boldsymbol{a}) \bullet f_u(\boldsymbol{b}) \bullet f_v(\boldsymbol{c}))].$$

We first look at the inner expectation. For $1 \leqslant p,q \in \dim(\rho)$, let $g_{pq} : G^R \to \mathbb{C}$ be defined as $g_{pq}(\boldsymbol{x}) := \rho(f_v(\boldsymbol{x}))_{pq}$. Also, let $h_{pq} : G^L \to \mathbb{C}$ be defined as $h_{pq}(\boldsymbol{y}) := \rho(f_u(\boldsymbol{y}))_{pq}$. We have

$$\begin{aligned}
\mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\chi_\rho(f_v(\boldsymbol{a}) f_u(\boldsymbol{b}) f_v(\boldsymbol{c}))] &= \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\mathsf{tr}(\rho(f_v(\boldsymbol{a}) \bullet f_u(\boldsymbol{b}) \bullet f_v(\boldsymbol{c})))] \\
(\rho \text{ is a homomorphism}) \qquad &= \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} [\mathsf{tr}(\rho(f_v(\boldsymbol{a})) \cdot \rho(f_u(\boldsymbol{b}) \cdot \rho(f_v(\boldsymbol{c}))] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \left[\sum_{1\leqslant p,q,r\leqslant \dim(\rho)} \rho(f_v(\boldsymbol{a}))_{pq} \cdot \rho(f_u(\boldsymbol{b})_{qr} \cdot \rho(f_v(\boldsymbol{c}))_{rp} \right] \\
&= \sum_{1\leqslant p,q,r\leqslant \dim(\rho)} \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \left[g_{pq}(\boldsymbol{a}) \cdot h_{qr}(\boldsymbol{b}) \cdot g_{rp}(\boldsymbol{c}) \right].
\end{aligned}$$

We now analyze the term $\Theta^e_{p,q,r} := \mathbf{E}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \left[g_{pq}(\boldsymbol{a}) h_{qr}(\boldsymbol{b}) g_{rp}(\boldsymbol{c}) \right]$ for a fixed (p,q,r). For the ease of notations, we write $g := g_{pq}, h := h_{qr}$ and $g' := g_{rp}$. Also, we use π for π_{uv}.

$$\mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b}} [g(\boldsymbol{a}) \cdot g'((\boldsymbol{b} \circ \pi)^{-1} \bullet \boldsymbol{a}^{-1})] = \mathop{\mathbf{E}}_{\boldsymbol{b}} [(g * g')((\boldsymbol{b} \circ \pi)^{-1})].$$

We now bound the expectation as follows:

$$\begin{aligned}
&\mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \left[g_{pq}(\boldsymbol{a}) \cdot h_{qr}(\boldsymbol{b}) \cdot g_{rp}(\boldsymbol{c}) \right] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \left[g(\boldsymbol{a}) \cdot h(\boldsymbol{b}) \cdot g'((\boldsymbol{b} \circ \pi)^{-1} \bullet \boldsymbol{a}^{-1}) \right] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{b}} \left[(g * g')(\boldsymbol{b}^{-1} \circ \pi) \cdot h(\boldsymbol{b}) \right] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{b}} \left[(g * g')(\boldsymbol{b} \circ \pi) \cdot h(\boldsymbol{b}^{-1}) \right] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{b}} \left[\left(\sum_\alpha \dim(\alpha) \mathsf{tr}(\hat{g}(\alpha) \hat{g}'(\alpha) \alpha(\boldsymbol{b} \circ \pi)) \right) \cdot \left(\sum_\beta \dim(\beta) \mathsf{tr}(\hat{h}(\beta) \beta(\boldsymbol{b}^{-1})) \right) \right]
\end{aligned}$$

$$= \sum_{\substack{\alpha,\beta,\\ \dim(\alpha),\dim(\beta)\geqslant 2}} \underbrace{\dim(\alpha)\dim(\beta)\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\mathrm{tr}(\hat{g}(\alpha)\hat{g}'(\alpha)\alpha(\boldsymbol{b}\circ\pi))\cdot\mathrm{tr}(\hat{h}(\beta)\beta(\boldsymbol{b}^{-1}))\right]}_{\mathbf{Term}^e(\alpha,\beta)},$$

where the last step uses the fact that the functions g, g' and h satisfy the condition of Lemma 2.29 and hence $\hat{g}(\alpha) = 0$ if $\dim(\alpha) = 1$ (same for $\hat{g}'(\alpha)$ and $\hat{h}(\beta)$).

We now break the sum into two parts:

$$\Theta^e_{p,q,r}(\mathsf{low}) := \sum_{\substack{\alpha,\beta,\\ \dim(\alpha),\dim(\beta)\geqslant 2,\\ \dim_{\geqslant 2}(\alpha)\leqslant C}} \mathbf{Term}^e(\alpha,\beta), \qquad \Theta^e_{p,q,r}(\mathsf{high}) := \sum_{\substack{\alpha,\beta,\\ \dim(\alpha),\dim(\beta)\geqslant 2,\\ \dim_{\geqslant 2}(\alpha)> C}} \mathbf{Term}^e(\alpha,\beta).$$

Recall, $\dim_{\geqslant 2}(\alpha)$ denotes the number of representations in $\alpha = (\rho_1, \rho_2, \ldots, \rho_R)$ which are of dimensions at least 2. With these notations, we have

$$\Theta := \sum_{p,q,r} \mathop{\mathbf{E}}_{e(u,v)\in E}[\Theta^e_{p,q,r}(\mathsf{low})] + \mathop{\mathbf{E}}_{e(u,v)\in E}[\Theta^e_{p,q,r}(\mathsf{high})].$$

The upper bound on Θ follows from Claim 4.4 and Claim 4.5 and triangle inequality (and also noting that p, q and r take at most $\sqrt{G}$ distinct values). □

Claim 4.4. *If $\mathcal{H}$ is at most $\frac{\delta^2}{10|G|^{10C}}$-satisfiable, then for every $\rho \in \mathsf{Irrep}(G)$ such that $\dim(\rho) \geqslant 2$, and every $1 \leqslant p, q, r \leqslant \dim(\rho)$,*

$$\left|\mathop{\mathbf{E}}_{e(u,v)\in E}[\Theta^e_{p,q,r}(\mathsf{low})]\right| \leqslant \frac{\delta}{2|G|^3}.$$

Claim 4.5. *Let C be a constant such that $C^{-d_0/2} \leqslant \frac{\delta^2}{12|G|^6}$, where d_0 is the constant from Theorem 2.2. For every $1 \leqslant p, q, r \leqslant \dim(\rho)$,*

$$\left|\mathop{\mathbf{E}}_{e(u,v)\in E}[\Theta^e_{p,q,r}(\mathsf{high})]\right| \leqslant \frac{\delta}{2|G|^3}.$$

