THEORY OF COMPUTING, Volume 16 (16), 2020, pp. 1–30
www.theoryofcomputing.org

A Characterization of Hard-to-Cover CSPs

Amey Bhangale* Prahladh Harsha† Girish Varma‡

Received October 13, 2016; Revised September 10, 2018; Published December 13, 2020

Abstract. We continue the study of the *covering complexity* of constraint satisfaction problems (CSPs) initiated by Guruswami, Håstad and Sudan [SIAM J. Comp. 2002] and Dinur and Kol [CCC'13]. The covering number of a CSP instance Φ is the smallest number of assignments to the variables of Φ, such that each constraint of Φ is satisfied by at least one of the assignments. We show the following results:

1. Assuming a covering variant of the Unique Games Conjecture, introduced by Dinur and Kol, we show that for every non-odd predicate P over any constant-size alphabet and every integer K, it is NP-hard to approximate the covering number within a factor of K. This yields a complete characterization of CSPs over constant-size alphabets that are hard to cover.

2. For a large class of predicates that are contained in the $2k$-LIN predicate, we show that it is quasi-NP-hard to distinguish between instances with covering number at most 2 and those with covering number at least $\Omega(\log\log n)$. This generalizes and improves the 4-LIN covering hardness result of Dinur and Kol.

A preliminary version of this paper appeared in *Proceedings of the 30th Computational Complexity Conference (CCC)*, 2015 [6].

*Research supported by the NSF grant CCF-1253886. Part of the work was done when the author was visiting TIFR.

†Supported in part by UGC-ISF Grant (6-2/2014(IC)) and the Swarnajayanti Fellowship Grant.

‡Supported by Google India under the Google India Ph. D. Fellowship Award.

ACM Classification: F.2.2

AMS Classification: 68Q25

Key words and phrases: hardness of approximation, covering complexity, odd predicates

 DOI: 10.4086/toc.2020.v016a016

1 Introduction

One of the central (yet unresolved) questions in inapproximability is the problem of coloring a (hyper)graph with as few colors as possible. A (hyper)graph $G = (V,E)$ is said to be k-colorable if there exists a coloring $c : V \to [k] := \{0,1,2,\dots,k-1\}$ of the vertices such that no (hyper)edge of G is monochromatic. The chromatic number of a (hyper)graph, denoted by $\chi(G)$, is the smallest k such that G is k-colorable. It is known that computing $\chi(G)$ to within a multiplicative factor of $n^{1-\varepsilon}$ on an n-vertex graph G is NP-hard for every $\varepsilon \in (0,1)$ [9, 23]. However, the complexity of the following problem is not yet completely understood: given a constant-colorable (hyper)graph, what is the minimum number of colors required to color the vertices of the graph efficiently such that every edge is non-monochromatic? The current best approximation algorithms for this problem require at least $n^{\Omega(1)}$ colors [17] while the hardness results are far from proving optimality of these approximation algorithms. (See Sec. 1.3 for a discussion on recent work in this area.)

The notion of *covering complexity* was introduced by Guruswami, Håstad and Sudan [11] and more formally by Dinur and Kol [8] to obtain a better understanding of the complexity of this problem. Let P be a predicate and Φ an instance of a constraint satisfaction problem (CSP) over n variables, where each constraint in Φ is a constraint of type P over the n variables and their negations. We will refer to such CSPs as P-CSPs. The *covering number* of Φ, denoted by $\nu(\Phi)$, is the smallest number of assignments to the variables such that each constraint of Φ is satisfied by at least one of the assignments, in which case we say that the set of assignments *covers* the instance Φ. If c assignments cover the instance Φ, we say that Φ is c-coverable or equivalently that the set of assignments form a c-covering for Φ. The covering number is a generalization of the notion of chromatic number (to be more precise, the logarithm of the the chromatic number) to all predicates in the following sense. Let G_Φ be the underlying *constraint (hyper)graph* of the instance Φ whose vertices are the variables of the instance Φ and (hyper)edges are in one-to-one correspondence with the constraints of Φ. Suppose P is the not-all-equal predicate NAE and the instance Φ has no negations in any of its constraints, then the covering number $\nu(\Phi)$ is exactly $\lceil \log \chi(G_\Phi) \rceil$ where G_Φ is the underlying constraint graph of the instance Φ.

Cover-P refers to the problem of finding the covering number of a given P-CSP instance. Finding the exact covering number for most interesting predicates P is NP-hard. We therefore study the problem of approximating the covering number. In particular, we would like to study the complexity of the following problem, denoted by COVERING-P-CSP(c,s), for some $1 \le c < s \in \mathbb{N}$: "given a c-coverable P-CSP instance Φ, find an s-covering for Φ". Similar problems have been studied for the Max-CSP setting: "for $0 < s < c \le 1$, "given a c-satisfiable P-CSP instance Φ, find an s-satisfying assignment for Φ". Max-CSPs and Cover-CSPs, as observed by Dinur and Kol [8], are very different problems. For instance, if P is an odd predicate, i.e, if for every assignment x, either x or its negation $x+\overline{1}$ satisfies P, then any P-CSP instance Φ has a trivial 2-covering any assignment and its negation. Thus, 3-LIN and 3-CNF[1], being odd predicates, are easy to cover though they are hard predicates in the Max-CSP setting. The main result of Dinur and Kol is that the 4-LIN predicate which accepts odd parities, in contrast to the above, is hard to cover: for every constant $t \ge 2$, COVERING-4-LIN-CSP$(2,t)$ is NP-hard. In fact, their arguments show that COVERING-4-LIN-CSP$(2,\Omega(\log\log\log n))$ is quasi-NP-hard.

[1] k-LIN : $\{0,1\}^k \to \{0,1\}$ refers to the k-bit predicate defined by $\mathsf{k\text{-}LIN}(x_1,x_2,\dots,x_k) := x_1 \oplus x_2 \oplus \cdots \oplus x_k$ while 3-CNF : $\{0,1\}^3 \to \{0,1\}$ refers to the 3-bit predicate defined by $\mathsf{3\text{-}CNF}(x_1,x_2,x_3) := x_1 \vee x_2 \vee x_3$

Having observed that CSPs based on odd predicates are easy to cover, Dinur and Kol proceeded to ask the question "are all non-odd-predicate CSPs hard to cover?" In a partial answer to this question, they showed that assuming a covering variant of the Unique Games Conjecture, COVERING-UGC(c), if a predicate P is not odd and there is a balanced pairwise independent distribution on its support, then for all constants k, COVERING-P-CSP$(2c,k)$ is NP-hard. (Here, c is a fixed constant that depends on the covering variant of the Unique Games Conjecture COVERING-UGC(c).) See Sec. 2 for the exact definition of the covering variant of the Unique Games Conjecture.

1.1 Our results

Our first result states that assuming the same covering variant of the Unique Games Conjecture, COVERING-UGC(c), of Dinur and Kol [8], one can in fact show the covering hardness of *all* non-odd predicates P over *any* constant-size alphabet $[q]$. The notion of odd predicate can be extended to any alphabet in the following natural way: a predicate $P \subseteq [q]^k$ is odd if for all assignments $x \in [q]^k$, there exists $a \in [q]$ such that the assignment $x+\overline{a}$ satisfies P.

Theorem 1.1 (Covering hardness of non-odd predicates)**.** *Assuming* COVERING-UGC(c)*, for any constant-size alphabet* $[q]$*, any constant* $k \in \mathbb{N}$ *and any* non-odd *predicate* $P \subseteq [q]^k$*, for all constants* $t \in \mathbb{N}$*, the* COVERING-P-CSP$(2cq,t)$ *problem is* NP*-hard.*

Since odd predicates $P \subseteq [q]^k$ are trivially coverable with q assignments, the above theorem, gives a *full characterization of hard-to-cover predicates* over any constant-size alphabet (modulo the covering variant of the Unique Games Conjecture): a predicate is hard to cover iff it is not odd.

We then ask if we can prove similar covering hardness results under more standard complexity assumptions (such as NP $\neq$ P or the exponential-time hypothesis (ETH)). Though we are not able to prove that every non-odd predicate is hard under these assumptions, we give sufficient conditions on the predicate P for the corresponding approximate covering problem to be quasi-NP-hard. Recall that $2\mathsf{k}\text{-}\mathsf{LIN} \subseteq \{0,1\}^{2k}$ is the predicate corresponding to the set of odd parity strings in $\{0,1\}^{2k}$.

Theorem 1.2 (NP hardness of Covering)**.** *Let* $k \geq 2$*. Let* $P \subseteq 2\mathsf{k}\text{-}\mathsf{LIN}$ *be any* $2k$*-bit predicate such there exist distributions* $\mathcal{P}_0, \mathcal{P}_1$ *supported on* $\{0,1\}^k$ *with the following properties:*

1. *the marginals of* $\mathcal{P}_0$ *and* $\mathcal{P}_1$ *on all* k *coordinates are uniform,*
2. *every* $a \in \mathsf{supp}(\mathcal{P}_0)$ *has even parity and every* $b \in \mathsf{supp}(\mathcal{P}_1)$ *has odd parity and furthermore, both* $a \diamond b, b \diamond a \in P$*, where* $a \diamond b$ *denotes the* $2k$*-bit string formed by the concatenation of strings* a *and* b*.*

Then for all $\varepsilon > 0, r \gg 1$*, there is a reduction from 3SAT to* COVERING-P-CSP *mapping a 3SAT instance* Ψ *on* n *variables to a* COVERING-P-CSP *instance* Φ *of size* $n^{O(r)}2^{2^{O(r)}}$ *in time* $n^{O(r)}2^{2^{O(r)}}$ *such that*

- *YES Case: If the 3SAT formula* Ψ *is satisfiable then there are* 2 *assignments each satisfying* $1-\varepsilon$ *of the constraints of* Φ*, that together cover the instance* Φ*.*

- *NO Case: if the 3SAT formula* Ψ *is not satisfiable, then the resulting instance* Φ *is not* $\Omega_k(r) - O_k(\log(1/\varepsilon))$ *coverable, even when considered as an instance of the (potentially larger) predicate* 2k-LIN.

In particular, unless $\mathrm{NP} \subseteq \mathrm{DTIME}(2^{\mathrm{poly}\log n})$*,* COVERING-$P$-CSP$(2, \Omega(\log\log n))$ *does not have a polynomial-time algorithm.*

If we assume $\mathrm{P} \neq \mathrm{NP}$ *then* COVERING-P-CSP$(2, C)$ *does not have a polynomial-time algorithm for any constant* $C > 2$.

The furthermore clause in the soundness guarantee is in fact a strengthening for the following reason: if two predicates P, Q satisfy $P \subseteq Q$ and Φ is a c-coverable P-CSP instance, then the Q-CSP instance $\Phi_{P\to Q}$ obtained by taking the constraint graph of Φ and replacing each P constraint with the weaker Q constraint, is also c-coverable.

The following is a simple corollary of the above theorem.

Corollary 1.3. *Let* $k \geq 2$ *be even,* $x, y \in \{0,1\}^k$ *be distinct strings having even and odd parity, respectively, and* $\bar{x}, \bar{y}$ *denote the complements of x and y, respectively. For any predicate P satisfying*

$$\text{2k-LIN} \supseteq P \supseteq \{x \diamond y,\ x \diamond \bar{y},\ \bar{x} \diamond y,\ \bar{x} \diamond \bar{y},\ y \diamond x,\ y \diamond \bar{x},\ \bar{y} \diamond x,\ \bar{y} \diamond \bar{x}\},$$

unless $\mathrm{NP} \subseteq \mathrm{DTIME}(2^{\mathrm{poly}\log n})$*, the problem* COVERING-P-CSP$(2, \Omega(\log\log n))$ *is not solvable in polynomial time.*

This corollary implies the covering hardness of 4-LIN predicate proved by Dinur and Kol [8] by setting $x := 00$ and $y := 01$. With respect to the covering hardness of 4-LIN, we note that we can considerably simplify the proof of Dinur and Kol and in fact obtain a even stronger soundness guarantee (see Theorem below). The stronger soundness guarantee in the theorem below states that there are no large ($\geq 1/\mathrm{poly}\log n$ fractional-size) independent sets in the constraint graph and hence, even the 4-NAE-CSP instance[2] with the same constraint graph as the given instance is not coverable using $\Omega(\log\log n)$ assignments. Both the Dinur–Kol result and the above corollary only guarantee (in the soundness case) that the 4-LIN-CSP instance is not coverable.

Theorem 1.4 (Hardness of Covering 4-LIN)**.** *Assuming that* $\mathrm{NP} \not\subseteq \mathrm{DTIME}(2^{\mathrm{poly}\log n})$*, for all* $\varepsilon \in (0,1)$*, there does not exist a polynomial-time algorithm that can distinguish between* 4-LIN-CSP *instances of the following two types:*

- *YES Case : There are* 2 *assignments such that each of them covers* $1 - \varepsilon$ *fraction of the constraints, and they together cover the entire instance.*

- *NO Case : The largest independent set in the constraint graph of the instance is of fractional size at most* $1/\mathrm{poly}\log n$.

[2]The k-NAE predicate over k bits is given by $k\text{-NAE} = \{0,1\}^k \setminus \{\bar{0}, \bar{1}\}$.

