Electronic Colloquium on Computational Complexity, Report No. 88 (2012)

Covering CSPs

Irit Dinur[*] Gillat Kol[†]

Abstract

We study the *covering complexity* of constraint satisfaction problems (CSPs). The *covering number* of a CSP instance $\mathcal{C}$, denoted $\nu(\mathcal{C})$, is the smallest number of assignments to the variables, such that each constraint is satisfied by at least one of the assignments. This covering notion describes situations in which we must satisfy all the constraints, and are willing to use more than one assignment to do so. At the same time, we want to minimize the number of assignments.

We study the covering problem for different constraint predicates. We first observe that if the predicate contains an odd predicate, then it is covered by any assignment and its negation. In particular, 3CNF and 3LIN, that are hard in the max-CSP sense, are easy to cover. However, the covering problem is hard for predicates that do not contain an odd predicate:

1. For the 4LIN predicate, it is $\mathcal{NP}$-hard to decide if a given instance $\mathcal{C}$ has $\nu(\mathcal{C}) \leq 2$ or $\nu(\mathcal{C}) \geq \omega(1)$.
2. (a) We propose a framework of *covering dictatorship tests*. We design and analyze such a dictatorship test for every predicate that supports a pairwise independent distribution.
   (b) We introduce a *covering unique games conjecture*, and use it to convert the covering dictatorship tests into conditional hardness results.
3. Finally, we study a hypothesis about the hardness of covering random instances that is similar to Feige's R3SAT hypothesis. We show the following somewhat surprising implication: If our hypothesis holds for dense enough instances, then it is hard to color an $O(1)$-colorable hypergraph with a *polynomial* number of colors.

[*] Weizmann Institute of Science, Israel, and Microsoft Research, Cambridge, MA. Research partly supported by ISF and BSF grants.

[†] Weizmann Institute of Science, Israel. Research supported by ISF grant. Part of the research was done while the author was visiting Microsoft Research, Cambridge, MA.

ISSN 1433-8092

1 Introduction

We study the *covering complexity* of constraint satisfaction problems (CSPs). Let φ be a predicate, and let $\mathcal{C}$ be a φ-CSP instance, which is a set of φ-constraints over n boolean variables and their negations. The *covering number* of $\mathcal{C}$, denoted $\nu(\mathcal{C})$, is the smallest number of assignments to the variables that "*covers*" all of the constraints, i.e., such that each constraint is satisfied by at least one of the assignments. We denote by cover-φ the problem of finding the covering number of a given φ-CSP instance.

The notion of cover-CSPs differs from the standard notion of max-CSPs, as they each operate under a different restriction and try optimize a different aspect of the problem given the restriction: The notion of max-CSPs is relevant when we restrict ourselves to a single assignment and want to maximize the fraction of satisfied constraints. In contrast, the notion of a covering number is of interest when we must satisfy all or nearly all of the constraints, and are willing to use more than one assignment to do so. Our goal is then to minimize the number of needed solutions.

One example of a situation described by the covering number is the dinner party problem: You are having some friends over for dinner, and each one has different dietary constraints. You want everyone to have at least something to eat, and at the same time would like to cook as few dishes as possible. Another example is when designing a system of health care centers, each offering different services, that will be accessible and will meet the needs of all patients.

Finding the exact covering number is $\mathcal{NP}$-hard for many interesting predicates φ. Therefore, we study the hardness of approximating this value, namely minimizing the number of solutions that together cover all of the constraints. Formally, we define the following gap problem:

gap-cover-$\varphi_{c,s}$ problem: Let $c < s \in \mathbb{N}$. Given a φ-CSP instance $\mathcal{C}$, decide between

- **Yes case:** $\nu\left(\mathcal{C}\right) \leq c$.
- **No case:** $\nu\left(\mathcal{C}\right) \geq s$.

As is done for the max-CSP case, we study the covering problem for different predicates φ, and seek a characterization of predicates that are covering-hard to approximate. It turns out that the set of predicates which are covering hard to approximate is very different from the set of predicates that are hard to approximate in the max-CSP sense. In fact we show that the sets are (in a sense) incomparable.

Covering and Coloring. Covering CSPs can be viewed as a generalization of graph (or hypergraph) coloring problems. A coloring problem is given by a system of not-equal (or not-all-equal) constraints on a set of vertices. It has already been observed by [7] that a graph (hypergraph) is 2^c colorable iff there are c assignments to the variables that cover all constraints. Our new notion of covering CSPs extends that of coloring as follows. It is natural to allow an algorithm "more" colors when attempting to legally color a graph, yet, in contrast, it is usually meaningless to allow "more" alphabet symbols for satisfying a φ-CSP for a general predicate φ. The covering formulation gives a natural way in which "more colors" can be used in satisfying a φ-CSP for any φ.

We mention that the paper [7] introduces the related notion of "covering PCPs" and proves hardness of approximate hypergraph coloring by analyzing the hardness of covering the not-all-equal predicate. Interestingly, our work reveals that understanding the hardness of covering the not-all-equal predicate is central for any covering-CSP problem.

1.1 Our Results

We first observe that odd predicates φ (i.e., predicates $\varphi : \{\pm 1\}^t \to \{\pm 1\}$ for which $\forall x : \varphi(x) = -\varphi(-x)$) are easy to cover: Any pair of an assignment a and its negation $-a$ will cover the entire instance, since always either a or $-a$ causes φ to be true. Moreover, let $\mathcal{O}$ be the set of predicates φ containing an odd predicate, all the predicates $\varphi \in \mathcal{O}$ are easy. Formally, we define $\mathcal{O}$ as follows (as is customary, we view a $(-1) = (-1)^1$ value as "true"):

$$\mathcal{O} = \left\{ \varphi : \{\pm 1\}^t \to \{\pm 1\} \,\middle|\, \forall x \in \{\pm 1\}^t : \varphi(x) = -1 \,\text{or}\, \varphi(-x) = -1 \right\}.$$

Observation. *Let $\varphi \in \mathcal{O}$, and let $\mathcal{C}$ be a φ-CSP instance. Then $\nu(\mathcal{C}) \le 2$.*

In particular, $\mathsf{3CNF}$ and $\mathsf{3LIN}$ which are both very hard to approximate in the max-CSP sense, are easy in the covering sense.

1.1.1 Covering Hardness of $\mathsf{4LIN}$

In contrast to $\mathsf{3LIN}$, we show that the predicate $\varphi = \mathsf{4LIN}$, that only accepts inputs with an odd number of 1s, is $\mathcal{NP}$-hard. Formally, for $\mathsf{4LIN} : \{\pm 1\}^4 \to \{\pm 1\}$, $\mathsf{4LIN}(x_1, x_2, x_3, x_4) = x_1 x_2 x_3 x_4$, we show:

Theorem 1. *gap-cover-$\mathsf{4LIN}_{2,k}$ is $\mathcal{NP}$-hard for every $k \in \mathbb{N}$.*
Furthermore, for sufficiently small $\epsilon > 0$, the following holds: In the yes case the instance is coverable by two assignments, each of which (seperatly) satisfies $1 - \epsilon$ fraction of the constraints. In the no case, no k assignments cover more than $1 - \frac{1}{2^k} + 20\sqrt{\epsilon}$ fraction of the constraints.

Observe that the problem gap-cover-$\mathsf{4LIN}_{1,k}$ is easy for every k, as we can run a Gaussian elimination process to check whether there exists a single assignment that satisfies all the constraints.

We mention that our result can be viewed as a strengthening of Håstad's hardness result [8] for linear predicates with even arity≥ 4, since in the yes case there is a solution that satisfies $1-\epsilon$ fraction of the constraints (actually, there are at least two such solutions). Furthermore, observe that k random assignments are expected to satisfy $1-\frac{1}{2^k}$ fraction of the constraints, thus the no case shows that no k assignments can cover significantly more constraints than random k assignments.

Our proof of Theorem 1 relies on a dictatorship test whose analysis extends the analysis of [7] of the hardness of covering the 4-not-all-equal predicate, using the language of the invariance principle developed by [15, 14, 16, 17].

1.1.2 Characterization of Covering-Hard Predicates

We conjecture that for every $\varphi \notin \mathcal{O}$ the cover-φ problem is hard to approximate, and are able to partially prove this conjecture. To do so, we offer a φ-based covering dictatorship test. We then suggest a covering conjecture that corresponds to the unique games conjecture, and show how to use the dictatorship test to obtain conditional hardness of covering results.

Covering dictatorship test for a general predicate φ

We develop a general framework for covering dictatorship tests using a given predicate φ. The completeness and soundness criteria are different in the covering world:

- In the yes case, two dictators perfectly cover the test's constraints.
- In the no case, any *regular* set of functions $F = \{f_1, \ldots, f_k\}$ fails to cover all of the test's constraints. F is called regular if for any $K \subseteq [k]$ the product function $f_K = \prod_{\ell \in K} f_\ell$ is far from a dictatorial function (i.e., all its influences are low). We mention that this involved soundness condition is inherent, see Section 1.2.2.

Following Austrin and Mossel [2] we prove

Theorem 2. *Let $\varphi \notin \mathcal{O}$, and assume that there exists a balanced, pairwise independent distribution on the support of φ. Then there exists a φ-based covering-dictatorship test with completeness 2 and soundness k, for every $k \in \mathbb{N}$.*

We remark that every predicate $\varphi \notin \mathcal{O}$ that does not have degree-1 and degree-2 terms in its Fourier expansion, satisfies the condition of the theorem.

Covering unique games hardness for a general predicate φ

We suggest the following covering conjecture that corresponds to the unique games conjecture:

Conjecture 3 (Covering Unique Games). *There exists $c \in \mathbb{N}$ such that for every sufficiently small $\delta > 0$ there exists $R \in \mathbb{N}$ such that the following holds. Given a bipartite label cover instance $\mathcal{LC}$ with permutation constraints over label set $[R]$ and vertex set $U \times V$, it is $\mathcal{NP}$-hard to decide between:*

- ***Yes case:*** *There exist c assignments such that for every vertex $u \in U$, at least one of the assignments satisfies all the edges touching u.*
- ***No case:*** $\mathsf{OPT}(\mathcal{LC}) \leq \delta$*. I.e., every assignment satisfies at most δ fraction of the edge constraints.*

We mention that Khot and Regev [12] consider a similar conjecture in the max-CSP setting: In the yes case they require a single assignment that for $1 - \delta$ fraction of the vertices $u \in U$, satisfies all the edges touching u. They show that their conjecture is equivalent to the unique games conjecture. See further discussion regarding the formulation of our covering conjecture in Section 7.1.

Our conjecture is clearly false with $c = 1$, but as far as we know may be true with even $c = 2$. The conjecture is incomparable to the unique games conjecture (our completeness does not require any single assignment to satisfy a large fraction of edges). However it clearly implies the unique games conjecture with completeness $\frac{1}{c}$ (instead of $1 - \epsilon$).

As usual, we say that a problem P is covering unique games hard, if it is hard assuming Conjecture 3.

Theorem 4. *Let $\varphi \notin \mathcal{O}$, and assume that there exists a balanced, pairwise independent distribution on the support of φ. Let c be the completeness constant from the covering unique games conjecture. Then gap-cover-$\varphi_{2c,k}$ is covering unique games-hard for every $k \in \mathbb{N}$.*

1.1.3 Hardness of Approximate Coloring and Covering Random CSP Instances

We now return to the problem of approximate coloring: Given an $O(1)$-colorable graph (or hypergraph), what is the smallest number of colors needed to color it in polynomial time? This is a notorious open question with an exponential gap between known upper [9, 3, 1] and lower [7, 10, 4] bounds. One might hope that viewing this classical problem in the broader context of covering-CSPs may shed new light on it.

We show some progress in this direction, if one is willing to assume hardness of covering *random* CSP instances.

In a seminal paper, Feige [5] hypothesizes that no polynomial time algorithm is able to distinguish between a random 3SAT and a satisfiable one, and shows that this implies various hardness of approximation results. However, since a 3SAT instance is always coverable by 2 assignments, it seems impossible to derive a hardness of coloring results from Feige's hypothesis.

We formulate an analogous hypothesis about the hardness of distinguishing between random and 2-coverable 4LIN-CSP instances. We prove that if our hypothesis holds with sufficient density, it implies hardness of approximate hypergraph coloring to within *polynomial* factors. For discussion of our hypothesis see Section 8.1.

Hypothesis 5 (**Covering 4LIN Hypothesis, with density parameter** Δ)**.** *There is no polynomial time algorithm that outputs* `typical` *for most* 4LIN-*CSP instances with* n *variables and* $m = \Delta \cdot n$ *clauses, and never outputs* `typical` *for a* 2*-coverable* 4LIN-*CSP instance.*

We point out that our $\mathcal{NP}$-hardness result (Theorem 1) implies that none of the currently known algorithmic techniques can refute this hypothesis. Furthermore, the best known algorithms can only refute instances with density at least $\Delta \geq n^{0.5}$ [6].

Theorem 6. *If Hypothesis 5 holds with density parameter* $\Delta = n^{\delta}$ *for some positive* $\delta > 0$*, then it is hard to decide if a* 4*-uniform hypergraph is* 4*-colorable or requires at least a polynomial number of colors.*

In Section 8 we formulate a (weaker) hypothesis for covering a general predicate φ, and show that it has the same implication.

1.2 Technique

We face two main challenges: The first is achieving *perfect* covering completeness (being able to cover *all* the constraints vs. covering $1 - \epsilon$ of them). We introduce a technique of "duplicating" the label cover instance and design an appropriate correlated-noise dictatorship test. The basic technique is explained below, variations of it are used in the first two parts of the work. The second challenge is in handling several assignments at once when proving the soundness property. This involves solving several different problems, some of which are very roughly described below. We next present a very informal discussion of our efforts.

1.2.1 Achieving perfect completeness

Hardness of approximation results for CSPs are usually obtained through a dictatorship test for a given function $f : \{\pm 1\}^R \to \{\pm 1\}$. A typical dictatorship test involves selecting a few points in $\{\pm 1\}^R$ and then querying the function f on slight perturbations of these points. The perturbation usually involves flipping the value of each coordinate independently with small probability ϵ. While the perturbation is very effective in killing the large Fourier coefficients of f, it also "ruins" the perfect completeness, causing even a perfect dictator to be accepted with probability $1 - \epsilon$.

To overcome this problem we offer the new notion of a *duplicated* label cover instance: Given a label cover instance, each constraint $\pi_{v,u} : [R] \to [R]$ will be extended to the "duplicated" constraint $\pi_{v,u} : [2R] \to [2R]$ by

$$\forall j \in [R], \quad \pi_{v,u}(j + R) = \pi_{v,u}(j) + R.$$

This notion of a duplicated label cover will be central in our work. Observe that if $L : V \to [R]$ satisfies the constraints in the original label cover, then both L and $L + R$ satisfy the constraints in the duplicated label cover. This allows us to design a dictatorship test with enough random noise to eliminate the large Fourier coefficients, without hurting the perfect completeness. The idea is that independently for each pair of coordinates $j, j + R$, noise will be applied to at most one of the two coordinates.

1.2.2 Dealing with several proofs

When proving covering soundness in a dictatorship test we have to analyze the test's behavior on several functions at once, which means an involved rejection probability expression. This expression is basically the product of the expressions for the individual functions.

One complication arises from the fact that the test might be completely covered even if none of the functions are "dictatorial". For example, suppose that f is a random function and $f' = f \cdot x_j$. Then always either $f(x)f(y)f(z)f(-xyz) = -1$ or $f'(x)f'(y)f'(z)f'(-xyz) = -1$. This means that the natural $\mathsf{4LIN}$ test will always pass while both f and f' are completely random functions. The reason this happens is because $f \cdot f'$ is a dictatorship, forcing our analysis to consider all possible products of the given functions.

This brings about another complication, which is that even if all given functions are "folded", or balanced, their product does not have to be. This means that the empty Fourier coefficient may be large, which complicates the analysis.

Covering hardness of $\mathsf{4LIN}$. Both of the above problems were faced by [7] when analyzing the covering soundness of the NAE_4 predicate. The technique of [7] does carry over (with some adaptation originating from our correlated noise) to proving covering soundness for the $\mathsf{4LIN}$ predicate. Indeed, we use a test very similar to theirs (this test was originally suggested by Håstad [8]). However, we analyze the test in the more recent framework of the invariance principle developed by [15, 14]. This technique follows our intuition of the problem better, and is less "tailor-made" for specific predicates (indeed, in the second part of the work we use the invariance principle to show a more general result).

We mention that we cannot use the invariance principle directly, and that the usage of the invariance principle to obtain $\mathcal{NP}$-hardness results (as opposed to conditional results) is challenging. Similar difficulties were recently faced by [16, 17], and we indeed use parts of their analysis.

Covering dictatorship test for a general predicate φ. Our starting point for analyzing a general predicate φ is the work of [2], who considered any predicate φ that contains a pairwise independent distribution in its support. Their test uses independent noise, has a simpler rejection expression, and also assumes folding.

To analyze our test we rely on the following observation: every predicate $\varphi \notin \mathcal{O}$ is contained in a shifted NAE predicate (see Claim 2.2). Equipped with this observation, we bound the rejection term by exploiting the symmetry of the Fourier expansion of NAE, extending a 'pairing' trick from [7]. We mention that we cannot simply reduce the covering soundness of NAE_4 to that of φ, as the distribution used by our dictatorship test must be supported on φ for maintaining completeness. Still, as it turns out, the key for analyzing the rejection term for a general predicate $\varphi \notin \mathcal{O}$ is analyzing the same term for NAE.

Covering unique games hardness for a general predicate φ. Having developed a dictatorship test, the "straightforward" path, following [11], is to analyze it for a function f that is the average of the long-code functions $f_v : \{\pm1\}^R \to \{\pm1\}$ for all neighbors v of a given u. An influential coordinate for f implies a consistent influential coordinate for many f_v's.

In our case, however, since we have k proofs, we also have k expected functions $f_1, \dots, f_k$ for the same vertex u. If these k functions cover the dictatorship test we can only deduce that there is a product of $f_1, \dots, f_k$ that has an influential coordinate. Constructing a good assignment for the label cover instance becomes non-trivial. To solve this we must analyze the dictatorship test in the more general multi-function setting.

1.3 Future Work

This work is a first step in studying the covering complexity of CSPs, and there are many interesting directions to pursue. Aside from the obvious directions of proving more $\mathcal{NP}$-hardness results, we mention a couple of directions:

Quantitative results. Our NP-hardness result for $\mathsf{4LIN}$ implies hardness of covering a 2-coverable instance with about $\Omega(\log\log\log n)$ assignments. For the special case of coloring, better quantitative results are known, corresponding to a gap between $O(1)$ and $\Omega(\log\log n)$ [10]. On the other hand, the best algorithms require $O(\log n)$ assignments corresponding to a polynomial number of colors. Our question is: Is there any predicate φ for which one can prove $\mathcal{NP}$-hardness with a gap of $O(1)$ and $\Omega(\log n)$? In fact, any gap between $O(1)$ and $\omega(\log\log n)$ would be interesting.

More general characterization. What is the complexity of covering predicates $\varphi \notin \mathcal{O}$ that do not support a pairwise independent distribution?

Covering unique games conjecture. What can be said about the covering unique games conjecture? Can it be related to other conjectures such as Khot's d-to-1 conjecture?

Reductions between covering problems. Can one devise 'direct' reductions between covering problems? For example, does cover-NAE reduce to cover-φ for some other predicate φ? Gadget reductions simply fail in this context, and it would be interesting to find alternatives.

1.4 Organization

We begin in Section 2 with preliminaries and definitions. The covering hardness of $\mathsf{4LIN}$ is proved in Sections 3-4. The characterization of covering-hard predicates can be found in Sections 5-7, where Sections 5-6 are devoted to the covering dictatorship test, and Section 7 is devoted to the covering unique games hardness result. Finally, the relations between random CSP instances and hardness of approximate coloring are discussed in Section 8.

2 Definitions and Preliminaries

2.1 Covering Problems

Let $X = \{x_1, ..., x_n\}$ be a set of n boolean variables, each taking a value in $\{\pm 1\}$. As is customary, we view a $(-1) = (-1)^1$ value as "*true*", and a $1 = (-1)^0$ value as "*false*" (e.g., $1 \wedge (-1) = 1$). Let $\varphi : \{\pm 1\}^t \to \{\pm 1\}$ be a predicate. A *φ-constraint* over X is an equation of the form $\varphi(\sigma_1 x_{i_1}, \ldots, \sigma_t x_{i_t}) = b$, where $i_1, \ldots, i_t \in [n]$ and $b, \sigma_1, \ldots, \sigma_t \in \{\pm 1\}$. A *$\varphi$-CSP instance* $\mathcal{C}$ is a set of φ-constraints over X.

Let $\mathcal{L} \subseteq \{\pm 1\}^n$ be a set of assignments for X. We say that $\mathcal{L}$ *covers* the instance $\mathcal{C}$ if for every constraint in $\mathcal{C}$, there exists an assignment in $\mathcal{L}$ that satisfies it. The *covering number* of $\mathcal{C}$, denoted $\nu(\mathcal{C})$, is the smallest number of assignments for X such that each constraint is satisfied by at least one of the assignments. We denote by cover-φ the problem of finding the covering number of a given CSP. The gap problem is define as follows

Problem 2.1 (gap-cover-φ)**.** Let $c < s \in \mathbb{N}$, and let φ be a predicate. Given a φ-CSP instance $\mathcal{C}$, decide between

- **Yes case:** $\nu(\mathcal{C}) \leq c$. I.e., there exists a set of at most c assignments that covers C.
- **No case:** $\nu(\mathcal{C}) \geq s$. I.e., no set of at most s assignments covers C.

2.1.1 Containment in NAE

The following claim shows that the support of any predicated $\varphi \notin \mathcal{O}$ is contained in the support of NAE, upto a "sign". The claim will be very useful to us, as it allows us to move from a general predicate φ to the specific predicate NAE. Recall -1 denotes acceptance:

Claim 2.2. *For every $\varphi \notin \mathcal{O}$, $\varphi : \{\pm 1\}^t \to \{\pm 1\}$, there is a "sign" $\sigma = (\sigma_1, \ldots, \sigma_t) \in \{\pm 1\}^t$ such that $\forall x \in \{\pm 1\}^t : \varphi(\sigma_1 x_1, \ldots, \sigma_t x_t) \geq \mathsf{NAE}_t(x_1, \ldots, x_t)$.*

The claim easily follows from the fact that for a predicate $\varphi \notin \mathcal{O}$ there exists an assignment a and its negation $-a$ that are both rejected by φ. Thus, by taking $\sigma = a$ we get that $\varphi(\sigma_1 x_1, \ldots, \sigma_t x_t)$ rejects both the assignment 1^t and $(-1)^t$, and thus its support is contained in the support of NAE_t.

2.2 Label Cover

A *bipartite label cover instance* is a tuple $\mathcal{LC} = (U, V, E, R_1, R_2, \Pi)$. Here U and V are the two vertex sets of a bipartite graph, and E is the set of edges between U and V. R_1 and

R_2 satisfy $R_1 \leq R_2 \in \mathbb{N}$. $[R_1]$ is the set of labels for vertices in U, and $[R_2]$ is the set of labels for vertices in V. Π is a collection of "*projections*", one for each edge in E. That is, $\Pi = \{\pi_{v,u} : [R_2] \to [R_1]\}_{(u,v)\in E}$.

Let L be an assignment for the vertices in $U \times V$, that assigns to each vertex in U a label from $[R_1]$, and to each vertex in V a label from $[R_2]$. Let $(u, v) \in E$ be an edge. We say that L *satisfies* the edge (u, v) if $\pi_{v,u}(L(v)) = L(u)$. The value of a label cover instance $\mathcal{LC}$, denoted $\mathsf{OPT}(\mathcal{LC})$, is the maximal fraction of satisfied edges over all assignments L. It is well know that it is $\mathcal{NP}$-hard to approximate the value of a given label cover instance.

2.2.1 Smooth Label Cover

A *smooth* label cover instance is a label cover instance that satisfies the following: Let $v \in V$. In expectation over neighbors u of v, every large set of assignments for v induces a large set of assignments for u. In other words, for a sufficiently large $A \subseteq [R_2]$, it holds that $\mathbb{E}_{u\in\Gamma(v)} |\pi_{v,u}(A)|$ is large.

The following lemma gives a construction of a smooth label cover instance given a $\mathsf{3SAT}$ formula. The lemma is implied by Theorem 2.2 and Lemma 2.3 of [13] (we use their construction with $T = \lceil \frac{1}{\epsilon^4} \rceil$ and $u = r$).

Lemma 2.3. *Let $\epsilon > 0$ be a sufficiently small constant and let $r \in \mathbb{N}$ be a sufficiently large constant. There exists an efficient transformation that maps an instance ψ of $\mathsf{3SAT}$ to an instance $\mathcal{LC}_{\epsilon,r} = (U, V, E, R_1, R_2, \Pi)$ of bipartite label cover such that*

- ***Completeness:*** *If ψ is satisfiable then $\mathsf{OPT}(\mathcal{LC}_{\epsilon,r}) = 1$.*

- ***Soundness:*** *If ψ is unsatisfiable then $\mathsf{OPT}(\mathcal{LC}_{\epsilon,r}) < c_0^r$, where $c_0 \in (0, 1)$ is an absolute constant.*

- ***Smoothness:*** *For every vertex $v \in V$ and any subset of labels $A \subseteq [R_2]$ satisfying $|A| \geq \frac{1}{\epsilon^3}$, it holds that*

$$\Pr_{u\in\Gamma(v)}\left[|\pi_{v,u}(A)| \geq \frac{1}{\epsilon^2}\right] \geq 1 - 2\epsilon.$$

2.2.2 Label Cover with Permutation Constraints

Of particular interest to us are bipartite label cover instances with *permutation constraints.* Namely, where $R_1 = R_2 = R \in \mathbb{N}$ (that is, the sets of labels for U and V are the same), and $\Pi = \{\pi_{v,u} : [R] \to [R]\}_{(u,v)\in E}$ is a collection of *permutations.*

2.2.3 Duplicated Label Cover

We define the new notion of a *duplicated* label cover instance, which will play a main role in our proofs. We assume to be given a bipartite label cover instance $\mathcal{LC}' = (U, V, E, R_1, R_2, \Pi')$ with $\Pi' = \left\{\pi'_{v,u} : [R_2] \to [R_1]\right\}_{(u,v)\in E}$. The *duplicated-$\mathcal{LC}'$* instance is a new bipartite label cover instance $\mathcal{LC} = (U, V, E, R_1, 2R_2, \Pi)$, where $\Pi = \{\pi_{v,u} : [2R_2] \to [R_1]\}_{(u,v)\in E}$, and for every $(u, v) \in E$ the projection $\pi_{v,u}$ is given by:

$$j \in [R_2] : \; \pi_{v,u}(j) = \pi_{v,u}(j + R_2) = {\pi'}_{v,u}(j) .$$

In other words, to construct the duplicated instance, we double V's labels set. The new labels added are of the form $j + R_2$ for $j \in [R_2]$, and each new label $j + R_2$ "behaves" like the original label j.