4.1.1 Bounding the low terms

Claim 4.6. *(Restatement of Claim 4.4) If $\mathcal{H}$ is at most $\frac{\delta^2}{10|G|^{10C}}$-satisfiable,then for every $\rho \in \mathsf{Irrep}(G)$ such that $\dim(\rho) \geqslant 2$, and every $1 \leqslant p, q, r \leqslant \dim(\rho)$,*

$$\left|\mathop{\mathbf{E}}_{e(u,v)\in E}[\Theta^e_{p,q,r}(\mathsf{low})]\right| \leqslant \frac{\delta}{2|G|^3}.$$

Proof. Fix any $\rho \in \mathsf{Irrep}(G)$ such that $\dim(\rho) \geqslant 2$. Assume towards contradiction that there exists $1 \leqslant p, q, r \leqslant \dim(\rho)$ such that

$$\mathop{\mathbf{E}}_{e(u,v)\in E}[|\Theta^e_{p,q,r}(\mathsf{low})|] \geqslant \left|\mathop{\mathbf{E}}_{e(u,v)\in E}[\Theta^e_{p,q,r}(\mathsf{low})]\right| > \frac{\delta}{2|G|^3}.$$

We show that in this case, $\mathcal{H}$ has a $> \frac{\delta^2}{10|G|^{10C}}$- satisfying assignment which is a contradiction. Consider the term $\mathbf{Term}^e(\alpha,\beta)$ when $\alpha = (\rho_1,\rho_2,\dots,\rho_R)$ and $\beta = (\tau_1,\tau_2,\dots,\tau_L)$.

$$\begin{aligned}
\mathbf{Term}^e(\alpha,\beta) &= \dim(\alpha)\dim(\beta)\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\mathrm{tr}(\hat{g}(\alpha)\hat{g}'(\alpha)\alpha(\boldsymbol{b}\circ\pi))\cdot\mathrm{tr}(\hat{h}(\beta)\beta(\boldsymbol{b}^{-1}))\right]\\
&= \dim(\alpha)\dim(\beta)\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\sum_{i,k}(\hat{g}(\alpha)\hat{g}'(\alpha))_{ik}\alpha(\boldsymbol{b}\circ\pi)_{ki}\cdot\sum_{i',k'}\hat{h}(\beta)_{i'k'}\beta(\boldsymbol{b}^{-1})_{k'i'}\right]\\
&= \dim(\alpha)\dim(\beta)\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\sum_{\substack{i,k\\i',k'}}(\hat{g}(\alpha)\hat{g}'(\alpha))_{ik}\alpha(\boldsymbol{b}\circ\pi)_{ki}\hat{h}(\beta)_{i'k'}\beta(\boldsymbol{b}^{-1})_{k'i'}\right]\\
&= \dim(\alpha)\dim(\beta)\sum_{\substack{i,k\\i',k'}}(\hat{g}(\alpha)\hat{g}'(\alpha))_{ik}\hat{h}(\beta)_{i'k'}\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\alpha(\boldsymbol{b}\circ\pi)_{ki}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right],
\end{aligned}$$

where (i,jk) are the tuples $i = (i_1,i_2,\dots,i_R)$ and $k = (k_1,k_2,\dots,k_R)$ such that for all $\ell\in[R]$, $1\leqslant i_\ell,j_\ell\leqslant\dim(\rho_\ell)$. Similarly, (i',k') are the tuples $i' = (i'_1,i'_2,\dots,i'_L)$ and $k' = (k'_1,k'_2,\dots,k'_L)$ such that for all $\ell'\in[L]$, $1\leqslant i'_{\ell'},j'_{\ell'}\leqslant\dim(\tau_{\ell'})$. Now,

$$\begin{aligned}
\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\alpha(\boldsymbol{b}\circ\pi))_{ki}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right] &= \mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\prod_{\ell'=1}^{L}\tau_{\ell'}(b_{\ell'}^{-1})_{k'_{\ell'},i'_{\ell'}}\cdot\prod_{\ell\in\pi^{-1}(\ell')}\rho_\ell(b_{\ell'})_{k_\ell i_\ell}\right]\\
&= \prod_{\ell'=1}^{L}\mathop{\mathbf{E}}_{b}\left[\tau_{\ell'}(b^{-1})_{k'_{\ell'},i'_{\ell'}}\cdot\prod_{\ell\in\pi^{-1}(\ell')}\rho_\ell(b)_{k_\ell i_\ell}\right].
\end{aligned}$$

Now suppose there exists ℓ' such that for all $\ell\in\pi^{-1}(\ell')$, $\dim(\rho_\ell) = 1$. Since product of dimension 1 representation is also a dimension 1 representation, for the expectation to be nonzero, $\dim(\tau_{\ell'})$ must be 1, by Proposition 2.10. Thus, if $\dim(\beta)\geqslant 2$, then there must exist an ℓ' such that $\dim(\tau_{\ell'})\geqslant 2$ and an $\ell\in\pi^{-1}(\ell')$ such that $\dim(\rho_\ell)\geqslant 2$ (again, for the expectation to be non-zero). Therefore, we conclude that the terms that are nonzero in the expression Θ_{low} are all (α,β) such that for all $\ell'\in[L]$ whenever $\dim(\tau_{\ell'})\geqslant 2$, there exists a $\ell\in\pi^{-1}(\ell')$ such that $\dim(\rho_\ell)\geqslant 2$. Let us define $\pi_{\geqslant 2}(\alpha) = \{\ell'\in L\mid\exists\ell\in\pi^{-1}(\ell'),\dim(\rho_\ell)\geqslant 2\}$ and $\beta_{\geqslant 2} = \{\ell'\mid\dim(\tau_{\ell'})\geqslant 2\}$. We get

$$\Theta^e_{p,q,r}(\mathsf{low}) = \sum_{\substack{\alpha,\beta,\\ \dim(\alpha)\geqslant 2,\dim(\beta)\geqslant 2,\\ \dim_{\geqslant 2}(\alpha)\leqslant C,\\ \beta_{\geqslant 2}\subseteq\pi_{\geqslant 2}(\alpha)}}\dim(\alpha)\dim(\beta)\sum_{i,k}(\hat{g}(\alpha)\hat{g}'(\alpha))_{ik}\sum_{i',k'}\hat{h}(\beta)_{i'k'}\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\alpha(\boldsymbol{b}\circ\pi))_{ki}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right].$$