1.2 Techniques

As one would expect, our proofs are very much inspired from the corresponding proofs in Dinur and Kol [8]. One of the main complications in the proof of Dinur and Kol [8] (as also in the earlier work of Guruswami, Håstad and Sudan [11]) was the one of handling several assignments simultaneously while proving the soundness analysis. For this purpose, both these works considered the rejection probability that all the assignments violated the constraint. This resulted in a very tedious expression for the rejection probability, which made the rest of the proof fairly involved. Holmerin [14] observed that this can be considerably simplified if one instead proved a stronger soundness guarantee that the largest independent set in the constraint graph is small (this might not always be doable, but in the cases when it is, it simplifies the analysis). We list below the further improvements in the proof that yield our Theorems 1.1, 1.2 and 1.4.

Covering-UG hardness for non-odd predicates (Theorem 1.1). Having observed that it suffices to prove an independent set analysis, we observed that only very mild conditions on the predicate are required to prove covering hardness. In particular, while Dinur and Kol used the Austrin–Mossel test [3] which required pairwise independence, we are able to import the long-code test of Bansal and Khot [4] which requires only 1-wise independence. We remark that the Bansal–Khot Test was designed for a specific predicate (hardness of finding independent sets in almost k-partite k-uniform hypergraphs) and had imperfect completeness. Our improvement comes from observing that their test requires only 1-wise independence and furthermore that their completeness condition, though imperfect, can be adapted to give a 2-cover composed of 2 nearly satisfying assignments using the *duplicate label technique* of Dinur–Kol. This enlarges the class of non-odd predicates for which one can prove covering hardness (see Theorem 3.1). We then perform a sequence of reductions from this class of CSP instances to CSP instances over all non-odd predicates to obtain the final result. Interestingly, one of the open problems mentioned in the work of Dinur and Kol [8] was to devise "direct" reductions between covering problems. The reductions we employ, strictly speaking, are not "direct" reductions between covering problems, since they rely on a stronger soundness guarantee for the source instance (namely, large covering number even for the NAE instance on the same constraint graph), which we are able to prove in Theorem 3.1.

We give an overview of the dictatorship test gadget which when composed with a covering-UG instance, gives the required covering hardness result. Let $P \subseteq [q]^k$ be a predicate such that there exists $a \in \mathsf{NAE}$ and

$$\mathsf{NAE} \supset P \supseteq \{a + \bar{b} \mid b \in [q]\},$$

i. e., P accepts all shifts of a particular assignment $a \in [q]^k$ where $a \in \mathsf{NAE}$. We are given a function $f : [q]^{2L} \to [q]$ and are interested in a k-query test, querying at $(x_1, x_2, \ldots, x_k)$ according to some distribution $\mathcal{D}$, which has the following three properties:

1. The accepting criteria of the test is $(f(x_1), f(x_2), \ldots, f(x_k)) \in P$
2. For every $i \in [L]$, the test should accept with probability 1 if f is either the i-th dictator or the $(i+L)$-th dictator.
3. If f is far from any dictator then the test, even with the predicate P replaced by NAE, should reject with significant probability.

We can think of the queries as a $k \times 2L$ matrix X where the rows represent $x_1, x_2, \ldots, x_k$. Here is a distribution $\mathcal{D}$ for which the test has all the above three properties: It will be a L-wise product distribution $\mu^{\otimes L}$, where μ is a distribution on $([q]^k)^2$ sampled uniformly from the set S,

$$S := \left\{ (y, y') \in [q]^k \times [q]^k \mid y \in \{a + \bar{b} \mid b \in [q]\} \vee y' \in \{a + \bar{b} \mid b \in [q]\} \right\}.$$

For each $i \in [L]$ we sample the i-th and $(i+L)$-th columns of X independently from μ. This completes the description of the distribution $\mathcal{D}$. It is clear from the construction that the test with accepting criteria (1) satisfies (2) as either the i-th column or the $(i+L)$-th column contains an accepting assignment. The argument that (3) also holds for this test crucially depends on the properties of the distribution μ – that each query x_i is distributed uniformly in $\{0,1\}^{2L}$ and the distribution μ is *connected* (see Definition 2.7), when viewed as a probability space $(([q]^2)^k, \mu)$. Using both these properties of the distribution μ, we can then apply the invariance principle to argue that the constrained (hyper)graph formed by the test distribution has a small independent set, which in turns imply (3).

Quasi-NP hardness result (Theorem 1.2). In this setting, we unfortunately are not able to use the simplification arising from using the independent set analysis and have to deal with the issue of several assignments. One of the steps in the 4-LIN proof of Dinur and Kol (as in several others results in this area) involves showing that a expression of the form $\mathbb{E}_{(X,Y)}[F(X)F(Y)]$ is not too negative where (X,Y) is not necessarily a product distribution but the marginals on the X and Y parts are identical. Observe that if (X,Y) was a product distribution, then the above expressions reduces to $(\mathbb{E}_X[F(X)])^2$, a positive quantity. Thus, the steps in the proof involve constructing a tailor-made distribution (X,Y) such that the error in going from the correlated probability space (X,Y) to the product distribution $(X \otimes Y)$ is not too much. More precisely, the quantity

$$\left| \mathop{\mathbb{E}}_{(X,Y)}[F(X)F(Y)] - \mathop{\mathbb{E}}_{X}[F(X)] \mathop{\mathbb{E}}_{Y}[F(Y)] \right|,$$

is small. Dinur and Kol used a distribution tailor-made for the 4-LIN predicate and used an invariance principle for correlated spaces to bound the error while transforming it to a product distribution. Our improvement comes from observing that one could use an alternate invariance principle (see Theorem 2.9) that works with milder restrictions and hence works for a wider class of predicates. This invariance principle for correlated spaces (Theorem 2.9) is an adaptation of invariance principles proved by Wenner [22] and Guruswami and Lee [12] in similar contexts. The rest of the proof is similar to the 4-LIN covering hardness proof of Dinur and Kol.

Covering hardness of 4-LIN (Theorem 1.4). The simplified proof of the covering hardness of 4-LIN follows directly from the above observation of using an independent set analysis instead of working with several assignments. In fact, this alternate proof eliminates the need for using results about correlated spaces [18], which was crucial in the Dinur–Kol setting. We further note that the quantitative improvement in the covering hardness ($\Omega(\log\log n)$ over $\Omega(\log\log\log n)$) comes from using a LABEL-COVER instance with a better smoothness property (see Theorem 2.5).

1.3 Recent work on approximate coloring

Besides the work on covering complexity, the works most related to our paper are the series of works that study the approximate coloring complexity question, stated in the beginning of the introduction. Saket [20] showed that unless $\mathrm{NP} \subseteq \mathrm{DTIME}(2^{\mathrm{poly}\log n})$, it is not possible to color a 2-colorable 4-uniform hypergraph with $\mathrm{poly}\log n$ colors. We remark that recently, with the discovery of the short code [5], there has been a sequence of works [7, 10, 16, 21, 15] which have considerably improved the status of the approximate coloring question. In particular, we know that it is quasi-NP-hard to color a 2-colorable 8-uniform hypergraph with $2^{(\log n)^c}$ colors for some constant $c \in (0,1)$. Stated in terms of covering number, this result states that it is quasi-NP-hard to cover a 1-coverable 8-NAE-CSP instance with $(\log n)^c$ assignments. It is to be noted that these results pertain to the covering complexity of specific predicates (such as NAE) whereas our results are concerned with classifying which predicates are hard to cover. It would be interesting if Theorems 1.2 and 1.4 can be improved to obtain similar hardness results (i. e., $\mathrm{poly}\log n$ as opposed to $\mathrm{poly}\log\log n$). The main bottleneck here seems to be reducing the uniformity parameter (namely, from 8).

Organization

The rest of the paper is organized as follows. We start with some preliminaries of LABEL-COVER, covering CSPs and Fourier analysis in Sec. 2. Theorems 1.1, 1.2 and 1.4 are proved in Sections 3, 4 and 5, respectively.

2 Preliminaries

2.1 Covering CSPs

We will denote the set $\{0,1,\cdots q-1\}$ by $[q]$. For $a \in [q], \bar{a} \in [q]^k$ is the element with a in all the k coordinates (where k and q will be implicit from the context).

Definition 2.1 (P-CSP)**.** For a predicate $P \subseteq [q]^k$, an instance of P-CSP is given by a (hyper)graph $G = (V,E)$, referred to as the *constraint graph*, and a literals function $L : E \to [q]^k$, where V is a set of variables and $E \subseteq V^k$ is a set of constraints. An assignment $f : V \to [q]$ is said to *cover* a constraint $e = (v_1, \cdots, v_k) \in E$, if $(f(v_1), \cdots, f(v_k)) + L(e) \in P$, where addition is coordinate-wise modulo q. A set of assignments $F = \{f_1, \cdots, f_c\}$ is said to *cover* (G,L), if for every $e \in E$, there is some $f_i \in F$ that covers e and F is said to be a *c-covering* for G. G is said to be *c-coverable* if there is a c-covering for G. If L is not specified then it is the constant function which maps E to $\bar{0}$.

Definition 2.2 (COVERING-P-CSP(c,s))**.** For $P \subseteq [q]^k$ and $c,s \in \mathbb{N}$, the COVERING-P-CSP(c,s) problem is, given a c-coverable instance $(G = (V,E), L)$ of P-CSP, find an s-covering.

Definition 2.3 (Odd)**.** A predicate $P \subseteq [q]^k$ is *odd* if $\forall x \in [q]^k, \exists a \in [q], x + \bar{a} \in P$, where addition is coordinate-wise modulo q.

For odd predicates the covering problem is *trivially solvable*, since any CSP instance on such a predicate is q-coverable by the q translates of any assignment, i. e., $\{x + \bar{a} \mid a \in [q]\}$ is a q-covering for any assignment $x \in [q]^k$.

2.2 Label Cover

Definition 2.4 (LABEL-COVER)**.** An instance $G = (U,V,E,L,R,\{\pi_e\}_{e\in E})$ of the LABEL-COVER constraint satisfaction problem consists of a bi-regular bipartite graph (U,V,E), two sets of alphabets L and R and a projection map $\pi_e : R \to L$ for every edge $e \in E$. Given a labeling $\ell : U \to L, \ell : V \to R$, an edge $e = (u,v)$ is said to be satisfied by ℓ if $\pi_e(\ell(v)) = \ell(u)$.

G is said to be *at most* δ*-satisfiable* if every labeling satisfies at most a δ fraction of the edges. G is said to be *c-coverable* if there exist c labelings such that for every vertex $u \in U$, one of the labelings satisfies all the edges incident on u.

An instance of UNIQUE-GAMES is a label cover instance where $L = R$ and the constraints π are permutations.

The hardness of LABEL-COVER stated below follows from the PCP Theorem [2, 1], Raz's Parallel Repetition Theorem [19] and a structural property proved by Håstad [13, Lemma 6.9].

Theorem 2.5 (Hardness of LABEL-COVER)**.** *For every $r \in \mathbb{N}$, there is a deterministic $n^{O(r)}$-time reduction from a 3-SAT instance of size n to an instance $G = (U,V,E,[L],[R],\{\pi_e\}_{e\in E})$ of* LABEL-COVER *with the following properties:*

1. $|U|,|V| \le n^{O(r)}$*; $L,R \le 2^{O(r)}$; G is bi-regular with degrees bounded by $2^{O(r)}$.*
2. *There exists a constant $c_0 \in (0,1/3)$ such that for any $v \in V$ and $\alpha \subseteq [R]$, for a random neighbor u,*
$$\mathop{\mathbb{E}}_{u}\left[|\pi_{uv}(\alpha)|^{-1}\right] \le |\alpha|^{-2c_0},$$
where $\pi_{uv}(\alpha) := \{i \in [L] \mid \exists j \in \alpha \text{ s.t. } \pi_{uv}(j) = i\}$. This implies that
$$\forall v, \alpha, \qquad Pr_u\left[|\pi_{uv}(\alpha)| < |\alpha|^{c_0}\right] \le \frac{1}{|\alpha|^{c_0}}.$$
3. *There is a constant $d_0 \in (0,1)$ such that,*
   - *YES Case : If the 3-SAT instance is satisfiable, then G is 1-coverable.*
   - *NO Case : If the 3-SAT instance is unsatisfiable, then G is at most $2^{-d_0 r}$-satisfiable.*

Our characterization of hardness of covering CSPs is based on the following conjecture due to Dinur and Kol [8].