When given a bipartite label cover instance $\mathcal{LC}' = (U, V, E, R, R, \Pi')$ with permutation constraints $\Pi' = \left\{\pi'_{v,u} : [R] \to [R]\right\}_{(u,v)\in E}$, we define the *unique games duplicated-$\mathcal{LC}'$* to be the new bipartite label cover instance $\mathcal{LC} = (U, V, E, 2R, 2R, \Pi)$ with permutation constraints, where $\Pi = \{\pi_{v,u} : [2R] \to [2R]\}_{(u,v)\in E}$, and for every $(u, v) \in E$ the permutation $\pi_{v,u}$ is given by:

$$j \in [R] : \; \pi_{v,u}(j) = {\pi'}_{v,u}(j) , \;\; \pi_{v,u}(j + R) = {\pi'}_{v,u}(j) + R.$$

2.3 Fourier Analysis

It is well knows that every function $f : \{\pm 1\}^n \to \mathbb{R}$ can be uniquely expressed as a multilinear polynomial (called the *Fourier expansion* of f), that is given by

$$f(x) = \sum_{S\subseteq[n]} \hat{f}(S)\, \chi_S(x) ,$$

where for every $S \subseteq [n]$ it holds that $\hat{f}(S) \in \mathbb{R}$, and χ_S is the function $\chi_S : \{\pm 1\}^n \to \{\pm 1\}$ given by $\chi_S(x) = \prod_{i\in S} x_i$.

2.3.1 Influences

Let $f : \{\pm 1\}^n \to \mathbb{R}$ be a function, and let $i \in [n]$ be a coordinate. The *influence* of coordinate i on the function f is

$$Inf_i(f) = \sum_{S:\, i\in S} \hat{f}^2(S) .$$

Let $d \in \mathbb{N}$. The *d-low-degree influence* of coordinate i on f is

$$Inf_i^{\leq d}(f) = \sum_{\substack{S:\, i \in S \\ |S| \leq d}} \hat{f}^2(S).$$

The influence $Inf_i(f)$ measures how much the function f depends on its i^{th} variable, while the low-degree influence $Inf_i^{\leq d}(f)$ measures this for the low degree part of f. An important property of low-degree influences is that the number of coordinates with a large low-degree influence must be small. In particular, we have the following claim:

Claim 2.4. *Let $d \in \mathbb{N}$, $\tau > 0$, and $f : \{\pm 1\}^n \to [-1, 1]$. It holds that*

$$\left| \left\{ i \in [n] \,\middle|\, Inf_i^{\leq d}(f) \geq \tau \right\} \right| \leq \frac{d}{\tau}.$$

2.3.2 The Bonami-Beckner Operator

We recall the *Bonami-Beckner operator* (noise operator) acting on boolean functions:

Definition 2.5. Let $\gamma \in [0, 1]$. The *Bonami-Beckner operator T_γ* is a linear operator mapping functions $f : \{\pm 1\}^n \to \mathbb{R}$ to functions $T_\gamma f : \{\pm 1\}^n \to \mathbb{R}$ via

$$T_\gamma f(x) = \mathop{\mathbb{E}}_{y}[f(xy)],$$

where in the expectation y is formed as follows: For every $i \in [n]$ (independently), we set $y_i = -1$ with probability $\frac{1}{2} - \frac{\gamma}{2}$, and set $y_i = 1$ with probability $\frac{1}{2} + \frac{\gamma}{2}$ (y_i has a bias of γ towards 1).

The operator T_γ can alternatively be defined by the following formula:

Claim 2.6. *For $\gamma \in [0, 1]$, and a function $f : \{\pm 1\}^n \to \mathbb{R}$, it holds that*

$$T_\gamma f(x) = \sum_{S \subseteq [n]} \gamma^{|S|} \hat{f}(S) \chi_S(x).$$

2.4 Correlated Probability Spaces

We say that $\left(\prod_{i \in [t]} \Omega_i, \mu\right)$ is a *finite correlated probability space* if μ is a distribution on the finite product set $\prod_{i \in [t]} \Omega_i$. Of a particular interest to us is the case where the correlated space is defined by a measure that is balanced and pairwise independent.

Definition 2.7 (**Balanceness**)**.** Let $\left(\prod_{i\in[t]}\Omega_i,\mu\right)$ be a finite correlated probability space. We say that μ is *balanced* if, for any $i\in[t]$ and $\omega\in\Omega_i$, it holds that

$$\Pr_{w\sim\mu}\left[w_i=\omega\right]=\frac{1}{|\Omega_i|}.$$

Definition 2.8 (**Pairwise Independence**)**.** Let $\left(\prod_{i\in[t]}\Omega_i,\mu\right)$ be a finite correlated probability space. We say that μ is *pairwise independent* if, for any $i\neq i'\in[t]$ and $\omega\in\Omega_i,\omega'\in\Omega_{i'}$, it holds that

$$\Pr_{w\sim\mu}\left[w_i=\omega\wedge w_{i'}=\omega'\right]=\Pr_{w\sim\mu}\left[w_i=\omega\right]\cdot\Pr_{w\sim\mu}\left[w_{i'}=\omega'\right].$$

We next recall the definition of *correlation* for correlated probability spaces, introduced by Mossel [14].

Definition 2.9 (**Correlation**)**.** Let $(\Omega\times\Psi,\mu)$ be a finite correlated probability space. Define the *correlation* between Ω and Ψ with respect to μ to be

$$\rho\left(\Omega,\Psi;\mu\right)=\max_{\substack{f:\Omega\to\mathbb{R}\\ g:\Psi\to\mathbb{R}}}\left\{\mathop{\mathbb{E}}_{(x,y)\sim\mu}\left[f\left(x\right)g\left(y\right)\right]\middle|\mathbb{E}\left[f\right]=\mathbb{E}\left[g\right]=0,\mathbb{E}\left[f^2\right]\leq1,\mathbb{E}\left[g^2\right]\leq1\right\},$$

where the expectations of f and f^2 are taken under μ's marginal on Ω, and the expectations of g and g^2 are taken under μ's marginal on Ψ.

Definition 2.10 (**Correlation**)**.** Let $\left(\prod_{i\in[t]}\Omega_i,\mu\right)$ be a finite correlated probability space. Define the *correlation* between $\Omega_1,\ldots,\Omega_t$ with respect to μ to be

$$\rho\left(\Omega_1,\ldots,\Omega_t;\mu\right)=\max_{i\in[t]}\left\{\rho\left(\Omega_i,\prod_{i'\in[t]\setminus\{i\}}\Omega_{i'};\mu\right)\right\}.$$

3 Covering Hardness of 4LIN

In this section we prove Theorem 1. For convenience we restate the theorem:

Theorem. *gap-cover-*$\mathsf{4LIN}_{2,k}$ *is $\mathcal{NP}$-hard for every $k\in\mathbb{N}$.*
Furthermore, for sufficiently small $\epsilon>0$, the following holds: In the yes case the instance is coverable by two assignments, each of which (seperatly) satisfies $1-\epsilon$ fraction of the constraints. In the no case, no k assignments cover more than $1-\frac{1}{2^k}+20\sqrt{\epsilon}$ fraction of the constraints.

3.1 PCP Verifier (Proof of Theorem 1)

As usual, we prove Theorem 1 by reduction from label cover. Specifically, we assume to be given a bipartite label cover instance $\mathcal{LC}' = \mathcal{LC}'_{\epsilon,r}$ constructed from a $\mathsf{3SAT}$ formula by the transformation described in Lemma 2.3, and construct a PCP verifier that checks proofs for $\mathcal{LC}'$ by only performing $\mathsf{4LIN}$ tests.

Let $\mathcal{LC}' = (U, V, E, R_1, R_2, \Pi')$, $\Pi' = \left\{\pi'_{v,u} : [R_2] \to [R_1]\right\}_{(u,v)\in E}$, be the given instance, and let $\mathcal{LC} = (U, V, E, R_1, 2R_2, \Pi)$, $\Pi = \{\pi_{v,u} : [2R_2] \to [R_1]\}_{(u,v)\in E}$, be the duplicated-$\mathcal{LC}'$ instance (see Section 2.2.3). A *proof* P for $\mathcal{LC}'$ consists of a collection of truth tables of boolean functions, one for each vertex $v \in V$. Formally, $P = (f_v)_{v\in V}$ where $f_v : \{\pm1\}^{2R_2} \to \{\pm1\}$. The function f_v is, supposedly, the long code encoding of the label assigned to v by a satisfying assignment for $\mathcal{LC}$.

Our verifier's algorithm for checking the proof P is found in Figure 1. The distributions $\mathcal{H}_{\epsilon,u,v,v'}$ on $\left(\{\pm1\}^{2R_2}\right)^4$ used by the verifier are specified in Section 3.2.2.

Algorithm 1 Ver_ϵ^P

- Randomly select an edge $(u, v) \in_R E$ and a neighbor $v' \in_R \Gamma(u) \subseteq V$.
- Generate a tuple $(x, y, z, w) \in \left(\{\pm1\}^{2R_2}\right)^4$ from the distribution $\mathcal{H}_{\epsilon,u,v,v'}$.
- Accept iff $f(x) f(y) g(z) g(w) = -1$,
  where f and g are the functions in P associated with vertices v and v' (respectively).

Let $\epsilon \in \left(0, \frac{1}{2}\right)$, $k \in \mathbb{N}$, and let $\mathcal{P} = \{P_1, ..., P_k\}$ be a set of any k proofs. Define $Rej\left(\mathsf{Ver}_\epsilon^{\mathcal{P}}\right)$ to be the indicator random variable for the rejection of the set of proofs $\mathcal{P}$ by Ver_ϵ. That is, $Rej\left(\mathsf{Ver}_\epsilon^{\mathcal{P}}\right)$ is 1 if none of the proofs in $\mathcal{P}$ satisfies the test selected by Ver_ϵ, and 0 if $\mathcal{P}$ contains a proof that satisfies the test.

We show that the verifier satisfies the following completeness and soundness conditions:

Lemma 3.1. Ver *satisfies the following properties:*

- ***Completeness:*** *Let* $\epsilon \in \left(0, \frac{1}{2}\right)$ *and* $r \in \mathbb{N}$. *If* $\mathsf{OPT}\left(\mathcal{LC}'_{\epsilon,r}\right) = 1$, *then there exist two proofs,* P *and* Q, *such that*

$$\Pr\left[Rej\left(\mathsf{Ver}_\epsilon^{\{P,Q\}}\right)\right] = 0.$$

That is, if there is a satisfying assignment for $\mathcal{LC}'_{\epsilon,r}$, *then there are* 2 *proofs that together cover all the tests performed by* Ver_ϵ.

Furthermore, each of the proofs P and Q is accepted by Ver_ϵ with probability $1-\epsilon$.

- ***Soundness:*** *For any sufficiently small $\epsilon \in \left(0, \frac{1}{2}\right)$ and sufficiently large $r \in \mathbb{N}$, there exist constants $\delta > 0$ and $\xi > 0$ that only depend on ϵ (e.g., $\delta = 20\sqrt{\epsilon}$ and $\xi = \epsilon^{14}$), such that for any $k \in \mathbb{N}$, the following holds: If there exists a set of proofs $\mathcal{P}$ of size at most k such that*
$$\Pr\left[Rej\left(\mathsf{Ver}_\epsilon^{\mathcal{P}}\right)\right] < \frac{1}{2^k} - \delta.$$
Then $\mathsf{OPT}\left(\mathcal{LC}'_{\epsilon,r}\right) > \xi$.
In particular, if $\mathsf{OPT}\left(\mathcal{LC}'_{\epsilon,r}\right) \leq \xi$, then there is no constant number of proofs that together cover all the tests performed by Ver_ϵ.

Note that the soundness property of Lemma 3.1 is tight in the sense that k random proofs are expected to cover all but $\frac{1}{2^k}$ fraction of the tests performed by the verifier. We show that no k proofs can do significantly better than k random proofs.

The proof of the completeness part of Lemma 3.1 can be found in Section 3.3, and the proof of soundness part can be found in Section 4. Theorem 1 follows easily from the last lemma:

Proof of Theorem 1 Let c_0 be the absolute constant from Lemma 2.3, let $\epsilon > 0$ be sufficiently small, and let $k \in \mathbb{N}$ be any constant. By taking a sufficiently large $r = r(\epsilon)$ such that $c_0^r \leq \xi = \xi(\epsilon)$, we get the following. Consider the $\mathsf{4LIN}$-CSP instance induced by the verifier when given a bipartite label cover instance $\mathcal{LC}'_{\epsilon,r}$ constructed by Lemma 2.3:

- If $\mathcal{LC}'_{\epsilon,r}$ was obtained from a satisfiable formula ψ, then $\mathsf{OPT}\left(\mathcal{LC}'_{\epsilon,r}\right) = 1$, and using the completeness property of the verifier, the required coverage by two assignments exists.

- If $\mathcal{LC}'_{\epsilon,r}$ was obtained from an unsatisfiable formula ψ, then $\mathsf{OPT}\left(\mathcal{LC}'_{\epsilon,r}\right) < c_0^r \leq \xi$, and using the soundness property of the verifier, no k proofs can cover more than $1 - \frac{1}{2^k} + \delta$ fraction of the constraints.

$\blacksquare$

3.2 Distributions

Consider the duplicated label cover instance $\mathcal{LC}$. Fix vertices $u \in U$, $v, v' \in \Gamma(u) \subseteq V$, and let $i \in [R_1]$. Let $\mathcal{X}^i, \mathcal{Y}^i = \{\pm 1\}^{\pi_{v,u}^{-1}(i)}$ and let $\mathcal{Z}^i, \mathcal{W}^i = \{\pm 1\}^{\pi_{v',u}^{-1}(i)}$. For every $i \in [R_1]$ we will have a distribution $\mathcal{H}^i_{\epsilon,u,v,v'}$ on $\Omega^i_{u,v,v'} = \mathcal{X}^i \times \mathcal{Y}^i \times \mathcal{Z}^i \times \mathcal{W}^i$. We think of this space as a correlated space in the sense of Mossel [14], written $\left(\Omega^i_{u,v,v'}; \mathcal{H}^i_{\epsilon,u,v,v'}\right)$.

We define $\mathcal{H}_{\epsilon,u,v,v'}$ to be the product distribution $\mathcal{H}_{\epsilon,u,v,v'} = \bigotimes_{i=1}^{R_1} \mathcal{H}^i_{\epsilon,u,v,v'}$ over the domain

$$\Omega_{u,v,v'} = \prod_{i=1}^{R_1} \left(\mathcal{X}^i \times \mathcal{Y}^i \times \mathcal{Z}^i \times \mathcal{W}^i\right) \cong \left(\prod_{i=1}^{R_1} \mathcal{X}^i\right) \times \left(\prod_{i=1}^{R_1} \mathcal{Y}^i\right) \times \left(\prod_{i=1}^{R_1} \mathcal{Z}^i\right) \times \left(\prod_{i=1}^{R_1} \mathcal{W}^i\right). \quad (1)$$

Again, we think of this space as a correlated space $(\Omega_{u,v,v'}; \mathcal{H}_{\epsilon,u,v,v'})$.

3.2.1 Our Noise Distribution $\mathcal{N}$

In order to define the distributions $\mathcal{H}^i_{\epsilon,u,v,v'}$, we use the following noise distribution. The distribution $\mathcal{N}_\epsilon(D)$ generates a $2D$-bits string x, such that every coordinate is 1 (noisy) with probability ϵ, but for every $j \in [D]$ it is never the case that both x_j and x_{j+D} are 1.

Definition 3.2. Let $\epsilon \in \left[0, \frac{1}{2}\right]$ and $D \in \mathbb{N}$. The distribution $\mathcal{N}_\epsilon(D)$ generates $x = (x_1, \ldots, x_{2D}) \in \{\pm 1\}^{2D}$ as follows: For every $j \in [D]$ independently,

- With probability $1 - 2\epsilon$ set $x_j = x_{j+D} = -1$.
- With probability ϵ set $x_j = -1$ and $x_{j+D} = 1$.
- With probability ϵ set $x_j = 1$ and $x_{j+D} = -1$.

The following claim bounds the noise expectation, and will be useful in the soundness analysis.

Claim 3.3. *Let $\epsilon \in \left(0, \frac{1}{2}\right)$, $D \in \mathbb{N}$ and $S \subseteq [2D]$. It holds that*

$$0 < \mathop{\mathbb{E}}_{x \sim \mathcal{N}_\epsilon(D)} \left[\chi_S(-x)\right] \leq (1 - 2\epsilon)^{|S|}.$$

Definition 3.4. Let $\epsilon \in \left(0, \frac{1}{2}\right)$. The *noise operator* N_ϵ is a linear operator mapping functions $f : \{\pm 1\}^{2n} \to \mathbb{R}$ to functions $N_\epsilon f : \{\pm 1\}^{2n} \to \mathbb{R}$ via

$$(N_\epsilon f)(x) = \mathop{\mathbb{E}}_{y \sim \mathcal{N}_{\frac{1}{2} - \frac{\epsilon}{2}}(n)} \left[f(-xy)\right].$$

We note that if a string $y \in \{\pm 1\}^{2n}$ is selected according to $\mathcal{N}_{\frac{1}{2} - \frac{\epsilon}{2}}$ then for every $i \in [n]$ (though *not* independently), it holds that $-y_i = -1$ with probability $\frac{1}{2} - \frac{\epsilon}{2}$, and $-y_i = 1$ with probability $\frac{1}{2} + \frac{\epsilon}{2}$ ($-y_i$ has a bias of ϵ towards 1).

Claim 3.5. *For $\epsilon \in \left(0, \frac{1}{2}\right)$ and a function $f : \{\pm 1\}^{2n} \to \mathbb{R}$, it holds that*

$$N_\epsilon f(x) = \sum_{S \subseteq [2n]} c_S \hat{f}(S) \chi_S(x),$$

where $0 < c_S \leq \epsilon^{|S|}$.

The proof of Claims 3.3 and 3.5 can be found in Section 3.4.

3.2.2 The Verifier's Distribution $\mathcal{H}$

Next we define the distribution $\mathcal{H}_{\epsilon,u,v,v'}$ to be used by the verifier: For $D \in \mathbb{N}$ and $a \in \{\pm 1\}$ we denote $a^D = \underbrace{a, \dots, a}_{D\,times}$ (the concatenation of a with itself D times). When given $D_1, D_2 \in \mathbb{N}$, we denote $\mathcal{X} = \mathcal{Y} = \{\pm 1\}^{2D_1}$ and $\mathcal{Z} = \mathcal{W} = \{\pm 1\}^{2D_2}$.

Definition 3.6. Let $\epsilon \in \left[0, \frac{1}{2}\right]$ and $D_1, D_2 \in \mathbb{N}$. The distribution $\mathcal{H}_\epsilon\left(D_1, D_2\right)$ generates

$$(x_1, \dots, x_{2D_1}, y_1, \dots, y_{2D_1}, z_1, \dots, z_{2D_2}, w_1, \dots, w_{2D_2}) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \times \mathcal{W}$$

as follows:

- Select the bits $x_1, \dots, x_{2D_1}, z_1, \dots, z_{2D_2}$, as well as the auxiliary bit a, independently and uniformly at random.
- Select the auxiliary bits $y'_1, \dots, y'_{2D_1}$ according to the distribution $\mathcal{N}_\epsilon\left(D_1\right)$.
  Select the auxiliary bits $w'_1, \dots, w'_{2D_2}$ according to the distribution $\mathcal{N}_\epsilon\left(D_2\right)$.
  (the bits $x_1, \dots, x_{2D_1}, z_1, \dots, z_{2D_2}, a, y'_1, \dots, y'_{D_1}, w'_1, \dots, w'_{D_2}$ are all independent).
- Set $y = -x\left(a^{2D_1} \wedge y'\right)$ and $w = -z\left(\left(-a^{2D_2}\right) \wedge w'\right)$. That is, for $j \in [2D_1]$ set $y_j = -x_j\left(a \wedge y'_j\right)$, and for $j \in [2D_2]$ set $w_j = -z_j\left(\left(-a\right) \wedge w'_j\right)$.

For $i \in [R_1]$ we define $\mathcal{H}^i_{\epsilon,u,v,v'} = \mathcal{H}_\epsilon\left(d_{i,u,v}, d_{i,u,v'}\right)$ where $d_{i,u,v} = \left|\pi^{-1}_{v,u}\left(i\right)\right|$ and $d_{i,u,v'} = \left|\pi^{-1}_{v',u}\left(i\right)\right|$. Observe that $\mathcal{H}^i_{\epsilon,u,v,v'}$ can be thought of as simply a distribution on $\left(\{\pm 1\}^{d_{i,v,u}}\right)^2 \times \left(\{\pm 1\}^{d_{i,u,v'}}\right)^2$. As mentioned above, the verifier's distribution is $\mathcal{H}_{\epsilon,u,v,v'} = \bigoplus_{i=1}^{R_1} \mathcal{H}^i_{\epsilon,u,v,v'}$.

3.2.3 The Invariant Distribution $\mathcal{I}$

Bounding the expectation of functions under the distribution $\mathcal{H}_\epsilon\left(D_1, D_2\right)$ turns out to be the key difficulty in the soundness analysis. The main reason is that there is a perfect correlation between $\mathcal{X} \times \mathcal{Y}$ and $\mathcal{Z} \times \mathcal{W}$: Given a draw (x, y, z, w) from $\mathcal{H}_\epsilon\left(D_1, D_2\right)$, one can guess the auxiliary bit a using only the pair (x, y) or using only the pair (z, w).

Our goal is to use the invariance principle to drive this correlation down to 0. To do that we pass to a distribution $\mathcal{I}_\epsilon\left(D_1, D_2\right)$ that has the same "1-wise" and "2-wise" correlations as $\mathcal{H}_\epsilon\left(D_1, D_2\right)$, but has no correlation between $\mathcal{X} \times \mathcal{Y}$ and $\mathcal{Z} \times \mathcal{W}$.

Definition 3.7. Let $\epsilon \in \left[0, \frac{1}{2}\right]$ and $D_1, D_2 \in \mathbb{N}$. The distribution $\mathcal{I}_\epsilon\left(D_1, D_2\right)$ generates

$$(x_1, \ldots, x_{D_1}, y_1, \ldots, y_{D_1}, z_1, \ldots, z_{D_2}, w_1, \ldots, w_{D_2}) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \times \mathcal{W}$$

as follows:

- Select the auxiliary bits a_1 and a_2 independently and uniformly at random.
- Select the bits $x_1, \ldots, x_{2D_1}, z_1, \ldots, z_{2D_2}, y'_1, \ldots, y'_{2D_1}, w'_1, \ldots, w'_{2D_2}$ as in $\mathcal{H}_\epsilon\left(D_1, D_2\right)$. (all bits are selected independently of a_1 and a_2).
- Set $y = -x\left(a_{\mathbf{1}}^{2D_1} \wedge y'\right)$ and $w = -z\left(a_{\mathbf{2}}^{2D_2} \wedge w'\right)$. That is, for $j \in [2D_1]$ set $y_j = -x_j\left(a_1 \wedge y'_j\right)$, and for $j \in [2D_2]$ set $w_j = -z_j\left(a_2 \wedge w'_j\right)$.

As before, for $i \in [R_1]$ we define $\mathcal{I}^i_{\epsilon,u,v,v'} = \mathcal{I}_\epsilon\left(d_{i,u,v}, d_{i,u,v'}\right)$ and $\mathcal{I}_{\epsilon,u,v,v'} = \bigoplus_{i=1}^{R_1} \mathcal{I}^i_{\epsilon,u,v,v'}$.

3.3 4LIN Completeness

In this section we prove the completeness property of Lemma 3.1. That is, we show that if there exists a satisfying assignment for $\mathcal{LC}'$, then there exist two proofs that together cover the tests of Ver.

Proof of Lemma 3.1 (Completeness) Let L be a satisfying assignment for $\mathcal{LC}'$. We construct the two proofs $P = \{f_1^v\}_{v\in V}$ and $Q = \{f_2^v\}_{v\in V}$ for Ver using the assignments L and $L + R_2$ (respectively). That is, $f_1^v, f_2^v : \{\pm 1\}^{2R_2} \to \{\pm 1\}$ satisfy $f_1^v = \chi_{L(v)}$ and $f_2^v = \chi_{L(v)+R_2}$.

Assume that the verifier selects the vertices $u \in U$ and $v, v' \in V$. Denote $i = L\left(u\right)$, $j = L\left(v\right)$ and $j' = L\left(v'\right)$. Recall that since L is a satisfying assignment it holds that $\pi_{v,u}\left(j\right) = \pi_{v',u}\left(j'\right) = i$. Let (x, y, z, w) be a possible draw from the distribution $\mathcal{H}_{\epsilon,u,v,v'}$, drawn with $\mathcal{H}^i_{\epsilon,u,v,v'}$ using the auxiliary bit a_i. Note that the tuple (x, y, z, w) induces a test "$f\left(x\right) f\left(y\right) g\left(z\right) g\left(w\right) = -1$". For $\ell \in \{1, 2\}$, denote $f_\ell = f_\ell^v$ and $g_\ell = f_\ell^{v'}$. Observe that:

- **If $a_i = 1$:** Then $f_\ell\left(x\right) \neq f_\ell\left(y\right)$ (for example, $f_1\left(x\right) = x_j \neq -x_j\left(a_i \wedge y'_j\right) = y_j = f_1\left(y\right)$). If additionally $w'_{j'} = -1$ then $g_1\left(z\right) = g_1\left(w\right)$, and if additionally $w'_{j'+R_2} = -1$ then $g_2\left(z\right) = g_2\left(w\right)$.
- **If $a_i = -1$:** Then $g_\ell\left(z\right) \neq g_\ell\left(w\right)$ (for example, $g_1\left(z\right) = z_{j'} \neq -z_{j'}\left(\left(-a_i\right) \wedge w'_{j'}\right) = w_{j'} = g_1\left(w\right)$). If additionally $y'_j = -1$ then $f_1\left(x\right) = f_1\left(y\right)$, and if additionally $y'_{j+R_2} = -1$ then $f_2\left(x\right) = f_2\left(y\right)$.