Now, let $F^{ik}_\alpha : G^L\to\mathbb{C}$ be the following function:

$$F^{ik}_\alpha(\boldsymbol{b}) := \alpha(\boldsymbol{b}^{-1}\circ\pi)_{ki}.$$

Note that,

$$\sum_k\|F^{ik}_\alpha\|_2^2 = \sum_k\mathop{\mathbf{E}}_{\boldsymbol{b}}[|\alpha(\boldsymbol{b}^{-1}\circ\pi)_{ki}|^2] = \mathop{\mathbf{E}}_{\boldsymbol{b}}\sum_k|\alpha(\boldsymbol{b}^{-1}\circ\pi)_{ki}|^2 = 1, \tag{7}$$

where the last equality uses the fact that the sum expression is exactly the norm of the column i of representation α, which is 1 ($\alpha(.)$ is unitary). We now analyze the expectation:

$$\begin{aligned}
\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\alpha(\boldsymbol{b}\circ\pi)_{ki}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right] &= \mathop{\mathbf{E}}_{\boldsymbol{b}}\left[F^{ik}_{\alpha}(\boldsymbol{b}^{-1})\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\sum_{\beta'}\dim(\beta')\mathsf{tr}(\hat{F^{ik}_{\alpha}}(\beta')\beta'(\boldsymbol{b}^{-1})^{\star})\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\sum_{\beta'}\dim(\beta')\mathsf{tr}(\hat{F^{ik}_{\alpha}}(\beta')\beta'(\boldsymbol{b}))\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\sum_{\beta'}\dim(\beta')\sum_{i'',k''}\hat{F^{ik}_{\alpha}}(\beta')_{i'',k''}\beta'(\boldsymbol{b})_{k'',i''}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right] \\
&= \sum_{\beta'}\dim(\beta')\sum_{i'',k''}\hat{F^{ik}_{\alpha}}(\beta')_{i'',k''}\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\beta'(\boldsymbol{b})_{k'',i''}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right].
\end{aligned}$$

By Proposition 2.10, the expectation is zero unless $\beta' = \beta$, $i'' = k'$ and $k'' = i'$, otherwise it is $1/\dim(\beta')$. Therefore,

$$\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\alpha(\boldsymbol{b}\circ\pi)_{ki}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right] = \hat{F^{ik}_{\alpha}}(\beta)_{k',i'}.$$

Plugging this into $\Theta^{e}_{p,q,r}(\mathsf{low})$, we get

$$\begin{aligned}
|\Theta^{e}_{p,q,r}(\mathsf{low})|^2 &= \left|\sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\sum_{\substack{i,k,\\ i',k'}}(\hat{g}(\alpha)\hat{g}'(\alpha))_{ik}\hat{h}(\beta)_{i'k'}\hat{F^{ik}_{\alpha}}(\beta)_{k',i'}\right|^2 \\
&= \left|\sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\sum_{\substack{i,j,k,\\ i',k'}}\hat{g}(\alpha)_{ij}\hat{g}'(\alpha)_{jk}\hat{h}(\beta)_{i'k'}\hat{F^{ik}_{\alpha}}(\beta)_{k',i'}\right|^2 \\
&\leqslant \left(\sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\sum_{\substack{i,j,k,\\ i',k'}}|\hat{g}'(\alpha)_{jk}|^2|\hat{h}(\beta)_{i'k'}|^2\right)\left(\sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\sum_{\substack{i,j,k,\\ i',k'}}|\hat{g}(\alpha)_{ij}|^2|\hat{F^{ik}_{\alpha}}(\beta)_{k',i'}|^2\right).
\end{aligned}$$

We can bound the second term as follows:

$$\begin{aligned}
\left(\sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\sum_{\substack{i,j,k,\\ i',k'}}|\hat{g}(\alpha)_{ij}|^2|\hat{F^{ik}_{\alpha}}(\beta)_{k',i'}|^2\right) &= \left(\sum_{\alpha}\dim(\alpha)\sum_{i,j}|\hat{g}(\alpha)_{ij}|^2\sum_{k}\sum_{\beta}\dim(\beta)\sum_{i',k'}|\hat{F^{ik}_{\alpha}}(\beta)_{k',i'}|^2\right) \\
&= \left(\sum_{\alpha}\dim(\alpha)\sum_{i,j}|\hat{g}(\alpha)_{ij}|^2\sum_{k}\|F^{ik}_{\alpha}\|_2^2\right)
\end{aligned}$$

$$= \left(\sum_{\alpha} \dim(\alpha) \sum_{i,j} |\hat{g}(\alpha)_{ij}|^2 \right) \quad \text{(Using Equation (7))}$$
$$= \|g\|_2^2 \leqslant 1.$$

Therefore,

$$|\Theta^e_{p,q,r}(\mathsf{low})|^2 \leqslant \sum_{\substack{\alpha,\beta,\\ \dim(\alpha),\dim(\beta)\geqslant 2,\\ \dim_{\geqslant 2}(\alpha)\leqslant C,\\ \beta_{\geqslant 2}\subseteq\pi_{\geqslant 2}(\alpha)}} \dim(\alpha)\dim(\beta) \sum_{\substack{i,j,k,\\ i',k'}} |\hat{g}'(\alpha)_{jk}|^2 |\hat{h}(\beta)_{i'k'}|^2$$
$$\leqslant |G|^C \sum_{\substack{\alpha,\beta,\\ \dim(\alpha),\dim(\beta)\geqslant 2,\\ \dim_{\geqslant 2}(\alpha)\leqslant C,\\ \beta_{\geqslant 2}\subseteq\pi_{\geqslant 2}(\alpha)}} \dim(\alpha)\dim(\beta) \sum_{\substack{j,k,\\ i',k'}} |\hat{g}'(\alpha)_{jk}|^2 |\hat{h}(\beta)_{i'k'}|^2$$
$$= |G|^C \sum_{\substack{\alpha,\beta,\\ \dim(\alpha),\dim(\beta)\geqslant 2,\\ \dim_{\geqslant 2}(\alpha)\leqslant C,\\ \beta_{\geqslant 2}\subseteq\pi_{\geqslant 2}(\alpha)}} \dim(\alpha)\dim(\beta) \|\hat{g}'(\alpha)\|_{\mathsf{HS}}^2 \|\hat{h}(\beta)\|_{\mathsf{HS}}^2.$$

We now show that if $\Theta^e_{p,q,r}(\mathsf{low})$ is large for a typical e, then it can be used to get a good labeling to the Label Cover instance $\mathcal{H}$.