Conjecture 2.6 (COVERING-UGC(c))**.** *There exists $c \in \mathbb{N}$ such that for every sufficiently small $\delta > 0$ there exists $L \in \mathbb{N}$ such that the following holds. Given an instance $G = (U,V,E,[L],[L],\{\pi_e\}_{e\in E})$ of* UNIQUE-GAMES *it is* NP*-hard to distinguish between the following two cases:*

- *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: Every assignment satisfies at most δ fraction of the edge constraints.*

2.3 Analysis of Boolean functions over probability spaces

For a function $f : \{0,1\}^L \to \mathbb{R}$, the *Fourier decomposition* of f is given by

$$f(x) = \sum_{\alpha \in \{0,1\}^L} \widehat{f}(\alpha)\chi_\alpha(x) \text{ where } \chi_\alpha(x) := (-1)^{\sum_{i=1}^{L} \alpha_i \cdot x_i} \text{ and } \widehat{f}(\alpha) := \mathop{\mathbb{E}}_{x \in \{0,1\}^L} f(x)\chi_\alpha(x).$$

We will use α, also to denote the subset of $[L]$ for which it is the characteristic vector. The *Efron–Stein decomposition* is a generalization of the Fourier decomposition to product distributions of arbitrary probability spaces. Let (Ω, μ) be a probability space and $(\Omega^L, \mu^{\otimes L})$ be the corresponding product space. For a function $f : \Omega^L \to \mathbb{R}$, the Efron–Stein decomposition of f with respect to the product space is given by

$$f(x_1, \cdots, x_L) = \sum_{\beta \subseteq [L]} f_\beta(x),$$

where f_β depends only on x_i for $i \in \beta$ and for all $\beta' \not\supseteq \beta, a \in \Omega^{\beta'}$, $\mathbb{E}_{x \in \mu^{\otimes R}}\left[f_\beta(x) \mid x_{\beta'} = a\right] = 0$. We will be dealing with functions of the form $f : \{0,1\}^{dL} \to \mathbb{R}$ for $d \in \mathbb{N}$ and d-to-1 functions $\pi : [dL] \to [L]$. We will also think of such functions as $f : \prod_{i \in L} \Omega_i \to \mathbb{R}$ where $\Omega_i = \{0,1\}^d$ consists of the d coordinates j such that $\pi(j) = i$. An Efron–Stein decomposition of $f : \prod_{i \in L} \Omega_i \to \mathbb{R}$ over the uniform distribution over $\{0,1\}^{dL}$, can be obtained from the Fourier decomposition as

$$f_\beta(x) = \sum_{\alpha \subseteq [dL] : \pi(\alpha) = \beta} \widehat{f}(\alpha)\chi_\alpha. \tag{2.1}$$

Let $\|f\|_2 := \mathbb{E}_{x \in \mu^{\otimes L}}[f(x)^2]^{1/2}$ and $\|f\|_\infty := \max_{x \in \Omega^{\otimes L}} |f(x)|$. For $i \in [L]$, the influence of the i-th coordinate on f is defined as follows.

$$\mathsf{Inf}_i[f] := \mathop{\mathbb{E}}_{x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_L} \mathsf{Var}_{x_i}[f(x_1, \cdots, x_L)] = \sum_{\beta : i \in \beta} \|f_\beta\|_2^2.$$

For an integer d, the degree d influence is defined as

$$\mathsf{Inf}_i^{\leq d}[f] := \sum_{\beta : i \in \beta, |\beta| \leq d} \|f_\beta\|_2^2.$$

It is easy to see that for Boolean functions, the sum of all the degree d influences is at most d.

Definition 2.7. Let (Ω^k, μ) be a probability space. Let $S = \{x \in \Omega^k \mid \mu(x) > 0\}$. We say that $S \subseteq \Omega^k$ is *connected* if for every $x, y \in S$, there is a sequence of strings starting with x and ending with y such that every element in the sequence is in S and every two adjacent elements differ in exactly one coordinate.

Let $\mu^{\otimes n}$ denote the n-wise product distribution of μ.

Theorem 2.8 ([18, Proposition 6.4])**.** *Let (Ω^k, μ) be a probability space such that the support of the distribution $\mathsf{supp}(\mu) \subseteq \Omega^k$ is connected and the minimum probability of every atom in $\mathsf{supp}(\mu)$ is at least α for some $\alpha \in (0, 1/2]$. Furthermore, assume that the marginal of μ on each of the k coordinates is uniform in Ω. Then there exist continuous functions $\overline{\Gamma} : (0,1) \to (0,1)$ and $\underline{\Gamma} : (0,1) \to (0,1)$ such* *that the following holds: For every $\varepsilon > 0$, there exists $\tau > 0$ and an integer d such that if a function $f : \Omega^L \to [0,1]$ satisfies*

$$\forall i \in [L], \mathsf{Inf}_i^{\leq d}(f) \leq \tau$$

then

$$\underline{\Gamma}\left(\mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu^{\otimes L}}[f(x_1)]\right) - \varepsilon \leq \mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu^{\otimes L}}\left[\prod_{j=1}^{k} f(x_j)\right] \leq \overline{\Gamma}\left(\mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu^{\otimes L}}[f(x_1)]\right) + \varepsilon.$$

There exists an absolute constant C such that one can take $\tau = \varepsilon^{C\frac{\log(1/\alpha)\log(1/\varepsilon)}{\varepsilon\alpha^2}}$ and $d = \log(1/\tau)\log(1/\alpha)$.

The following invariance principle for correlated spaces proved in Section 6 is an adaptation of similar invariance principles (c.f., [22, Theorem 3.12], [12, Lemma B.3]) to our setting.

Theorem 2.9 (Invariance Principle for correlated spaces)**.** *Let $(\Omega_1^k \times \Omega_2^k, \mu)$ be a correlated probability space such that the marginal of μ on any pair of coordinates one each from Ω_1 and Ω_2 is a product distribution. Let μ_1, μ_2 be the marginals of μ on Ω_1^k and Ω_2^k, respectively. Let X, Y be two random $k \times L$ dimensional matrices chosen as follows. Independently for every $i \in [L]$, the pair of columns $(x^i, y^i) \in \Omega_1^k \times \Omega_2^k$ is chosen from μ. Let x_i, y_i denote the i-th rows of X and Y, respectively. If $F : \Omega_1^L \to [-1,+1]$ and $G : \Omega_2^L \to [-1,+1]$ are functions such that*

$$\tau := \sqrt{\sum_{i\in[L]} \mathsf{Inf}_i[F]\cdot\mathsf{Inf}_i[G]} \;\; \textit{and} \;\; \Gamma := \max\left\{\sqrt{\sum_{i\in[L]}\mathsf{Inf}_i[F]}, \sqrt{\sum_{i\in[L]}\mathsf{Inf}_i[G]}\right\},$$

then

$$\left|\mathop{\mathbb{E}}_{(X,Y)\in\mu^{\otimes L}}\left[\prod_{i\in[k]} F(x_i)\cdot G(y_i)\right] - \mathop{\mathbb{E}}_{X\in\mu_1^{\otimes L}}\left[\prod_{i\in[k]} F(x_i)\right]\cdot \mathop{\mathbb{E}}_{Y\in\mu_2^{\otimes L}}\left[\prod_{i\in[k]} G(y_i)\right]\right| \leq 2^{O(k)}\Gamma\tau\,. \qquad (2.2)$$

3 Covering-UG Hardness of Covering CSPs

In this section, we prove the following theorem, which in turn implies Theorem 1.1 (see below for proof).

Theorem 3.1. *Let $[q]$ be any constant-size alphabet and $k \geq 2$. Recall that $\mathsf{NAE} := [q]^k \setminus \{\bar{b} \mid b \in [q]\}$. Let $P \subseteq [q]^k$ be a predicate such that there exists $a \in \mathsf{NAE}$ and $\mathsf{NAE} \supset P \supseteq \{a + \bar{b} \mid b \in [q]\}$. Assuming* COVERING-UGC$(c)$*, for every sufficiently small constant $\delta > 0$ it is* NP*-hard to distinguish between P-*CSP *instances $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of the following two cases:*

- *YES Case : $\mathcal{G}$ is $2c$-coverable.*
- *NO Case : $\mathcal{G}$ does not have an independent set of fractional size δ.*

Proof of Theorem 1.1. Let Q be an arbitrary non-odd predicate i.e, $Q \subseteq [q]^k \setminus \{h + \bar{b} \mid b \in [q]\}$ for some $h \in [q]^k$. Consider the predicate $Q' \subseteq [q]^k$ defined as $Q' := Q - h := \{g - h \mid g \in Q\}$, where the operation '$-$' refers to coordinate-wise substraction performed $\pmod q$. Observe that $Q' \subseteq \mathsf{NAE}$. Given any

Q'-CSP instance Φ with literals function $L(e) = \overline{0}$, consider the Q-CSP instance $\Phi_{Q' \to Q}$ with literals function M given by $M(e) := \overline{h}, \forall e$. It has the same constraint graph as Φ. Clearly, Φ is c-coverable iff $\Phi_{Q' \to Q}$ is c-coverable. Thus, it suffices to prove the result for any predicate $Q' \subseteq \mathsf{NAE}$ with literals function $L(e) = \overline{0}$[3]. We will consider two cases, both of which will follow from Theorem 3.1.

Suppose the predicate Q' satisfies $Q' \supseteq \{a + \bar{b} \mid b \in [q]\}$ for some $a \in [q]^k$. Then this predicate Q' satisfies the hypothesis of Theorem 3.1 and the theorem follows if we show that the soundness guarantee of Theorem 3.1 implies that in Theorem 1.1. Any instance in the NO case of Theorem 3.1, is not $t := \log_q(1/\delta)$-coverable even on the NAE-CSP instance with the same constraint graph. This is because any t-covering for the NAE-CSP instance gives a coloring of the constraint graph using q^t colors, by choosing the color of every variable to be a string of length t and having the corresponding assignments in each position in $[t]$. Hence the Q'-CSP instance is also not t-coverable.

Suppose $Q' \not\supseteq \{a + \bar{b} \mid b \in [q]\}$ for all $a \in [q]^k$. Then consider the predicate $P = \{a + \bar{b} \mid a \in Q', b \in [q]\} \subseteq \mathsf{NAE}$. Notice that P satisfies the conditions of Theorem 3.1 and if the P-CSP instance is t-coverable then the Q'-CSP instance is qt-coverable. Hence a YES instance of Theorem 3.1 maps to a $2cq$-coverable Q-CSP instance and NO instance maps to an instance with covering number at least $\log_q(1/\delta)$, where the latter follows from the fact that the covering number of the instance as a Q'-CSP is at least the covering number of it as a P-CSP. □

We now prove Theorem 3.1 by giving a reduction from an instance $G = (U, V, E, [L], [L], \{\pi_e\}_{e \in E})$ of UNIQUE-GAMES as in Definition 2.4, to an instance $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of a P-CSP for any predicate P that satisfies the conditions mentioned. As stated in the introduction, we adapt the long-code test of Bansal and Khot [4] for proving the hardness of finding independent sets in almost k-partite k-uniform hypergraphs to our setting. The set of variables $\mathcal{V}$ is $V \times [q]^{2L}$. Any assignment to $\mathcal{V}$ is given by a set of functions $f_v : [q]^{2L} \to [q]$, for each $v \in V$. The set of constraints $\mathcal{E}$ is given by the following test which checks whether f_v's are long codes of a good labeling to V. There is a constraint corresponding to all the variables that are queried together by the test.

Long Code Test $\mathcal{T}_1$

1. Choose $u \in U$ uniformly and k neighbors $w_1, \dots, w_k \in V$ of u uniformly and independently at random.

2. Choose a random matrix X of dimension $k \times 2L$ as follows. Let X^i denote the i-th column of X. Independently for each $i \in [L]$, choose (X^i, X^{i+L}) uniformly at random from the set

$$S := \left\{ (y, y') \in [q]^k \times [q]^k \mid y \in \{a + \bar{b} \mid b \in [q]\} \vee y' \in \{a + \bar{b} \mid b \in [q]\} \right\}. \tag{3.1}$$

3. Let $x_1, \cdots, x_k$ be the rows of matrix X. Accept iff

$$(f_{w_1}(x_1 \circ \pi_{uw_1}), f_{w_2}(x_2 \circ \pi_{uw_2}), \cdots, f_{w_k}(x_k \circ \pi_{uw_k})) \in P,$$

where $x \circ \pi$ is the string defined as $(x \circ \pi)(i) := x_{\pi(i)}$ for $i \in [L]$ and $(x \circ \pi)(i) := x_{\pi(i-L)+L}$ otherwise.

[3] This observation [8] that the cover-Q problem for any non-odd predicate Q is equivalent to the cover-Q' problem where $Q' \subseteq \mathsf{NAE}$ shows the centrality of the NAE predicate in understanding the covering complexity of any non-odd predicate.

Before plunging into the formal analysis of the reduction, let us see the intuition behind the test. The test is designed so that if the functions $f_{w_1}, f_{w_2}, \ldots, f_{w_k}$ are dictator functions satisfying the UG-constraints associated with the common neighbor u or their L *shifts*, then the test passes. This is obvious as the bit pattern from the locations queried is either y or y', one of which belongs to the predicate P. This gives a 2-covering of the instance: one corresponds to the actual dictator functions satisfying the UG-constraints and another consists of L shifts of those dictator functions. Another property of the set S that is used in the test is that it defines a probability space that is *connected*. This will be used in the soundness analysis of the test. We now prove the completeness and the soundness of the reduction.

Lemma 3.2 (Completeness)**.** *If the* UNIQUE-GAMES *instance G is c-coverable then the P-*CSP *instance $\mathcal{G}$ is $2c$-coverable.*

Proof. Let $\ell_1, \ldots, \ell_c : U \cup V \to [L]$ be a c-covering for G as described in Definition 2.4. We will show that the $2c$ assignments given by $f_v^i(x) := x_{\ell_i(v)}, g_v^i(x) := x_{\ell_i(v)+L}, i = 1, \ldots, c$ form a $2c$-covering of $\mathcal{G}$. Consider any $u \in U$ and let ℓ_i be the labeling that covers all the edges incident on u. For any $(u, w_j)_{j \in \{1,\cdots,k\}} \in E$ and X chosen by the long code test $\mathcal{T}_1$, the vector $(f_{w_1}^i(x_1 \circ \pi_{uw_1}), \cdots, f_{w_k}^i(x_k \circ \pi_{uw_k}))$ gives the $\ell_i(u)$-th column of X. Similarly the above expression corresponding to g^i gives the $(\ell_i(u)+L)$-th column of the matrix X. Since, for all $i \in [L]$, either i-th column or $(i+L)$-th column of X contains element from $\{a + \bar{b} \mid b \in [q]\} \subseteq P$, either $(f_{w_1}^i(x_1 \circ \pi_{uw_1}), \cdots, f_{w_k}^i(x_k \circ \pi_{uw_k})) \in P$ or $(g_{w_1}^i(x_1 \circ \pi_{uw_1}), \cdots, g_{w_k}^i(x_k \circ \pi_{uw_k})) \in P$. Hence the set of $2c$ assignments $\{f_v^i, g_v^i\}_{i \in \{1,\cdots,c\}}$ covers all constraints in $\mathcal{G}$. □

To prove soundness, we show that the set S, as defined in Equation (3.1), is connected, so that Theorem 2.8 is applicable. For this, we view $S \subseteq [q]^k \times [q]^k$ as a subset of $([q]^2)^k$ as follows: the element $(y, y') \in S$ is mapped to the element $((y_1, y_1'), \cdots, (y_k, y_k')) \in ([q]^2)^k$.