The above implies that P satisfies the test (x, y, z, w), unless ($a_i = 1$ and $w'_{j'} = 1$) or ($a_i = -1$ and $y'_j = 1$). Similarly, Q satisfies the test (x, y, z, w), unless ($a_i = 1$ and $w'_{j'+R_2} = 1$) or ($a_i = -1$ and $y'_{j+R_2} = 1$). We conclude that each of the proofs P and Q is accepted with probability $1 - \epsilon$.

We now show that at least one of the proofs P and Q satisfies the test (x, y, z, w). We assume that P does not satisfy the test and show that Q does. Assume without loss of generality that $a_i = 1$. Since P does not satisfy the test it must hold that $w'_{j'} = 1$. But since w' is selected according to the distribution $\mathcal{N}_\epsilon\left(R_2\right)$, it cannot be the case that both $w'_{j'}$ and $w'_{j'+R_2}$ are 1, thus $w'_{j'+R_2} = -1$. Since $a_i = 1$, this implies $g_2\left(z\right) = g_2\left(w\right)$ and $f_2\left(x\right) \neq f_2\left(y\right)$. We conclude that Q satisfies the test (x, y, z, w) and the assertion follows. ■

3.4 Fourier Expansion of Noise (Proof of Claims 3.3 and 3.5)

In this section we prove Claims 3.3 and 3.5. We first prove Claim 3.3. Namely, we show that for every $\epsilon \in \left(0, \frac{1}{2}\right)$, $D \in \mathbb{N}$ and $S \subseteq [2D]$ it holds that

$$0 < \mathop{\mathbb{E}}_{x \sim \mathcal{N}_\epsilon(D)} \left[\chi_S\left(-x\right)\right] \leq \left(1 - 2\epsilon\right)^{|S|}.$$

We denote $\mathcal{N} = \mathcal{N}_\epsilon\left(D\right)$. The proof uses the following definitions: For $S \subseteq [2D]$, let

$$\begin{aligned} \mathsf{p}\left(S\right) &= \left\{j \in [D] \,|\, j, j + D \in S\right\}, \\ \mathsf{s}\left(S\right) &= S \setminus \left(\mathsf{p}\left(S\right) \cup \left(\mathsf{p}\left(S\right) + D\right)\right). \end{aligned}$$

Here p stands for “pairs”, and s stands for “singles”.

Proof of Claim 3.3 It holds that

$$\mathop{\mathbb{E}}_{x \sim \mathcal{N}} \left[\chi_S\left(-x\right)\right] = \mathop{\mathbb{E}}_{x \sim \mathcal{N}} \left[\prod_{j \in S} \left(-x_j\right)\right] = \mathop{\mathbb{E}}_{x \sim \mathcal{N}} \left[\prod_{j \in \mathsf{p}(S)} \left(-x_j\right) \cdot \left(-x_{j+D}\right) \cdot \prod_{j \in \mathsf{s}(S)} \left(-x_j\right)\right].$$

Since for every $j \neq j' \in [D]$ it holds that the pair of bits $\left\{x_j, x_{j+D}\right\}$ is selected independently of the pair $\{x_{j'}, x_{j'+D}\}$, the last term can be written as

$$\begin{aligned}
&\prod_{j\in\mathsf{p}(S)} \mathop{\mathbb{E}}_{x\sim\mathcal{N}}[x_j \cdot x_{j+D}] \cdot \prod_{j\in\mathsf{s}(S)} \mathop{\mathbb{E}}_{x\sim\mathcal{N}}[-x_j] \\
= &\prod_{j\in\mathsf{p}(S)} \left[(1-2\epsilon)\cdot(-1)^2 + 2\epsilon\cdot(-1)\cdot 1\right] \cdot \prod_{j\in\mathsf{s}(S)} \left[-\left((1-\epsilon)(-1)+\epsilon\cdot 1\right)\right] \\
= &\prod_{j\in\mathsf{p}(S)} (1-4\epsilon) \cdot \prod_{j\in\mathsf{s}(S)} (1-2\epsilon) \\
= &(1-4\epsilon)^{|\mathsf{p}(S)|}(1-2\epsilon)^{|\mathsf{s}(S)|}
\end{aligned}$$

Clearly, $\mathbb{E}_{x\sim\mathcal{N}}[\chi_S(-x)] > 0$. On the other hand, observe that $1-4\epsilon \le (1-2\epsilon)^2$ and that $2|\mathsf{p}(S)| + |\mathsf{s}(S)| = |S|$. We conclude that

$$\mathop{\mathbb{E}}_{x\sim\mathcal{N}}[\chi_S(-x)] \le (1-2\epsilon)^{2|\mathsf{p}(S)|+|\mathsf{s}(S)|} \le (1-2\epsilon)^{|S|}.$$

■

We now prove Claim 3.5. That is, we show that $N_\epsilon f(x) = \sum_{S\subseteq[2n]} c_S \hat{f}(S)\chi_S(x)$ for some constants $0 < c_S \le \epsilon^{|S|}$.

Proof of Claim 3.5 For $S \subseteq [2n]$, let

$$c_s = \mathop{\mathbb{E}}_{y\sim\mathcal{N}_{\frac{1}{2}-\frac{\epsilon}{2}}(n)}[\chi_S(-y)].$$

Using the Fourier expansion of f we can write

$$\begin{aligned}
N_\epsilon f(x) &= \mathop{\mathbb{E}}_{y\sim\mathcal{N}_{\frac{1}{2}-\frac{\epsilon}{2}}(n)}[f(-xy)] \\
&= \mathop{\mathbb{E}}_{y\sim\mathcal{N}_{\frac{1}{2}-\frac{\epsilon}{2}}(n)}\left[\sum_{S\subseteq[2n]} \hat{f}(S)\cdot\chi_S(-xy)\right] \\
&= \mathop{\mathbb{E}}_{y\sim\mathcal{N}_{\frac{1}{2}-\frac{\epsilon}{2}}(n)}\left[\sum_{S\subseteq[2n]} \hat{f}(S)\cdot\chi_S(x)\cdot\chi_S(-y)\right] \\
&= \sum_{S\subseteq[2n]} \hat{f}(S)\cdot\chi_S(x)\cdot \mathop{\mathbb{E}}_{y\sim\mathcal{N}_{\frac{1}{2}-\frac{\epsilon}{2}}(n)}[\chi_S(-y)] \\
&= \sum_{S\subseteq[2n]} c_S\hat{f}(S)\chi_S(x),
\end{aligned}$$

Using Claim 3.3, it holds that

$$0 < c_s = \mathop{\mathbb{E}}_{y \sim \mathcal{N}_{\frac{1}{2}-\frac{\epsilon}{2}}(n)} \left[\chi_S(-y)\right] \leq \left(1 - 2\left(\frac{1}{2} - \frac{\epsilon}{2}\right)\right)^{|S|} = \epsilon^{|S|}.$$

■

4 4LIN Soundness

In this section we prove the soundness property of Lemma 3.1. That is, we show that if there exists a small set of proofs that cover almost all the tests performed by Ver, then $\mathsf{OPT}(\mathcal{LC}')$ is high.

For the rest of the proof we fix a sufficiently small $\epsilon \in \left(0, \frac{1}{2}\right)$ and $k, r \in \mathbb{N}$, and denote $\mathsf{Ver} = \mathsf{Ver}_\epsilon$. We use the values $\gamma = \epsilon^4$, $\tau = \epsilon$, $\delta = 20\sqrt{\epsilon}$ and $\xi = \epsilon^{14}$. For simplicity of notation, from now on we denote $\mathcal{H} = \mathcal{H}_{\epsilon,u,v,v'}$ and $\mathcal{I} = \mathcal{I}_{\epsilon,u,v,v'}$ (unless stated otherwise). When we write $\mathbb{E}_{u,v,v'}$, we mean that the expectation should be taken over the random selection of vertices u, v, v' by the verifier.

The main ingredient in the soundness proof is the following Lemma 4.2. The lemma shows that the expectation $\mathbb{E}_{u,v,v'} \mathbb{E}_{\mathcal{H}}\left[f(x) f(y) g(z) g(w)\right]$ cannot be too small, unless there is a coordinate $i \in [R_1]$ such that its pre-image by $\pi_{v,u}$ is influential for f, *and* its pre-image by $\pi_{v',u}$ is influential for g. Informally, this means that the encodings of v and v' "agree" on a label for u. In order to state Lemma 4.2, we follow the lines of [17], and define a following notion of an *influence of a set of coordinates* on a function f:

Definition 4.1. For a function $f : \{\pm 1\}^n \to \mathbb{R}$ and a subset $S \subseteq [n]$, the *influence of S on f* is

$$Inf^*_S(f) = \sum_{\substack{R \subseteq [n] \\ R \cap S \neq \phi}} \hat{f}^2(R).$$

We note that this definition is non-standard, expect for the case where $S = \{i\}$ is a singleton. In this case, $Inf^*_{\{i\}}(f) = Inf_i(f)$.

Lemma 4.2. *Let P be a proof for* Ver*, and let f and g be the functions in P associated with vertices v and v' (respectively). Assume that with probability $1 - \epsilon^5$ over the selection of vertices u, v, v' by* Ver *it holds that*

$$JointInf_{u,v,v'} \equiv \sum_{i \in [R_1]} Inf^*_{\pi^{-1}_{v,u}(i)}\left(T_{1-\gamma} f\right) \cdot Inf^*_{\pi^{-1}_{v',u}(i)}\left(T_{1-\gamma} g\right) \leq \tau.$$

Then

$$\mathop{\mathbb{E}}_{u,v,v'}\mathop{\mathbb{E}}_{\mathcal{H}}\left[f\left(x\right)f\left(y\right)g\left(z\right)g\left(w\right)\right]\geq-\delta.$$

Lemma 4.2 is proved in Section 4.1. Next, we prove the soundness property of Lemma 3.1 given Lemma 4.2.

Proof of Lemma 3.1 (Soundness) Assume that there exists a set proofs $\mathcal{P}=\{P_1,\ldots,P_k\}$ for which $Rej=\Pr\left[Rej\left(\mathsf{Ver}^{\mathcal{P}}\right)\right]<\frac{1}{2^k}-\delta$. We wish to show that $\mathsf{OPT}\left(\mathcal{LC}'\right)>\xi$.

As usual, we first arithmetize the probability that $\mathcal{P}$ passes the test. For $v,v'\in V$ and $\ell\in[k]$, let f_ℓ and g_ℓ be the functions in P_ℓ associated with vertices v and v' (respectively). For a subset $K\subseteq[k]$, let

$$f_K=\prod_{\ell\in K}f_\ell.$$

The function f_ϕ is the all 1's function, i.e., for every $x\in\{\pm1\}^{R_2}$ it holds that $f_\phi\left(x\right)=1$. We remark that we cannot use the "standard" folding technique in this proof as even if the functions $f_1,\ldots,f_k$ are all folded (i.e., odd), the function f_K may not be (e.g., if $K=\{1,2\}$ then $f_K\left(-x\right)=f_1\left(-x\right)f_2\left(-x\right)=\left(-f_1\left(x\right)\right)\cdot\left(-f_2\left(x\right)\right)=f_K\left(x\right)$).

It holds that

$$\begin{aligned}Rej &= \mathop{\mathbb{E}}_{u,v,v'}\mathop{\mathbb{E}}_{\mathcal{H}}\left[\prod_{\ell\in[k]}\frac{1}{2}\left(f_\ell\left(x\right)f_\ell\left(y\right)g_\ell\left(z\right)g_\ell\left(w\right)+1\right)\right]\\ &= \frac{1}{2^k}\mathop{\mathbb{E}}_{u,v,v'}\mathop{\mathbb{E}}_{\mathcal{H}}\left[\sum_{K\subseteq[k]}f_K\left(x\right)f_K\left(y\right)g_K\left(z\right)g_K\left(w\right)\right]\\ &= \frac{1}{2^k}\sum_{K\subseteq[k]}\mathop{\mathbb{E}}_{u,v,v'}\mathop{\mathbb{E}}_{\mathcal{H}}\left[f_K\left(x\right)f_K\left(y\right)g_K\left(z\right)g_K\left(w\right)\right].\end{aligned}$$

Let us write

$$Term_K=\mathop{\mathbb{E}}_{u,v,v'}\mathop{\mathbb{E}}_{\mathcal{H}}\left[f_K\left(x\right)f_K\left(y\right)g_K\left(z\right)g_K\left(w\right)\right],$$

and get that

$$Rej=\frac{1}{2^k}\sum_{K\subseteq[k]}Term_K.$$

For $K=\phi$ it holds that $Term_\phi=1$. Therefore, if it was the case that for every $K\neq\phi$ it holds that $Term_K\geq-\delta$ then

$$Rej=\frac{1}{2^k}\left(Term_\phi+\sum_{\phi\neq K\subseteq[k]}Term_K\right)\geq\frac{1}{2^k}\left(1+\left(2^k-1\right)\left(-\delta\right)\right)>\frac{1}{2^k}-\delta.$$

Since we assume $Rej < \frac{1}{2^k} - \delta$, there must exist $\phi \neq K \subseteq [k]$ such that $Term_K < -\delta$. We abuse notation, omit the sub-K, and write f and g instead of f_K and g_K (respectively). Using the new notations, we have

$$\mathop{\mathbb{E}}_{u,v,v'} \mathop{\mathbb{E}}_{\mathcal{H}} \left[f\left(x\right) f\left(y\right) g\left(z\right) g\left(w\right)\right] < -\delta.$$

Using Lemma 4.2, with probability at least ϵ^5 over the selection of vertices u, v, v' by the verifier, it holds that $JointInf_{u,v,v'} > \tau$.

Obtaining a good labeling. We next construct a good labeling for the duplicated label cover instance $\mathcal{LC}$. Since every assignment for $\mathcal{LC}$ naturally induces an assignment for $\mathcal{LC}'$ with the same value, the claim of the lemma follows.

Consider the following labeling L for $\mathcal{LC}$: To label $v \in V$ we select a set $S \subseteq [2R_2]$ with probability $\hat{f}^2\left(S\right)$, and set $L\left(v\right)$ to a random element of S (or an arbitrary label if $S = \phi$). To label $u \in U$ we randomly select a neighbor $v' \in_R \Gamma\left(u\right)$, and set $L\left(u\right)$ to $\pi_{v',u}\left(L\left(v'\right)\right)$.

We next show that our strategy for assigning labels satisfies a constant ξ fraction of the constraints of $\mathcal{LC}$. Let $v \in V$, and let $A \subseteq [2R_2]$ be a subset. We use the following argument of [17] to lower bound the probability that $L\left(v\right)$ is in A: It is easy to prove that for every $r \in \mathbb{R}^+$ and $\gamma \in [0,1]$ it holds that $r \geq \gamma\left(1-\gamma\right)^{\frac{1}{r}}$. Using this fact and Claim 2.6, it holds that

$$\begin{aligned}\Pr\left[L\left(v\right) \in A\right] &\geq \sum_{\substack{S \subseteq [2R_2] \\ S \cap A \neq \phi}} \hat{f}^2\left(S\right) \cdot \frac{|S \cap A|}{|S|} \geq \sum_{\substack{S \subseteq [2R_2] \\ S \cap A \neq \phi}} \hat{f}^2\left(S\right) \cdot \gamma\left(1-\gamma\right)^{\frac{|S|}{|S \cap A|}} \\ &\geq \gamma \sum_{\substack{S \subseteq [2R_2] \\ S \cap A \neq \phi}} \hat{f}^2\left(S\right) \cdot \left(1-\gamma\right)^{|S|} \geq \gamma \sum_{\substack{S \subseteq [2R_2] \\ S \cap A \neq \phi}} \left(\widehat{T_{1-\gamma} f}\right)^2\left(S\right) = \gamma Inf^*_A\left(T_{1-\gamma} f\right).\end{aligned}$$

Let $(u, v) \in E$ be an edge of $\mathcal{LC}$. The edge (u, v) is satisfied by L with probability

$$\begin{aligned}
& \Pr_{L's\, choices}\left[\pi_{v,u}\left(L\left(v\right)\right) = L\left(u\right)\right] \\
= & \mathop{\mathbb{E}}_{v' \in \Gamma(u)}\left[\sum_{i\in[R_1]} \Pr\left[L\left(v\right) \in \pi_{v,u}^{-1}\left(i\right) \wedge L\left(v'\right) \in \pi_{v',u}^{-1}\left(i\right)\right]\right] \\
= & \mathop{\mathbb{E}}_{v' \in \Gamma(u)}\left[\sum_{i\in[R_1]} \Pr\left[L\left(v\right) \in \pi_{v,u}^{-1}\left(i\right)\right] \cdot \Pr\left[L\left(v'\right) \in \pi_{v',u}^{-1}\left(i\right)\right]\right] \\
\geq & \gamma^2 \mathop{\mathbb{E}}_{v' \in \Gamma(u)}\left[\sum_{i\in[R_1]} Inf^*_{\pi_{v,u}^{-1}(i)}\left(T_{1-\gamma}f\right) \cdot Inf^*_{\pi_{v',u}^{-1}(i)}\left(T_{1-\gamma}g\right)\right] \\
= & \gamma^2 \mathop{\mathbb{E}}_{v' \in \Gamma(u)}\left[JointInf_{u,v,v'}\right].
\end{aligned}$$

Recall that with probability at least ϵ^5 over the selection of vertices u, v, v' by the verifier, it holds that $JointInf_{u,v,v'} > \tau$. We get that the probability that a random edge of $\mathcal{LC}$ is satisfied by L is

$$\mathop{\mathbb{E}}_{(u,v)\in E}\left[\Pr_{L's\, choices}\left[\pi'_{v,u}\left(L\left(v\right)\right) = L\left(u\right)\right]\right] \geq \gamma^2 \cdot \mathop{\mathbb{E}}_{(u,v)\in E}\mathop{\mathbb{E}}_{v'\in\Gamma(u)}\left[JointInf_{u,v,v'}\right] =$$
$$\gamma^2 \cdot \mathop{\mathbb{E}}_{u,v,v' \sim \mathsf{Ver}}\left[JointInf_{u,v,v'}\right] > \gamma^2\epsilon^5\tau = \epsilon^{14} = \xi.$$

■

4.1 Existence of a Joint Influential Coordinate (Proof of Lemma 4.2)

In this section we prove Lemma 4.2. Denote

$$\begin{aligned}
A_{u,v,v'} &= \mathop{\mathbb{E}}_{\mathcal{H}}\left[f\left(x\right)f\left(y\right)g\left(z\right)g\left(w\right)\right] - \mathop{\mathbb{E}}_{\mathcal{H}}\left[T_{1-\gamma}f\left(x\right)T_{1-\gamma}f\left(y\right)T_{1-\gamma}g\left(z\right)T_{1-\gamma}g\left(w\right)\right], \\
B_{u,v,v'} &= \mathop{\mathbb{E}}_{\mathcal{H}}\left[T_{1-\gamma}f\left(x\right)T_{1-\gamma}f\left(y\right)T_{1-\gamma}g\left(z\right)T_{1-\gamma}g\left(w\right)\right] - \\
&\quad \mathop{\mathbb{E}}_{\mathcal{I}}\left[T_{1-\gamma}f\left(x\right)T_{1-\gamma}f\left(y\right)T_{1-\gamma}g\left(z\right)T_{1-\gamma}g\left(w\right)\right], \\
C_{u,v,v'} &= \mathop{\mathbb{E}}_{\mathcal{I}}\left[T_{1-\gamma}f\left(x\right)T_{1-\gamma}f\left(y\right)T_{1-\gamma}g\left(z\right)T_{1-\gamma}g\left(w\right)\right].
\end{aligned}$$

Observe that

$$\begin{aligned}
\mathop{\mathbb{E}}_{u,v,v'}\mathop{\mathbb{E}}_{\mathcal{H}}\left[f\left(x\right)f\left(y\right)g\left(z\right)g\left(w\right)\right] &= \mathop{\mathbb{E}}_{u,v,v'}\mathop{\mathbb{E}}_{\mathcal{H}}\left[A_{u,v,v'} + B_{u,v,v'} + C_{u,v,v'}\right] \\
&\geq -\left|\mathop{\mathbb{E}}_{u,v,v'}\left[A_{u,v,v'}\right]\right| - \left|\mathop{\mathbb{E}}_{u,v,v'}\left[B_{u,v,v'}\right]\right| + \mathop{\mathbb{E}}_{u,v,v'}\left[C_{u,v,v'}\right].
\end{aligned}$$

To bound $\mathbb{E}_{u,v,v'}\,\mathbb{E}_{\mathcal{H}}\left[f(x)\,f(y)\,g(z)\,g(w)\right]$ we use the following three lemmas and claim. The lemmas bound each of $A_{u,v,v'}$, $B_{u,v,v'}$ and $C_{u,v,v'}$ separately.

Lemma 4.3. $\left|\mathbb{E}_{u,v,v'}\left[A_{u,v,v'}\right]\right| \leq 12\sqrt{\epsilon}$.

Lemma 4.4. *For every $u \in U$ and $v, v' \in \Gamma(u)$, it holds that $\left|B_{u,v,v'}\right| \leq JointInf_{u,v,v'}$.*

Lemma 4.5. $\mathbb{E}_{u,v,v'}\left[C_{u,v,v'}\right] \geq 0$.

Claim 4.6. *Let $u \in U$, $v \in \Gamma(u)$, and let $f : \{\pm 1\}^{2R_2} \to \{\pm 1\}$ be any function. For $\gamma \in (0,1)$, it holds that*

$$\sum_{i \in [R_1]} Inf^*_{\pi^{-1}_{v,u}(i)}\left(T_{1-\gamma} f\right) \leq \frac{1}{\gamma}.$$

Lemma 4.3 is proved in Section 4.3, and Lemma 4.4 is proved in Section 4.4. The proof of Lemma 4.5, as well as the proof of Claim 4.6 can be found at the end of this subsection. We turn to prove Lemma 4.2 using the above lemmas and claim.

Proof of Lemma 4.2 Recall that we assume that with probability $1-\epsilon^5$ over the selection of vertices u, v, v' by the verifier, it holds that $JointInf_{u,v,v'} \leq \tau$. For every function $f : \{\pm 1\}^n \to [-1,1]$ and every $S \subseteq [n]$ it holds that $0 \leq Inf^*_S(f) \leq \sum_{R \subseteq [n]} \hat{f}^2(R) \leq 1$. Therefore, using Claim 4.6, for every u, v, v', it is the case that

$$\begin{aligned} 0 \leq JointInf_{u,v,v'} &= \sum_{i \in [R_1]} Inf^*_{\pi^{-1}_{v,u}(i)}\left(T_{1-\gamma} f\right) \cdot Inf^*_{\pi^{-1}_{v',u}(i)}\left(T_{1-\gamma} g\right) \\ &\leq \sum_{i \in [R_1]} Inf^*_{\pi^{-1}_{v,u}(i)}\left(T_{1-\gamma} f\right) \leq \frac{1}{\gamma}. \end{aligned}$$

Using Lemma 4.4 (recall that we set $\gamma = \epsilon^4$ and $\tau = \epsilon$), we get

$$\begin{aligned} \left|\mathop{\mathbb{E}}_{u,v,v'}\left[B_{u,v,v'}\right]\right| &\leq \mathop{\mathbb{E}}_{u,v,v'}\left[JointInf_{u,v,v'}\right] \leq \left(1-\epsilon^5\right) \cdot \tau + \epsilon^5 \cdot \frac{1}{\gamma} \\ &< \tau + \epsilon^5 \cdot \frac{1}{\gamma} = \epsilon + \epsilon^5 \cdot \epsilon^{-4} = 2\epsilon. \end{aligned}$$

Now, using Lemmas 4.3 and 4.5 we get that

$$\begin{aligned} \mathop{\mathbb{E}}_{u,v,v'}\mathop{\mathbb{E}}_{\mathcal{H}}\left[f(x)\,f(y)\,g(z)\,g(w)\right] &\geq -\left|\mathop{\mathbb{E}}_{u,v,v'}\left[A_{u,v,v'}\right]\right| - \left|\mathop{\mathbb{E}}_{u,v,v'}\left[B_{u,v,v'}\right]\right| + \mathop{\mathbb{E}}_{u,v,v'}\left[C_{u,v,v'}\right] \\ &\geq -12\sqrt{\epsilon} - 2\epsilon + 0 \geq -20\sqrt{\epsilon} = -\delta. \end{aligned}$$

■

Proof of Lemma 4.5 Observe that for a tuple (x, y, z, w) drawn from $\mathcal{I}$ it holds that the pairs (x, y) and (z, w) are selected independently and from the same distribution. Therefore,

$$\begin{aligned}
\mathbb{E}\left[C_{u,v,v'}\right] &= \mathop{\mathbb{E}}_{u,v,v'}\left[\mathop{\mathbb{E}}_{\mathcal{I}}\left[T_{1-\gamma}f(x)\,T_{1-\gamma}f(y)\,T_{1-\gamma}g(z)\,T_{1-\gamma}g(w)\right]\right] \\
&= \mathop{\mathbb{E}}_{u,v,v'}\left[\mathop{\mathbb{E}}_{\mathcal{I}}\left[T_{1-\gamma}f(x)\,T_{1-\gamma}f(y)\right]\cdot\mathop{\mathbb{E}}_{\mathcal{I}}\left[T_{1-\gamma}g(z)\,T_{1-\gamma}g(w)\right]\right] \\
&= \mathop{\mathbb{E}}_{u}\left[\mathop{\mathbb{E}}_{v\in\Gamma(u)}\left[\mathop{\mathbb{E}}_{\mathcal{I}}\left[T_{1-\gamma}f(x)\,T_{1-\gamma}f(y)\right]\right]\cdot\mathop{\mathbb{E}}_{v'\in\Gamma(u)}\left[\mathop{\mathbb{E}}_{\mathcal{I}}\left[T_{1-\gamma}g(z)\,T_{1-\gamma}g(w)\right]\right]\right] \\
&= \mathop{\mathbb{E}}_{u}\left[\left(\mathop{\mathbb{E}}_{v\in\Gamma(u)}\left[\mathop{\mathbb{E}}_{\mathcal{I}}\left[T_{1-\gamma}f(x)\,T_{1-\gamma}f(y)\right]\right]\right)^2\right] \\
&\geq 0.
\end{aligned}$$

$\blacksquare$

Proof of Claim 4.6 Recall that for every $r \in \mathbb{R}^+$ and $\gamma \in (0,1)$ it holds that $r \geq \gamma(1-\gamma)^{\frac{1}{r}}$. Therefore it also holds that $\frac{1}{r} \geq \gamma(1-\gamma)^r \geq \gamma(1-\gamma)^{2r}$, implying $\frac{1}{\gamma} \geq r \cdot (1-\gamma)^{2r}$. Using this last inequality and Claim 2.6, we get

$$\begin{aligned}
&\sum_{i\in[R_1]} Inf^*_{\pi^{-1}_{v,u}(i)}\left(T_{1-\gamma}f\right) = \\
&\sum_{i\in[R_1]}\sum_{\substack{R\subseteq[2R_2]\\ R\cap\pi^{-1}_{v,u}(i)\neq\phi}}\left(\widehat{T_{1-\gamma}f}\right)^2(R) = \sum_{i\in[R_1]}\sum_{\substack{R\subseteq[2R_2]\\ i\in\pi_{v,u}(R)}}\left(\widehat{T_{1-\gamma}f}\right)^2(R) = \\
&\sum_{R\subseteq[2R_2]}\left(\sum_{i\in\pi_{v,u}(R)}\left(\widehat{T_{1-\gamma}f}\right)^2(R)\right) = \sum_{R\subseteq[2R_2]}\left|\pi_{v,u}(R)\right|\left(\widehat{T_{1-\gamma}f}\right)^2(R) \leq \\
&\sum_{R\subseteq[2R_2]}|R|\cdot(1-\gamma)^{2|R|}\cdot\widehat{f}^2(R) \leq \frac{1}{\gamma}\sum_{R\subseteq[2R_2]}\widehat{f}^2(R) \leq \frac{1}{\gamma}.
\end{aligned}$$

$\blacksquare$

4.2 Correlation Under $\mathcal{H}$

In this section we prove that the spaces $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, $\mathcal{W}$ are not completely correlated under $\mathcal{H}$. Formally, we prove:

Lemma 4.7. *Let $\epsilon \in \left(0, \frac{1}{2}\right)$ and $D_1, D_2 \in \mathbb{N}$. It holds that*

$$\rho\left(\mathcal{X},\mathcal{Y},\mathcal{Z},\mathcal{W};\mathcal{H}_\epsilon\left(D_1, D_2\right)\right) \leq \sqrt{1-\epsilon}.$$

The proof of Lemma 4.7 uses the following definition and lemma by Mossel [14].