Randomized labeling: Consider the following randomized labeling. For each $v \in \mathcal{V}$, consider $g_{rp} : G^R \to \mathbb{C}$ which is defined as $g_{rp}(\boldsymbol{x}) := \rho(f_v(\boldsymbol{x}))_{rp}$. Select $\boldsymbol{\alpha} = (\rho_1, \rho_2, \ldots, \rho_R)$ with probability $\dim(\alpha)\|\hat{g}_{rp}(\alpha)\|_{\mathsf{HS}}^2$. Select a uniformly random $\ell_v \in [R]$ such that $\dim(\rho_{\ell_v}) \geqslant 2$ and assign the label ℓ_v to v.

For each $u \in \mathcal{U}$, consider $h_{qr} : G^L \to \mathbb{C}$ be defined as $h_{qr}(\boldsymbol{y}) := \rho(f_u(\boldsymbol{y}))_{qr}$. Select $\beta = (\tau_1, \tau_2, \ldots, \tau_L)$ with probability $\dim(\beta)\|\hat{h}_{qr}(\beta)\|_{\mathsf{HS}}^2$. Select a uniformly random $\ell_u \in [L]$ such that $\dim(\tau_{\ell_u}) \geqslant 2$ and assign the label ℓ_u to u.

Now fix an edge $e(u, v)$. The probability p_e that this edge is satisfied by the randomized labeling, is lower bounded by:

$$p_e \geqslant \sum_{\substack{\alpha,\beta,\\ \dim(\alpha),\dim(\beta)\geqslant 2,\\ \dim_{\geqslant 2}(\alpha)\leqslant C,\\ \beta_{\geqslant 2}\subseteq\pi_{\geqslant 2}(\alpha)}} \dim(\alpha)\dim(\beta) \|\hat{g}'(\alpha)\|_{\mathsf{HS}}^2 \|\hat{h}(\beta)\|_{\mathsf{HS}}^2 \cdot \frac{1}{\dim_{\geqslant 2}(\alpha)}$$
$$\geqslant \frac{1}{C} \sum_{\substack{\alpha,\beta,\\ \dim(\alpha),\dim(\beta)\geqslant 2,\\ \dim_{\geqslant 2}(\alpha)\leqslant C,\\ \beta_{\geqslant 2}\subseteq\pi_{\geqslant 2}(\alpha)}} \dim(\alpha)\dim(\beta) \|\hat{g}'(\alpha)\|_{\mathsf{HS}}^2 \|\hat{h}(\beta)\|_{\mathsf{HS}}^2$$
$$\geqslant \frac{1}{C \cdot |G|^C} |\Theta^e_{p,q,r}(\mathsf{low})|^2.$$

Therefore the expected number of edges satisfied by the randomized labeling is lower bounded by

$$\begin{aligned}
\mathop{\mathbf{E}}_{e\in E}[p_e] &\geqslant \mathop{\mathbf{E}}_{e}\left[\frac{1}{C\cdot|G|^C}|\Theta^e_{p,q,r}(\mathsf{low})|^2\right] \\
&= \frac{1}{C\cdot|G|^C}\mathop{\mathbf{E}}_{e}\left[|\Theta^e_{p,q,r}(\mathsf{low})|^2\right] \\
&\geqslant \frac{1}{C\cdot|G|^C}\mathop{\mathbf{E}}_{e}\left[|\Theta^e_{p,q,r}(\mathsf{low})|\right]^2 && \text{(Using convexity)} \\
&\geqslant \frac{1}{C\cdot|G|^C}\cdot\frac{\delta^2}{4|G|^6} \\
&> \frac{\delta^2}{10|G|^{10C}}.
\end{aligned}$$

Since the expected fraction of the edges that are satisfied is strictly greater than $\frac{\delta^2}{10|G|^{10C}}$, by conditional expectation, there exists a labeling to the Label Cover instance $\mathcal{H}$ that satisfies more than $\frac{\delta^2}{10|G|^{10C}}$ fraction of the edges, which is a contradiction. □

4.1.2 Bounding the high terms

We now show the following claim:

Claim 4.7. *(Restatement of Claim 4.5) Let C be a constant such that $C^{-d_0/2} \leqslant \frac{\delta^2}{12|G|^6}$, where d_0 is the constant from Theorem 2.2. For every $1 \leqslant p,q,r \leqslant \dim(\rho)$,*

$$\left|\mathop{\mathbf{E}}_{e(u,v)\in E}[\Theta^e_{p,q,r}(\mathsf{high})]\right| \leqslant \frac{\delta}{2|G|^3}.$$

Proof. Recall,

$$\Theta^e_{p,q,r}(\mathsf{high}) := \sum_{\substack{\alpha,\beta,\\ \dim(\alpha),\dim(\beta)\geqslant 2,\\ \dim_{\geqslant 2}(\alpha)>C}} \mathbf{Term}^e(\alpha,\beta).$$

Let's analyze the expression $\Theta^e_{p,q,r}(\mathsf{high})$ more carefully. For the notational convenience, we suppress the conditions on α,β and simply write the sum over pairs α,β. We will analyze the complete sum with the extra conditions on α,β, once we simplify the expression.

Let $U(e,\alpha)$ be the transformation (i.e., change of basis) which takes a representation $\alpha(.)$ and converts it into a direct sum of irreducible representations of $\pi_e(G^R) := \{\boldsymbol{x}\circ\pi_e \mid x\in G^L\}$ which is a subgroup of G^R isomorphic to the group G^L. For simplicity, we denote this unitary matrix by U. Recall that the decomposition is unique.

We extend the definition of the *block diagonal matrices* to include any permutation of columns of a block diagonal matrix. For clarity, we call such general matrices *block matrices*. Note that with this extended definition, it still makes sense to talk about the 'blocks', except that the blocks are

not contiguous and not necessarily along the diagonal. For a given (e,α), we apply a column-permutation matrix $P(e,\alpha)$ to the block diagonal matrix $U\alpha U^\star$. We will get back to the specific choice of $P(e,\alpha)$ later in the proof, but for now just write the permutation matrix as P for notational convenience.

$$\begin{aligned}\mathsf{tr}(\hat{g}(\alpha)\hat{g}'(\alpha)\alpha(\boldsymbol{b}\circ\pi)) &= \mathsf{tr}(U\hat{g}(\alpha)\hat{g}'(\alpha)\alpha(\boldsymbol{b}\circ\pi)U^\star) \qquad \text{(cyclic property of tr, and } UU^\star = I)\\ &= \mathsf{tr}(U\hat{g}(\alpha)\hat{g}'(\alpha)U^\star U\alpha(\boldsymbol{b}\circ\pi)U^\star)\\ &= \mathsf{tr}(U\hat{g}(\alpha)\hat{g}'(\alpha)U^\star P^{-1}PU\alpha(\boldsymbol{b}\circ\pi)U^\star) \qquad (8)\end{aligned}$$

In this last expression, $U\alpha(\boldsymbol{b}\circ\pi)U^\star$ is a block diagonal matrix, whereas $PU\alpha(\boldsymbol{b}\circ\pi)U^\star$ is a block matrix.