Claim 3.3. *Let $\Omega = [q]^2$. The set $S \subset \Omega^k$ is connected.*

Proof. Consider any $x := (x^1, x^2), y := (y^1, y^2) \in S \subset [q]^k \times [q]^k$. Suppose both $x^1, y^1 \in \{a + \bar{b} \mid b \in [q]\}$, then it is easy to come up with a sequence of strings belonging to S, starting with x and ending with y such that consecutive strings differ in at most 1 coordinate,. Now suppose $x^1, y^2 \in \{a + \bar{b} \mid b \in [q]\}$. First we come up with a sequence from x to $z := (z^1, z^2)$ such that $z^1 := x^1$ and $z^2 = y^2$, and then another sequence for z to y. □

Lemma 3.4 (Soundness)**.** *For every constant $\delta > 0$, there exists a constant s such that, if G is at most s-satisfiable then $\mathcal{G}$ does not have an independent set of size δ.*

Proof. Let $I \subseteq \mathcal{V}$ be an independent set of fractional size δ in the constraint graph. For every variable $v \in V$, let $f_v : [q]^{2L} \to \{0, 1\}$ be the indicator function of the independent set restricted to the vertices that correspond to v. For a vertex $u \in U$, let $N(u) \subseteq V$ be the set of neighbors of u and define $f_u(x) := \mathbb{E}_{w \in N(u)}[f_w(x \circ \pi_{uw})]$. Since I is an independent set, we have

$$0 = \mathop{\mathbb{E}}_{u, w_i, \ldots, w_k} \mathop{\mathbb{E}}_{X \sim \mathcal{T}_1} \left[\prod_{i=1}^{k} f_{w_i}(x_i \circ \pi_{uw_i}) \right] = \mathop{\mathbb{E}}_{u} \mathop{\mathbb{E}}_{X \sim \mathcal{T}_1} \left[\prod_{i=1}^{k} f_u(x_i) \right]. \tag{3.2}$$

Since the bipartite graph (U, V, E) is left regular and $|I| \geq \delta |V|$, we have $\mathbb{E}_{u,x}[f_u(x)] \geq \delta$. By an averaging argument, for at least $\delta/2$ fraction of the vertices $u \in U$, $\mathbb{E}_x[f_u(x)] \geq \delta/2$. Call a vertex $u \in U$ *good* if it satisfies this property. A string $x \in [q]^{2L}$ can be thought as an element from $([q]^2)^L$ by grouping the pair of coordinates x_i, x_{i+L}. Let $\bar{x} \in ([q]^2)^L$ denotes this grouping of x, i. e., j-th coordinate of $\bar{x}$ is (x_j, x_{j+L}) is distributed u.a.r. in $[q]^2$. With this grouping, the function f_u can be viewed as $f_u : ([q]^2)^L \to \{0,1\}$. From Equation (3.2), we have that for any $u \in U$,

$$\mathop{\mathbb{E}}_{X \sim \mathcal{T}_1} \left[\prod_{i=1}^{k} f_u(\bar{x}_i) \right] = 0.$$

By Claim 3.3, for all $j \in [L]$ the tuple $((\bar{x}_1)_j, \ldots, (\bar{x}_k)_j)$ (corresponding to columns (X^j, X^{j+L}) of X) is sampled from a distribution whose support is a connected set. Hence for a good vertex $u \in U$, we can apply Theorem 2.8 with $\varepsilon = \underline{\Gamma}(\delta/2)/2$ to get that there exists $j \in [L], d \in \mathbb{N}, \tau > 0$ such that $\mathsf{Inf}_j^{\leq d}(f_u) > \tau$. We will use this fact to give a randomized labeling for G. Labels for vertices $w \in V, u \in U$ will be chosen uniformly and independently from the sets

$$\mathsf{Lab}(w) := \left\{ i \in [L] \mid \mathsf{Inf}_i^{\leq d}(f_w) \geq \frac{\tau}{2} \right\}, \mathsf{Lab}(u) := \left\{ i \in [L] \mid \mathsf{Inf}_i^{\leq d}(f_u) \geq \tau \right\}.$$

By the above argument (using Theorem 2.8), we have that for a good vertex u, $\mathsf{Lab}(u) \neq \emptyset$. Furthermore, since the sum of degree d influences is at most d, the above sets have size at most $2d/\tau$. Now, for any $j \in \mathsf{Lab}(u)$, we have

$$\tau < \mathsf{Inf}_j^{\leq d}[f_u] = \sum_{S: j \in S, |S| \leq d} \|f_{u,S}\|^2 = \sum_{S: j \in S, |S| \leq d} \left\| \mathop{\mathbb{E}}_{w \in N(u)} \left[f_{w, \pi_{uw}^{-1}(S)} \right] \right\|^2 \quad \text{(By Definition.)}$$

$$\leq \sum_{S: j \in S, |S| \leq d} \mathop{\mathbb{E}}_{w \in N(u)} \left\| f_{w, \pi_{uw}^{-1}(S)} \right\|^2 \quad = \mathop{\mathbb{E}}_{w \in N(u)} \mathsf{Inf}_{\pi_{uw}^{-1}(j)}^{\leq d}[f_w]. \quad \text{(By Convexity of square.)}$$

Hence, by another averaging argument, there exists at least $\tau/2$ fraction of neighbors w of u such that $\mathsf{Inf}_{\pi_{uw}^{-1}(j)}^{\leq d}(f_w) \geq \tau/2$ and hence $\pi_{uw}^{-1}(j) \in \mathsf{Lab}(w)$. Therefore, for a good vertex $u \in U$, at least $\tau/2 \cdot \tau/2d$ fraction of edges incident on u are satisfied in expectation. Also, at least $\delta/2$ fraction of vertices in U are good, it follows that the expected fraction of edges that are satisfied by this random labeling is at least $\delta/2 \cdot \tau/2 \cdot \tau/2d$. Choosing $s < \delta\tau^2/8d$ completes the proof. □

4 NP Hardness of Covering CSPs

In this section, we prove Theorem 1.2. We give a reduction from an instance of a LABEL-COVER, $G = (U, V, E, [L], [R], \{\pi_e\}_{e \in E})$ as in Definition 2.4, to a P-CSP instance $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ for any predicate P that satisfies the conditions mentioned in Theorem 1.2. The reduction and proof is similar to that of Dinur and Kol [8]. The main difference is that they used a test and invariance principle very specific to the 4-LIN predicate, while we show that a similar analysis can be performed under milder conditions on the test distribution.

We assume that $R = dL$ and $\forall i \in [L], e \in E, |\pi_e^{-1}(i)| = d$. This is done just for simplifying the notation and the proof does not depend upon it. The set of variables $\mathcal{V}$ is $V \times \{0,1\}^{2R}$. Any assignment to $\mathcal{V}$ is given by a set of functions $f_v : \{0,1\}^{2R} \to \{0,1\}$, for each $v \in V$. The set of constraints $\mathcal{E}$ is given by the following test, which checks whether f_v's are long codes of a good labeling to V.

Long Code test $\mathcal{T}_2$

1. Choose $u \in U$ uniformly and $v, w \in V$ neighbors of u uniformly and independently at random. For $i \in [L]$, define the sets $B_{uv}(i) := \pi_{uv}^{-1}(i), B'_{uv}(i) := R + \pi_{uv}^{-1}(i)$ and similarly for w.

2. Choose matrices X, Y of dimension $k \times 2dL$ as follows. For $S \subseteq [2dL]$, we denote by $X|_S$ the submatrix of X restricted to the columns S. Independently for each $i \in [L]$, choose $c_1 \in \{0, 1\}$ uniformly and

   (a) if $c_1 = 0$, choose $\left(X|_{B_{uv}(i) \cup B'_{uv}(i)}, Y|_{B_{uw}(i) \cup B'_{uw}(i)}\right)$ from $\mathcal{P}_0^{\otimes 2d} \otimes \mathcal{P}_1^{\otimes 2d}$,

   (b) if $c_1 = 1$, choose $\left(X|_{B_{uv}(i) \cup B'_{uv}(i)}, Y|_{B_{uw}(i) \cup B'_{uw}(i)}\right)$ from $\mathcal{P}_1^{\otimes 2d} \otimes \mathcal{P}_0^{\otimes 2d}$.

3. Perturb X, Y as follows. Independently for each $i \in [L]$, choose $c_2 \in \{*, 0, 1\}$ as follows:

   $$\Pr[c_2 = *] = 1 - 2\varepsilon, \text{ and } \Pr[c_2 = 1] = \Pr[c_2 = 0] = \varepsilon.$$

   Perturb the i-th matrix block $\left(X|_{B_{uv}(i) \cup B'_{uv}(i)}, Y|_{B_{uw}(i) \cup B'_{uw}(i)}\right)$ as follows:

   (a) if $c_2 = *$, leave the matrix block $\left(X|_{B_{uv}(i) \cup B'_{uv}(i)}, Y|_{B_{uw}(i) \cup B'_{uw}(i)}\right)$ unperturbed,

   (b) if $c_2 = 0$, choose $\left(X|_{B'_{uv}(i)}, Y|_{B'_{uw}(i)}\right)$ uniformly from $\{0,1\}^{k \times d} \times \{0,1\}^{k \times d}$,

   (c) if $c_2 = 1$, choose $\left(X|_{B_{uv}(i)}, Y|_{B_{uw}(i)}\right)$ uniformly from $\{0,1\}^{k \times d} \times \{0,1\}^{k \times d}$.

4. Let $x_1, \cdots, x_k$ and $y_1, \cdots, y_k$ be t he rows of the matrices X and Y, respectively. Accept if

   $$(f_v(x_1), \cdots, f_v(x_k), f_w(y_1), \cdots, f_w(y_k)) \in P.$$

Lemma 4.1 (Completeness)**.** *If G is an YES instance of* LABEL-COVER*, then there exists f, g such that each of them covers $1 - \varepsilon$ fraction of $\mathcal{E}$ and they together cover all of $\mathcal{E}$.*

Proof. Let $\ell : U \cup V \to [L] \cup [R]$ be a labeling to G that satisfies all the constraints. Consider the assignments $f_v(x) := x_{\ell(v)}$ and $g_v(x) := x_{R+\ell(v)}$ for each $v \in V$. First consider the assignment f. For any $(u, v), (u, w) \in E$ and $x_1, \cdots, x_k, y_1, \cdots, y_k$ chosen by the long code test $\mathcal{T}_2$, $(f_v(x_1), \cdots, f_v(x_k))$, $(f_w(y_1), \cdots, f_w(y_k))$ gives the $\ell(v)$-th and $\ell(w)$-th column of the matrices X and Y, respectively. Since $\pi_{uv}(\ell(v)) = \pi_{uw}(\ell(w))$, they are jointly distributed either according to $\mathcal{P}_0 \otimes \mathcal{P}_1$ or $\mathcal{P}_1 \otimes \mathcal{P}_0$ after Step 2. The probability that these rows are perturbed in Step 3c is at most ε. Hence with probability $1 - \varepsilon$ over the test distribution, f is accepted. A similar argument shows that the test accepts g with probability $1 - \varepsilon$. Note that in Step 3, the columns given by f, g, are never re-sampled uniformly together. Hence they together cover $\mathcal{G}$. □

Now we will show that if G is a NO instance of LABEL-COVER then no t assignments can cover the 2k-LIN-CSP with constraint hypergraph $\mathcal{G}$. For the rest of the analysis, we will use $+1, -1$ instead of the symbols $0, 1$. Suppose for contradiction, there exist t assignments $f_1, \cdots, f_t : \{\pm 1\}^{2R} \to \{\pm 1\}$ that form a t-cover to $\mathcal{G}$. The probability that all the t assignments are rejected in Step 4 is

$$\mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[\prod_{i=1}^{t} \frac{1}{2} \left(\prod_{j=1}^{k} f_{i,v}(x_j) f_{i,w}(y_j) + 1 \right) \right] = \frac{1}{2^t} + \frac{1}{2^t} \sum_{\emptyset \subset S \subseteq \{1, \cdots, t\}} \mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[\prod_{j=1}^{k} f_{S,v}(x_j) f_{S,w}(y_j) \right]. \tag{4.1}$$

where $f_{S,v}(x) := \prod_{i\in S} f_{i,v}(x)$. Since the t assignments form a t-cover, the LHS in Equation (4.1) is 0 and hence, there exists an $S \neq \emptyset$ such that

$$\mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{j=1}^{k} f_{S,v}(x_j) f_{S,w}(y_j)\right] \le -1/(2^t-1). \tag{4.2}$$

The following lemma shows that this is not possible if t is not too large, thus proving that there does not a exist t-cover.