Definition 4.8 (**Markov Operator, Mossel [14], Definition 2.1**)**.** Let $(\Omega \times \Psi, \mu)$ be a finite correlated probability space. The *conditional expectation operator*, or *Markov operator,* associated with the space, denoted U, maps functions $g : \Psi \to \mathbb{R}$ to functions $Ug : \Omega \to \mathbb{R}$ by

$$(Ug)(x) = \mathop{\mathbb{E}}_{(X,Y)\sim\mu} [g(Y) | X = x].$$

Lemma 4.9 (**Mossel [14], Lemma 2.8**)**.** *Let $(\Omega \times \Psi, \mu)$ be a finite correlated probability space. Let $g : \Psi \to \mathbb{R}$ be such that $\mathbb{E}_{(x,y)\sim\mu}[g(y)] = 0$ and $\mathbb{E}_{(x,y)\sim\mu}[g^2(y)] \le 1$. Then, among all the functions $f : \Omega \to \mathbb{R}$ satisfying $\mathbb{E}_{(x,y)\sim\mu}[f^2(x)] \le 1$, the maximal value of $\left|\mathbb{E}_{(x,y)\sim\mu}[f(x)g(y)]\right|$ is given by*

$$\left|\mathop{\mathbb{E}}_{(x,y)\sim\mu}[f(x)g(y)]\right| = \sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu}\left[(Ug(x))^2\right]}.$$

We are now ready to prove Lemma 4.7.

Proof of Lemma 4.7 For the sake of this proof we denote $\mathcal{H} = \mathcal{H}_\epsilon(D_1, D_2)$. We show that the correlation between $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ and $\mathcal{W}$ under $\mathcal{H}$ is not perfect, specifically that

$$\rho(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \mathcal{W}; \mathcal{H}) \le \sqrt{1-\epsilon}.$$

Since the marginal of $\mathcal{H}$ on any product of three out of the four spaces $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, $\mathcal{W}$ (e.g., $\mathcal{X} \times \mathcal{Z} \times \mathcal{W}$) is the same (up-to the order of the coordinates), the claim of the lemma follows.

Denote $\mathcal{A} = \{\pm 1\}$. We first note that $\mathcal{H}$ naturally induces a distribution $\mathcal{H}'$ on 5-tuple $(a, x, y, z, w) \in \mathcal{A} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \times \mathcal{W}$ (a is a bit, while x, y are $2D_1$-bits strings and z, w are $2D_2$-bits strings). The distribution $\mathcal{H}'$ is obtained by adding the auxiliary bit a used by $\mathcal{H}$ to the 4-tuple (x, y, z, w) drawn.

Denote

$$\begin{aligned}
\mathcal{F} &= \left\{ f : \mathcal{A} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to \mathbb{R} \,\middle|\, \mathbb{E}[f] = 0, \mathbb{E}\left[f^2\right] \le 1 \right\}, \\
\mathcal{F}' &= \{ f \in \mathcal{F} \,|\, f(a, x, y, z) \text{ only depends on } x, y, z \}, \\
\mathcal{G} &= \left\{ g : \mathcal{W} \to \mathbb{R} \,\middle|\, \mathbb{E}[g] = 0, \mathbb{E}\left[g^2\right] \le 1 \right\}.
\end{aligned}$$

For $f \in \mathcal{F}$ and $g \in \mathcal{G}$ we define

$$E_{f,g} = \mathop{\mathbb{E}}_{\mathcal{H}'}[f(a, x, y, z)\, g(w)].$$

We observe that

$$\rho\left(\mathcal{A}\times\mathcal{X}\times\mathcal{Y}\times\mathcal{Z},\mathcal{W};\mathcal{H}'\right)=\max_{f\in\mathcal{F},g\in\mathcal{G}}\left\{E_{f,g}\right\}\geq\max_{f\in\boldsymbol{\mathcal{F}'},g\in\mathcal{G}}\left\{E_{f,g}\right\}=\rho\left(\mathcal{X}\times\mathcal{Y}\times\mathcal{Z},\mathcal{W};\mathcal{H}\right).$$

We next bound $\rho\left(\mathcal{A}\times\mathcal{X}\times\mathcal{Y}\times\mathcal{Z},\mathcal{W};\mathcal{H}'\right)$. Let $f\in\mathcal{F}$ and $g\in\mathcal{G}$. Let U be the Markov operator associated with the correlated probability space $\left(\left(\mathcal{A}\times\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}\right)\times\mathcal{W},\mathcal{H}'\right)$, mapping functions on the latter space $\mathcal{W}$, to functions on the former space $\mathcal{A}\times\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$. Using Lemma 4.9, it holds that

$$E_{f,g}^2 \;=\; \left(\mathop{\mathbb{E}}_{\mathcal{H}'}\left[f\left(a,x,y,z\right)g\left(w\right)\right]\right)^2\leq\mathop{\mathbb{E}}_{\mathcal{H}'}\left[\left(Ug\left(a,x,y,z\right)\right)^2\right]. \tag{2}$$

Below, we use the notation $\left(a,x,y,z,w',w\right)\sim\mathcal{H}'$ to indicate that the 5-tuple $\left(a,x,y,z,w\right)$ was selected according to $\mathcal{H}'$ using the auxiliary string w'. Recall that $\mathcal{H}'$ selects w' according to the distribution $\mathcal{N}_\epsilon\left(D_2\right)$ and independently of a,x,y,z, and then sets $w=-z\left(\left(-a^{2D_2}\right)\wedge w'\right)$. In addition, observe that

$$\left(N_{1-2\epsilon}g\right)\left(z\right)=\mathop{\mathbb{E}}_{w'\sim\mathcal{N}_\epsilon\left(D_2\right)}\left[g\left(-zw'\right)\right].$$

Fix a tuple $\left(a,x,y,z,\cdot\right)$ in the support of $\mathcal{H}'$. We next bound the term $Ug\left(a,x,y,z\right)$. We consider the following two cases:

- **If $a=1$:**

$$\begin{aligned}
Ug\left(1,x,y,z\right) &= \mathop{\mathbb{E}}_{\left(A,X,Y,Z,W\right)\sim\mathcal{H}'}\left[g\left(W\right)\mid\left(A,X,Y,Z\right)=\left(1,x,y,z\right)\right]\\
&= \mathop{\mathbb{E}}_{\left(A,X,Y,Z,W',W\right)\sim\mathcal{H}'}\left[g\left(-z\left(\left(-1^{2D_2}\right)\wedge W'\right)\right)\mid\left(A,X,Y,Z\right)=\left(1,x,y,z\right)\right]\\
&= \mathop{\mathbb{E}}_{W'\sim\mathcal{N}_\epsilon\left(D_2\right)}\left[g\left(-z\cdot W'\right)\right]\\
&= N_{1-2\epsilon}g\left(z\right).
\end{aligned}$$

- **If $a=-1$:**

$$\begin{aligned}
Ug\left(-1,x,y,z\right) &= \mathop{\mathbb{E}}_{\left(A,X,Y,Z,W\right)\sim\mathcal{H}'}\left[g\left(W\right)\mid\left(A,X,Y,Z\right)=\left(-1,x,y,z\right)\right]\\
&= \mathop{\mathbb{E}}_{\left(A,X,Y,Z,W',W\right)\sim\mathcal{H}'}\left[g\left(-z\left(1^{2D_2}\wedge W'\right)\right)\mid\left(A,X,Y,Z\right)=\left(-1,x,y,z\right)\right]\\
&= g\left(-z\right).
\end{aligned}$$

Recall that $\hat{g}\left(\phi\right)=\mathbb{E}\left[g\right]=0$, and $\left\|g\right\|_2^2=\mathbb{E}\left[g^2\right]\leq1$ (the marginal of $\mathcal{H}$ on $\mathcal{W}$ is uniform).

We next bound $\|N_{1-2\epsilon}g\|_2^2$ using Claim 3.5 and Parseval's Theorem:

$$\begin{aligned}\|N_{1-2\epsilon}g\|_2^2 &\leq \sum_{S\subseteq[2D_2]} \left(\widehat{N_{1-2\epsilon}g}\,(S)\right)^2 \leq \sum_{S\subseteq[2D_2]} \hat{g}^2\,(S)\,(1-2\epsilon)^{2|S|} = \sum_{\substack{S\subseteq[2D_2]\\ S\neq\phi}} \hat{g}^2\,(S)\,(1-2\epsilon)^{2|S|} \\ &\leq (1-2\epsilon)^2 \sum_{S\subseteq[2D_2]} \hat{g}^2\,(S) = (1-2\epsilon)^2\,\|g\|_2^2 \leq (1-2\epsilon)^2 \leq 1-2\epsilon.\end{aligned}$$

We are now able to bound $E_{f,g}^2$. Let $\mathcal{U}$ be the uniform distribution over $2D_2$ bits. Using Equation 2 we have

$$\begin{aligned}E_{f,g}^2 &\leq \mathop{\mathbb{E}}_{\mathcal{H}'}\left[(Ug\,(a,x,y,z))^2\right] \\ &= \mathop{\mathbb{E}}_{\mathcal{H}'}\left[\frac{1+a}{2}\cdot(Ug\,(1,x,y,z))^2 + \frac{1-a}{2}\cdot(Ug\,(-1,x,y,z))^2\right] \\ &= \frac{1}{2}\mathop{\mathbb{E}}_{z\sim\mathcal{U}}\left[(N_{1-2\epsilon}g\,(z))^2\right] + \frac{1}{2}\mathop{\mathbb{E}}_{z\sim\mathcal{U}}\left[(g\,(-z))^2\right] \\ &\leq \frac{1}{2}\|N_{1-2\epsilon}g\|_2^2 + \frac{1}{2}\|g\|_2^2 \leq \frac{1}{2}(1-2\epsilon) + \frac{1}{2} = 1-\epsilon.\end{aligned}$$

Therefore

$$\rho\,(\mathcal{X}\times\mathcal{Y}\times\mathcal{Z},\mathcal{W};\mathcal{H}) \leq \sqrt{1-\epsilon}.$$

■

4.3 Applying the Bonami-Beckner Operator

In this section we prove Lemma 4.3. That is, we show that for every selection of vertices u, v, v' by the verifier, it holds that

$$\left|\mathop{\mathbb{E}}_{u,v,v'}\left[\mathop{\mathbb{E}}_{\mathcal{H}}\left[f\,(x)\,f\,(y)\,g\,(z)\,g\,(w)\right] - \mathop{\mathbb{E}}_{\mathcal{H}}\left[T_{1-\gamma}f\,(x)\,T_{1-\gamma}f\,(y)\,T_{1-\gamma}g\,(z)\,T_{1-\gamma}g\,(w)\right]\right]\right| \leq 12\sqrt{\epsilon}.$$

The lemma is in the spirit of Lemma 6.2 in [14]. Unfortunately we cannot invoke that lemma as a black box because the Bonami-Beckner operators $T_{1-\gamma}$ differ: in both our cases the operator has the form $T^{\otimes R_1}$; but the T's themselves differ. In our case T itself is an operator on a product space of some $d \in \mathbb{N}$ coordinates, and it randomizes *each* coordinate (independently) with probability γ; whereas in [14], T is an operator that leaves the input as is, or rerandomizes completely with probability γ.

A similar reason caused [16] to reprove a version of this Lemma 6.2 (they also had a second reason - the correlation between their spaces is perfect, at least according to the definition of [14]). We cannot use their proof either since it is tailored to their distribution and it uses the d-to-1 conjecture. We mention that we overcome the need of the d-to-1 conjecture by using the smoothness property of our label cover.

The proof of Lemma 4.3 follows the lines of [16], and inserts the Bonami-Beckner operators one by one. The proof also assumes knowledge of the *Efron-Stein decomposition*. Given a string x and a subset of indices S, let $x_S = (x_i)_{i \in S}$ (the bits of x that are in the set S).

Theorem 4.10 (**Efron-Stein decomposition**). *Let $(\Omega_1, \mu_1), \ldots, (\Omega_n, \mu_n)$ be n finite probability spaces, and let $(\Omega, \mu) = \prod_{i \in [n]} (\Omega_i, \mu_i)$. Every function $f : \Omega \to \mathbb{R}$ admits the following unique decomposition (called the Efron-Stein decomposition):*

$$f(x) = \sum_{S \subseteq [n]} f_S(x_S),$$

where the functions f_S satisfy

- *The function f_S depends only on x_S.*
- *For every $S \nsubseteq S'$, and every $x_{S'} \in \{\pm 1\}^{S'}$, it holds that*

$$\mathop{\mathbb{E}}_{X \sim \mu} [f_S(X_S) | X_{S'} = x_{S'}] = 0.$$

Lemma 4.3 is implied by the following lemma:

Lemma 4.11. *Each of the following four terms is bounded by $3\sqrt{\epsilon}$*

$$\left| \mathop{\mathbb{E}}_{u,v,v'} \left[\mathop{\mathbb{E}}_{\mathcal{H}} [f(x) f(y) g(z) g(w)] - \mathop{\mathbb{E}}_{\mathcal{H}} [f(x) f(y) g(z) T_{1-\gamma} g(w)] \right] \right|,$$
$$\left| \mathop{\mathbb{E}}_{u,v,v'} \left[\mathop{\mathbb{E}}_{\mathcal{H}} [f(x) f(y) g(z) T_{1-\gamma} g(w)] - \mathop{\mathbb{E}}_{\mathcal{H}} [f(x) f(y) T_{1-\gamma} g(z) T_{1-\gamma} g(w)] \right] \right|,$$
$$\left| \mathop{\mathbb{E}}_{u,v,v'} \left[\mathop{\mathbb{E}}_{\mathcal{H}} [f(x) f(y) T_{1-\gamma} g(z) T_{1-\gamma} g(w)] - \mathop{\mathbb{E}}_{\mathcal{H}} [f(x) T_{1-\gamma} f(y) T_{1-\gamma} g(z) T_{1-\gamma} g(w)] \right] \right|,$$
$$\left| \mathop{\mathbb{E}}_{u,v,v'} \left[\mathop{\mathbb{E}}_{\mathcal{H}} [f(x) T_{1-\gamma} f(y) T_{1-\gamma} g(z) T_{1-\gamma} g(w)] - \mathop{\mathbb{E}}_{\mathcal{H}} [T_{1-\gamma} f(x) T_{1-\gamma} f(y) T_{1-\gamma} g(z) T_{1-\gamma} g(w)] \right] \right|.$$

Proof We show that the first of the four terms is bounded by $3\sqrt{\epsilon}$. For $u \in U$ and $v, v' \in \Gamma(u)$, denote

$$\Delta_{u,v,v'} = \left| \mathop{\mathbb{E}}_{\mathcal{H}} [f(x) f(y) g(z) g(w)] - \mathop{\mathbb{E}}_{\mathcal{H}} [f(x) f(y) g(z) T_{1-\gamma} g(w)] \right|,$$

where f and g are the functions associated with vertices v and v' (respectively). We wish to bound $\left|\mathbb{E}_{u,v,v'}\left[\Delta_{u,v,v'}\right]\right|$.

Fix $u \in U$ and $v, v' \in \Gamma(u)$. We start by bounding $\Delta_{u,v,v'}$. Recall from Section 3.2 that $\mathcal{H}$ is a distribution on the space $\prod_{i=1}^{R_1}\left(\mathcal{X}^i \times \mathcal{Y}^i \times \mathcal{Z}^i \times \mathcal{W}^i\right)$ (Equation 1). Denote

$$\Omega = \prod_{i=1}^{R_1}\left(\mathcal{X}^i \times \mathcal{Y}^i \times \mathcal{Z}^i\right) \cong \left(\prod_{i=1}^{R_1}\mathcal{X}^i\right) \times \left(\prod_{i=1}^{R_1}\mathcal{Y}^i\right) \times \left(\prod_{i=1}^{R_1}\mathcal{Z}^i\right) \cong \left(\{\pm 1\}^{2R_2}\right)^3,$$

and

$$\Psi = \prod_{i=1}^{R_1}\mathcal{W}^i \cong \{\pm 1\}^{2R_2}.$$

Again, let us denote by U the Markov operator for the correlated probability space $(\Omega \times \Psi, \mathcal{H})$, mapping functions on Ψ to functions on Ω. It holds that

$$\begin{aligned}
\Delta_{u,v,v'} &= \left|\mathbb{E}_{\mathcal{H}}\left[f(x) f(y) g(z) \cdot \left(\left(id - T_{1-\gamma}\right) g\right)(w)\right]\right| \\
&= \left|\mathbb{E}_{(x,y,z,\cdot)\sim\mathcal{H}}\left[f(x) f(y) g(z) \cdot \mathbb{E}_{(X,Y,Z,W)\sim\mathcal{H}}\left[\left(\left(id - T_{1-\gamma}\right) g\right)(W) \,|(X,Y,Z) = (x,y,z)\right]\right]\right| \\
&= \left|\mathbb{E}_{\mathcal{H}}\left[f(x) f(y) g(z) \cdot \left(U\left(\left(id - T_{1-\gamma}\right) g\right)\right)(x,y,z)\right]\right|.
\end{aligned}$$

We now consider the quantity inside the last expectation to be a product of two functions on Ω, namely

$$F(x,y,z) = f(x) f(y) g(z)$$

and

$$G(x,y,z) = \left(U\left(\left(id - T_{1-\gamma}\right) g\right)\right)(x,y,z).$$

We take the *Efron-Stein decomposition* of these two functions with respect to the marginal of the (product) distribution $\mathcal{H}$ on the space Ω: $F(x,y,z) = \sum_{S\subseteq[R_1]} F_S(x,y,z)$ and $G(x,y,z) = \sum_{S\subseteq[R_1]} G_S(x,y,z)$. Then, by the orthogonality of the Efron-Stein decomposition and Cauchy-Schwarz, we get

$$\Delta_{u,v,v'} = \left|\sum_{S\subseteq[R_1]}\mathbb{E}_{\mathcal{H}}\left[F_S(x,y,z)\, G_S(x,y,z)\right]\right| \leq \sqrt{\sum_{S\subseteq[R_1]}\|F_S\|_2^2}\sqrt{\sum_{S\subseteq[R_1]}\|G_S\|_2^2},$$

where the 2-norms are with respect to $\mathcal{H}$'s marginal on Ω. By orthogonality again, the quantity $\sum_{S\subseteq[R_1]}\|F_S\|_2^2$ is just $\|F\|_2^2$, which is exactly 1 because F's range is $\{\pm 1\}$. Hence, we have

$$\Delta_{u,v,v'}^{2} \leq \sum_{S\subseteq[R_1]} \|G_S\|_2^2 . \tag{3}$$

As Mossel shows in Proposition 2.11 of [14], the Markov operator U commutes with taking the Efron-Stein decomposition. Therefore, if we set $G' = (id - T_{1-\gamma})\, g$ then $G = UG'$ and $G_S = (UG')_S = U\left(G'_S\right)$. Note that when we consider G'_S, the Efron-Stein decomposition is with respect to $\mathcal{H}$'s marginal distribution on Ψ, i.e., the uniform distribution. It is also easy to check that this Efron-Stein decomposition of g has

$$g_S = \sum_{\substack{R\subseteq[2R_2] \\ \pi_{v',u}(R)=S}} \hat{g}\,(R)\,\chi_R.$$

It follows that applying the Bonami-Beckner operator $T_{1-\gamma}$ to g also commutes with taking the Efron-Stein decomposition. Hence we have

$$G_S = U\left(G'_S\right) = U\left((id - T_{1-\gamma})\, g_S\right).$$

Using Equation 3, this implies

$$\Delta_{u,v,v'}^{2} \leq \sum_{S\subseteq[R_1]} \|U\left((id - T_{1-\gamma})\, g_S\right)\|_2^2 . \tag{4}$$

Proposition 2.13 of Mossel [14] shows that the correlation of a product of correlated probability spaces is equal to the maximum correlation among the individual correlated spaces. Hence, using Lemma 4.7, we get

$$\rho\,(\Omega, \Psi; \mathcal{H}) \leq \sqrt{1-\epsilon}.$$

Using Mossel's Proposition 2.12 we conclude that for each $S \subseteq [R_1]$,

$$\begin{aligned} \|U\left((id - T_{1-\gamma})\, g_S\right)\|_2^2 &\leq \left(\sqrt{1-\epsilon}\right)^{2|S|} \|(id - T_{1-\gamma})\, g_S\|_2^2 \\ &= (1-\epsilon)^{|S|} \sum_{\substack{R\subseteq[2R_2] \\ \pi_{v',u}(R)=S}} \left(1 - (1-\gamma)^{2|R|}\right) \hat{g}^2\,(R)\,, \end{aligned} \tag{5}$$

where the 2-norm in the first term is with respect to $\mathcal{H}$'s marginal on Ω, and the 2-norm in the second term is with respect to $\mathcal{H}$'s marginal on Ψ (i.e., the uniform distribution).

For $S \subseteq [R_1]$, $u \in U$ and $v' \in \Gamma(u)$ we define:

$$
\begin{aligned}
Term_{S,v',u} &= \sum_{\substack{R\subseteq[2R_2]\\ \pi_{v',u}(R)=S}} \left(1-(1-\gamma)^{2|R|}\right)\hat{g}^2(R),\\
Term^{small}_{S,v',u} &= \sum_{\substack{R\subseteq[2R_2]\\ |\mathbf{R}|\le\frac{\mathbf{2}}{\epsilon^{\mathbf{3}}},\,\pi_{v',u}(R)=S}} \left(1-(1-\gamma)^{2|R|}\right)\hat{g}^2(R),\\
Term^{large}_{S,v',u} &= \sum_{\substack{R\subseteq[2R_2]\\ |\mathbf{R}|>\frac{\mathbf{2}}{\epsilon^{\mathbf{3}}},\,\pi_{v',u}(R)=S}} \left(1-(1-\gamma)^{2|R|}\right)\hat{g}^2(R).
\end{aligned}
$$

Observe that for every u and v' it holds that

$$
0 \le \sum_{S\subseteq[R_1]} Term_{S,v',u} \le \sum_{S\subseteq[R_1]} \sum_{\substack{R\subseteq[2R_2]\\ \pi_{v',u}(R)=S}} \hat{g}^2(R) = \sum_{R\subseteq[2R_2]} \hat{g}^2(R) = 1.
$$

We now turn to bound $\left|\mathbb{E}_{u,v,v'}\left[\Delta_{u,v,v'}\right]\right|$. Due to convexity and Equations 4 and 5, it holds that

$$
\begin{aligned}
&\left(\mathop{\mathbb{E}}_{u,v,v'}\left[\Delta_{u,v,v'}\right]\right)^2 \le \mathop{\mathbb{E}}_{u,v,v'}\left[\Delta^2_{u,v,v'}\right] \le\\
&\mathop{\mathbb{E}}_{u,v,v'}\left[\sum_{S\subseteq[R_1]}(1-\epsilon)^{|S|}\sum_{\substack{R\subseteq[2R_2]\\ \pi_{v',u}(R)=S}}\left(1-(1-\gamma)^{2|R|}\right)\hat{g}^2(R)\right] \le\\
&\mathop{\mathbb{E}}_{u,v,v'}\left[\sum_{\substack{S\subseteq[R_1]\\ |S|\ge\frac{1}{\epsilon^2}}}(1-\epsilon)^{|S|}Term_{S,v',u}\right] + \mathop{\mathbb{E}}_{u,v,v'}\left[\sum_{\substack{S\subseteq[R_1]\\ |S|<\frac{1}{\epsilon^2}}}Term^{small}_{S,v',u}\right] + \mathop{\mathbb{E}}_{u,v,v'}\left[\sum_{\substack{S\subseteq[R_1]\\ |S|<\frac{1}{\epsilon^2}}}Term^{large}_{S,v',u}\right].
\end{aligned}
$$

We bound each of the last three expectations separately.