We reiterate that the identity in Equation (8) holds for any unitary matrix U and column-permutation matrix P. For a fixed (e,α), we will be using an arbitrary fixed $U(e,\alpha)$ (any unitary transformation which converts the representation α into a block diagonal matrix). The choice of $P(e,\alpha)$ will be delicate and in Claim 4.8, we will show an existence of a permutation matrix $P(e,\alpha)$ using which we can bound $\Theta^e_{p,q,r}(\mathsf{high})$ effectively.

From this point onward, the choice of the unitary matrix does not matter as long as it converts $\alpha(\cdot)$ into a block diagonal matrix (also the arrangement of blocks along the diagonal does not matter). We are going to suppress the use of U and write:

$$U\hat{g}(\alpha) = A(\alpha), \qquad \hat{g}'(\alpha)U^\star P^{-1} = A'(\alpha) \qquad \text{and} \qquad U\alpha(\boldsymbol{b}\circ\pi)U^\star = B(\alpha)(\boldsymbol{b}).$$

Note that by Claim 2.23, the $\|\cdot\|_{\mathsf{HS}}$ of the matrices are preserved, i.e., $\|A(\alpha)\|_{\mathsf{HS}} = \|\hat{g}(\alpha)\|_{\mathsf{HS}}$ and $\|A'(\alpha)\|_{\mathsf{HS}} = \|\hat{g}'(\alpha)\|_{\mathsf{HS}}$. Coming back to the task of simplifying the expression:

$$\begin{aligned}\Theta^e_{p,q,r}(\mathsf{high}) &= \sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\mathsf{tr}(\hat{g}(\alpha)\hat{g}'(\alpha)\alpha(\boldsymbol{b}\circ\pi))\cdot\mathsf{tr}(\hat{h}(\beta)\beta(\boldsymbol{b}^{-1}))\right]\\ &= \sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\mathsf{tr}(U\hat{g}(\alpha)\hat{g}'(\alpha)U^\star P^{-1}PU\alpha(\boldsymbol{b}\circ\pi)U^\star)\cdot\mathsf{tr}(\hat{h}(\beta)\beta(\boldsymbol{b}^{-1}))\right]\\ &= \sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\mathsf{tr}(A(\alpha)A'(\alpha)PB(\alpha)(\boldsymbol{b}))\cdot\mathsf{tr}(\hat{h}(\beta)\beta(\boldsymbol{b}^{-1}))\right]\\ &= \sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[\sum_{i,k}(A(\alpha)A'(\alpha))_{ik}\cdot(PB(\alpha)(\boldsymbol{b}))_{ki}\cdot\sum_{i',k'}\hat{h}(\beta)_{i'k'}\beta(\boldsymbol{b}^{-1})_{k'i'}\right]\\ &= \sum_{\alpha,\beta}\dim(\alpha)\dim(\beta)\sum_{i,k}(A(\alpha)A'(\alpha))_{ik}\cdot\sum_{i',k'}\hat{h}(\beta)_{i'k'}\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[(PB(\alpha)(\boldsymbol{b}))_{ki}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right]\\ &= \sum_{\alpha}\dim(\alpha)\sum_{i,k}(A(\alpha)A'(\alpha))_{ik}\cdot\sum_{\beta}\dim(\beta)\sum_{i',k'}\hat{h}(\beta)_{i'k'}\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[(P(\oplus_{m=1}^t n_m\beta_m(\boldsymbol{b})))_{ki}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right],\end{aligned}$$

where β_ms are the block along the diagonal of the block diagonal matrix $B(\alpha)(\boldsymbol{b})$ with multiplicity n_m.

Consider the expectation:

$$\mathop{\mathbf{E}}_{\boldsymbol{b}}\left[(P(\oplus_{m=1}^t n_m\beta_m(\boldsymbol{b})))_{ki}\cdot\beta(\boldsymbol{b}^{-1})_{k'i'}\right].$$

For a block matrix $PB(\alpha)(\cdot)$, let $\mathcal{B}(PB(\alpha))$ be the indices (i,j) that belong to the blocks in the matrix. Let $\beta^{\mathsf{row}}_{P,U,\alpha,k}$ ($\beta^{\mathsf{col}}_{P,U,\alpha,i}$) denotes the irreducible representation of G^L present in the i^{th} column (k^{th} row) of the block matrix $PB(\alpha)(\cdot)$.[8] For a fixed (α,k) following are the only scenarios when the expectation is nonzero:

- i must be such that (k,i) belongs to some block β_m in the block matrix $PB(\alpha)(\cdot)$, i.e., $(k,i) \in \mathcal{B}(PB(\alpha))$ (as otherwise $(P(\oplus_{m=1}^{t} n_m \beta_m(\boldsymbol{b})))_{ki} = 0$).
- β must be equal to $\beta^{\mathsf{col}}_{P,U,\alpha,i}$ (Proposition 2.10). Furthermore, the entry of the matrix β_m given by (k,i) must be the "transpose" of (k',i') (Proposition 2.10). This also means that if we vary (i,k) inside a block β_m, then we get distinct (k',i') (i.e., transpose of (k,i) in that block) for which the expectation is non-zero (we will use this fact later). Thus, (α,i,k) uniquely determines (i',k') for which the expectation is non-zero. We denote this map by $(i',k') \leftarrow (\alpha,i,k)$.
- If both the above conditions are true, then the expectation is $\frac{1}{\dim(\beta_m)}$ (again, using Proposition 2.10).