Lemma 4.2 (Soundness)**.** *Let $c_0 \in (0,1)$ be the constant from Theorem 2.5 and $S \subseteq \{1,\cdots,t\}, |S| > 0$. If G is at most s-satisfiable then*

$$\mathop{\mathbb{E}}_{u,v,w}\ \mathop{\mathbb{E}}_{X,Y\in\mathcal{T}_2}\left[\prod_{i=1}^{k} f_{S,v}(x_i) f_{S,w}(y_i)\right] \ge -O(ks^{c_0/8}) - 2^{O(k)}\frac{s^{(1-3c_0)/8}}{\varepsilon^{3/2c_0}}.$$

Proof. Notice that for a fixed u, the distribution of X and Y have identical marginals. Hence the value of the above expectation, if calculated according to a distribution that is the direct product of the marginals, is positive. We will first show that the expectation can change by at most $O(ks^{c_0/8})$ in moving to an *attenuated* version of the functions (see Claim 4.3). Then we will show that the error incurred by changing the distribution to the product distribution of the marginals has absolute value at most $2^{O(k)}\frac{s^{(1-3c_0)/8}}{\varepsilon^{3/2c_0}}$ (see Claim 4.5). This is done by showing that there is a labeling to G that satisfies an s fraction of the constraints if the error is more than $2^{O(k)}\frac{s^{(1-3c_0)/8}}{\varepsilon^{3/2c_0}}$.

For the rest of the analysis, we write f_v and f_w instead of $f_{S,v}$ and $f_{S,w}$, respectively. Let $f_v = \sum_{\alpha\subseteq[2R]} \widehat{f_v}(\alpha)\chi_\alpha$ be the Fourier decomposition of the function and for $\gamma \in (0,1)$, let $T_{1-\gamma} f_v := \sum_{\alpha\subseteq[2R]}(1-\gamma)^{|\alpha|}\widehat{f_v}(\alpha)\chi_\alpha$. The following claim is similar to a lemma of Dinur and Kol [8, Lemma 4.11]. The only difference in the proof is that, we use the *smoothness* from Property 2 of Theorem 2.5 (which was shown by Håstad [13, Lemma 6.9]).

Claim 4.3. *Let $\gamma := s^{(c_0+1)/4}\varepsilon^{1/c_0}$ where c_0 is the constant from Theorem 2.5.*

$$\left|\mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\underbrace{\prod_{i=1}^{k} f_v(x_i) f_w(y_i)}_{\Delta_0}\right] - \mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\underbrace{\prod_{i=1}^{k} T_{1-\gamma} f_v(x_i) T_{1-\gamma} f_w(y_i)}_{\Delta_1}\right]\right| \le O(ks^{c_0/8}).$$

Proof. The claim bounds the change in the expectation when we change the expression Δ_0 to Δ_1. The expression Δ_0 is a product of $2k$ functions and Δ_1 is the product of the same functions after applying the $T_{1-\gamma}$ operator to each of these functions. We prove the claim by bounding the error with $O(s^{c_0/8})$ when we add an extra $T_{1-\gamma}$ operator each time. Thus, the total error will be $O(ks^{c_0/8})$ by doing the telescoping sum and using the triangle inequality.

For notational convenience, we bound the error when we add the first $T_{1-\gamma}$. The effect of adding all the remaining subsequent $T_{1-\gamma}$ operators can be analyzed in a similar way.

$$\left|\mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^{k} f_v(x_i) f_w(y_i)\right] - \mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\left(\prod_{i=1}^{k-1} f_v(x_i) f_w(y_i)\right) f_v(x_k) T_{1-\gamma} f_w(y_k)\right]\right| \le O(s^{c_0/8}). \tag{4.3}$$

Recall that X,Y denote the matrices chosen by test $\mathcal{T}_2$. Let Y_{-k} be the matrix obtained from Y by removing the k-th row and $F^{u,v,w}(X,Y_{-k}) := \left(\prod_{i=1}^{k-1} f_v(x_i) f_w(y_i)\right) f_v(x_k)$. Then, Eq. (4.3) can be rewritten as

$$\left| \mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[F^{u,v,w}(X,Y_{-k}) \left(I - T_{1-\gamma}\right) f_w(y_k) \right] \right| \le O(s^{c_0/8}). \tag{4.4}$$

Let U be the operator that maps functions on the variable y_k, to one on the variables (X,Y_{-k}) defined by

$$(Uf)(X,Y_{-k}) := \mathop{\mathbb{E}}_{y_k | X, Y_{-k}} f(y_k).$$

Let $G^{u,v,w}(X,Y_{-k}) := \left(U(I-T_{1-\gamma})f_w\right)(X,Y_{-k})$. Note that $\mathbb{E}_{(X,Y)\sim\mathcal{T}_2}\, G^{u,v,w}(X,Y_{-k}) = 0$. This is because $\mathbb{E}_{(X,Y)\sim\mathcal{T}_2}\, G^{u,v,w}(X,Y_{-k}) = \mathbb{E}_{y_k\sim\{0,1\}^{2L}}((I-T_{1-\gamma})f_w)(y_k) = (\widehat{(I-T_{1-\gamma})}f_w)(\emptyset)$, where the marginal distribution on y_k is uniform in $\{0,1\}^{2L}$. Finally, by construction, $\mathbb{E}_{(X,Y)\sim\mathcal{T}_2}\, G^{u,v,w}(X,Y_{-k}) = 0$ follows, since f_w is an odd function. The domain of $G^{u,v,w}$ can be thought of as $(\{0,1\}^{2k-1})^{2dL}$ and the test distribution on any row is independent across the blocks $\{B_{uv}(i) \cup B'_{uv}(i)\}_{i\in[L]}$. We now think of $G^{u,v,w}$ as having domain $\prod_{i\in[L]} \Omega_i$ where $\Omega_i = (\{0,1\}^{2k-1})^{2d}$ corresponds to the set of rows in $B_{uv}(i) \cup B'_{uv}(i)$. Let the following be the Efron–Stein decomposition of $G^{u,v,w}$ with respect to $\mathcal{T}_2$,

$$G^{u,v,w}(X,Y_{-k}) = \sum_{\alpha\subseteq[L]} G^{u,v,w}_\alpha(X,Y_{-k}).$$

The following technical claim follows from a result similar to [8, Lemma 4.7] and then using [18, Proposition 2.12]. We defer its proof to Section 4.1. Here we use the role of the random variable c_2 in $\mathcal{T}_2$, which helps to break the perfect correlation between one row and rest of the rows restricted to the columns $B_{uv}(i) \cup B'_{uv}(i)$ for all $i \in [L]$.

Claim 4.4. *For $\alpha \subseteq [L]$*

$$\left\|G^{u,v,w}_\alpha\right\|^2 \le (1-\varepsilon)^{|\alpha|} \sum_{\beta\subseteq[2R]:\widetilde{\pi}_{uw}(\beta)=\alpha} \left(1-(1-\gamma)^{2|\beta|}\right) \widehat{f_w}(\beta)^2 \tag{4.5}$$

where $\widetilde{\pi}_{uw}(\beta) := \{i \in [L] : \exists j \in [R], (j \in \beta \vee j+R \in \beta) \wedge \pi_{uv}(j) = i\}$.

Substituting the Efron–Stein decomposition of $G^{u,v,w}, F^{u,v,w}$ into the LHS of Eq. (4.3) gives

$$\begin{aligned}
\left| \mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[F^{u,v,w}(X,Y_{-k}) \left(I - T_{1-\gamma}\right) f_w(y_k) \right] \right| &= \left| \mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} F^{u,v,w}(X,Y_{-k}) G^{u,v,w}(X,Y_{-k}) \right| \\
\text{(by orthonormality of Efron–Stein decomposition)} \quad &= \left| \mathop{\mathbb{E}}_{u,v,w} \sum_{\alpha\subseteq[L]} \mathop{\mathbb{E}}_{\mathcal{T}_2} F^{u,v,w}_\alpha(X,Y_{-k}) G^{u,v,w}_\alpha(X,Y_{-k}) \right| \\
\text{(by Cauchy–Schwarz inequality)} \quad &\le \mathop{\mathbb{E}}_{u,v,w} \sqrt{\sum_{\alpha\subseteq[L]} \|F^{u,v,w}_\alpha\|^2} \cdot \sqrt{\sum_{\alpha\subseteq[L]} \|G^{u,v,w}_\alpha\|^2} \\
\text{(Using } \sum_{\alpha\subseteq[L]} \|F^{u,v,w}_\alpha\|^2 = \|F^{u,v,w}\|_2^2 = 1) \quad &\le \mathop{\mathbb{E}}_{u,w} \sqrt{\sum_{\alpha\subseteq[L]} \|G^{u,v,w}_\alpha\|^2}.
\end{aligned}$$

Using concavity of square root and substituting for $\|G_\alpha^{u,v,w}\|^2$ from Equation (4.5), we get that the above is not greater than

$$\sqrt{\mathbb{E}_{u,w} \sum_{\alpha\subseteq[L]} \sum_{\substack{\beta\subseteq[2R]:\\ \widetilde{\pi}_{uw}(\beta)=\alpha}} \underbrace{(1-\varepsilon)^{|\alpha|}\left(1-(1-\gamma)^{2|\beta|}\right)\widehat{f}_w(\beta)^2}_{=:\mathsf{Term}_{u,w}(\alpha,\beta)}}.$$

We will now break the above summation into three different parts and bound each part separately.

$$\Theta_0 := \mathbb{E}_{u,w} \sum_{\alpha,\beta:|\alpha|\geq\frac{1}{\varepsilon s^{c_0/4}}} \mathsf{Term}_{u,w}(\alpha,\beta), \qquad \Theta_1 := \mathbb{E}_{u,w} \sum_{\substack{\alpha,\beta:|\alpha|<\frac{1}{\varepsilon s^{c_0/4}}\\ |\beta|\leq\frac{2}{s^{1/4}\varepsilon^{1/c_0}}}} \mathsf{Term}_{u,w}(\alpha,\beta),$$

$$\Theta_2 := \mathbb{E}_{u,w} \sum_{\substack{\alpha,\beta:|\alpha|<\frac{1}{\varepsilon s^{c_0/4}}\\ |\beta|>\frac{2}{s^{1/4}\varepsilon^{1/c_0}}}} \mathsf{Term}_{u,w}(\alpha,\beta).$$

Upper bound for Θ_0. When $|\alpha| > \dfrac{1}{\varepsilon s^{c_0/4}}$, $(1-\varepsilon)^{|\alpha|} < s^{c_0/4}$. Also since f_w is $\{+1,-1\}$ valued, sum of squares of Fourier coefficient is 1. Hence $|\Theta_0| < s^{c_0/4}$.

Upper bound for Θ_1. When $|\beta| \leq \dfrac{2}{s^{1/4}\varepsilon^{1/c_0}}$,

$$1-(1-\gamma)^{2|\beta|} \leq 1-\left(1-\frac{4}{s^{1/4}\varepsilon^{1/c_0}}\gamma\right) = \frac{4}{s^{1/4}\varepsilon^{1/c_0}}\gamma = 4s^{c_0/4}.$$

Again since the sum of squares of Fourier coefficients is 1, $|\Theta_1| \leq 4s^{c_0/4}$.

Upper bound for Θ_2. From Property 2 of Theorem 2.5, we have that for any $v \in V$ and β with $|\beta| > \dfrac{2}{s^{1/4}\varepsilon^{1/c_0}}$, the probability that $|\widetilde{\pi}_{uv}(\beta)| < 1/\varepsilon s^{c_0/4}$, for a random neighbor u, is at most $\varepsilon s^{c_0/4}$. Hence $|\Theta_2| \leq s^{c_0/4}$.

□

Fix u,v,w chosen by the test. Recall that we thought of f_v as having domain $\prod_{i\in[L]}\Omega_i$ where $\Omega_i = \{0,1\}^{2d}$ corresponds to the set of coordinates in $B_{uv}(i)\cup B'_{uv}(i)$. Since the grouping of coordinates depends on u, we define $\overline{\mathsf{Inf}}_i^u[f_v] := \mathsf{Inf}_i[f_v]$ where $i\in[L]$ for explicitness. From Equation (2.1),

$$\overline{\mathsf{Inf}}_i^u[f_v] = \sum_{\alpha\subseteq[2dL]:i\in\widetilde{\pi}_{uv}(\alpha)} \widehat{f}_v(\alpha)^2,$$

where $\widetilde{\pi}_{uv}(\alpha) := \{i\in[L] : \exists j\in[R], (j\in\alpha \vee j+R\in\alpha)\wedge \pi_{uv}(j)=i\}$.