Bounding the First Expectation: Roughly speaking, in this term the sets S are large, thus $(1-\epsilon)^{|S|}$ is small, and so is the expectation. Formally, for every u and v' it holds that

$$
\sum_{\substack{S\subseteq[R_1]\\ |S|\ge\frac{1}{\epsilon^2}}}(1-\epsilon)^{|S|}Term_{S,v',u} \le (1-\epsilon)^{\frac{1}{\epsilon^2}}\sum_{\substack{S\subseteq[R_1]\\ |S|\ge\frac{1}{\epsilon^2}}}Term_{S,v',u} \le \left((1-\epsilon)^{\frac{1}{\epsilon}}\right)^{\frac{1}{\epsilon}}\cdot 1 \le \epsilon.
$$

Bounding the Second Expectation: Roughly speaking, in this term the sets R are small, thus $\left(1-(1-\gamma)^{2|R|}\right)$ is small, and so is the expectation. Formally, for every u and v' it holds that

$$\begin{aligned}
\sum_{\substack{S\subseteq[R_1]\\|S|<\frac{1}{\epsilon^2}}} Term^{small}_{S,v',u} &= \sum_{\substack{S\subseteq[R_1]\\|S|<\frac{1}{\epsilon^2}}} \sum_{\substack{R\subseteq[2R_2]\\|R|\leq\frac{2}{\epsilon^3},\pi_{v',u}(R)=S}} \left(1-(1-\gamma)^{2|R|}\right)\hat{g}^2(R)\\
&\leq \left(1-(1-\gamma)^{\frac{4}{\epsilon^3}}\right)\sum_{R\subseteq[2R_2]}\hat{g}^2(R) \leq \left(1-(1-\gamma)^{\frac{4}{\epsilon^3}}\right)\cdot 1\\
&\leq 1-\left(1-\frac{4}{\epsilon^3}\gamma\right) = \frac{4}{\epsilon^3}\cdot\epsilon^4 = 4\epsilon.
\end{aligned}$$

(Recall that for every $a\in\mathbb{R}^+$ it holds that $(1-\gamma)^a \geq 1-a\gamma$ by the generalized Bernoulli's inequality).

Bounding the Third Expectation: Roughly speaking, in this term the sets R are large, but their images $S=\pi_{v',u}(R)$ are small. Due to the smoothness property, this case is rare.

Formally, fix $v'\in V$, and let $R\subseteq[2R_2]$ be such that $|R|>\frac{2}{\epsilon^3}$. We partition R into two disjoint sets, $R'=\{j\in[R_2]\,|\,j\in R\}$ and $R''=\{j\in[R_2]\,|\,j+R_2\in R\}$. Note that either $|R'|>\frac{1}{\epsilon^3}$ or $|R''|>\frac{1}{\epsilon^3}$, and assume without loss of generality that $|R''|>\frac{1}{\epsilon^3}$. Using the smoothness property promised by Lemma 2.3, it holds that

$$\Pr_{u\in\Gamma(v')}\left[|\pi_{v',u}(R)|<\frac{1}{\epsilon^2}\right] \leq \Pr_{u\in\Gamma(v')}\left[|\pi'_{v',u}(R'')|<\frac{1}{\epsilon^2}\right] \leq 2\epsilon.$$

Therefore,

$$\begin{aligned}
\mathop{\mathbb{E}}_{u\in\Gamma(v')}\left[\sum_{\substack{S\subseteq[R_1]\\|S|<\frac{1}{\epsilon^2}}}\left[Term^{large}_{S,v',u}\right]\right] &= \mathop{\mathbb{E}}_{u\in\Gamma(v')}\left[\sum_{\substack{S\subseteq[R_1]\\|S|<\frac{1}{\epsilon^2}}}\;\sum_{\substack{R\subseteq[2R_2]\\||R|>\frac{2}{\epsilon^3},\,\pi_{v',u}(R)=S}}\left(1-(1-\gamma)^{2|R|}\right)\hat{g}^2(R)\right]\\
&\le \mathop{\mathbb{E}}_{u\in\Gamma(v')}\left[\sum_{\substack{R\subseteq[2R_2]\\|R|>\frac{2}{\epsilon^3},\,\left|\pi_{v',u}(R)\right|<\frac{1}{\epsilon^2}}}\hat{g}^2(R)\right]\\
&\le \sum_{\substack{R\subseteq[2R_2]\\|R|>\frac{2}{\epsilon^3}}}\left(\hat{g}^2(R)\cdot\Pr_{u\in\Gamma(v')}\left[\left|\pi_{v',u}(R)\right|<\frac{1}{\epsilon^2}\right]\right)\\
&\le 2\epsilon\sum_{R\subseteq[2R_2]}\hat{g}^2(R)=2\epsilon.
\end{aligned}$$

The last three bounds imply

$$\mathop{\mathbb{E}}_{u,v,v'}\left[\Delta_{u,v,v'}\right]\le\sqrt{\epsilon+4\epsilon+2\epsilon}<3\sqrt{\epsilon}.$$

This concludes the proof that the first term is bounded by $3\sqrt{\epsilon}$.

One can show that the other three terms are bounded by $3\sqrt{\epsilon}$ in a similar way. For example, in order to show that the second term is bounded by $3\sqrt{\epsilon}$ we do the following: First, we interchange the roles of w and z. The proof still goes through as the marginal of $\mathcal{H}$ on (x,y,z) is the same as its marginal on (x,y,w). Second, we use $F(x,y,w)=f(x)f(y)T_{1-\gamma}g(w)$, which still has $\|F\|_2^2\le1$, since $T_{1-\gamma}g$ is bounded in $[-1,1]$. ■

4.4 Applying the Invariance Principle (Proof of Lemma 4.4)

In this section we prove Lemma 4.4. That is, we show that for every selection of vertices u,v,v' by the verifier, it holds that

$$\begin{aligned}
&\left|\mathop{\mathbb{E}}_{\mathcal{H}}\left[T_{1-\gamma}f(x)\,T_{1-\gamma}f(y)\,T_{1-\gamma}g(z)\,T_{1-\gamma}g(w)\right]-\mathop{\mathbb{E}}_{\mathcal{I}}\left[T_{1-\gamma}f(x)\,T_{1-\gamma}f(y)\,T_{1-\gamma}g(z)\,T_{1-\gamma}g(w)\right]\right|\\
&\le JointInf_{u,v,v'}=\sum_{i\in[R_1]}Inf^*_{\pi^{-1}_{v,u}(i)}\left(T_{1-\gamma}f\right)\cdot Inf^*_{\pi^{-1}_{v',u}(i)}\left(T_{1-\gamma}g\right),
\end{aligned}$$

where f and g are the functions associated with vertices v and v' (respectively).

Again, for simplicity of notation we write $\mathcal{H}=\mathcal{H}_{\epsilon,u,v,v'}$, $\mathcal{H}^i=\mathcal{H}^i_{\epsilon,u,v,v'}$, $\mathcal{I}=\mathcal{I}_{\epsilon,u,v,v'}$ and

$\mathcal{I}^i = \mathcal{I}^i_{\epsilon,u,v,v'}$. The proof inductively changes the distribution from $\mathcal{H} = \mathcal{H}^1 \otimes \ldots \otimes \mathcal{H}^{R_1}$ to $\mathcal{I} = \mathcal{I}^1 \otimes \ldots \otimes \mathcal{I}^{R_1}$, one component at a time. For $i \in [R_1]$ and a pair of functions $f, g : \{\pm 1\}^{2R_2} \to \{\pm 1\}$, define

$$\Delta_i (f, g) = \left| \underset{\left(\bigotimes_{t=1}^{i-1} \mathcal{I}^t\right) \otimes \left(\bigotimes_{t=i}^{R_1} \mathcal{H}^t\right)}{\mathbb{E}} \left[f(x) f(y) g(z) g(w) \right] - \underset{\left(\bigotimes_{t=1}^{i} \mathcal{I}^t\right) \otimes \left(\bigotimes_{t=i+1}^{R_1} \mathcal{H}^t\right)}{\mathbb{E}} \left[f(x) f(y) g(z) g(w) \right] \right|.$$

Observe that in order to prove Lemma 4.4 it suffices to show that for every selection of vertices u, v, v' it holds that

$$\sum_{i \in [R_1]} \Delta_i (T_{1-\gamma} f, T_{1-\gamma} g) \leq \sum_{i \in [R_1]} Inf^*_{\pi^{-1}_{v,u}(i)} (T_{1-\gamma} f) \cdot Inf^*_{\pi^{-1}_{v',u}(i)} (T_{1-\gamma} g).$$

Therefore, Lemma 4.4 is implied by the following Lemma 4.12:

Lemma 4.12. *Let $u \in U$, $v, v' \in \Gamma(u)$, $i \in [R_1]$ and let $f, g : \{\pm 1\}^{2R_2} \to \{\pm 1\}$ be any pair of functions. It holds that*

$$\Delta_i (f, g) \leq Inf^*_{\pi^{-1}_{v,u}(i)} (f) \cdot Inf^*_{\pi^{-1}_{v',u}(i)} (g).$$

The rest of this subsection is devoted to proving Lemma 4.12. The proof of Lemma 4.12 uses the following definitions: Let $f : \{\pm 1\}^n \to \mathbb{R}$, $x \in \{\pm 1\}^n$ and $A = \left\{a_1, \ldots, a_{|A|}\right\} \subseteq [n]$. We define $x_A = (x_i)_{i \in A}$ and $x_{-A} = (x_i)_{i \in [n] \setminus A}$. We break the Fourier expansion of f according to its dependence on the bits x_A:

$$\begin{aligned} f(x) &= \sum_{R \subseteq [n]} \hat{f}(R) \chi_R(x) = \sum_{S \subseteq A} \left(\sum_{\substack{R \subseteq [n] \\ R \cap A = S}} \hat{f}(R) \chi_R(x_A, x_{-A}) \right) \\ &= \sum_{S \subseteq A} \chi_S(x_A) \left(\sum_{\substack{R \subseteq [n] \\ R \cap A = S}} \hat{f}(R) \chi_{R \setminus S}(x_{-A}) \right), \end{aligned}$$

where in the last term $\chi_S : \{\pm 1\}^A \to \{\pm 1\}$ and $\chi_{R \setminus S} : \{\pm 1\}^{[n] \setminus A} \to \{\pm 1\}$.

For any two subsets $S \subseteq A \subseteq [n]$ we define the function $f_{S,A} : \{\pm 1\}^{[n] \setminus A} \to \mathbb{R}$ by

$$f_{S,A}(x_{-A}) = \sum_{\substack{R \subseteq [n] \\ R \cap A = S}} \hat{f}(R) \chi_{R \setminus S}(x_{-A}). \qquad (6)$$

Note that

$$\left\| f_{S,A} \right\|_2^2 = \sum_{\substack{R \subseteq [n] \\ R \cap A = S}} \hat{f}^2 \left(R \right). \tag{7}$$

Using the functions $f_{S,A}$, we get the following decomposition:

$$f \left(x \right) = f \left(x_A, x_{-A} \right) = \sum_{S \subseteq A} f_{S,A} \left(x_{-A} \right) \chi_S \left(x_A \right). \tag{8}$$

Observe that $f_{S,A} \left(x_{-A} \right)$ can be viewed as the Fourier coefficient associated with the set S, where $f \left(x \right) = f \left(x_A, x_{-A} \right)$ is viewed as a function of x_A, and where x_{-A} is fixed.

The proof of Lemma 4.12 also uses the following Lemma 4.13, proved in Section 4.4.1.

Lemma 4.13. *Let $\epsilon \in \left(0, \frac{1}{2} \right)$ and $D_1, D_2 \in \mathbb{N}$. For $\phi \neq S \subseteq [2D_1]$ and $\phi \neq T \subseteq [2D_2]$, there exist constants $c_{T,S} \in \mathbb{R}$, $\left| c_{S,T} \right| < 1$, such that the following holds:*
For every four functions $f : \mathcal{X} \to \mathbb{R}$, $g : \mathcal{Y} \to \mathbb{R}$, $h : \mathcal{Z} \to \mathbb{R}$, $k : \mathcal{W} \to \mathbb{R}$, it holds that

$$\mathop{\mathbb{E}}_{\mathcal{H}_\epsilon(D_1,D_2)} \left[f \left(x \right) g \left(y \right) h \left(z \right) k \left(w \right) \right] - \mathop{\mathbb{E}}_{\mathcal{I}_\epsilon(D_1,D_2)} \left[f \left(x \right) g \left(y \right) h \left(z \right) k \left(w \right) \right] =$$
$$\sum_{\substack{S \subseteq [2D_1], T \subseteq [2D_2] \\ S,T \neq \phi}} c_{S,T} \hat{f} \left(S \right) \hat{g} \left(S \right) \hat{h} \left(T \right) \hat{k} \left(T \right).$$

We are now ready to prove Lemma 4.12.

Proof of Lemma 4.12 To avoid notational complication, we prove the claim just for the case $i = 1$. We write $\mathcal{H}' = \bigotimes_{t=\mathbf{2}}^{R_1} \mathcal{H}^t$, $A = \pi_{v,u}^{-1} \left(1 \right) \subseteq [2R_2]$ and $B = \pi_{v',u}^{-1} \left(1 \right) \subseteq [2R_2]$. It holds that

$$\begin{aligned} \Delta_1 \left(f, g \right) \quad \leq \quad & \left| \mathop{\mathbb{E}}_{\mathcal{H}'} \left[\mathop{\mathbb{E}}_{\mathcal{H}^1} \left[f \left(x_A, x_{-A} \right) f \left(y_A, y_{-A} \right) g \left(z_B, z_{-B} \right) g \left(w_B, w_{-B} \right) \right] \right. \right. \\ & \left. \left. - \mathop{\mathbb{E}}_{\mathcal{I}^1} \left[f \left(x_A, x_{-A} \right) f \left(y_A, y_{-A} \right) g \left(z_B, z_{-B} \right) g \left(w_B, w_{-B} \right) \right] \right] \right|. \end{aligned}$$

Note that a draw from $\mathcal{H}'$ is a tuple $\left(x_{-A}, y_{-A}, z_{-B}, w_{-B} \right)$. Using the decomposition of Equation 8,

$$f \left(x_A, x_{-A} \right) = \sum_{S \subseteq A} f_{S,A} \left(x_{-A} \right) \chi_S \left(x_A \right),$$

and Lemma 4.13 (note that $|c_{S,T}| < 1$), it holds that

$$\begin{aligned}
\Delta_1(f,g) &\leq \left| \mathop{\mathbb{E}}_{\mathcal{H}'} \left[\sum_{\substack{S \subseteq A, T \subseteq B \\ S,T \neq \phi}} c_{S,T} \left[f_{S,A}(x_{-A}) f_{S,A}(y_{-A}) g_{T,B}(z_{-B}) g_{T,B}(w_{-B}) \right] \right] \right| \\
&\leq \sum_{\substack{S \subseteq A, T \subseteq B \\ S,T \neq \phi}} \left| \mathop{\mathbb{E}}_{\mathcal{H}'} \left[f_{S,A}(x_{-A}) f_{S,A}(y_{-A}) g_{T,B}(z_{-B}) g_{T,B}(w_{-B}) \right] \right| . \qquad (9)
\end{aligned}$$

Observe that x_{-A} is independent of z_{-A}. Furthermore, observe that $y_{-A} = x_{-A} y''_{-A}$ and $w_{-B} = z_{-B} w''_{-B}$, for some y''_{-A} and w''_{-B} that are independent of both x_{-A} and z_{-B}. We consider the Fourier expansions of $f_{S,A}$ and $g_{T,B}$ using Equation 6:

$$\begin{aligned}
f_{S,A}(x_{-A}) &= \sum_{\substack{R \subseteq [2R_2] \\ R \cap A = S}} \hat{f}(R) \chi_{R \setminus S}(x_{-A}) \\
g_{T,B}(x_{-B}) &= \sum_{\substack{Q \subseteq [2R_2] \\ Q \cap B = T}} \hat{f}(Q) \chi_{Q \setminus T}(x_{-B}) .
\end{aligned}$$

For sets $R, R' \subseteq [2R_2] \setminus A$ and $Q, Q' \subseteq [2R_2] \setminus B$, we denote

$$C_{R,R',Q,Q'} = \widehat{f_{S,A}}(R) \widehat{f_{S,A}}(R') \widehat{g_{T,B}}(Q) \widehat{g_{T,B}}(Q') .$$

By substituting in Equation 9 the functions $f_{S,A}$ and $g_{T,B}$ by their Fourier expansions, we get

$$\begin{aligned}
&\Delta_1(f,g) \\
\leq\ & \sum_{\substack{S \subseteq A, T \subseteq B \\ S,T \neq \phi}} \left| \mathop{\mathbb{E}}_{\mathcal{H}'} \left[\sum_{\substack{R,R' \subseteq [2R_2] \setminus A \\ Q,Q' \subseteq [2R_2] \setminus B}} C_{R,R',Q,Q'} \cdot \chi_R(x_{-A}) \chi_{R'}(y_{-A}) \chi_Q(z_{-B}) \chi_{Q'}(w_{-B}) \right] \right| \\
=\ & \sum_{\substack{S \subseteq A, T \subseteq B \\ S,T \neq \phi}} \left| \mathop{\mathbb{E}}_{\mathcal{H}'} \left[\sum_{\substack{R,R' \subseteq [2R_2] \setminus A \\ Q,Q' \subseteq [2R_2] \setminus B}} C_{R,R',Q,Q'} \cdot \chi_{R \triangle R'}(x_{-A}) \chi_{Q \triangle Q'}(z_{-B}) \chi_{R'}\left(y''_{-A}\right) \chi_{Q'}\left(w''_{-B}\right) \right] \right| \\
=\ & \sum_{\substack{S \subseteq A, T \subseteq B \\ S,T \neq \phi}} \left| \sum_{\substack{R,R' \subseteq [2R_2] \setminus A \\ Q,Q' \subseteq [2R_2] \setminus B}} C_{R,R',Q,Q'} \cdot \mathop{\mathbb{E}}_{\mathcal{H}'} \left[\chi_{R \triangle R'}(x_{-A}) \right] \mathop{\mathbb{E}}_{\mathcal{H}'} \left[\chi_{Q \triangle Q'}(z_{-B}) \right] \mathop{\mathbb{E}}_{\mathcal{H}'} \left[\chi_{R'}\left(y''_{-A}\right) \chi_{Q'}\left(w''_{-B}\right) \right] \right| .
\end{aligned}$$

Since x_{-A} and z_{-A} are chosen uniformly at random (the marginals of $\mathcal{H}'$ are uniform), the term inside the inner sum is 0 unless $R = R'$ and $Q = Q'$. Since $\left|\mathbb{E}_{\mathcal{H}'}\left[\chi_R\left(y''_{-A}\right)\chi_Q\left(w''_{-B}\right)\right]\right| \leq 1$, we get

$$\begin{aligned}
\Delta_1(f,g) &\leq \sum_{\substack{S\subseteq A, T\subseteq B\\ S,T\neq\phi}} \sum_{\substack{R\subseteq[2R_2]\setminus A\\ Q\subseteq[2R_2]\setminus B}} \widehat{f_{S,A}}^2(R)\,\widehat{g_{T,B}}^2(Q)\\
&= \sum_{\substack{S\subseteq A, T\subseteq B\\ S,T\neq\phi}} \left(\sum_{R\subseteq[2R_2]\setminus A} \widehat{f_{S,A}}^2(R)\right)\left(\sum_{Q\subseteq[2R_2]\setminus B} \widehat{g_{T,B}}^2(Q)\right)\\
&= \sum_{\substack{S\subseteq A, T\subseteq B\\ S,T\neq\phi}} \|f_{S,A}\|_2^2\,\|g_{T,B}\|_2^2\\
&= \left(\sum_{\substack{S\subseteq A\\ S\neq\phi}} \|f_{S,A}\|_2^2\right)\left(\sum_{\substack{T\subseteq B\\ T\neq\phi}} \|g_{T,B}\|_2^2\right).
\end{aligned}$$

Observe that using Equation 7 and Definition 4.1, it holds that

$$\sum_{\substack{S\subseteq A\\ S\neq\phi}} \|f_{S,A}\|_2^2 = \sum_{\substack{S\subseteq A\\ S\neq\phi}} \sum_{\substack{R\subseteq[2R_2]\\ R\cap A=S}} \hat{f}^2(R) = \sum_{\substack{R\subseteq[2R_2]\\ R\cap A\neq\phi}} \hat{f}^2(R) = Inf_A^*(f).$$

Similarly, $\sum_{\substack{T\subseteq B\\ T\neq\phi}} \|g_{T,B}\|_2^2 = Inf_B^*(g)$, and the assertion follows. ■

4.4.1 Applying Invariance Principle to a Single Coordinate (Proof of Lemma 4.13)

In this section we prove Lemma 4.13. Fix $\epsilon \in \left(0, \frac{1}{2}\right)$ and $D_1, D_2 \in \mathbb{N}$. For the sake of this subsection we denote $\mathcal{I} = \mathcal{I}_\epsilon(D_1, D_2)$, $\mathcal{H} = \mathcal{H}_\epsilon(D_1, D_2)$, and

$$\begin{aligned}
N_S &= \mathop{\mathbb{E}}_{x\sim\mathcal{N}_\epsilon(D_1)}\left[\chi_S(-x)\right],\\
N_T &= \mathop{\mathbb{E}}_{x\sim\mathcal{N}_\epsilon(D_2)}\left[\chi_T(-x)\right],\\
E_{\mathcal{H},S,R,T,Q} &= \mathop{\mathbb{E}}_{\mathcal{H}}\left[\chi_S(x)\,\chi_R(y)\,\chi_T(z)\,\chi_Q(w)\right],\\
E_{\mathcal{I},S,R,T,Q} &= \mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_S(x)\,\chi_R(y)\,\chi_T(z)\,\chi_Q(w)\right].
\end{aligned}$$

The proof of Lemma 4.13 uses the following two claims:

Claim 4.14. *For $S, R \subseteq [2D_1]$ and $T, Q \subseteq [2D_2]$ it holds that:*

- *If $S \neq R$ or $T \neq Q$ then $E_{\mathcal{H},S,R,T,Q} = 0$*

- *If $S = R$ and $T = Q$ then* $E_{\mathcal{H},S,R,T,Q} = \frac{1}{2}\left((-1)^{|S|} N_T + (-1)^{|T|} N_S\right)$.

Claim 4.15. *For $S, R \subseteq [2D_1]$ and $T, Q \subseteq [2D_2]$ it holds that:*

- *If $S \neq R$ or $T \neq Q$ then* $E_{\mathcal{I},S,R,T,Q} = 0$.
- *If $S = R$ and $T = Q$ then* $E_{\mathcal{I},S,R,T,Q} = \frac{1}{4}\left((-1)^{|S|} + N_S\right)\left((-1)^{|T|} + N_T\right)$.

The proofs of Lemmas 4.14 and 4.15 are deferred to the end of this subsection. We are now ready to prove Lemma 4.13.

Proof of Lemma 4.13 For $S \subseteq [2D_1]$ and $T \subseteq [2D_2]$ we define $c_{S,T} \in \mathbb{R}$ to be

$$c_{S,T} = -\frac{1}{4}\left((-1)^{|S|} - N_S\right)\left((-1)^{|T|} - N_T\right).$$

Observe that for $S \neq \phi$ and $T \neq \phi$ it holds that $|c_{S,T}| < 1$, as Claim 3.3 shows that $0 < N_S \leq (1-2\epsilon)^{|S|}$ and $0 < N_T \leq (1-2\epsilon)^{|T|}$. Also, if $S = \phi$ or $T = \phi$ then $N_S = N_T = 1$ and $c_{S,T} = 0$.