Therefore, we have

$$\Theta^{e}_{p,q,r}(\mathsf{high}) = \sum_{\alpha} \dim(\alpha) \sum_{i,k} (A(\alpha)A'(\alpha))_{ik} \sum_{\beta} \dim(\beta) \sum_{i',k'} \hat{h}(\beta)_{i'k'} \mathop{\mathbf{E}}_{\boldsymbol{b}} \left[(P(\oplus_{m=1}^{t} n_m \beta_m(\boldsymbol{b})))_{ki} \cdot \beta(\boldsymbol{b}^{-1})_{k'i'} \right]$$

$$= \sum_{\alpha} \dim(\alpha) \sum_{\substack{k, \\ i|(k,i)\in\mathcal{B}(PB(\alpha)), \\ (i',k')\leftarrow(\alpha,i,k)}} (A(\alpha)A'(\alpha))_{ik} \cdot \hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,i})_{i'k'}$$

$$= \sum_{\alpha} \dim(\alpha) \sum_{\substack{k, \\ i|(k,i)\in\mathcal{B}(PB(\alpha)), \\ (i',k')\leftarrow(\alpha,i,k)}} \sum_{j} A(\alpha)_{ij} \cdot A'(\alpha)_{jk} \cdot \hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,i})_{i'k'}$$

$$= \sum_{\alpha} \sum_{j,k} \dim(\alpha) A'(\alpha)_{jk} \cdot \sum_{i|(k,i)\in\mathcal{B}(PB(\alpha))} A(\alpha)_{ij} \cdot \hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,i})_{i'k'}. \qquad \text{(rearranging)}$$

We now apply the Cauchy-Schwartz inequality twice to simplify the expression.

$$|\Theta^{e}_{p,q,r}(\mathsf{high})|^2 \leqslant \left(\sum_{\alpha} \sum_{j,k} \dim(\alpha) |A'(\alpha)_{jk}|^2 \right) \cdot \sum_{\alpha} \sum_{j,k} \dim(\alpha) \left| \sum_{\substack{i|(k,i)\in\mathcal{B}(PB(\alpha)), \\ (i',k')\leftarrow(\alpha,i,k)}} A(\alpha)_{ij} \cdot \hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,i})_{i'k'} \right|^2$$

[8] Since we allow permutation of columns of a block diagonal matrices, $\beta^{\mathsf{row}}_{P,U,\alpha,i}$ and $\beta^{\mathsf{col}}_{P,U,\alpha,i}$ may be different, and hence the superscript.

Consider the first summation,

$$\sum_{\alpha}\sum_{j,k}\dim(\alpha)|A'(\alpha)_{jk}|^2 = \sum_{\alpha}\dim(\alpha)\sum_{j,k}|A'(\alpha)_{jk}|^2 = \sum_{\alpha}\dim(\alpha)\|A'\|_{\mathsf{HS}}^2 = \sum_{\alpha}\dim(\alpha)\|\hat{g}'(\alpha)'\|_{\mathsf{HS}}^2,$$

where in the last step we use the fact that $\|A'(\alpha)\|_{\mathsf{HS}} = \|\hat{g}'(\alpha)\|_{\mathsf{HS}}$. Using Proposition 2.21, this is upper bounded by $\|g'\|_2^2$ which is at most 1. Therefore,

$$|\Theta^e_{p,q,r}(\mathsf{high})|^2 \leqslant \sum_{\alpha}\sum_{j,k}\dim(\alpha)\left|\sum_{\substack{i|(k,i)\in\mathcal{B}(PB(\alpha)),\\(i',k')\leftarrow(\alpha,i,k)}} A(\alpha)_{ij}\cdot\hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,i})_{i'k'}\right|^2.$$

By applying Cauchy-Schwartz inequality to the innermost summation, we get

$$\begin{aligned}|\Theta^e_{p,q,r}(\mathsf{high})|^2 &\leqslant \sum_{\alpha}\sum_{j,k}\dim(\alpha)\left(\sum_{\substack{i|(k,i)\in\mathcal{B}(PB(\alpha)),\\(i',k')\leftarrow(\alpha,i,k)}}|A(\alpha)_{ij}|^2\right)\left(\sum_{\substack{\tilde{i}|(k,\tilde{i})\in\mathcal{B}(PB(\alpha)),\\(i',k')\leftarrow(\alpha,\tilde{i},k)}}|\hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,\tilde{i}})_{i'k'}|^2\right)\\ &\leqslant \sum_{\alpha}\sum_{j,k}\dim(\alpha)\left(\sum_{i|(k,i)\in\mathcal{B}(PB(\alpha))}|A(\alpha)_{ij}|^2\right)\left(\sum_{\substack{\tilde{i}|(k,\tilde{i})\in\mathcal{B}(PB(\alpha)),\\(i',k')\leftarrow(\alpha,\tilde{i},k)}}|\hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,\tilde{i}})_{i'k'}|^2\right).\end{aligned}$$

Now, let's look carefully at the summation. Fix the term $|A(\alpha)_{ij}|^2$. Note that this term appears for every k such that $(k,i)\in\mathcal{B}(PB(\alpha))$. On rearranging the summation,

$$\begin{aligned}|\Theta^e_{p,q,r}(\mathsf{high})|^2 &\leqslant \sum_{\alpha}\sum_{i,j}\dim(\alpha)\sum_{k|(k,i)\in\mathcal{B}(PB(\alpha))}|A(\alpha)_{ij}|^2\left(\sum_{\substack{\tilde{i}|(k,\tilde{i})\in\mathcal{B}(PB(\alpha)),\\(i',k')\leftarrow(\alpha,\tilde{i},k)}}|\hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,\tilde{i}})_{i'k'}|^2\right)\\ &\leqslant \sum_{\alpha}\sum_{i,j}\dim(\alpha)\cdot|A(\alpha)_{ij}|^2\sum_{k|(k,i)\in\mathcal{B}(PB(\alpha))}\sum_{\substack{\tilde{i}|(k,\tilde{i})\in\mathcal{B}(PB(\alpha)),\\(i',k')\leftarrow(\alpha,\tilde{i},k)}}|\hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,\tilde{i}})_{i'k'}|^2.\end{aligned}$$

Note that in the above expression, $\beta^{\mathsf{col}}_{P,U,\alpha,\tilde{i}} = \beta^{\mathsf{col}}_{P,U,\alpha,i}$ (because of the block matrix nature of $PB(\alpha)(\cdot)$). Therefore, we have

$$|\Theta^e_{p,q,r}(\mathsf{high})|^2 \leqslant \sum_{\alpha}\sum_{i,j}\dim(\alpha)\cdot|A(\alpha)_{ij}|^2\sum_{\substack{k|(k,i)\in\mathcal{B}(PB(\alpha)),\\\tilde{i}|(k,\tilde{i})\in\mathcal{B}(PB(\alpha)),\\(i',k')\leftarrow(\alpha,\tilde{i},k)}}|\hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,i})_{i'k'}|^2.$$

As mentioned earlier, if we vary $(k,\tilde{i})$ inside a block β_m of $(P(\oplus_{m=1}^t n_m\beta_m(.)))$, then we get distinct (k',i') under the map $(i',k')\leftarrow(\alpha,\tilde{i},k)$. The last sum is precisely varying inside one of the blocks (for a fixed (α,i,j)))! Therefore,

$$|\Theta^e_{p,q,r}(\mathsf{high})|^2 \leqslant \sum_{\alpha}\sum_{i,j}\dim(\alpha)|A(\alpha)_{ij}|^2\sum_{1\leqslant i',k'\leqslant\dim(\beta^{\mathsf{col}}_{P,U,\alpha,i})}|\hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,i})_{i'k'}|^2$$

$$= \sum_{\alpha} \dim(\alpha) \sum_{i,j} |U\hat{g}(\alpha)_{ij}|^2 \cdot \|\hat{h}(\beta^{\mathsf{col}}_{P,U,\alpha,i})\|^2_{\mathsf{HS}}.$$

By taking a closer look at the expression above, it is not hard to see that there can be multiple scenarios when the expression is large. For instance, it might happen that some $\beta^{\mathsf{col}}_{P,U,\alpha,i}$s have small dimension and in this case we will not be able to get the advantage that we saw in the dictatorship test.