Claim 4.5. *Let* $\tau_{u,v,w} := \sum_{i\in[L]} \overline{\mathsf{Inf}}_i^u[T_{1-\gamma}f_v] \cdot \overline{\mathsf{Inf}}_i^u[T_{1-\gamma}f_w]$.

$$\mathop{\mathbb{E}}_{u,v,w}\left|\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^k T_{1-\gamma}f_v(x_i)T_{1-\gamma}f_w(y_i)\right] - \mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^k T_{1-\gamma}f_v(x_i)\right]\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^k T_{1-\gamma}f_w(y_i)\right]\right| \leq 2^{O(k)}\sqrt{\frac{\mathbb{E}_{u,v,w}\,\tau_{u,v,w}}{\gamma}}.$$

Proof. It is easy to check that $\sum_{i\in[L]} \overline{\mathsf{Inf}}_i^u[T_{1-\gamma}f_v] \leq 1/\gamma$ (c.f., [22, Lemma 1.13]). For any u,v,w, since the test distribution satisfies the conditions of Theorem 2.9, we get

$$\left|\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^k T_{1-\gamma}f_v(x_i)T_{1-\gamma}f_w(y_i)\right] - \mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^k T_{1-\gamma}f_v(x_i)\right]\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^k T_{1-\gamma}f_w(y_i)\right]\right| \leq 2^{O(k)}\sqrt{\frac{\tau_{u,v,w}}{\gamma}}.$$

The claim follows by taking expectation over u,v,w and using the concavity of square root. □

From Claims 4.5 and 4.3 and using the fact the the marginals of the test distribution $\mathcal{T}_2$ on $(x_1,\ldots,x_k)$ is the same as marginals on $(y_1,\ldots,y_k)$, for $\gamma := s^{(c_0+1)/4}\varepsilon^{1/c_0}$, we get

$$\mathop{\mathbb{E}}_{u,v,w}\ \mathop{\mathbb{E}}_{X,Y\in\mathcal{T}_2}\left[\prod_{i=1}^k f_v(x_i)f_w(y_i)\right] \geq -O(ks^{c_0/8}) - 2^{O(k)}\sqrt{\frac{\mathbb{E}_{u,v,w}\,\tau_{u,v,w}}{\gamma}} + \mathop{\mathbb{E}}_u\left(\mathop{\mathbb{E}}_v\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^k T_{1-\gamma}f_v(x_i)\right]\right)^2. \tag{4.6}$$

If $\tau_{u,v,w}$ in expectation is large, there is a standard way of decoding the assignments to a labeling to the label cover instance, as shown in Claim 4.6.

Claim 4.6. *If G is an at most s-satisfiable instance of* LABEL-COVER *then*

$$\mathop{\mathbb{E}}_{u,v,w}\tau_{u,v,w} \leq \frac{s}{\gamma^2}.$$

Proof. Note that $\sum_{\alpha\subseteq[2R]}(1-\gamma)^{|\alpha|}\widehat{f_v}(\alpha)^2 \leq 1$. We will give a randomized labeling to the LABEL-COVER instance. For each $v\in V$, choose a random $\alpha\subseteq[2R]$ with probability $(1-\gamma)^{|\alpha|}\widehat{f_v}(\alpha)^2$ and assign a uniformly random label j in α to v; if the label $j\geq R$, change the label to $j-R$ and with the remaining probability assign an arbitrary label. For $u\in U$, choose a random neighbor $w\in V$ and a random $\beta\subseteq[2R]$ with probability $(1-\gamma)^{|\beta|}\widehat{f_w}(\beta)^2$, choose a random label ℓ in β and assign the label $\widetilde{\pi}_{uw}(\ell)$ to u. With the remaining probability, assign an arbitrary label. The fraction of edges satisfied by this labeling is at least

$$\mathop{\mathbb{E}}_{u,v,w}\sum_{i\in[L]}\ \sum_{(\alpha,\beta):i\in\widetilde{\pi}_{uv}(\alpha),i\in\widetilde{\pi}_{uw}(\beta)}\frac{(1-\gamma)^{|\alpha|+|\beta|}}{|\alpha|\cdot|\beta|}\widehat{f_v}(\alpha)^2\widehat{f_w}(\beta)^2.$$

Using the fact that $1/r \geq \gamma(1-\gamma)^r$ for every $r>0$ and $\gamma\in[0,1]$, we lower bound $1/|\alpha|$ and $1/|\beta|$ by $\gamma(1-\gamma)^{|\alpha|}$ and $\gamma(1-\gamma)^{|\beta|}$, respectively. The above is then not less than

$$\gamma^2\mathop{\mathbb{E}}_{u,v,w}\sum_{i\in[L]}\left(\sum_{\alpha:i\in\widetilde{\pi}_{uv}(\alpha)}(1-\gamma)^{2|\alpha|}\widehat{f_v}(\alpha)^2\right)\left(\sum_{\beta:i\in\widetilde{\pi}_{uw}(\beta)}(1-\gamma)^{2|\beta|}\widehat{f_w}(\beta)^2\right) = \gamma^2\mathop{\mathbb{E}}_{u,v,w}\tau_{u,v,w}.$$

Since G is at most s-satisfiable, the labeling can satisfy at most an s fraction of constraints and the right-hand side of the above equation is at most s. □

Lemma 4.2 follows from the above claim and Equation (4.6). □

Proof of Theorem 1.2. Using Theorem 2.5, the size of the CSP instance $\mathcal{G}$ produced by the reduction is $N = n^r 2^{2^{O(r)}}$ and the parameter $s \le 2^{-d_0 r}$. Setting $r = \Theta(\log\log n)$, gives that $N = 2^{\mathrm{poly}\log n}$ for a constant k. Lemma 4.2 and Equation (4.2) imply that

$$O(ks^{c_0/8}) + 2^{O(k)}\frac{s^{(1-3c_0)/8}}{\varepsilon^{3/2c_0}} \ge \frac{1}{2^t - 1}.$$

Since k is a constant, this gives that $t = \Omega(\log\log n)$.

For every constant $C > 2$, by choosing r a large enough constant, we get the hardness result assuming $\mathrm{P} \ne \mathrm{NP}$. □

4.1 Proof of Claim 4.4

We will be reusing the notation introduced in the long code test $\mathcal{T}_2$. We denote the $k \times 2d$ dimensional matrix $X|_{B(i)\cup B'(i)}$ by X^i and $Y|_{B(i)\cup B'(i)}$ by Y^i. Also by X^i_j, we mean the j-th row of the matrix X^i and Y^i_{-k} is the first $k-1$ rows of Y^i. The spaces of the random variables X^i, X^i_j, Y^i_{-k} will be denoted by $\mathcal{X}^i, \mathcal{X}^i_j, \mathcal{Y}^i_{-k}$.

Before we proceed to the proof of claim, we need a few definitions and lemmas related to correlated spaces defined by Mossel [18].

Definition 4.7. Let $(\Omega_1 \times \Omega_2, \mu)$ be a finite correlated space, the correlation between Ω_1 and Ω_2 with respect to μ us defined as

$$\rho(\Omega_1, \Omega_2; \mu) := \max_{\substack{f:\Omega_1\to\mathbb{R}, \mathbb{E}[f]=0, \mathbb{E}[f^2]\le 1 \\ g:\Omega_2\to\mathbb{R}, \mathbb{E}[g]=0, \mathbb{E}[g^2]\le 1}} \mathop{\mathbb{E}}_{(x,y)\sim\mu}[|f(x)g(y)|].$$

Definition 4.8 (Markov Operator)**.** Let $(\Omega_1 \times \Omega_2, \mu)$ be a finite correlated space, the Markov operator, associated with this space, denoted by U, maps a function $g : \Omega_2 \to \mathbb{R}$ to functions $Ug : \Omega_1 \to \mathbb{R}$ by the following map:

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

The following results (due to Mossel [18]) provide a way to give an upper bound on the correlation of correlated spaces.

Lemma 4.9 ([18, Lemma 2.8])**.** *Let $(\Omega_1 \times \Omega_2, \mu)$ be a finite correlated space. Let $g : \Omega_2 \to \mathbb{R}$ be such that $\mathbb{E}_{(x,y)\sim\mu}[g(y)] = 0$ and $\mathbb{E}_{(x,y)\sim\mu}[g(y)^2] \le 1$. Then, among all functions $f : \Omega_1 \to \mathbb{R}$ that satisfy $\mathbb{E}_{(x,y)\sim\mu}[f(x)^2] \le 1$, the maximum value of $|\mathbb{E}[f(x)g(y)]|$ is given as:*

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

Proposition 4.10 ([18, Proposition 2.11])**.** *Let $(\prod_{i=1}^n \Omega_i^{(1)} \times \prod_{i=1}^n \Omega_i^{(2)}, \prod_{i=1}^n \mu_i)$ be a product correlated space. Let $g : \prod_{i=1}^n \Omega_i^{(2)} \to \mathbb{R}$ be a function and U be the Markov operator mapping functions from* *the space* $\prod_{i=1}^{n}\Omega_i^{(2)}$ *to functions on space* $\prod_{i=1}^{n}\Omega_i^{(1)}$. *If* $g=\sum_{S\subseteq[n]}g_S$ *and* $Ug=\sum_{S\subseteq[n]}(Ug)_S$ *be the Efron–Stein decompositions of* g *and* Ug*, respectively, then,*

$$(Ug)_S = U(g_S)$$

i. e., the Efron–Stein decomposition commutes with Markov operators.

Proposition 4.11 ([18, Proposition 2.12])**.** *Assume the setting of Proposition 4.10 and furthermore assume that* $\rho(\Omega_i^{(1)},\Omega_i^{(2)};\mu_i)\le\rho$ *for all* $i\in[n]$. *Then for all* g *it holds that*

$$\|U(g_S)\|_2 \le \rho^{|S|}\|g_S\|_2.$$

We will prove the following claim.

Claim 4.12. *For each* $i\in[L]$,

$$\rho\left(\mathcal{X}^i\times\mathcal{Y}^i_{-k},\mathcal{Y}^i_k;\mathcal{T}_2^i\right)\le\sqrt{1-\varepsilon}.$$

Before proving this claim, first let's see how it leads to the proof of Claim 4.4.

Proof of Claim 4.4. Proposition 4.10 shows that the Markov operator U commutes with taking the Efron–Stein decomposition. Hence, $G_\alpha^{u,v,w}:=(U((I-T_{1-\gamma})f_w))_\alpha=U((I-T_{1-\gamma})(f_w)_\alpha)$, where $(f_w)_\alpha$ is the Efron–Stein decomposition of f_w w.r.t. the marginal distribution of $\mathcal{T}_2$ on $\prod_{i=1}^{L}\mathcal{Y}^i_k$, which is a uniform distribution. Therefore, $(f_w)_\alpha=\sum_{\substack{\beta\subseteq[2R],\\ \tilde{\pi}_{uw}(\beta)=\alpha}}\hat{f}_w(\beta)\chi_\beta$. Using Proposition 4.11 and Claim 4.12, we have

$$\begin{aligned}\|G_\alpha^{u,v,w}\|_2^2=\|U((I-T_{1-\gamma})(f_w)_\alpha)\|_2^2 &\le(\sqrt{1-\varepsilon})^{2|\alpha|}\|(I-T_{1-\gamma})(f_w)_\alpha\|_2^2\\ &=(1-\varepsilon)^{|\alpha|}\sum_{\beta\subseteq[2R]:\tilde{\pi}_{uw}(\beta)=\alpha}\left(1-(1-\gamma)^{2|\beta|}\right)\hat{f}_w(\beta)^2,\end{aligned}$$

where the norms are with respect to the marginals of $\mathcal{T}_2$ in the corresponding spaces. □

Proof of Claim 4.12. Recall the random variable $c_2\in\{*,0,1\}$ defined in Step 3 of test $\mathcal{T}_2$. Let g and f be the functions that satisfies $\mathbb{E}[g]=\mathbb{E}[f]=0$ and $\mathbb{E}[g^2],\mathbb{E}[f^2]\le1$ such that $\rho\left(\mathcal{X}^i\times\mathcal{Y}^i_{-k},\mathcal{Y}^i_k;\mathcal{T}_2^i\right)=\mathbb{E}[|fg|]$. Define the *Markov Operator*

$$Ug(X^i,Y^i_{-k})=\mathop{\mathbb{E}}_{(\tilde{X},\tilde{Y})\sim\mathcal{T}_2^i}[g(\tilde{Y}_k)\mid(\tilde{X},\tilde{Y}_{-k})=(X^i,Y^i_{-k})].$$

By Lemma 4.9, we have

$$\begin{aligned}\rho\left(\mathcal{X}^i\times\mathcal{Y}^i_{-k},\mathcal{Y}^i_k;\mathcal{T}_2^i\right)^2 &\le\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2]\\ &=(1-2\varepsilon)\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2\mid c_2=*]+\varepsilon\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2\mid c_2=0]+\\ &\qquad\varepsilon\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2\mid c_2=1]\\ &\le(1-2\varepsilon)+\varepsilon\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2\mid c_2=0]+\varepsilon\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2\mid c_2=1],\end{aligned}$$

where the last inequality uses the fact that $\mathbb{E}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2 \mid c_2 = *] = \mathbb{E}[g^2]$, which is at most 1. Consider the case when $c_2 = 0$. By definition, we have

$$\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2 \mid c_2 = 0] = \mathop{\mathbb{E}}_{\binom{X^i,}{Y^i_{-k}} \sim \mathcal{T}_2^i} \left(\mathop{\mathbb{E}}_{(\tilde{X},\tilde{Y}) \sim \mathcal{T}_2^i}[g(\tilde{Y}_k) \mid (\tilde{X},\tilde{Y}_{-k}) = (X^i,Y^i_{-k}) \wedge c_2 = 0] \right)^2.$$

Under the conditioning, for any fixed value of X^i, Y^i_{-k}, the value of $\tilde{Y}_k|_{B'(i)}$ is a uniformly random string whereas $\tilde{Y}_k|_{B(i)}$ is a fixed string (since the *parity* of all columns in $B(i)$ is 1). Let $\mathcal{U}$ be the uniform distribution on $\{-1,+1\}^d$ and $\mathcal{P}(X^i,Y^i_{-k}) \in \{+1,-1\}^d$ denotes the column wise parities of $\begin{bmatrix} X^i|_{B(i)} \\ Y^i_{-k}|_{B(i)} \end{bmatrix}$.