It is easy to see that

$$\begin{aligned}&\mathop{\mathbb{E}}_{\mathcal{H}}\left[f(x)\,g(y)\,h(z)\,k(w)\right] - \mathop{\mathbb{E}}_{\mathcal{I}}\left[f(x)\,g(y)\,h(z)\,k(w)\right] = \\ &\sum_{\substack{S,R\subseteq[2D_1]\\ T,Q\subseteq[2D_2]}} \hat{f}(S)\,\hat{g}(R)\,\hat{h}(T)\,\hat{k}(Q)\left(E_{\mathcal{H},S,R,T,Q} - E_{\mathcal{I},S,R,T,Q}\right).\end{aligned}$$

Using Claims 4.14 and 4.15, the inner term $E_{\mathcal{H},S,R,T,Q} - E_{\mathcal{I},S,R,T,Q}$ is 0 unless $S = R$ and $T = Q$. If $S = R$ and $T = Q$ then the following easy calculation shows that $E_{\mathcal{H},S,R,T,Q} - E_{\mathcal{I},S,R,T,Q} = c_{T,Q}$:

$$\begin{aligned}
&E_{\mathcal{H},S,R,T,Q} - E_{\mathcal{I},S,R,T,Q}\\
= \;& \frac{1}{2}\left((-1)^{|S|}\cdot N_T + (-1)^{|T|}\cdot N_S\right) - \frac{1}{4}\left((-1)^{|S|} + N_S\right)\left((-1)^{|T|} + N_T\right)\\
= \;& \frac{1}{4}\left(2\cdot(-1)^{|S|}\cdot N_T + 2\cdot(-1)^{|T|}\cdot N_S - \left((-1)^{|S|+|T|} + (-1)^{|S|} N_T + (-1)^{|T|} N_S + N_S N_T\right)\right)\\
= \;& -\frac{1}{4}\left((-1)^{|S|+|T|} - (-1)^{|S|} N_T - (-1)^{|T|} N_S + N_S N_T\right)\\
= \;& -\frac{1}{4}\left((-1)^{|S|} - N_S\right)\left((-1)^{|T|} - N_T\right)\\
= \;& c_{T,Q}
\end{aligned}$$

■

Proof of Claim 4.14 Recall if (x, y, z, w) was drawn from $\mathcal{H}$ then $y = -x\left(a^{2D_1} \wedge y'\right)$ and $w = -z\left(\left(-a^{2D_2}\right) \wedge w'\right)$, where x, z, a, y', w' are all independent. Therefore

$$\begin{aligned}
E_{\mathcal{H},S,R,T,Q} &= \mathop{\mathbb{E}}_{\mathcal{H}}\left[\chi_S(x)\,\chi_R(y)\,\chi_T(z)\,\chi_Q(w)\right] \\
&= \mathop{\mathbb{E}}_{\mathcal{H}}\left[\chi_S(x)\,\chi_R\left(-x\left(a^{2D_1} \wedge y'\right)\right)\chi_T(z)\,\chi_Q\left(-z\left(\left(-a^{2D_2}\right) \wedge w'\right)\right)\right] \\
&= \mathop{\mathbb{E}}_{\mathcal{H}}\left[\chi_{S\triangle R}(x)\,\chi_R\left(-\left(a^{2D_1} \wedge y'\right)\right)\chi_{T\triangle Q}(z)\,\chi_Q\left(-\left(\left(-a^{2D_2}\right) \wedge w'\right)\right)\right] \\
&= \mathop{\mathbb{E}}_{\mathcal{H}}\left[\chi_{S\triangle R}(x)\right]\cdot\mathop{\mathbb{E}}_{\mathcal{H}}\left[\chi_{T\triangle Q}(z)\right]\cdot\mathop{\mathbb{E}}_{\mathcal{H}}\left[\chi_R\left(-\left(a^{2D_1} \wedge y'\right)\right)\chi_Q\left(-\left(\left(-a^{2D_2}\right) \wedge w'\right)\right)\right].
\end{aligned}$$

Since the marginals of $\mathcal{H}$ are uniform on both the $\mathcal{X}$ and $\mathcal{Z}$ spaces, it holds that the last term is 0 unless $S = R$ and $T = Q$. When $S = R$ and $T = Q$ it holds that

$$\begin{aligned}
E_{\mathcal{H},S,R,T,Q} &= \mathop{\mathbb{E}}_{\mathcal{H}}\left[\chi_S\left(-\left(a^{2D_1} \wedge y'\right)\right)\chi_T\left(-\left(\left(-a^{2D_2}\right) \wedge w'\right)\right)\right] \\
&= \mathop{\mathbb{E}}_{\mathcal{H}}\left[\frac{1+a}{2}\cdot\chi_S\left((-1)^{2D_1}\right)\cdot\chi_T\left(-w'\right) + \frac{1-a}{2}\cdot\chi_S\left(-y'\right)\cdot\chi_T\left((-1)^{2D_2}\right)\right] \\
&= \frac{1}{2}\left((-1)^{|S|}\cdot\mathop{\mathbb{E}}_{w'\sim\mathcal{N}_\epsilon(D_2)}\left[\chi_T\left(-w'\right)\right] + (-1)^{|T|}\cdot\mathop{\mathbb{E}}_{y'\sim\mathcal{N}_\epsilon(D_1)}\left[\chi_S\left(-y'\right)\right]\right) \\
&= \frac{1}{2}\left((-1)^{|S|}\cdot N_T + (-1)^{|T|}\cdot N_S\right).
\end{aligned}$$

■

Proof of Claim 4.15 Recall if (x, y, z, w) was drawn from $\mathcal{I}$ then $y = -x\left(a_1^{2D_1} \wedge y'\right)$ and $w = -z\left(a_2^{2D_2} \wedge w'\right)$, when x, z, a_1, a_2, y', w' are all independent. Therefore

$$\begin{aligned}
E_{\mathcal{I},S,R,T,Q} &= \mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_S(x)\,\chi_R(y)\,\chi_T(z)\,\chi_Q(w)\right] \\
&= \mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_S(x)\,\chi_R\left(-x\left(a_1^{2D_1} \wedge y'\right)\right)\chi_T(z)\,\chi_Q\left(-z\left(a_2^{2D_2} \wedge w'\right)\right)\right] \\
&= \mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_{S\triangle R}(x)\,\chi_R\left(-\left(a_1^{2D_1} \wedge y'\right)\right)\chi_{T\triangle Q}(z)\,\chi_Q\left(-\left(a_2^{2D_2} \wedge w'\right)\right)\right] \\
&= \mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_{S\triangle R}(x)\right]\cdot\mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_{T\triangle Q}(z)\right]\cdot\mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_R\left(-\left(a_1^{2D_1} \wedge y'\right)\right)\right]\cdot\mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_Q\left(-\left(a_2^{2D_2} \wedge w'\right)\right)\right].
\end{aligned}$$

Since the marginals of $\mathcal{I}$ are uniform on both the $\mathcal{X}$ and $\mathcal{Z}$ spaces, it holds that the last term is 0 unless $S = R$ and $T = Q$. When $S = R$ and $T = Q$ it holds that

$$\begin{aligned}
E_{\mathcal{I},S,R,T,Q} &= \mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_S\left(-\left(a_1^{2D_1} \wedge y'\right)\right)\right] \mathop{\mathbb{E}}_{\mathcal{I}}\left[\chi_T\left(-\left(a_2^{2D_2} \wedge w'\right)\right)\right] \\
&= \mathop{\mathbb{E}}_{\mathcal{I}}\left[\frac{1+a_1}{2} \cdot \chi_S\left(\left(-1\right)^{2D_1}\right) + \frac{1-a_1}{2} \cdot \chi_S\left(-y'\right)\right] \cdot \\
&\quad \mathop{\mathbb{E}}_{\mathcal{I}}\left[\frac{1+a_2}{2} \cdot \chi_T\left(\left(-1\right)^{2D_2}\right) + \frac{1-a_2}{2} \cdot \chi_T\left(-w'\right)\right] \\
&= \frac{1}{2}\left(\left(-1\right)^{|S|} + \mathop{\mathbb{E}}_{y' \sim \mathcal{N}_\epsilon(D_1)}\left[\chi_S\left(-y'\right)\right]\right) \cdot \frac{1}{2}\left(\left(-1\right)^{|T|} + \mathop{\mathbb{E}}_{w' \sim \mathcal{N}_\epsilon(D_2)}\left[\chi_S\left(-w'\right)\right]\right) \\
&= \frac{1}{4}\left(\left(-1\right)^{|S|} + N_S\right)\left(\left(-1\right)^{|T|} + N_T\right).
\end{aligned}$$

■

5 Covering Dictatorship Test for General Predicates

In this section we develop a general framework for covering dictatorship tests using a given predicate φ, for a large subset of the predicates $\varphi \notin \mathcal{O}$. In particular, in Section 5.1 we prove Theorem 2. In Section 5.2 we prove a more general version of Theorem 2 (see Lemma 5.6), offering a dictatorship test in the multi-function setting. The general version is used in Section 7 to obtained a conditional characterization of covering hard predicates. For convenience we restate Theorem 2:

Theorem. *Let $\varphi \notin \mathcal{O}$, and assume that there exists a balanced, pairwise independent distribution on the support of φ. Then there exists a φ-based covering-dictatorship test with completeness 2 and soundness k, for every $k \in \mathbb{N}$.*

5.1 Single Function Dictatorship Test

In this subsection we assume that φ and η are as in Theorem 2, and construct a φ-based covering dictatorship test $\mathsf{DICT1}$, using the distribution η. The test assumes to have an oracle access to the function $f : \{\pm 1\}^{2n} \to \{\pm 1\}$ being tested.

Roughly speaking, we show that our dictatorship test satisfies the following properties: There exist two dictatorships f and g that together cover all the tests made by $\mathsf{DICT1}$. However, any constant number of functions whose every product is "far" from a dictatorship do not cover all the tests made by $\mathsf{DICT1}$. In other words, if a constant number of functions cover all the tests, then there is a subset of these functions whose product is "close" to a dictatorship.

5.1.1 The Test

Our dictatorship test uses the distribution $\mathcal{D}_{\epsilon,\eta}$ that is specified next. Let us first define the following noisy version of η:

Definition 5.1. Let $\epsilon \in [0,1]$, let η be a distribution over $\{\pm 1\}^t$, and let $\mathcal{U}$ be the uniform distribution over $\{\pm 1\}^t$. Define the distribution η'_ϵ generating $y \in \{\pm 1\}^t$ as follows:

$$\forall y: \ \ \eta'_\epsilon (y) = (1-\epsilon)\, \eta (y) + \epsilon \cdot \mathcal{U} (y) .$$

That is, in order to draw a t-bits string y from η'_ϵ: With probability $1-\epsilon$ draw y from η, and with probability ϵ draw a random y.

Definition 5.2. Let $\epsilon \in [0,1]$, and let η be a distribution over $\{\pm 1\}^t$. Define the distribution $\mathcal{D}_{\epsilon,\eta}$ generating $w = (y,z) \in \left(\{\pm 1\}^t\right)^2$ as follows:

$$\forall (y,z): \ \ \mathcal{D}_{\epsilon,\eta} (y,z) = \frac{1}{2} \eta (y)\, \eta'_\epsilon (z) + \frac{1}{2} \eta'_\epsilon (y)\, \eta (z) .$$

That is, in order to draw a pair (y,z) of t-bits strings from $\mathcal{D}_{\epsilon,\eta}$: With probability $\frac{1}{2}$ draw y from η and z from η'_ϵ, and with probability $\frac{1}{2}$ draw y from η'_ϵ and z from η.

Our dictatorship test is found in Figure 2. For a string s_i, we use the notation $s_{i,j}$ to indicate the j^{th} coordinate of s_i.

Algorithm 2 $\mathsf{DICT1}_\epsilon^f$

- Select $w_1 = (y_1, z_1), \ldots, w_n = (y_n, z_n) \in \left(\{\pm 1\}^t\right)^2$ according to the distribution $\mathcal{D}_{\epsilon,\eta}$.
- For $i \in [t]$, let $x_i = y_{1,i}, \ldots, y_{n,i}, z_{1,i}, \ldots, z_{n,i} \in \{\pm 1\}^{2n}$.
- Accept iff $\varphi (f (x_1), \ldots, f (x_t)) = -1$.

It will be convenient for us to view the dictatorship test in a matrix notation. For a matrix M, we denote by M_i the i^{th} row of M, and by M^j the j^{th} column of M. Consider the following $2n \times t$ matrix M: The first n rows of the matrix are $y_1, \ldots, y_n$, that is $M_1 = y_1, \ldots, M_n = y_n$. The following n rows are $z_1, \ldots, z_n$, that is $M_{n+1} = z_1, \ldots, M_{2n} = z_n$. Note that the t columns of the obtained matrix are $x_1, \ldots, x_t$, that is $M^1 = x_1, \ldots, M^t = x_t$.

5.1.2 Definitions

In order to analyze the dictatorship test we need the following definitions: Let $\epsilon \in (0,1)$, $k \in \mathbb{N}$ and let $\mathcal{F} = \{f_1, \ldots, f_k\}$ be a set of functions $f_\ell : \{\pm 1\}^{2n} \rightarrow \{\pm 1\}$.

We denote by $Rej\left(\mathsf{DICT1}_\epsilon^{\mathcal{F}}\right)$ the indicator random variable for the rejection of the set $\mathcal{P}$ by $\mathsf{DICT1}_\epsilon$. That is, $Rej\left(\mathsf{DICT1}_\epsilon^{\mathcal{F}}\right)$ is 1 if none of the functions f_ℓ in $\mathcal{F}$ passes the test selected by $\mathsf{DICT1}$, and 0 if there exists a function f_ℓ in $\mathcal{F}$ that passes the test.

For $K \subseteq [k]$, we define the product function $f_K : \{\pm 1\}^{2n} \to \{\pm 1\}$ by

$$f_K = \prod_{\ell \in K} f_\ell.$$

The function f_ϕ is the all 1's functions, i.e., for every $x \in \{\pm 1\}^{2n}$ it holds that $f_\phi(x) = 1$.

We say that a function is *regular* if none of its coordinates has high influence:

Definition 5.3 (**Regularity**). Let $d \in \mathbb{N}$, $\tau \in [0,1]$, and let $f : \{\pm 1\}^n \to \{\pm 1\}$ be a function. We say that f is (d,τ)*-regular*, if

$$\max_{j \in [n]} \left\{ Inf_j^{\leq d}(f) \right\} \leq \tau.$$

Let $\mathcal{F} = \{f_1, \ldots, f_k\}$ be a set of functions $f_\ell : \{\pm 1\}^n \to \{\pm 1\}$. We say that $\mathcal{F}$ is (d,τ)*-regular* if for every subset $K \subseteq [k]$ it holds that f_K is (d,τ)-regular.

5.1.3 Test Analysis

We are now ready to state our result regarding $\mathsf{DICT1}$. The following lemma clearly implies Theorem 2.

Lemma 5.4. *Let $\varphi : \{\pm 1\}^t \to \{\pm 1\}$ be a predicate satisfying $\varphi \notin \mathcal{O}$. Assume that there exists a balanced, pairwise independent distribution η on the support of φ.*
Then, $\mathsf{DICT1}$ *satisfies the following properties:*

- ***Completeness:*** *For any $\epsilon \in (0,1)$ the following holds. Let $j \in [n]$ and let $f, g : \{\pm 1\}^{2n} \to \{\pm 1\}$ be the two dictatorships $f = \chi_j$ and $g = \chi_{j+n}$. Then*

$$\Pr\left[Rej\left(\mathsf{DICT1}_\epsilon^{\{f,g\}}\right)\right] = 0.$$

  In particular, there exist two dictatorships that cover all the tests performed by $\mathsf{DICT1}$. *Furthermore, each of the functions f and g is accepted by $\mathsf{DICT1}_\epsilon$ with probability $1 - \frac{\epsilon}{2}$.*

- ***Soundness:*** *For any $\epsilon \in (0,1)$ and $k \in \mathbb{N}$, there exist constants $d \in \mathbb{N}$ and $\tau > 0$ that only depend on ϵ, t and k, such that the following holds. Let $\mathcal{F} = \{f_1, \ldots, f_k\}$ be a set of functions $f_\ell : \{\pm 1\}^{2n} \to \{\pm 1\}$. Assume that $\mathcal{F}$ is (d,τ)-regular. Then*

$$\Pr\left[Rej\left(\mathsf{DICT1}_\epsilon^{\mathcal{F}}\right)\right] > \frac{1}{2^{10kt}}.$$

In particular, if a set of a constant number of functions covers all the tests performed by $\mathsf{DICT1}$*, then the set is not* (d,τ)*-regular.*

5.2 General Dictatorship Test

In this section we offer a more general covering dictatorship test. Instead of getting oracle access to a single function $f:\{\pm 1\}^{2n}\rightarrow\{\pm 1\}$, the general dictatorship test gets access to a family of functions $F=\{f^1,\ldots,f^r\}$, $r\geq t\in\mathbb{N}$, $f:\{\pm 1\}^{2n}\rightarrow\{\pm 1\}$. Intuitively, the test aims at checking whether all the functions $f^1,\ldots,f^r$ are the same dictatorship.

The general dictatorship test is used in Section 7. Roughly speaking, the $f^1,\ldots,f^r$ functions considered by Section 7 are the long code encodings of the $r\gg t$ neighbors $v_1,\ldots,v_r$ of a single vertex u in a label cover instance.

5.2.1 The Test

Our general dictatorship test DICT is found in Figure 3. It is easy to see that if F contains r copies of the same function f, i.e., $f^1=\cdots=f^r=f$, then DICT_ϵ^F operates the same as $\mathsf{DICT1}_\epsilon^f$.

Algorithm 3 $\mathsf{DICT}_\epsilon^{F=\left\{f^1,\ldots,f^r\right\}}$

- Select $w_1=(y_1,z_1),\ldots,w_n=(y_n,z_n)\in\left(\{\pm 1\}^t\right)^2$ according to the distribution $\mathcal{D}_{\epsilon,\eta}$.
- For $i\in[t]$, let $x_i=y_{1,i},\ldots,y_{n,i},z_{1,i},\ldots,z_{n,i}\in\{\pm 1\}^{2n}$.
- Select a random set of t *different* indices $I=\{i_1,\ldots,i_t\}\subseteq[r]$ (selection with no repetitions).
- Accept iff $\varphi\left(f^{i_1}(x_1),\ldots,f^{i_t}(x_t)\right)=-1$.

5.2.2 Definitions

In order to analyze the dictatorship test we need the following definitions: Let $\epsilon\in(0,1)$, $k,r\in\mathbb{N}$, and let $\mathcal{F}$ be a set of sets of functions $\mathcal{F}=\{F_1,\ldots,F_k\}$, $F_\ell=\{f_\ell^1,\ldots,f_\ell^r\}$, $f_\ell^i:\{\pm 1\}^{2n}\rightarrow\{\pm 1\}$.

We denote by $Rej\left(\mathsf{DICT}_\epsilon^{\mathcal{F}}\right)$ the indicator random variable for the rejection of the set $\mathcal{F}$ by DICT. That is, $Rej\left(\mathsf{DICT}_\epsilon^{\mathcal{F}}\right)$ is 1 if none of the sets F_ℓ in $\mathcal{F}$ passes the test selected by DICT, and 0 if there exists a set F_ℓ in $\mathcal{F}$ that passes the test.

We define the *cross influence* of a pair of function as follows:

Definition 5.5 (**Cross Influence**)**.** Let $d \in \mathbb{N}$, and let $f, g : \{\pm 1\}^n \to \{\pm 1\}$ be a pair of function. We denote by $XInf_d(f, g)$ the *d-cross influence* of f and g:

$$XInf_d(f, g) = \max_{j \in [n]} \left\{ \min \left\{ Inf_j^{\leq d}(g), Inf_j^{\leq d}(f) \right\} \right\}.$$

For $i \in [r]$ and $K \subseteq [k]$ we define the product function $f_K^i : \{\pm 1\}^{2n} \to \{\pm 1\}$ by

$$f_K^i = \prod_{\ell \in K} f_K^i.$$

The function f_ϕ^i is the all 1's functions, i.e., for every $x \in \{\pm 1\}^{2n}$ it holds that $f_\phi^i(x) = 1$.

Let $d \in \mathbb{N}$ and $\tau \in [0, 1]$. Let $(i, i') \in [r]^2$, $i \neq i'$, be a pair of indices. We say that (i, i') is $(d, \tau, \mathcal{F})$-*cross regular*, if for every two sets $K, K' \subseteq [k]$ it holds that

$$XInf_d\left(f_K^i, f_{K'}^{i'}\right) \leq \tau.$$

Let $I \subseteq [r]$ be a set of indices. We say I is $(d, \tau, \mathcal{F})$-*cross regular* if every pair $(i, i') \in I^2$, $i \neq i'$, is $(d, \tau, \mathcal{F})$-cross regular.

5.2.3 Test Analysis

We are now ready to state our result regarding DICT:

Lemma 5.6. *Let $\varphi : \{\pm 1\}^t \to \{\pm 1\}$ be a predicate satisfying $\varphi \notin \mathcal{O}$. Assume that there exists a balanced, pairwise independent distribution η on the support of φ.*
Then, DICT *satisfies the following properties:*

- ***Completeness:*** *For any $\epsilon \in (0, 1)$ the following holds. Let $j \in [n]$ and let $f^i, g^i : \{\pm 1\}^{2n} \to \{\pm 1\}$, $i \in [r]$, be the following functions*

  $$f^1 = \cdots = f^r = \chi_j, \quad g^1 = \cdots = g^r = \chi_{j+n}.$$

  Let $F = \{f^1, \ldots, f^r\}$ and $G = \{g^1, \ldots, g^r\}$. Then

  $$\Pr\left[Rej\left(\mathsf{DICT}_\epsilon^{\{F,G\}}\right)\right] = 0.$$

  Furthermore, each of the sets F and G is accepted by DICT_ϵ with probability $1 - \frac{\epsilon}{2}$.

- ***Soundness:*** *For any $\epsilon \in (0, 1)$ and $k \in \mathbb{N}$, there exist $r, d \in \mathbb{N}$ and $\tau > 0$ that only depend on ϵ, t, and k, such that the following holds. Let $\mathcal{F}$ be a set of sets of functions*

$\mathcal{F} = \{F_1, \ldots, F_k\}$, $F_\ell = \{f_\ell^1, \ldots, f_\ell^r\}$, $f_\ell^i : \{\pm 1\}^{2n} \rightarrow \{\pm 1\}$. Let $\alpha = \frac{1}{2^{10kt}}$, and assume that at least $1 - \alpha$ fraction of the t-elements sets $I \subseteq [r]$ are *$(d, \tau, \mathcal{F})$-cross regular. Then*

$$\Pr\left[Rej\left(\mathsf{DICT}_\epsilon^{\mathcal{F}}\right)\right] > \frac{1}{2^{10kt}}.$$

It is easy to see that Lemma 5.4 is a special case of Lemma 5.6. Thus, we will only prove Lemma 5.6. The proof of the completeness part of Lemma 5.6 can be found in Section 5.3, and the proof of soundness part can be found in Section 6. For the rest of the text we fix t, ϵ, φ and η. We omit the ϵ and η sub-indices and write $\mathsf{DICT} = \mathsf{DICT}_\epsilon$, $\eta' = \eta'_\epsilon$ and $\mathcal{D} = \mathcal{D}_{\epsilon,\eta}$.

5.3 Completeness

In this section we prove the completeness property of Lemma 5.6.

Proof of Lemma 5.6 (Completeness) Recall the matrix M defined at the end of Subsection 5.1.1, after the algorithm $\mathsf{DICT1}$. For $i \in [t]$, it holds that x_i is the i^{th} column of M. For $j \in [n]$ it holds that y_j is the j^{th} row of M, and z_j is row number $j + n$ of M. Also recall that there exists $j \in [n]$ such that for every $i \in [r]$ it holds that $f^i = \chi_j$ and $g^i = \chi_{j+n}$.

When running DICT^F we compare the following value to -1:

$$\varphi\left(f^{i_1}(x_1), \ldots, f^{i_t}(x_t)\right) = \varphi\left(\chi_j(x_1), \ldots, \chi_j(x_t)\right) = \varphi\left(M_{j,1}, \ldots, M_{j,t}\right) = \varphi\left(M_j\right) = \varphi\left(y_j\right).$$

When running DICT^G we compare the following value to -1:

$$\begin{aligned}\varphi\left(g^{i_1}(x_1), \ldots, g^{i_t}(x_t)\right) &= \varphi\left(\chi_{j+n}(x_1), \ldots, \chi_{j+n}(x_t)\right) \\ &= \varphi\left(M_{j+n,1}, \ldots, M_{j+n,t}\right) = \varphi\left(M_{j+n}\right) = \varphi\left(z_j\right)\end{aligned}$$

Recall that (y_j, z_j) was drawn from $\mathcal{D}$. Thus, either y_j or z_j was drawn from η, implying that at least one of them is in support of φ. Hence, either $\varphi(y_j) = -1$ or $\varphi(z_j) = -1$, and at least one of F and G is accepted by DICT.

Moreover, observe that F is only rejected by DICT if y_j is not in the support of φ. This can only happen if $\mathcal{D}$ samples y_i using η', and η' samples y_i using the uniform distribution (instead of using η). The probability of this event is $\frac{\epsilon}{2}$. Therefore, F (and similarly also G) is accepted with probability at least $1 - \frac{\epsilon}{2}$. ■

6 Dictatorship Test Soundness

In this section we prove the soundness property of Lemma 5.6.

6.1 Properties of $\mathcal{D}$

We prove that the distribution $\mathcal{D}$ satisfies several properties that will turn useful for the soundness proof in Section 6.2. Observe that one can identify $(y,z) \in \left(\{\pm 1\}^t\right)^2$ with $w \in \left(\{\pm 1\}^2\right)^t$ given by $\forall i \in [t] : w_i = (y_i, z_i)$. For the rest of the section we view $\mathcal{D}$ as a distribution on $w \in \left(\{\pm 1\}^2\right)^t$ instead of $(y,z) \in \left(\{\pm 1\}^t\right)^2$ (unless otherwise stated). That is, we view the generated w as a t-symbols string where every symbol is a pair of bits.

Claim 6.1. *$\mathcal{D}$ is balanced.*

Proof Recall that we assume that η is balanced. The distribution η' is a sum of the two balanced distributions, η and the uniform distribution, therefore η' is also balanced.

Let $i \in [t]$ and $(\omega, \omega') \in \{\pm 1\}^2$. We denote $w_i = (y_i, z_i) \in \{\pm 1\}^2$, and get

$$\begin{aligned}
\Pr_{w \sim \mathcal{D}}\left[w_i = (\omega, \omega')\right] &= \frac{1}{2} \Pr_{(y,z)\sim \eta \times \eta'}\left[y_i = \omega \wedge z_i = \omega'\right] + \frac{1}{2} \Pr_{(y,z)\sim \eta' \times \eta}\left[y_i = \omega \wedge z_i = \omega'\right] \\
&= \frac{1}{2} \Pr_{y \sim \eta}\left[y_i = \omega\right] \cdot \Pr_{z \sim \eta'}\left[z_i = \omega'\right] + \frac{1}{2} \Pr_{y \sim \eta'}\left[y_i = \omega\right] \cdot \Pr_{z \sim \eta}\left[z_i = \omega'\right] \\
&= \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 = \frac{1}{4} = \frac{1}{\left|\{\pm 1\}^2\right|}.
\end{aligned}$$

■

Claim 6.2. *$\mathcal{D}$ is pairwise independent.*

Proof Recall that we assume that η is pairwise independent. The distribution η' is a sum of the two pairwise independent distributions, η and the uniform distribution, therefore η' is also pairwise independent.

Let $i, i' \in [t]$ and $(\omega, \omega'), (\sigma, \sigma') \in \{\pm 1\}^2$. We denote $w_i = (y_i, z_i), w_{i'} = (y_{i'}, z_{i'}) \in \{\pm 1\}^2$, and get

$$\begin{aligned}
& \Pr_{w \sim \mathcal{D}}\left[w_i = (\omega, \omega') \wedge w_{i'} = (\sigma, \sigma')\right] \\
= \; & \frac{1}{2} \Pr_{(y,z)\sim \eta \times \eta'}\left[(y_i = \omega \wedge y_{i'} = \sigma) \bigwedge (z_i = \omega' \wedge z_{i'} = \sigma')\right] + \\
& \frac{1}{2} \Pr_{(y,z)\sim \eta' \times \eta}\left[(y_i = \omega \wedge y_{i'} = \sigma) \bigwedge (z_i = \omega' \wedge z_{i'} = \sigma')\right] \\
= \; & \frac{1}{2} \Pr_{y \sim \eta}\left[y_i = \omega \wedge y_{i'} = \sigma\right] \cdot \Pr_{z \sim \eta'}\left[z_i = \omega' \wedge z_{i'} = \sigma'\right] + \\
& \frac{1}{2} \Pr_{y \sim \eta'}\left[y_i = \omega \wedge y_{i'} = \sigma\right] \cdot \Pr_{z \sim \eta}\left[z_i = \omega' \wedge z_{i'} = \sigma'\right]
\end{aligned}$$

Since η and η' are both pairwise independent and balanced, the last term is

$$\begin{aligned} = \; & \frac{1}{2}\Pr_{y\sim\eta}\left[y_i=\omega\right]\cdot\Pr_{y\sim\eta}\left[y_{i'}=\sigma\right]\cdot\Pr_{z\sim\eta'}\left[z_i=\omega'\right]\cdot\Pr_{z\sim\eta'}\left[z_{i'}=\sigma'\right]+ \\ & \frac{1}{2}\Pr_{y\sim\eta'}\left[y_i=\omega\right]\cdot\Pr_{y\sim\eta'}\left[y_{i'}=\sigma\right]\cdot\Pr_{z\sim\eta}\left[z_i=\omega'\right]\cdot\Pr_{z\sim\eta}\left[z_{i'}=\sigma'\right] \\ = \; & \left(\frac{1}{2}\right)^5+\left(\frac{1}{2}\right)^5=\frac{1}{16}. \end{aligned}$$

Since $\mathcal{D}$ is balanced, it holds that

$$\Pr_{w\sim\mathcal{D}}\left[w_i=(\omega,\omega')\right]\cdot\Pr_{w\sim\mathcal{D}}\left[w_{i'}=(\sigma,\sigma')\right]=\frac{1}{4}\cdot\frac{1}{4}=\frac{1}{16},$$

and the assertion follows. ■

The following lemma shows that the probability space $\left(\{\pm1\}^2\right)^t$ is not completely correlated under $\mathcal{D}$.