We avoid the above mentioned scenario by noting that when this happens then it must be the case that many distinct $\beta^{\mathsf{col}}_{P,U,\alpha,i}$s occur in the expression as we vary i. Thus, on average we can efficiently upper bound the expression, by using a careful choice of P given by the following claim, as long as $|(\pi_{uv})_{\geqslant 2}(\alpha)|$ is large.

Claim 4.8. *Let $\varepsilon_0 \in (0, \frac{1}{2}]$. Suppose $\alpha, e(u,v)$ are such that $|(\pi_{uv})_{\geqslant 2}(\alpha)| \geqslant c$, where $c \geqslant 10|G|\log(\frac{1}{\varepsilon_0})$, then there exists a column-permutation matrix $\tilde{P}$ such that*

$$\sum_{i,j} |U\hat{g}(\alpha)_{ij}|^2 \cdot \|\hat{h}(\beta^{\mathsf{col}}_{\tilde{P},U,\alpha,i})\|^2_{\mathsf{HS}} \leqslant \|\hat{g}(\alpha)\|^2_{\mathsf{HS}} \cdot \left(\varepsilon_0 + \sqrt{\max_{\beta|\dim(\beta)\geqslant c} .\|\hat{h}(\beta)\|^2_{\mathsf{HS}}}\right).$$

Proof. Fix an edge $e(u,v)$ and $\alpha = (\rho_1, \rho_2, \ldots, \rho_R)$ such that $|(\pi_{uv})_{\geqslant 2}(\alpha)| \geqslant c$. We can write α as the direct sum of irreducible representations of G^L as follows:

$$\bigotimes_{i=1}^{R} \rho_i = \bigotimes_{\ell=1}^{L} \left(\bigotimes_{j \in \pi_{uv}^{-1}(\ell)} \rho_j \right) \cong \bigotimes_{\ell=1}^{L} \underbrace{\left(\oplus_{k=1}^{t_\ell} \rho_k^\ell\right)}_{B_\ell} = \oplus_m n_m \beta_m =: U\alpha U^\star,$$

where U is an arbitrary unitary matrix which converts α into direct sum of representations in $\mathsf{Irrep}(G^L)$, and ρ_j, ρ_k^ℓ are the irreducible representations of G. The last equality is by taking tensors of one representation from each of the blocks B_ℓ. We now show that if we pick a random permutation of the columns of $U\alpha U^\star$ then it gives the desired bound.

Take a random permutation $\tilde{P}$ of the columns of $U\alpha U^\star$. For brevity, we use d_m to denote $\dim(\beta_m)$. For any fixed $i \in \dim(\alpha)$, we have

$$\begin{aligned}
\mathop{\mathbf{E}}_{\tilde{U}} \left[\|\hat{h}(\beta^{\mathsf{col}}_{\tilde{P},U,\alpha,i})\|^2_{\mathsf{HS}} \right] &= \frac{\sum_m n_m d_m \|\hat{h}(\beta_m)\|^2_{\mathsf{HS}}}{\sum_m n_m d_m} \\
&\leqslant \frac{\sqrt{\sum_m n_m d_m} \sqrt{\sum_m n_m d_m \cdot \|\hat{h}(\beta_m)\|^4_{\mathsf{HS}}}}{\sum_m n_m d_m} && \text{(Using Cauchy-Schwartz)} \\
&= \sqrt{\frac{\sum_m n_m d_m \cdot \|\hat{h}(\beta_m)\|^4_{\mathsf{HS}}}{\sum_m n_m d_m}}.
\end{aligned}$$

Since we know that $\sum_m d_m \cdot \|\hat{h}(\beta_m)\|^2_{\mathsf{HS}} \leqslant \|h\|_2^2 \leqslant 1$, we can upper bound the expression as follows:

$$\mathop{\mathbf{E}}_{\tilde{P}} \left[\|\hat{h}(\beta^{\mathsf{col}}_{\tilde{P},U,\alpha,i})\|^2_{\mathsf{HS}} \right] \leqslant \sqrt{\max_m \frac{n_m \cdot \|\hat{h}(\beta_m)\|^2_{\mathsf{HS}}}{\sum_m n_m d_m}}$$

$$\leqslant \sqrt{\max_m \left\{ \min \left\{ \frac{n_m}{\sum_m n_m d_m}, \|\hat{h}(\beta_m)\|_{\mathsf{HS}}^2 \right\} \right\}}.$$

Using Lemma 2.30, for each m, we have either $d_m \geqslant c$ or $n_m \leqslant \varepsilon_0^2 \cdot \dim(\alpha)$. Therefore, we get

$$\begin{aligned}
\mathop{\mathbf{E}}_{\tilde{P}} \left[\|\hat{h}(\beta^{\mathsf{col}}_{\tilde{P},U,\alpha,i})\|_{\mathsf{HS}}^2 \right] &\leqslant \sqrt{\varepsilon_0^2 + \max_{\beta | \dim(\beta) \geqslant c} \|\hat{h}(\beta)\|_{\mathsf{HS}}^2} \\
&\leqslant \varepsilon_0 + \sqrt{\max_{\beta | \dim(\beta) \geqslant c} \|\hat{h}(\beta)\|_{\mathsf{HS}}^2}.
\end{aligned}$$

By linearity of expectation,

$$\begin{aligned}
\mathop{\mathbf{E}}_{\tilde{P}} \left[\sum_{i,j} |U\hat{g}(\alpha)_{ij}|^2 \cdot \|\hat{h}(\beta^{\mathsf{col}}_{\tilde{P},U,\alpha,i})\|_{\mathsf{HS}}^2 \right] &= \sum_{i,j} |U\hat{g}(\alpha)_{ij}|^2 \cdot \mathop{\mathbf{E}}_{U} \left[\|\hat{h}(\beta^{\mathsf{col}}_{\tilde{P},U,\alpha,i})\|_{\mathsf{HS}}^2 \right] \\
&\leqslant \|U\hat{g}(\alpha)\|_{\mathsf{HS}}^2 \cdot \left(\varepsilon_0 + \sqrt{\max_{\beta | \dim(\beta) \geqslant c} . \|\hat{h}(\beta)\|_{\mathsf{HS}}^2} \right) \\
&= \|\hat{g}(\alpha)\|_{\mathsf{HS}}^2 \cdot \left(\varepsilon_0 + \sqrt{\max_{\beta | \dim(\beta) \geqslant c} . \|\hat{h}(\beta)\|_{\mathsf{HS}}^2} \right)
\end{aligned}$$