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2 \mid c_2 = 0] &= \mathop{\mathbb{E}}_{X^i,Y^i_{-k} \sim \mathcal{T}_2^i} \left(\mathop{\mathbb{E}}_{(\tilde{X},\tilde{Y}) \sim \mathcal{T}_2^i} \left[g(\tilde{Y}_k) \mid {}^{(\tilde{X},\tilde{Y}_{-k}) = (X^i,Y^i_{-k}) \wedge}_{c_2 = 0} \right] \right)^2 \\
&= \mathop{\mathbb{E}}_{\substack{X^i,Y^i_{-k} \sim \mathcal{T}_2^i, \\ z = \mathcal{P}(X^i,Y^i_{-k})}} \left(\mathop{\mathbb{E}}_{r \sim \mathcal{U}}[g(-z,r)] \right)^2 \\
&= \mathop{\mathbb{E}}_{z \sim \mathcal{U}} \left(\mathop{\mathbb{E}}_{r \sim \mathcal{U}}[g(z,r)] \right)^2 \quad \text{(since marginal on } z \text{ is uniform)} \\
&= \mathop{\mathbb{E}}_{z \sim \mathcal{U}} \left(\mathop{\mathbb{E}}_{r \in \mathcal{U}} \sum_{\alpha \subseteq B(i) \cup B'(i)} \hat{g}(\alpha) \chi_\alpha(z,r) \right)^2 \\
&= \mathop{\mathbb{E}}_{z \sim \mathcal{U}} \left(\sum_{\alpha \subseteq B(i) \cup B'(i)} \hat{g}(\alpha) \mathop{\mathbb{E}}_{r \in \mathcal{U}}[\chi_\alpha(z,r)] \right)^2 \\
&= \mathop{\mathbb{E}}_{z \sim \mathcal{U}} \left(\sum_{\alpha \subseteq B(i)} \hat{g}(\alpha) \chi_\alpha(z) \right)^2 \qquad = \qquad \sum_{\alpha \subseteq B(i)} \hat{g}(\alpha)^2.
\end{aligned}$$

Similarly we have,

$$\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2 \mid c_2 = 1] = \sum_{\alpha \subseteq B'(i)} \hat{g}(\alpha)^2.$$

Now we can bound the correlation as follows.

$$\begin{aligned}
\rho\left(\mathcal{X}^i \times \mathcal{Y}^i_{-k}, \mathcal{Y}^i_k; \mathcal{T}_2^i\right)^2 &\leq (1-2\varepsilon) + \varepsilon \sum_{\alpha \subseteq B(i)} \hat{g}(\alpha)^2 + \varepsilon \sum_{\alpha \subseteq B'(i)} \hat{g}(\alpha)^2 \\
&\leq (1-2\varepsilon) + \varepsilon \sum_{\alpha \subseteq B(i) \cup B'(i)} \hat{g}(\alpha)^2 \quad (\text{Using } \hat{g}(\phi) = \mathbb{E}[g] = 0) \\
&\leq (1-\varepsilon). \qquad (\text{Using } \mathbb{E}[g^2] \leq 1 \text{ and Parseval's Identity})
\end{aligned}$$

□

5 Improvement to covering hardness of 4-LIN

In this section, we prove Theorem 1.4. We give a reduction from an instance of LABEL-COVER, $G = (U,V,E,[L],[R],\{\pi_e\}_{e\in E})$ as in Definition 2.4, to a 4-LIN-CSP instance $\mathcal{G} = (\mathcal{V},\mathcal{E})$. The set of variables $\mathcal{V}$ is $V \times \{0,1\}^{2R}$. Any assignment to $\mathcal{V}$ is given by a set of functions $f_v : \{0,1\}^{2R} \to \{0,1\}$, for each $v \in V$. The set of constraints $\mathcal{E}$ is given by the following test, which checks whether f_v's are long codes of a good labeling to V.

Long Code test $\mathcal{T}_3$

1. Choose $u \in U$ uniformly and neighbors $v,w \in V$ of u uniformly and independently at random.
2. Choose x,x',z,z' uniformly and independently from $\{0,1\}^{2R}$ and y from $\{0,1\}^{2L}$. Choose $(\eta,\eta') \in \{0,1\}^{2L} \times \{0,1\}^{2L}$ as follows. Independently for each $i \in [L]$, set $(\eta_i,\eta_{L+i},\eta'_i,\eta'_{L+i})$ to
   (a) $(0,0,0,0)$ with probability $1-2\varepsilon$,
   (b) $(1,0,1,0)$ with probability ε and
   (c) $(0,1,0,1)$ with probability ε.
3. For $y \in \{0,1\}^{2L}$, let $y\circ\pi_{uv} \in \{0,1\}^{2R}$ be the string such that $(y\circ\pi_{uv})_i := y_{\pi_{uv}(i)}$ for $i \in [R]$ and $(y\circ\pi_{uv})_i := y_{\pi_{uv}(i-R)+L}$ otherwise. Given $\eta \in \{0,1\}^{2L}, z \in \{0,1\}^{2R}$, the string $\eta\circ\pi_{uv}\cdot z \in \{0,1\}^{2R}$ is obtained by taking coordinate-wise product of $\eta\circ\pi_{uv}$ and z. Accept iff
$$f_v(x) + f_v(x + y\circ\pi_{uv} + \eta\circ\pi_{uv}\cdot z) + f_w(x') + f_w(x' + y\circ\pi_{uw} + \eta'\circ\pi_{uw}\cdot z' + \overline{1}) = 1 \pmod 2. \quad (5.1)$$
   (Here by addition of strings, we mean the coordinate-wise sum modulo 2.)

Lemma 5.1 (Completeness)**.** *If G is an YES instance of* LABEL-COVER*, then there exists f,g such that each of them covers $1-\varepsilon$ fraction of $\mathcal{E}$ and they together cover all of $\mathcal{E}$.*

Proof. Let $\ell : U \cup V \to [L]\cup[R]$ be a labeling to G that satisfies all the constraints. Consider the assignments given by $f_v(x) := x_{\ell(v)}$ and $g_v(x) := x_{R+\ell(v)}$ for each $v \in V$. On input f_v, for any pair of edges $(u,v),(u,w) \in E$, and x,x',z,z',η,η',y chosen by the long code test $\mathcal{T}_3$, the LHS in Eq. (5.1) evaluates to
$$x_{\ell(v)} + x_{\ell(v)} + y_{\ell(u)} + \eta_{\ell(u)}z_{\ell(v)} + x'_{\ell(w)} + x'_{\ell(w)} + y_{\ell(u)} + \eta'_{\ell(u)}z'_{\ell(w)} + 1 = \eta_{\ell(u)}z_{\ell(v)} + \eta'_{\ell(u)}z'_{\ell(w)} + 1.$$
Similarly for g_v, the expression evaluates to $\eta_{L+\ell(u)}z_{R+\ell(v)} + \eta'_{L+\ell(u)}z'_{R+\ell(w)} + 1$. Since $(\eta_i,\eta'_i) = (0,0)$ with probability $1-\varepsilon$, each of f,g covers $1-\varepsilon$ fraction of $\mathcal{E}$. Also for $i \in [L]$ whenever $(\eta_i,\eta'_i) = (1,1)$, $(\eta_{L+i},\eta'_{L+i}) = (0,0)$ and vice versa. So one of the two evaluations above is $1 \pmod 2$. Hence the pair of assignments f,g cover $\mathcal{E}$. □

Lemma 5.2 (Soundness)**.** *Let c_0 be the constant from Theorem 2.5. If G is at most s-satisfiable with $s < \delta^{10/c_0+5}/4$, then any independent set in $\mathcal{G}$ has fractional size at most δ.*

Proof. Let $I \subseteq \mathcal{V}$ be an independent set of fractional size δ in the constraint graph $\mathcal{G}$. For every variable $v \in V$, let $f_v : \{0,1\}^{2R} \to \{0,1\}$ be the indicator function of the independent set restricted to the vertices that correspond to v. Since I is an independent set, we have

$$\mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\substack{x,x',\\ z,z',\\ \eta,\eta',y}} \left[f_v(x) f_v(x + y \circ \pi_{uv} + \eta \circ \pi_{uv} \cdot z) f_w(x') f_w(x' + y \circ \pi_{uw} + \eta' \circ \pi_{uw} \cdot z' + 1) \right] = 0. \tag{5.2}$$

For $\alpha \subseteq [2R]$, let $\pi^{\oplus}_{uv}(\alpha) \subseteq [2L]$ be the set containing elements $i \in [2L]$ such that if $i < L$ there are an odd number of $j \in [R] \cap \alpha$ with $\pi_{uv}(j) = i$ and if $i \geq L$ there are an odd number of $j \in ([2R] \setminus [R]) \cap \alpha$ with $\pi_{uv}(j - R) = i - L$. It is easy to see that $\chi_\alpha(y \circ \pi_{uw}) = \chi_{\pi^{\oplus}_{uv}(\alpha)}(y)$. Expanding f_v in the Fourier basis and taking expectation over x, x' and y, we get that

$$\mathop{\mathbb{E}}_{u,v,w} \sum_{\alpha,\beta \subseteq [2R] : \pi^{\oplus}_{uv}(\alpha) = \pi^{\oplus}_{uw}(\beta)} \widehat{f_v}(\alpha)^2 \widehat{f_w}(\beta)^2 (-1)^{|\beta|} \mathop{\mathbb{E}}_{z,z',\eta,\eta'} \left[\chi_\alpha(\eta \circ \pi_{uv} \cdot z) \chi_\beta(\eta' \circ \pi_{uw} \cdot z') \right] = 0. \tag{5.3}$$

Now the expectation over z, z' simplifies as

$$\mathop{\mathbb{E}}_{u,v,w} \sum_{\alpha,\beta \subseteq [2R] : \pi^{\oplus}_{uv}(\alpha) = \pi^{\oplus}_{uw}(\beta)} \underbrace{\widehat{f_v}(\alpha)^2 \widehat{f_w}(\beta)^2 (-1)^{|\beta|} \Pr_{\eta,\eta'}[\alpha \cdot (\eta \circ \pi_{uv}) = \beta \cdot (\eta' \circ \pi_{uw}) = \bar{0}]}_{=: \mathsf{Term}_{u,v,w}(\alpha,\beta)} = 0, \tag{5.4}$$

where we think of α, β as the characteristic vectors in $\{0,1\}^{2R}$ of the corresponding sets. We will now break up the above summation into different parts and bound each part separately. For a projection $\pi : [R] \to [L]$, define $\widetilde{\pi}(\alpha) := \{i \in [L] : \exists j \in [R], (j \in \alpha \vee j + R \in \alpha) \wedge (\pi(j) = i)\}$. We divide the space of (α, β) into 4 sets as follows.

$$\begin{aligned}
E_0 &:= \left\{ (\alpha,\beta) \middle| \pi^{\oplus}_{uv}(\alpha) = \pi^{\oplus}_{uw}(\beta) = \emptyset \right\}, \\
E_1 &:= \left\{ (\alpha,\beta) \middle| \pi^{\oplus}_{uv}(\alpha) = \pi^{\oplus}_{uw}(\beta) \neq \emptyset, \max\{|\alpha|,|\beta|\} \leq 2/\delta^{5/c_0} \right\}, \\
E_2 &:= \left\{ (\alpha,\beta) \middle| \pi^{\oplus}_{uv}(\alpha) = \pi^{\oplus}_{uw}(\beta) \neq \emptyset, \max\{|\widetilde{\pi}_{uv}(\alpha)|, |\widetilde{\pi}_{uw}(\beta)|\} \geq 1/\delta^5 \right\}, \\
E_3 &:= \left\{ (\alpha,\beta) \middle| \pi^{\oplus}_{uv}(\alpha) = \pi^{\oplus}_{uw}(\beta) \neq \emptyset, \max\{|\alpha|,|\beta|\} > 2/\delta^{5/c_0}, \max\{|\widetilde{\pi}_{uv}(\alpha)|, |\widetilde{\pi}_{uw}(\beta)|\} < 1/\delta^5 \right\}.
\end{aligned}$$

And define the following quantities for $i \in \{0,1,2,3\}$.

$$\Theta_i := \sum_{(\alpha,\beta) \in E_i} \mathop{\mathbb{E}}_{u,v,w} \mathsf{Term}_{u,v,w}(\alpha,\beta).$$

Lower bound for Θ_0. If $\pi^{\oplus}_{uw}(\beta) = \emptyset$, then $|\beta|$ is even. Hence, all the terms in Θ_0 are positive and

$$\Theta_0 \geq \mathop{\mathbb{E}}_{u,v,w} \mathsf{Term}_{u,v,w}(0,0) = \mathop{\mathbb{E}}_{u} \left(\mathop{\mathbb{E}}_{v} \widehat{f_v}(0)^2 \right)^2 \geq \left(\mathop{\mathbb{E}}_{u,v} \widehat{f_v}(0) \right)^4 = \delta^4.$$

Upper bound for Θ_1. Consider the following strategy for labeling vertices $u \in U$ and $v \in V$. For $u \in U$, pick a random neighbor v, choose α with probability $\widehat{f_v}(\alpha)^2$ and set its label to a random element in $\widetilde{\pi}_{uv}(\alpha)$. For $w \in V$, choose β with probability $\widehat{f_w}(\beta)^2$ and set its label to a random element of β. If the label $j \geq R$, change the label to $j - R$. The probability that a random edge (u,w) of the label cover is satisfied by this labeling is

$$\mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\alpha,\beta:\\ \widetilde{\pi}_{uv}(\alpha)\cap\widetilde{\pi}_{uw}(\beta)\neq\emptyset}} \widehat{f_v}(\alpha)^2 \widehat{f_w}(\beta)^2 \frac{1}{|\widetilde{\pi}_{uv}(\alpha)|\cdot|\beta|} \geq \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\alpha,\beta:\\ \pi^{\oplus}_{uv}(\alpha)=\pi^{\oplus}_{uw}(\beta)\neq\emptyset\\ \max\{|\alpha|,|\beta|\}\leq 2/\delta^{5/c_0}}} \widehat{f_v}(\alpha)^2 \widehat{f_w}(\beta)^2 \frac{\delta^{10/c_0}}{4}$$
$$\geq |\Theta_1| \cdot \frac{\delta^{10/c_0}}{4}.$$

Since the instance is at most s-satisfiable, the above is not greater than s. Choosing $s < \delta^{10/c_0+5}/4$, will imply $|\Theta_1| \leq \delta^5$.