Lemma 6.3. *$\rho\left(\{\pm1\}^2,\left(\{\pm1\}^2\right)^{t-1};\mathcal{D}\right)=\rho<1$, where ρ is a constant that only depends on t and ϵ.*

The proof of Lemma 6.3 uses the following lemma by Mossel [14], that gives a criteria under which a probability space is not completely correlated.

Lemma 6.4 (**Mossel [14], Lemma 2.9**). *Let $(\Omega\times\Psi,\mu)$ be a finite correlated probability space, such that the probability of the smallest atom in $\Omega\times\Psi$ is at least $\gamma>0$. That is, for every $(a,b)\in\Omega\times\Psi$ in the support of μ, it holds that $\mu(a,b)\geq\gamma$.*

Define the bipartite graph $G=(\Omega,\Psi,E)$ where $(a,b)\in\Omega\times\Psi$ satisfies $(a,b)\in E$ if $\mu(a,b)>0$. Then, if G is connected then $\rho(\Omega,\Psi;\mu)\leq 1-\frac{\gamma^2}{2}$.

Proof of Lemma 6.3 Let Ψ be the support of the marginal distribution of $\mathcal{D}$ on $\left(\{\pm1\}^2\right)^{t-1}$, and denote $\Omega=\{\pm1\}^2$ (the support of the marginal distribution of $\mathcal{D}$ on $\{\pm1\}^2$ is $\{\pm1\}^2$ itself, as $\mathcal{D}$ is balanced). We show that $\rho(\Omega,\Psi;\mathcal{D})=\rho'<1$ (ρ' is a constant that only depends on t and ϵ) by applying Lemma 6.4.

Consider the bipartite graph $G=(\Omega,\Psi,E)$ where $(a,b)\in\Omega\times\Psi$ satisfies $(a,b)\in E$ if $\mathcal{D}(a,b)>0$. Our goal is to show that G is connected.

Let $a\in\Omega$ and $b\in\Psi$. Denote $a=(y_a,z_a)$ and $b=(y'_b,z'_b)$ where $y_a,z_a\in\{\pm1\}$ and $y'_b,z'_b\in\{\pm1\}^{t-1}$. Since η is balanced, y_a agrees with a word y in the support of η, and z_a agrees with word in the support of η. In addition, observe that there exists $w'=(y',z')\in\left(\{\pm1\}^t\right)^2$ in the support of $\mathcal{D}$ that agrees with b. By the way $\mathcal{D}$ was constructed, either y'

or z' was selected according to η. Therefore, either y'_b agrees with a word y' on the support of η, or z'_b agrees with a word z' on the support of η.

Assume without loss of generality that y'_b agrees with the word y' on the support of η, and denote by $y'_a \in \{\pm 1\}$ the bit that "completes" y'_b to y' (i.e., $y' = y'_a \circ y'_b$). Recall that there exists a word $y \in \{\pm 1\}^t$ in the support of η that agrees with y_a, and let $y_b \in \{\pm 1\}^{t-1}$ be the string that "completes" y_a to y (i.e., $y = y_a \circ y_b$).

Let $c \in \Omega$ and $d \in \Psi$ be two vertices in G. Denote $c = (y_c, z_c)$ and $d = (y_d, z_d)$ where $y_d, z_d \in \{\pm 1\}$ and $y_d, z_d \in \{\pm 1\}^{t-1}$. Let $y = y_c \circ y_d$ and $z = z_c \circ z_d$. Recall that there is an edge between c and d in G if $w = (y, z)$ is in the support of $\mathcal{D}$. In other words, if either y or z are in the support of η.

The following is a path in G connecting a and b:

$$a = (y_a, z_a) \to b' = (y_b, y_b) \to a' = (y'_a, y_a) \to b = (y'_b, z'_b)\,.$$

The edge $a \to b'$ exists as the concatenation of the first components of vertices a, b' is $y_a \circ y_b = y$, which is in the support of η. The edge $b' \to a'$ exists as the concatenation of the second components of vertices a', b' is $y_a \circ y_b = y$, which is in the support of η. The edge $a' \to b$ exists as the concatenation of the first components of vertices a', b is $y'_a \circ y'_b = y'$, which is in the support of η. ■

6.2 Soundness

In this section we prove the soundness property of Lemma 5.6. The proof uses the following invariance principle theorem, which is an easy corollary of Theorem 6.6 and Lemma 6.9 of Mossel [14], obtained using the triangle inequality.

Theorem 6.5 (Invariance Principle). *Let $\left(\prod_{i=1}^t \Omega_i, \mu\right)$ and $\left(\prod_{i=1}^t \Omega_i, \mu'\right)$ be two a finite correlated probability spaces. Assume that μ and μ' have the same marginal distribution on each coordinate. Furthermore, assume that both satisfy the following properties:*
For $\mu^* \in \{\mu, \mu'\}$,

- *μ^* is pairwise independent.*
- $\forall i \in [t], \omega \in \Omega_i : \Pr_{w \sim \mu^*}[w_i = \omega] > 0$.
- $\rho(\Omega_1, \ldots, \Omega_t; \mu^*) < 1$.

Then, for every $\delta > 0$ there exist $d \in \mathbb{N}$ and $\tau > 0$ such that the following holds for all $n \in \mathbb{N}$:

Let $t \in \mathbb{N}$, and let $f_1, \ldots, f_t$, $f_i : \Omega_i^n \to [-1,1]$, be t functions satisfying

$$\forall i \neq i' \in [t] : XInf_d\left(f_i, f_{i'}\right) \leq \tau.$$

Then,

$$\left| \prod_{i \in [t]} \mathop{\mathbb{E}}_{w_1,\ldots,w_n \sim \mu}\left[f_i\left(w_{1,i}, \ldots, w_{n,i}\right)\right] - \prod_{i \in [t]} \mathop{\mathbb{E}}_{w_1,\ldots,w_n \sim \mu'}\left[f_i\left(w_{1,i}, \ldots, w_{n,i}\right)\right] \right| \leq \delta.$$

We are now ready to prove the soundness property of Lemma 5.6:

Proof of Lemma 5.6 (Soundness) Fix $\epsilon \in (0,1)$ and $k \in \mathbb{N}$, and let $r, d \in \mathbb{N}$ and $\tau > 0$ be the constants determined below. Let $\mathcal{F}$ be a set of sets of functions $\mathcal{F} = \{F_1, \ldots, F_k\}$, $F_\ell = \{f_\ell^1, \ldots, f_\ell^r\}$, $f_\ell^i : \{\pm 1\}^{2n} \to \{\pm 1\}$. We assume that at least $1 - \alpha$ fraction of the t-elements sets $I \subseteq [r]$ are $(d, \tau, \mathcal{F})$-cross regular. We want to show that

$$Rej = \Pr\left[Rej\left(\mathsf{DICT}^{\mathcal{F}}\right)\right] > \frac{1}{2^{10kt}}.$$

Arithmetizing the rejection probability by moving from φ to NAE. We assume without loss of generality that the support of φ is contained in the support of NAE_t, i.e., $\forall x \in \{\pm 1\}^t : \varphi(x) \geq \mathsf{NAE}_t(x)$. Claim 2.2 shows that this is correct "upto a sign", i.e., there always exists a sign $\sigma = (\sigma_1, \ldots, \sigma_t) \in \{\pm 1\}^t$ such that $\varphi(\sigma_1 x_1, \ldots, \sigma_t x_t) \geq \mathsf{NAE}_t(x_1, \ldots, x_t)$. Define the predicate φ^* by $\varphi^*(x_1, \ldots, x_t) \equiv \varphi(\sigma_1 x_1, \ldots, \sigma_t x_t)$. Observe that when constructing a φ-based dictatorship test, we have the liberality of negating any of the variables before applying φ. In particular, we are free to apply φ^* instead of φ. For this reason, we only need to deal with dictatorship tests based on predicates that are contained in NAE.

Since we assume $\forall x \in \{\pm 1\}^t : \varphi(x) \geq \mathsf{NAE}_t(x)$, we get

$$\begin{aligned} Rej &= \mathop{\mathbb{E}}_{\substack{i_1 < \ldots < i_t \in [r] \\ w_1,\ldots,w_n \sim \mathcal{D}}}\left[\prod_{\ell \in [k]} \frac{1}{2}\left(\varphi\left(f_\ell^{i_1}(x_1), \ldots, f_\ell^{i_t}(x_t)\right) + 1\right)\right] \\ &\geq \mathop{\mathbb{E}}_{\substack{i_1 < \ldots < i_t \in [r] \\ w_1,\ldots,w_n \sim \mathcal{D}}}\left[\frac{1}{2^k} \prod_{\ell \in [k]} \left(\mathsf{NAE}_t\left(f_\ell^{i_1}(x_1), \ldots, f_\ell^{i_t}(x_t)\right) + 1\right)\right]. \end{aligned}$$

A calculation shows that the Fourier expansion of the NAE predicate is

$$\forall x \in \{\pm 1\}^t : \mathsf{NAE}_t(x) = \frac{1}{2^{t-2}} \left(\sum_{\substack{S \subseteq [t] \\ |S|\, even}} \chi_S(x) \right) - 1.$$

Therefore,

$$Rej \geq \mathop{\mathbb{E}}_{\substack{i_1 < \ldots < i_t \in [r] \\ w_1, \ldots, w_n \sim \mathcal{D}}} \left[\frac{1}{2^{k(t-1)}} \prod_{\ell \in [k]} \left(\sum_{\substack{S \subseteq [t] \\ |S|\, even}} \chi_S\left(f_\ell^{i_1}(x_1), \ldots, f_\ell^{i_t}(x_t) \right) \right) \right]. \tag{10}$$

We identify a subset $A \subseteq [m]$, $m \in \mathbb{N}$, with a vector $B \in \{0,1\}^m$ in the natural way: For every $i \in [m]$ it holds that $B_i = 1$ iff $i \in A$. Let $A_1, \ldots, A_t \subseteq [m]$ be t sets, and let $B_1, \ldots, B_t \in \{0,1\}^m$ be the associated vectors. We denote by $A_1 \oplus \ldots \oplus A_t \subseteq [m]$ the set associated with the vector $B_1 + \ldots + B_t$, where the addition is vector addition modulo 2. In other words, $A_1 \oplus \ldots \oplus A_t$ contains the elements in $[m]$ that appear in an odd number of the sets $A_1, \ldots, A_t$.

Consider the term on the right hand side of Equation 10. Each of the monomial in the expectation is a product of k terms of the form $\chi_{S_\ell}\left(f_\ell^{i_1}(x_1), \ldots, f_\ell^{i_t}(x_t) \right)$, where $\ell \in [k]$, and $S_\ell \subseteq [t]$ is of even size. Therefore, each monomial can be specified by k sets $S_1, \ldots S_k \subseteq [t]$. Consider the set of $t \times k$ matrices K with entries in $\{0,1\}$ that satisfy the following property: The rows of K sum up to 0^k (vector addition modulo 2), i.e., $K_1 + \ldots + K_t = 0^k$. We claim that there is a one-to-one mapping between this set of matrices and the set of monomials:

- Given a matrix K, view its k columns as k sets $S_1, \ldots S_k$. Since the rows sum up to 0^k, each of the columns sums up to 0, and thus each of the sets S_ℓ is of even size.

- Given k sets $S_1, \ldots S_k$ of even size, construct the matrix K whose columns are $S_1, \ldots S_k$. Since each column contains an even number of 1's, each column sums up to 0, and thus the rows sum up to 0^k.

Conclude that

$$Rej \;\; \geq \mathop{\mathbb{E}}_{i_1 < \ldots < i_t \in [r]} \left[\frac{1}{2^{k(t-1)}} \sum_{\substack{K_1, \ldots, K_t \subseteq [k] \\ K_1 \oplus \ldots \oplus K_t = \phi}} \mathop{\mathbb{E}}_{w_1, \ldots, w_n \sim \mathcal{D}} \left[f_{K_1}^{i_1}(x_1) \cdot \ldots \cdot f_{K_t}^{i_t}(x_t) \right] \right]. \tag{11}$$

Applying the invariance principle. We next apply the invariance principle (Theorem 6.5) and pass from the distribution $\mathcal{D}$ to the uniform distribution $\mathcal{U}'$ on $\left(\{\pm1\}^2\right)^t$. We first claim that the probability spaces $\left(\left(\{\pm1\}^2\right)^t, \mathcal{D}\right)$ and $\left(\left(\{\pm1\}^2\right)^t, \mathcal{U}'\right)$ satisfy the conditions of Theorem 6.5:

- Claim 6.1 shows that $\mathcal{D}$ is balanced, thus $\mathcal{D}$ and $\mathcal{U}'$ have the same marginal distribution on each coordinate.

- Clearly, $\mathcal{U}'$ satisfies the requirements of Theorem 6.5. Claims 6.2 and 6.1, and Lemma 6.3 show that $\mathcal{D}$ also satisfies the requirements of Theorem 6.5.

Let $d \in \mathbb{N}$ and $\tau > 0$ be the constants promised by Theorem 6.5 for the probability spaces $\left(\left(\{\pm1\}^2\right)^t, \mathcal{D}\right)$ and $\left(\left(\{\pm1\}^2\right)^t, \mathcal{U}'\right)$ and the constant $\delta = \frac{1}{2^{10kt}}$. We assume that for at least $1-\alpha$ fraction of the t-elements sets $I \subseteq [r]$ are $(d, \tau, \mathcal{F})$-cross regular. We claim that for such sets I we can use the invariance principle to move from the term $\mathbb{E}_{w_1,\ldots,w_n\sim\mathcal{D}}\left[f_{K_1}^{i_1}(x_1)\cdot\ldots\cdot f_{K_t}^{i_t}(x_t)\right]$ to the term $\mathbb{E}_{w_1,\ldots,w_n\sim\mathcal{U}'}\left[f_{K_1}^{i_1}(x_1)\cdot\ldots\cdot f_{K_t}^{i_t}(x_t)\right]$ with little loss. The reason is that for every t subsets $K_1,\ldots,K_t \subseteq [k]$, and any two functions $f \neq g \in \left\{f_{K_1}^{i_1},\ldots,f_{K_t}^{i_t}\right\}$, it holds that $XInf_d\{f,g\} \leq \tau$.

Since the term in the outer expectation of Equation 11 is bounded in $[0,1]$ (it is a rejection probability), after applying the invariance principle we get

$$Rej \;\geq\; \mathop{\mathbb{E}}_{i_1<\ldots<i_t\in[r]}\left[\frac{1}{2^{k(t-1)}}\sum_{\substack{K_1,\ldots,K_t\subseteq[k]\\ K_1\oplus\ldots\oplus K_t=\phi}}\left(\mathop{\mathbb{E}}_{w_1,\ldots,w_n\sim\mathcal{U}'}\left[f_{K_1}^{i_1}(x_1)\cdot\ldots\cdot f_{K_t}^{i_t}(x_t)\right]-\delta\right)\right]-\alpha.$$

Denote by $\mathcal{U}$ is the uniform distribution on $\{\pm1\}^{2n}$, and get

$$Rej \;\geq\; \mathop{\mathbb{E}}_{i_1<\ldots<i_t\in[r]}\left[\frac{1}{2^{k(t-1)}}\sum_{\substack{K_1,\ldots,K_t\subseteq[k]\\ K_1\oplus\ldots\oplus K_t=\phi}}\mathop{\mathbb{E}}_{x_1,\ldots,x_t\sim\mathcal{U}}\left[f_{K_1}^{i_1}(x_1)\cdot\ldots\cdot f_{K_t}^{i_t}(x_t)\right]\right]-(\delta+\alpha)\,.$$

Now assume that the t indices $i_1,\ldots,i_t \in_R [r]$ were selected independently (with repetitions). Then, for a large enough constant r (a function of k and t), the probability that an index was selected twice ($\exists s \neq s' \in [t] : i_s = i_{s'}$), is at most $\beta = \frac{1}{2^{10kt}}$. Denote $\gamma = \delta+\alpha+\beta$, and get

$$Rej \;\geq\; \mathop{\mathbb{E}}_{\substack{i_1,\ldots,i_t\in[r]\\ x_1,\ldots,x_t\sim\mathcal{U}}}\left[\frac{1}{2^{k(t-1)}}\sum_{\substack{K_1,\ldots,K_t\subseteq[k]\\ K_1\oplus\ldots\oplus K_t=\phi}} f_{K_1}^{i_1}(x_1)\cdot\ldots\cdot f_{K_t}^{i_t}(x_t)\right]-\gamma$$

Increasing t to power of 2 and moving to expected functions. Let $t' = 2^{s'}$ ($s' \in \mathbb{N}$) be the smallest power of 2 that is larger or equals to t. Observe that for every $x \in \{\pm 1\}^t$ and $x' \in \{\pm 1\}^{t'}$ that agree on the first t bits, it holds that $\mathsf{NAE}_t(x) \geq \mathsf{NAE}_{t'}(x')$. Observe that the term in the last expectation is the expansion of the term

$$\frac{1}{2^k} \prod_{\ell \in [k]} \left(\mathsf{NAE}_t \left(f_\ell^{i_1}(x_1), \ldots, f_\ell^{i_t}(x_t) \right) + 1 \right)$$

describing the reject probability of a set of k proofs by the NAE_t predicate. By replacing t by t', the term only decreases. Hence,

$$Rej \;\; \geq \; \mathop{\mathbb{E}}_{\substack{i_1,\ldots,i_{t'} \in [r] \\ x_1,\ldots,x_{t'} \sim \mathcal{U}}} \left[\frac{1}{2^{k(t'-1)}} \sum_{\substack{K_1,\ldots,K_{t'} \subseteq [k] \\ K_1 \oplus \ldots \oplus K_{t'} = \phi}} f_{K_1}^{i_1}(x_1) \cdot \ldots \cdot f_{K_{t'}}^{i_{t'}}(x_{t'}) \right] - \gamma.$$

For $K \subseteq [k]$, define the *expected function* $f_K : \{\pm 1\}^{2n} \to [-1, 1]$ by

$$f_K(x) = \mathop{\mathbb{E}}_{i \in [r]} \left[f_K^i(x) \right].$$

Since the indices $i_1, \ldots, i_{t'} \in [r]$ are selected independently, it holds that

$$\begin{aligned} Rej \;\; &\geq \; \mathop{\mathbb{E}}_{x_1,\ldots,x_{t'} \sim \mathcal{U}} \left[\frac{1}{2^{k(t'-1)}} \sum_{\substack{K_1,\ldots,K_{t'} \subseteq [k] \\ K_1 \oplus \ldots \oplus K_{t'} = \phi}} \mathop{\mathbb{E}}_{i_1 \in [r]} \left[f_{K_1}^{i_1}(x_1) \right] \cdot \ldots \cdot \mathop{\mathbb{E}}_{i_{t'} \in [r]} \left[f_{K_{t'}}^{i_{t'}}(x_{t'}) \right] \right] - \gamma \\ &= \; \frac{1}{2^{k(t'-1)}} \mathop{\mathbb{E}}_{x_1,\ldots,x_{t'} \sim \mathcal{U}} \left[\sum_{\substack{K_1,\ldots,K_{t'} \subseteq [k] \\ K_1 \oplus \ldots \oplus K_{t'} = \phi}} f_{K_1}(x_1) \cdot \ldots \cdot f_{K_{t'}}(x_{t'}) \right] - \gamma. \end{aligned}$$

Applying an extension of the pairing technique of [7]. To analyze the last term, we use a generalization of the "pairing" technique of [7]. For $K \subseteq [k]$ and $i \leq i' \in [t']$ define the sum function $G_{i,i'}^K : \left(\{\pm 1\}^{2n} \right)^{i'-i+1} \to [-1, 1]$ by

$$G_{i,i'}^K(x_i, \ldots, x_{i'}) = \sum_{\substack{K_i,\ldots,K_{i'} \subseteq [k] \\ K_i \oplus \ldots \oplus K_{i'} = K}} f_{K_i}(x_i) \cdot \ldots \cdot f_{K_{i'}}(x_{i'}).$$

For $s \in \{0, \dots, s'\}$ (that is, $2^s \leq 2^{s'} = t'$), denote

$$Term_s = \mathop{\mathbb{E}}_{x_1,\dots,x_{2^s} \sim \mathcal{U}} \left[G^{\phi}_{1,2^s}\left(x_1, \dots, x_{2^s}\right) \right].$$

Using this new notation, our goal is to bound

$$Rej \geq \frac{1}{2^{k(t'-1)}} \cdot Term_{s'} - \gamma. \tag{12}$$

Let us bound $Term_s$ for $s \geq 1$:

$$\begin{aligned}
&Term_s = \\
&\mathop{\mathbb{E}}_{x_1,\dots,x_{2^s} \sim \mathcal{U}} \left[G^{\phi}_{1,2^s}\left(x_1, \dots, x_{s^s}\right) \right] = \\
&\mathop{\mathbb{E}}_{x_1,\dots,x_{2^s} \sim \mathcal{U}} \left[\sum_{K \subseteq [k]} G^{K}_{1,2^{s-1}}\left(x_1, \dots, x_{2^{s-1}}\right) G^{K}_{2^{s-1}+1,2^s}\left(x_{2^{s-1}+1}, \dots, x_{2^s}\right) \right] = \\
&\sum_{K \subseteq [k]} \mathop{\mathbb{E}}_{x_1,\dots,x_{2^{s-1}} \sim \mathcal{U}} \left[G^{K}_{1,2^{s-1}}\left(x_1, \dots, x_{2^{s-1}}\right) \right] \cdot \mathop{\mathbb{E}}_{x_{2^{s-1}+1},\dots,x_{2^s} \sim \mathcal{U}} \left[G^{K}_{2^{s-1}+1,2^s}\left(x_{2^{s-1}+1}, \dots, x_{2^s}\right) \right].
\end{aligned}$$

Observe that $G^{K}_{1,2^{s-1}}$ and $G^{K}_{2^{s-1}+1,2^s}$ are the same function on 2^{s-1} input strings, and that $x_1, \dots, x_{2^{s-1}}$ are distributed the same as $x_{2^{s-1}+1}, \dots, x_{2^s}$ (specifically, both are uniformly distributed). Therefore,

$$\begin{aligned}
Term_s &= \sum_{K \subseteq [k]} \left(\mathop{\mathbb{E}}_{x_1,\dots,x_{2^{s-1}} \sim \mathcal{U}} \left[G^{K}_{1,2^{s-1}}\left(x_1, \dots, x_{2^{s-1}}\right) \right] \right)^2 \\
&\geq \left(\mathop{\mathbb{E}}_{x_1,\dots,x_{2^{s-1}} \sim \mathcal{U}} \left[G^{\phi}_{1,2^{s-1}}\left(x_1, \dots, x_{2^{s-1}}\right) \right] \right)^2 = Term^2_{s-1}.
\end{aligned}$$

By using the last equation recursively, we get

$$Term_{s'} \geq Term^2_{s'-1} \geq \cdots \geq Term_0^{2^{s'}} = \left(\mathop{\mathbb{E}}_{x_1 \sim \mathcal{U}} \left[{G^{\phi}}_{1,1}\left(x_1\right) \right] \right)^{t'} = \left(\mathop{\mathbb{E}}_{x_1 \sim \mathcal{U}} \left[f_{\phi}\left(x_1\right) \right] \right)^{t'} = 1.$$

Using Equation 12 (recall that we use $\delta, \alpha, \beta = \frac{1}{2^{10kt}}$ and $\gamma = \delta + \alpha + \beta$), we get

$$Rej \geq \frac{1}{2^{k(t'-1)}} \cdot 1 - \gamma \geq \frac{1}{2^{2kt}} - 3 \cdot \frac{1}{2^{10kt}} \geq \frac{1}{2^{10kt}}.$$

■

7 Characterization of Covering-Hard Predicates

In this section we prove covering unique games hardness for a large subset of the predicates $\varphi \notin \mathcal{O}$. Formally, we prove Theorem 4 under Conjecture 3. For convenience we restate the conjecture and theorem:

Conjecture (Covering Unique Games). *There exists $c \in \mathbb{N}$ such that for every sufficiently small $\delta > 0$ there exists $R \in \mathbb{N}$ such that the following holds. Given a bipartite label cover instance $\mathcal{LC}$ with permutation constraints over label set $[R]$ and vertex set $U \times V$, it is $\mathcal{NP}$-hard to decide between:*

- ***Yes case:*** *There exist c assignments such that for every vertex $u \in U$, at least one of the assignments satisfies all the edges touching u.*
- ***No case:*** $\mathsf{OPT}\left(\mathcal{LC}\right) \leq \delta$.

Theorem. *Let $\varphi \notin \mathcal{O}$, and assume that there exists a balanced, pairwise independent distribution on the support of φ. Let c be the completeness constant from the covering unique games conjecture. Then gap-cover-$\varphi_{2c,k}$ is covering unique games-hard for every $k \in \mathbb{N}$.*

7.1 Discussion of our Covering Unique Games Conjecture

We would like to follow the lines of [2] and get a conditional hardness result using our dictatorship test. It appears impossible to derive a hardness-of-covering result by combining a dictatorship test with the unique games conjecture because of its inherent imperfect completeness. Additionally, the 2-to-1 conjecture, that does have perfect completeness, does not seem to be nicely suited for transferring a dictatorship test to hardness in a generic way. Thus, we devise our own (covering) variant of the unique games conjecture.