The existence of $\tilde{P}$, as claimed, follows from above and using the conditional expectation. □

Finishing the proof: We now proceed to upper bound the high terms. Let $\eta := C^{-d_0}$ where d_0 is the constant given in Theorem 2.2, $c := \frac{1}{\eta}$ and $\varepsilon_0 := \sqrt{\eta}$. Note that the condition on C and d_0 implies that $c \geqslant 10|G| \log(\frac{1}{\varepsilon_0})$. Next, we use the smoothness property of our Label Cover instance in order to apply Claim 4.8. With these settings of the parameters, the property says that for every $v \in \mathcal{V}$ and α such that $\dim_{\geqslant 2}(\alpha) > C$, for at least $(1-\eta)$ fraction of the neighbors $u \sim v$ of v, $|(\pi_{uv})_{\geqslant 2}(\alpha)| \geqslant c$. In what follows, we use the column-permutation matrix $\tilde{P} = P(e, \alpha)$, given by the Claim 4.8, for this setting of c and ϵ_0.

$$\begin{aligned}
\mathop{\mathbf{E}}_{(u,v)\in E} \left[|\Theta^e_{p,q,r}(\mathsf{high})|^2 \right] &\leqslant \mathop{\mathbf{E}}_{(u,v)\in E} \left[\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) > C}} \dim(\alpha) \sum_{i,j} |U\hat{g}(\alpha)_{ij}|^2 \cdot \|\hat{h}(\beta^{\mathsf{col}}_{\tilde{P},U,\alpha,i})\|_{\mathsf{HS}}^2 \right] \\
&\leqslant \mathop{\mathbf{E}}_{(u,v)\in E} \left[\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) > C}} \dim(\alpha) \cdot \|\hat{g}(\alpha)\|_{\mathsf{HS}}^2 \cdot \left(\varepsilon_0 + \sqrt{\max_{\beta | \dim(\beta) \geqslant c} \|\hat{h}(\beta)\|_{\mathsf{HS}}^2}. \right) \right] + \eta. \\
&\leqslant \mathop{\mathbf{E}}_{(u,v)\in E} \left[\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) > C}} \dim(\alpha) \cdot \|\hat{g}(\alpha)\|_{\mathsf{HS}}^2 \cdot \left(\sqrt{\max_{\beta | \dim(\beta) \geqslant c} \|\hat{h}(\beta)\|_{\mathsf{HS}}^2}. \right) \right] + \eta + \varepsilon_0 \|g\|_2^2.
\end{aligned}$$

Consider the summation,

$$\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) > C}} \dim(\alpha) \cdot \|\hat{g}(\alpha)\|_{\mathsf{HS}}^2 \cdot \left(\sqrt{\max_{\beta | \dim(\beta) \geqslant c} \|\hat{h}(\beta)\|_{\mathsf{HS}}^2}. \right)$$

$$\leqslant \sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) > C}} \dim(\alpha) \cdot \|\hat{g}(\alpha)\|_{\mathsf{HS}}^2 \cdot \left(\sqrt{\sum_{\beta | \dim(\beta) \geqslant c} \|\hat{h}(\beta)\|_{\mathsf{HS}}^2} \cdot \right)$$

$$\leqslant \sqrt{\frac{1}{c}} \left(\sum_{\substack{\alpha, \\ \dim_{\geqslant 2}(\alpha) > C}} \dim(\alpha) \cdot \|\hat{g}(\alpha)\|_{\mathsf{HS}}^2 \right) \cdot \left(\sqrt{\sum_{\beta | \dim(\beta) \geqslant c} \dim(\beta) \cdot \|\hat{h}(\beta)\|_{\mathsf{HS}}^2} \cdot \right)$$

$$\leqslant \sqrt{\frac{1}{c}} \cdot \|g\|_2^2 \cdot \|h\|_2$$

Therefore using the fact that $\|g\|_2, \|h\|_2 \leqslant 1$, we get

$$\mathop{\mathbf{E}}_{(u,v) \in E} \left[|\Theta^e_{p,q,r}(\mathsf{high})|^2 \right] \leqslant \sqrt{\frac{1}{c}} + \eta + \varepsilon_0 \leqslant 3\sqrt{\eta} \leqslant 3C^{-d_0/2} \leqslant \left(\frac{\delta}{2|G|^3} \right)^2,$$

where the last inequality follows from the choice of C. This implies,

$$\mathop{\mathbf{E}}_{(u,v) \in E} \left[|\Theta^e_{p,q,r}(\mathsf{high})| \right] \leqslant \frac{\delta}{2|G|^3},$$

as required. □

References

[AB09] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.

[ALM+98] 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.

[AM09] Per Austrin and Elchanan Mossel. Approximation resistant predicates from pairwise independence. *Computational Complexity*, 18(2):249–271, 2009.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM (JACM)*, 45(1):70–122, 1998.

[BLR93] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. volume 47, pages 549–595. Elsevier, 1993.

[BNP08] László Babai, Nikolay Nikolov, and László Pyber. Product growth and mixing in finite groups. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 248–257, 2008.

[EH05] Lars Engebretsen and Jonas Holmerin. Three-query pcps with perfect completeness over non-boolean domains. *Random Structures & Algorithms*, 27(1):46–75, 2005.

[EHR04] Lars Engebretsen, Jonas Holmerin, and Alexander Russell. Inapproximability results for equations over finite groups. *Theoretical Computer Science*, 312(1):17–45, 2004.

[FGL+96] 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.

[Gow08] W.T. Gowers. Quasirandom groups. *Combinatorics, Probability and Computing*, 17(3):363–387, 2008.

[GR02] Mikael Goldmann and Alexander Russell. The complexity of solving equations over finite groups. *Information and Computation*, 178(1):253–262, 2002.

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

[Mos10] Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis*, 19(6):1713–1756, 2010.

[O'D14] Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, 1998.

[Tan09] Linqing Tang. Conditional hardness of approximating satisfiable Max 3CSP-q. In *International Symposium on Algorithms and Computation*, pages 923–932. Springer, 2009.

[Ter99] Audrey Terras. *Fourier analysis on finite groups and applications*. Number 43. Cambridge University Press, 1999.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il