A natural attempt at formulating the covering conjecture would be to require in the yes case the existence of c assignments that together cover all the edges of the given label cover instance, where c is some absolute constant. Unfortunately, we were only able to derive a hardness result using a stronger version of the conjecture. Specifically, in the yes case we require the existence of c assignments such that for every vertex $u \in U$, at least one of the assignments satisfies all the edges touching u. We mention that Khot and Regev [12] show that a similar conjecture in the max-CSP setting is equivalent to the unique games conjecture:

Conjecture 7.1 (Unique Games [12]). *For every sufficiently small $\delta > 0$ there exists $R \in \mathbb{N}$ such that the following holds. Given a bipartite label cover instance $\mathcal{LC}$ with permutation constraints over label set $[R]$ and vertex set $U \times V$, it is $\mathcal{NP}$-hard to decide between:*

- ***Yes case:*** *There exists an assignment that for $1-\delta$ fraction of the vertices $u \in U$, satisfies all the edges touching u.*

- ***No case:*** $\mathsf{OPT}\left(\mathcal{LC}\right) \leq \delta$.

Our conjecture is clearly false with $c = 1$, but as far as we know may be true with even $c = 2$. The conjecture is incomparable to the unique games conjecture (our completeness does not require any single assignment to satisfy a large fraction of edges). However it clearly implies the unique games conjecture with completeness $\frac{1}{c}$ (instead of $1-\epsilon$).

7.2 PCP Verifier (Proof of Theorem 4)

As usual, we prove Theorem 4 by reduction from the covering unique games conjecture (Conjecture 3). Specifically, we assume to be given a bipartite label cover instance $\mathcal{LC}'$ with permutation constraints, and construct a PCP verifier that checks proofs for $\mathcal{LC}'$ by only performing φ-tests.

Let $\mathcal{LC}' = (U, V, E, R, R, \Pi')$, $\Pi' = \left\{\pi'_{v,u} : [R] \to [R]\right\}_{(u,v)\in E}$, be the given instance, and let $\mathcal{LC} = (U, V, E, 2R, 2R, \Pi)$, $\Pi = \{\pi_{v,u} : [2R] \to [2R]\}_{(u,v)\in E}$, be the unique games duplicated-$\mathcal{LC}'$ instance (see Section 2.2.3). A *proof* P for $\mathcal{LC}'$ consists of a collection of truth tables of boolean functions, one for each vertex $v \in V$. Formally, $P = (f_v)_{v\in V}$ where $f_v : \{\pm1\}^{2R} \to \{\pm1\}$. The function f_v is, supposedly, the long code encoding of the label assigned to v by a satisfying assignment for $\mathcal{LC}$.

Our verifier's algorithm for checking the proof P is found in Figure 4. The algorithm uses the following definition. For a function $f : \{\pm1\}^{2R} \to \{\pm1\}$ and a permutation $\pi : [2R] \to [2R]$ we define the function $f\pi : \{\pm1\}^{2R} \to \{\pm1\}$ by

$$f\pi\left(x\right) = f\left(x_{\pi(1)}, \ldots, x_{\pi(2R)}\right).$$

Algorithm 4 Ver_ϵ^P

- Randomly select a vertex $u \in_R U$.
- Run DICT_ϵ^F for $F = \{f^v \pi_{v,u}\}_{v\in\Gamma(u)}$,
  where f^v is the function in P associated with vertex v.

As before, we define $Rej\left(\mathsf{Ver}_\epsilon^{\mathcal{P}}\right)$ to be the indicator random variable for the rejection of the set of proofs $\mathcal{P} = \{P_1, ..., P_k\}$ by Ver_ϵ. Theorem 4 is an easy corollary of the following lemma:

Lemma 7.2. *Let $c \in \mathbb{N}$ be the constant promised by the covering unique games conjecture (Conjecture 3). Let $\varphi : \{\pm 1\}^t \to \{\pm 1\}$ be a predicate satisfying $\varphi \notin \mathcal{O}$. Assume that there exists a balanced, pairwise independent distribution η on the support of φ. Then,* Ver *satisfies the following properties:*

- ***Completeness:*** *Assume that there exist c assignments such that for every vertex $u \in U$, at least one of the assignments satisfies all the edges touching u. Then, there exists a set $\mathcal{P}$ of at most $2c$ proofs such that*

$$\Pr\left[Rej\left(\mathsf{Ver}_\epsilon^{\mathcal{P}}\right)\right] = 0.$$

  In particular, if c assignments cover all the edges of $\mathcal{LC}'$ (in the above sense), then there are $2c$ proofs that together cover all the tests performed by Ver_ϵ.

- ***Soundness:*** *For a sufficiently small $\epsilon > 0$ and $k \in \mathbb{N}$, there exists a constant $\xi > 0$ that only depends on ϵ, t and k, such that the following holds: Assume that there exists a set $\mathcal{P}$ of at most k proofs such that*

$$\Pr\left[Rej\left(\mathsf{Ver}_\epsilon^{\mathcal{P}}\right)\right] < \frac{1}{2 \cdot 2^{10kt}}.$$

  Then, $\mathsf{OPT}\left(\mathcal{LC}'\right) > \xi$.
  In particular, if $\mathsf{OPT}\left(\mathcal{LC}'\right) \leq \xi$, *then there is no constant number of proofs that together cover all the tests performed by* Ver_ϵ.

Proof of Lemma 7.2 (Completeness) Let $L_1, \ldots, L_c : U \cup V \to [R]$ be the c promised assignments for $\mathcal{LC}'$. Specifically, for every $u \in U$ there exists an assignment L_ℓ, $\ell \in [c]$, that satisfies all the edges touching u.

For each $\ell \in [c]$, we construct the two proofs $P_\ell = \{f_\ell^v\}_{v \in V}$ and $Q_\ell = \{g_\ell^v\}_{v \in V}$ for Ver using the assignments L_ℓ and $L_\ell + R$ (respectively). That is, $f_\ell^v, g_\ell^v : \{\pm 1\}^{2R} \to \{\pm 1\}$ satisfy $f_\ell^v = \chi_{L_\ell(v)}$ and $g_\ell^v = \chi_{L_\ell(v)+R}$. We denote $\mathcal{P} = \{P_\ell, Q_\ell\}_{\ell \in [c]}$. We next show that $\mathsf{Ver}^{\mathcal{P}}$ always accepts.

Fix a vertex $u \in U$, and assume that u was selected during the execution of Ver. Let L_ℓ, $\ell \in [c]$, be an assignment that satisfies all the edges touching u. When running $\mathsf{Ver}^{\{P_\ell, Q_\ell\}}$, the verifier runs $\mathsf{DICT}^{\{F_{\ell,u}, G_{\ell,u}\}}$ for the sets $F_{\ell,u} = \{f_\ell^v \pi_{v,u}\}_{v \in \Gamma(u)}$ and $G_{\ell,u} = \{g_\ell^v \pi_{v,u}\}_{v \in \Gamma(u)}$. For every $v \in \Gamma(u)$ it holds that

$$f_\ell^v \pi_{v,u} = \chi_{L_\ell(v)} \pi_{v,u} = \chi_{\pi_{v,u}(L_\ell(v))} = \chi_{L_\ell(u)},$$

and

$$g_\ell^v \pi_{v,u} = \chi_{L_\ell(v)+R} \pi_{v,u} = \chi_{\pi_{v,u}(L_\ell(v)+R)} = \chi_{L_\ell(u)+R}.$$

Using the completeness property of Lemma 5.6, $\mathsf{DICT}^{\{F_{\ell,u},G_{\ell,u}\}}$ always accepts. This implies that $\mathsf{Ver}^{\{P_\ell,Q_\ell\}}$ accepts whenever u is chosen. Therefore, $\mathsf{Ver}^{\mathcal{P}} = \mathsf{Ver}^{\{P_\ell,Q_\ell\}_{\ell\in[c]}}$ always accepts, and the assertion follows. ■

Proof of Lemma 7.2 (Soundness) Let $k \in \mathbb{N}$ and let $\mathcal{P}$ be a set k proofs $\mathcal{P} = \{P_1, \ldots, P_k\}$, $P_\ell = \{f_\ell^v\}_{v\in V}$, $f_\ell^v : \{\pm 1\}^{2R} \to \{\pm 1\}$, for which $\Pr\left[Rej\left(\mathsf{Ver}^{\mathcal{P}}\right)\right] < \frac{1}{2\cdot 2^{10kt}}$. We wish to show that $\mathsf{OPT}\left(\mathcal{LC}'\right) > \xi$, for some constant ξ that only depends on ϵ, t and k.

For simplicity of exposition we assume that $\mathcal{LC}'$ (and therefore also $\mathcal{LC}$) is U-regular, that is $\forall u, u' \in U : |\Gamma(u)| = |\Gamma(u')| = r$ and that r is sufficiently large (as required by soundness property of Lemma 5.6).

Fix a vertex $u \in U$, and let $v_1, \ldots, v_r \in \Gamma(u)$ be its r neighbors. For $i \in [r]$ and $\ell \in [k]$, denote $g_\ell^{v_i} = f_\ell^{v_i} \pi_{v_i,u}$. Let $F_\ell^u = \{g_\ell^{v_1}, \ldots, g_\ell^{v_r}\}$ and $\mathcal{F}^u = \{F_1^u, \ldots, F_k^u\}$. It holds that

$$\Pr\left[Rej\left(\mathsf{Ver}^{\mathcal{P}}\right)\right] = \mathop{\mathbb{E}}_{u\in U}\left[\Pr\left[Rej\left(\mathsf{DICT}^{\mathcal{F}^u}\right)\right]\right].$$

Since $\Pr\left[Rej\left(\mathsf{Ver}^{\mathcal{P}}\right)\right] < \frac{1}{2\cdot 2^{10kt}}$ there exists a subset $U' \subseteq U$, $|U'| \geq \frac{1}{2}|U|$, such that for every $u \in U'$ it holds that $\Pr\left[Rej\left(\mathsf{DICT}^{\mathcal{F}^u}\right)\right] \leq \frac{1}{2^{10kt}}$. We call the vertices in U' *good* vertices.

Fix a good vertex $u \in U'$. Using the soundness property of Lemma 5.6, for some $d \in \mathbb{N}$ and $\tau > 0$ (functions of ϵ, t and k), it holds that at least $\alpha = \frac{1}{2^{10kt}}$ fraction of the t-elements sets $I \subseteq [r]$ are *not* $(d, \tau, \mathcal{F}^u)$-cross regular. Meaning that there exists a pair $(i, i') \in I^2$, $i \neq i'$, that is not $(d, \tau, \mathcal{F}^u)$-cross regular. We say that such a pair is *cross influential for I* with respect to $\mathcal{F}^u$. We call a pair *cross influential* with respect to $\mathcal{F}^u$, if it cross influential with respect to $\mathcal{F}^u$ for at least one set I.

Denote by $XInfPairs^u \subseteq [r]^2$ the set of cross influential pairs with respect to $\mathcal{F}^u$. Formally,

$$XInfPairs^u = \left\{(i, i') \in [r]^2 \,|\, \exists K, K' \subseteq [k] : XInf_d\left(g_K^{v_i}, g_{K'}^{v_{i'}}\right) \geq \tau\right\}.$$

We claim that the set $XInfPairs^u$ contains at least $\frac{\alpha}{t^2}$ fractions of the pairs in $[r]^2$. The following is a way of choosing a random pair $(i, i') \in [r]^2$: First select a t-elements set $I \subseteq [r]$, then select a random pair $(i, i') \in I^2$. The selected set I has a cross influential pair with probability is at least α. Each of the pairs in I is selected with probability $\frac{1}{t^2}$. Thus, the selected pair is cross influential with probability at least $\frac{\alpha}{t^2}$.

Obtaining a good labeling. We next construct a good labeling for the duplicated label cover instance $\mathcal{LC}$. Since every assignment for $\mathcal{LC}$ naturally induces an assignment for $\mathcal{LC}'$ with at least the same value, the claim of the lemma follows.

Consider the following labeling L for $\mathcal{LC}$: The set of candidate assignments for vertex $v \in V$ is given by

$$C(v) = \left\{ j \in [2R] \,|\, \exists K \subseteq [k] : Inf_j^{\leq d}(f_K^v) \geq \tau \right\}.$$

Note that, using Claim 2.4, $|C(v)| \leq \frac{d}{\tau} \cdot 2^k$. Define a labeling L by picking, for each $v \in V$ a label $L(v)$ uniformly at random from $C(v)$ (or an arbitrary label if $C(v)$ is empty). For $u \in U$, randomly select $v' \in_R \Gamma(u)$ and set $L(u)$ to $\pi_{v',u}(L(v'))$.

Let (u, v) be an edge of $\mathcal{LC}$, where $u \in U'$ is good. The probability that the edge (u, v) is satisfied by L is

$$\Pr[L(u) = \pi_{v,u}(L(v))] = \Pr[\pi_{v',u}(L(v')) = \pi_{v,u}(L(v))].$$

Recall that with probability at least $\frac{\alpha}{t^2}$ it holds that (v, v') is a cross influential pair with respect to $\mathcal{F}^u$, i.e., $v = v_i$, $v' = v_{i'}$, and $(i, i') \in XInfPairs^u$. Therefore, with probability at least $\frac{\alpha}{t^2}$, it holds that $\pi_{v',u}(C(v')) \cap \pi_{v,u}(C(v)) \neq \phi$. Conclude that the edge (u, v) is satisfied with probability at least $\frac{\alpha}{t^2} \cdot \frac{1}{|C(v)||C(v')|} \geq \frac{\alpha}{t^2} \left(\frac{\tau}{d \cdot 2^k}\right)^2$.

Since we assume that $\mathcal{LC}$ is U-regular, and since $|U'| \geq \frac{1}{2}|U|$, it holds that $\frac{1}{2}$ of the edges (u, v) have $u \in U'$. Therefore, a random edge of $\mathcal{LC}$ is satisfied by L with probability at least $\frac{1}{2} \cdot \frac{\alpha}{t^2} \cdot \left(\frac{\tau}{d \cdot 2^k}\right)^2 \geq \frac{\tau^2}{d^2 \cdot 2^{20kt}} = \xi$. ■

8 Hardness of Approximate Coloring and Covering Random CSP Instances

An outstanding open question is to approximate the number of colors required to color a given $O(1)$-colorable graph or hypergraph. While it is known to be hard to color a $O(1)$-colorable hypergraph with a polylogarithmic number of colors, the best known algorithm requires a polynomial number of colors. Thus, there is an exponential gap between the best lower and upper bounds. In the covering language this is almost[1] equivalent to the question of approximating the covering number of an $O(1)$-coverable NAE instance. We next study this question in relation to the hardness of random CSP instances.

[1] It is not exactly equivalent since the NAE formulation allows negations of variables whereas the coloring formulation does not.

In a seminal paper, Feige [5] studies the relation between hardness of random instances of 3SAT and the hardness of approximation problems, including some notorious problems for which neither algorithms nor hardness are known. In that paper he states a hypothesis about no polynomial time algorithm being able to distinguish between a random 3SAT and a satisfiable one. More accurately,

Hypothesis 8.1 (**Feige's Hypothesis 1** [5])**.** *There is no polynomial time algorithm that outputs* `typical` *for most* 3CNF-*CSP instances with n variables and $m = \Delta \cdot n$ clauses, and never outputs* `typical` *on a satisfiable instance; even when Δ is an arbitrarily large constant independent of n.*

We formulate an analogous hypothesis about the hardness of distinguishing between random and 2-coverable 4LIN-CSP instances (Hypothesis 8.2, which is a restatement of Hypothesis 5), and a weaker hypothesis about φ-CSP instances for some predicate φ (Hypothesis 8.3). We prove that both of these hypotheses imply the hardness of approximate coloring of hypergraphs. We show a direct connection between the density Δ in the hypothesis and the inapproximability factor in the result. When our 4LIN hypothesis is pushed to extreme, it implies hardness of approximate coloring to within polynomial factors.

Hypothesis 8.2 (**Covering 4LIN Hypothesis, with density parameter** Δ)**.** *There is no polynomial time algorithm that outputs* `typical` *for most* 4LIN-*CSP instances with n variables and $m = \Delta \cdot n$ clauses, and never outputs* `typical` *for a 2-coverable instance.*

The following hypothesis is about a general predicate and is weaker than the former hypothesis since it is implied by it.

Hypothesis 8.3 (**Covering φ Hypothesis, with density parameter** Δ)**.** *There are some universal constants $t, c \in \mathbb{N}$ and a predicate $\varphi : \{\pm 1\}^t \to \{\pm 1\}$ such that no polynomial time algorithm outputs* `typical` *for most φ-CSP instances with n variables and $m = \Delta \cdot n$ clauses, and never outputs* `typical` *for a c-coverable φ-CSP instance.*

Our main theorem of this section is Theorem 8.4 (generalized restatement of Theorem 6):

Theorem 8.4. *If Hypothesis 8.3 holds with parameters c, t and density Δ such that $c \ll \log \Delta$ then it is hard to distinguish if a given t-uniform hypergraph is 2^c-colorable or $\Delta^{\Omega(1)}$ colorable.*

In particular, Hypothesis 8.2 with density parameter $\Delta = n^\delta$ for some positive $\delta > 0$ implies that it is hard to decide if a 4-uniform hypergraph is 4-colorable or requires at least a polynomial number of colors.

We first show that the covering number of a random φ-CSP is proportional to its log-density, as long as $\varphi \notin \mathcal{O}$. (Recall that for any $\varphi \in \mathcal{O}$, the covering number of any φ-CSP is at most 2).

Claim 8.5. *Let $\varphi \notin \mathcal{O}$ and let $\mathcal{C}$ be a random instance of φ-CSP, with n variables and $m = \Delta \cdot n$ constraints. Then $\nu(\mathcal{C}) \geq \Omega(\log \Delta)$, except with probability exponentially small in n.*

Proof Let us first assume that φ is the NAE predicate on t variables and that all occurrences of φ are unsigned, i.e., without negations of variables. Fix an CSP instance $\mathcal{C}$, and let $L_1, \ldots, L_k \in \{\pm 1\}^n$ be any set of $k \in \mathbb{N}$ assignments for $\mathcal{C}$. It is not hard to see that there must be a subset $S \subseteq [n]$, such that each assignment L_ℓ, $\ell \in [k]$, is constant on S (either all 1s or all -1s), and such that $|S| \geq n \cdot 2^{-k}$. The reason is that each of the assignments L_ℓ partitions the n variables into two sets: Variables that are assigned the value 1, and variable that are assigned the value -1.

If the given instance $\mathcal{C}$ has a constraint fully contained in S then $L_1, \ldots, L_k$ do not cover it. The probability that a randomly chosen constraint is contained in a set of size $n \cdot 2^{-k}$ is 2^{-kt} where t is the arity of the constraint. The probability that out of m constraints of $\mathcal{C}$ none landed inside S is

$$\left(1 - 2^{-kt}\right)^m \approx \exp\left(\frac{-m}{2^{kt}}\right),$$

and if we multiply this by the number 2^{kn} of possibilities to choose k assignments and using a union bound we get

$$\Pr_I\left[\nu(I) \leq k\right] \leq \exp\left(\frac{-m}{2^{kt}}\right) \cdot 2^{kn} = \exp\left(-n \cdot \left(\frac{\Delta}{2^{kt}} - k\right)\right).$$

Clearly if $\Delta > 2^{2kt}$ then $\frac{\Delta}{2^{kt}} - k > 1$ which causes the above probability to be exponentially small. In our case t is fixed, and so this proves that $\nu(\mathcal{C}) \geq \Omega(\log \Delta)$ with high probability.

It remains to justify the assumption that φ is the NAE predicate. This simply follows from the fact that for every $\varphi \notin \mathcal{O}$ there is some signed-NAE predicate that contains it, see Claim 2.2. The unsigned assumption means that we've proven that even covering the unsigned part of the instance is already hard, assuming that there are many unsigned constraints. But this is indeed the case as the number of unsigned constraints is expected to be $m \cdot 2^{-t}$. ■

The proof of Theorem 8.4 now follows.

Proof of Theorem 8.4 Assume Hypothesis 8.3 with density parameter Δ, and let $c, t \in \mathbb{N}$ and $\varphi : \{\pm 1\}^t \to \{\pm 1\}$ be such that no algorithm can decide if a given φ-CSP instance $\mathcal{C}$ is random or whether $\nu(\mathcal{C}) \leq c$. We can assume that $\varphi \notin \mathcal{O}$ otherwise the hypothesis is clearly false since $\nu(\mathcal{C}) \leq 2$ for all φ-CSP instances $\mathcal{C}$.

We reduce the problem of deciding if a given φ-CSP instance is random or has covering number at most c, to the problem of deciding if a given hypergraph has chromatic number at least $\Delta^{\Omega(1)}$ or at most 2^c.

Let $\mathcal{C}$ be a given φ-CSP instance. By Claim 2.2 there is some sign $\sigma = (\sigma_1, \ldots, \sigma_t) \in \{\pm 1\}^t$ such that the support of $\varphi(\sigma_1 x_1, \ldots, \sigma_t x_t)$ is contained in the support of $\mathsf{NAE}(x_1, \ldots, x_t)$. Let $\mathcal{C}' \subseteq \mathcal{C}$ be the sub-instance of $\mathcal{C}$ consisting only of the clauses that occur with the sign σ. In other words, recall that each clause is obtained by applying φ to t *literal*s. We denote by $\mathcal{C}'$ the subset of $\mathcal{C}$ containing clauses of the form $\varphi(\sigma_1 x_{i_1}, \ldots, \sigma_i x_{i_t})$, where $i_1, \ldots, i_t \in [n]$. If $\mathcal{C}$ is a random φ-CSP instance, $\mathcal{C}'$ is expected to have density 2^{-t} in $\mathcal{C}$.

We construct a t-uniform hypergraph H over the vertex set $[n]$ by turning each variable into a vertex and each constraint in $\mathcal{C}'$ into a hyperedge. We prove that if $\nu(\mathcal{C}) = c$ then H is 2^c-colorable, and if $\mathcal{C}$ is random, then when viewing H as a NAE-CSP instance it holds that $\nu(H) \geq \Omega(\log \Delta)$ with high probability.

Suppose $\nu(\mathcal{C}) \leq c$, and let $L_1, \ldots, L_c : [n] \to \{\pm 1\}$ be c covering assignments. We can color the vertices of this hypergraph by 2^c colors by assigning each vertex $v \in [n]$ the color $(L_1(v), \ldots, L_c(v)) \in \{\pm 1\}^c$. No hyperedge $\{v_1, \ldots, v_t\} \in H$ is monochromatic since that would mean that

$$\forall i \in [c], \quad L_i(v_1) = \cdots = L_i(v_t).$$

But this means that all c assignments violate the NAE constraint on $\{v_1, \ldots, v_t\}$, and therefore also violate the φ constraint (with sign σ) on $\{v_1, \ldots, v_t\}$.

If, on the other hand, $\mathcal{C}$ is a random instance, then H is just a random hypergraph, and by Claim 8.5 its covering number is at least $\Omega(\log \Delta)$. ■

8.1 Discussion of our Hypotheses

Our hypothesis 8.3 differs from Feige's on two counts. First, the choice of predicate in our hypothesis is not a $\mathsf{3CNF}$. Our hypothesis would clearly be false for $\mathsf{3CNF}$ simply because it is in $\mathcal{O}$ and easily coverable by two assignments. It seems to us that there is nothing particularly special about $\mathsf{3CNF}$ and that Feige's hypothesis, if true, should be true with many other predicates, including ones that are not in $\mathcal{O}$. Unfortunately, there are virtually[2] no direct reductions between random instances of one predicate to another. If Feige's hypothesis were to hold for any predicate $\varphi \notin \mathcal{O}$ it would immediately imply our hypothesis for the same Δ, and with $c = 1$.

Whether or not these hypotheses are true for higher values of Δ is less clear. Feige only relies on arbitrarily large constants Δ, and does hypothesize about larger-density formulas.

[2] Excluding 'trivial' cases in which one predicate is contained in another predicate.

The current state-of-the-art algorithms are able to refute random 3CNF formulas only if the density is at least $\Delta \geq n^{0.5}$ [6], so even the strongest form of our Hypothesis 8.2 is consistent with the current knowledge.

Finally, we note that our $\mathcal{NP}$-hardness result for approximating the cover number of 4LIN (Theorem 1) can be taken as some evidence supporting our hypotheses. First, if our hypothesis were true then one would expect such an $\mathcal{NP}$-hardness result to hold. Second, this shows at the very least that if $\mathcal{P} \neq \mathcal{NP}$ we don't expect any of the known algorithmic techniques (essentially, SDPs) to refute the hypotheses.

References

[1] Sanjeev Arora, Eden Chlamtac, and Moses Charikar. New approximation guarantee for chromatic number. In *STOC*, pages 215–224, 2006.

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

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

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

[5] Uriel Feige. Relations between average case complexity and approximation complexity. In *STOC*, pages 534–543, 2002.

[6] Uriel Feige and Eran Ofek. Easily refutable subformulas of large random 3CNF formulas. *Theory of Computing*, 3(1):25–43, 2007.

[7] Venkatesan Guruswami, Johan Håstad, and Madhu Sudan. Hardness of approximate hypergraph coloring. *SIAM Journal on Computing*, 31(6):1663–1686, 2002.

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

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

[10] Subhash Khot. Hardness results for approximate hypergraph coloring. In *STOC*, pages 351–359, 2002.

[11] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007.

[12] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within 2-epsilon. *Journal of Computer and System Sciences*, 74(3):335–349, 2008.

[13] Subhash Khot and Rishi Saket. A 3-query non-adaptive pcp with perfect completeness. In *IEEE Conference on Computational Complexity*, pages 159–169, 2006.

[14] Elchanan Mossel. Gaussian bounds for noise correlation of functions and tight analysis of long codes. In *FOCS*, pages 156–165, 2008.

[15] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. In *FOCS*, pages 21–30, 2005.

[16] Ryan O'Donnell and John Wright. A new point of NP-hardness for unique games. In *STOC*, pages 289–306, 2012.

[17] Ryan O'Donnell and Yi Wu. Conditional hardness for satisfiable 3-CSPs. In *STOC*, pages 493–502, 2009.