Upper bound for Θ_2. Suppose $|\widetilde{\pi}_{uv}(\alpha)| \geq 1/\delta^5$, then note that

$$\Pr_{\eta,\eta'}[\alpha\cdot(\eta\circ\pi_{uv}) = \beta\cdot(\eta'\circ\pi_{uw}) = 0] \leq \Pr_{\eta}[\alpha\cdot(\eta\circ\pi_{uv}) = 0] \leq (1-\varepsilon)^{|\widetilde{\pi}_{uv}(\alpha)|} \leq (1-\varepsilon)^{1/\delta^5}.$$

Since the sum of squares of Fourier coefficients of f is less than 1 and ε is a constant, we get that $|\Theta_2| \leq 1/2^{\Omega(1/\delta^5)} < O(\delta^5)$.

Upper bound for Θ_3. From the third property of Theorem 2.5, we have that for any $v \in V$ and $\alpha \subseteq [2R]$ with $|\alpha| > 2/\delta^{5/c_0}$, the probability that $|\widetilde{\pi}_{uv}(\alpha)| < 1/\delta^5$, for a random neighbor u of v, is at most δ^5. Hence $|\Theta_3| \leq \delta^5$.

On substituting the above bounds in Equation (5.4), we get that $\delta^4 - O(\delta^5) \leq 0$, which gives a contradiction for small enough δ. Hence there is no independent set in $\mathcal{G}$ of size δ. □

Proof of Theorem 1.4. From Theorem 2.5, the size of the CSP instance $\mathcal{G}$ produced by the reduction is $N = n^r 2^{2^{O(r)}}$ and the parameter $s \leq 2^{-d_0 r}$. Setting $r = \Theta(\log\log n)$, gives that $N = 2^{\operatorname{poly}\log n}$ and the size of the largest independent set $\delta = 1/\operatorname{poly}\log n = 1/\operatorname{poly}\log N$. □

6 Invariance Principle for correlated spaces

Theorem 2.9 (Invariance Principle for correlated spaces) [Restated] *Let $(\Omega_1^k \times \Omega_2^k, \mu)$ be a correlated probability space such that the marginal of μ on any pair of coordinates one each from Ω_1 and Ω_2 is a product distribution. Let μ_1, μ_2 be the marginals of μ on Ω_1^k and Ω_2^k, respectively. Let X, Y be two random $k \times L$ dimensional matrices chosen as follows. Independently for every $i \in [L]$, the pair of* *columns* $(x^i, y^i) \in \Omega_1^k \times \Omega_2^k$ *is chosen from* μ. *Let* x_i, y_i *denote the i-th rows of X and Y, respectively. If* $F : \Omega_1^L \to [-1,+1]$ *and* $G : \Omega_2^L \to [-1,+1]$ *are functions such that*

$$\tau := \sqrt{\sum_{i\in[L]} \mathsf{Inf}_i[F]\cdot \mathsf{Inf}_i[G]} \ \ \textit{and} \ \ \Gamma := \max\left\{\sqrt{\sum_{i\in[L]}\mathsf{Inf}_i[F]}, \sqrt{\sum_{i\in[L]}\mathsf{Inf}_i[G]}\right\},$$

then

$$\left|\mathop{\mathbb{E}}_{(X,Y)\in\mu^{\otimes L}}\left[\prod_{i\in[k]} F(x_i)G(y_i)\right] - \mathop{\mathbb{E}}_{X\in\mu_1^{\otimes L}}\left[\prod_{i\in[k]} F(x_i)\right]\mathop{\mathbb{E}}_{Y\in\mu_2^{\otimes L}}\left[\prod_{i\in[k]} G(y_i)\right]\right| \le 2^{O(k)}\Gamma\tau. \tag{6.1}$$

Proof. We will prove the theorem by using the hybrid argument. For $i \in [L+1]$, let $X^{(i)}, Y^{(i)}$ be distributed according to $(\mu_1 \otimes \mu_2)^{\otimes i} \otimes \mu^{\otimes L-i}$. Thus, $(X^{(0)}, Y^{(0)}) = (X,Y)$ is distributed according to $\mu^{\otimes L}$ while $(X^{(L)}, Y^{(L)})$ is distributed according to $(\mu_1 \otimes \mu_2)^{\otimes L}$. For $i \in [L]$, define

$$\mathsf{err}_i := \left|\mathop{\mathbb{E}}_{X^{(i)},Y^{(i)}}\left[\prod_{j=1}^{k} F(x_j^{(i)})G(y_j^{(i)})\right] - \mathop{\mathbb{E}}_{X^{(i+1)},Y^{(i+1)}}\left[\prod_{j=1}^{k} F(x_j^{(i+1)})G(y_j^{(i+1)})\right]\right|. \tag{6.2}$$

The left-hand side of Equation (2.2) is not greater than $\sum_{i\in[L]} \mathsf{err}_i$. Now for a fixed i, we will bound err_i. We use the Efron–Stein decomposition of F, G to split them into two parts: the part that depends on the i-th input and the part independent of the i-th input.

$$F = F_0 + F_1 \text{ where } F_0 := \sum_{\alpha : i\notin\alpha} F_\alpha \text{ and } F_1 := \sum_{\alpha : i\in\alpha} F_\alpha.$$

$$G = G_0 + G_1 \text{ where } G_0 := \sum_{\beta : i\notin\beta} G_\beta \text{ and } G_1 := \sum_{\beta : i\in\beta} G_\beta.$$

Note that $\mathsf{Inf}_i[F] = \|F_1\|_2^2$ and $\mathsf{Inf}_i[G] = \|G_1\|_2^2$. Furthermore, the functions F_0 and F_1 are bounded since $F_0(x) = \mathbb{E}_{x'}[F(x')|x'_{[L]\setminus i} = x_{[L]\setminus i}] \in [-1,+1]$ and $F_1(x) = F(x) - F_0(x) \in [-2,+2]$. For $a \in \{0,1\}^k$, let $F_a(X) := \prod_{j=1}^k F_{a_j}(x_j)$. Similarly G_0, G_1 are bounded and G_a defined analogously. Substituting these definitions in Equation (6.2) and expanding the products gives

$$\mathsf{err}_i = \left|\sum_{a,b\in\{0,1\}^k}\left(\mathop{\mathbb{E}}_{X^{(i)},Y^{(i)}}\left[F_a(X^{(i)})G_b(Y^{(i)})\right] - \mathop{\mathbb{E}}_{X^{(i+1)},Y^{(i+1)}}\left[F_a(X^{(i+1)})G_b(Y^{(i+1)})\right]\right)\right|.$$

Since both the distributions are identical on $(\Omega_1^k)^{\otimes L}$ and $(\Omega_2^k)^{\otimes L}$, all terms with $a = \bar{0}$ or $b = \bar{0}$ are zero. For instance when $a = \bar{0}$, F_a does not depend on the i-th coordinate. Therefore, in both the distributions $(X^{(i)}, Y^{(i)})$ and $(X^{(i+1)}, Y^{(i+1)})$, the i-th column of X can be dropped. Now, the distributions of $Y^{(i)}$ and $Y^{(i+1)}$ are identical conditioned on the X with the i-th column dropped. Thus, the expectation is 0.

Since μ is uniform on any pair of coordinates on each from the Ω_1 and Ω_2 sides, terms with $|a| = |b| = 1$ also evaluates to zero using a similar argument as above. Now consider the remaining terms with $|a|,|b| \geq 1, |a|+|b| > 2$. Consider one such term where $a_1, a_2 = 1$ and $b_1 = 1$. In this case, by the Cauchy–Schwarz inequality we have that

$$\left| \mathop{\mathbb{E}}_{X^{(i-1)},Y^{(i-1)}} \left[F_a(X^{(i-1)}) G_b(Y^{(i-1)}) \right] \right| \leq \sqrt{\mathbb{E}\, F_1(x_1)^2 G_1(y_1)^2} \cdot \|F_1\|_2 \cdot \left\| \prod_{j>2} F_{a_j} \right\|_\infty \cdot \left\| \prod_{j>1} G_{b_j} \right\|_\infty.$$

From the facts that the marginal of μ to any pair of coordinates one each from Ω_1 and Ω_2 sides are uniform, $\mathsf{Inf}_i[F] = \|F_1\|_2^2$ and $|F_0(x)|, |F_1(x)|, |G_0(x)|, |G_1(x)|$ are all bounded by 2, the right side of above becomes

$$\sqrt{\mathbb{E}\, F_1(x_1)^2} \sqrt{\mathbb{E}\, G_1(y_1)^2} \cdot \|F_1\|_2 \cdot \left\| \prod_{j>2} F_{a_j} \right\|_\infty \cdot \left\| \prod_{j>1} G_{b_j} \right\|_\infty \leq \sqrt{\mathsf{Inf}_i[F]^2 \mathsf{Inf}_i[G]} \cdot 2^{2k}.$$

All the other terms corresponding to other pairs (a,b), which are at most 2^{2k} in number, are bounded analogously. Hence,

$$\begin{aligned} \sum_{i\in[L]} \mathsf{err}_i &\leq 2^{4k} \sum_{i\in[L]} \left(\sqrt{\mathsf{Inf}_i[F]^2 \mathsf{Inf}_i[G]} + \sqrt{\mathsf{Inf}_i[F] \mathsf{Inf}_i[G]^2} \right) \\ &= 2^{4k} \sum_{i\in[L]} \sqrt{\mathsf{Inf}_i[F] \mathsf{Inf}_i[G]} \left(\sqrt{\mathsf{Inf}_i[F]} + \sqrt{\mathsf{Inf}_i[G]} \right). \end{aligned}$$

Applying the Cauchy–Schwarz inequality, followed by a triangle inequality, we obtain

$$\sum_{i\in[L]} \mathsf{err}_i \leq 2^{4k} \sqrt{\sum_{i\in[L]} \mathsf{Inf}_i[F] \mathsf{Inf}_i[G]} \left(\sqrt{\sum_{i\in[L]} \mathsf{Inf}_i[F]} + \sqrt{\sum_{i\in[L]} \mathsf{Inf}_i[G]} \right).$$

This completes the proof. □

References

[1] SANJEEV ARORA, CARSTEN LUND, RAJEEV MOTWANI, MADHU SUDAN, AND MARIO SZEGEDY: Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998. Preliminary version in FOCS'92. [doi:10.1145/278298.278306, ECCC:TR98-008] 8

[2] SANJEEV ARORA AND SHMUEL SAFRA: Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998. Preliminary version in FOCS'92. [doi:10.1145/273865.273901] 8

[3] PER AUSTRIN AND ELCHANAN MOSSEL: Approximation resistant predicates from pairwise independence. *Comput. Complexity*, 18(2):249–271, 2009. Preliminary version in CCC'08. [doi:10.1007/s00037-009-0272-6, arXiv:0802.2300, ECCC:TR08-009] 5

[4] NIKHIL BANSAL AND SUBHASH KHOT: Inapproximability of hypergraph vertex cover and applications to scheduling problems. In *Proc. 37th Internat. Colloq. Automata, Lang., Programming (ICALP'10)*, pp. 250–261. Springer, 2010. [doi:10.1007/978-3-642-14165-2_22] 5, 11

[5] BOAZ BARAK, PARIKSHIT GOPALAN, JOHAN HÅSTAD, RAGHU MEKA, PRASAD RAGHAVENDRA, AND DAVID STEURER: Making the long code shorter. *SIAM J. Comput.*, 44(5):1287–1324, 2015. Preliminary version in FOCS'12. [doi:10.1137/130929394, arXiv:1111.0405, ECCC:TR11-142] 7

[6] AMEY BHANGALE, PRAHLADH HARSHA, AND GIRISH VARMA: A characterization of hard-to-cover CSPs. In *Proc. 30th Comput. Complexity Conf. (CCC'15)*, pp. 280–303. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015. [doi:10.4230/LIPIcs.CCC.2015.280, arXiv:1411.7747] 1

[7] IRIT DINUR AND VENKATESAN GURUSWAMI: PCPs via the low-degree long code and hardness for constrained hypergraph coloring. *Israel J. Math.*, 209:611–649, 2015. Preliminary version in FOCS'13. [doi:10.1007/s11856-015-1231-3, ECCC:TR13-122] 7

[8] IRIT DINUR AND GILLAT KOL: Covering CSPs. In *Proc. 28th IEEE Conf. on Comput. Complexity (CCC'13)*, pp. 207–218. IEEE Comp. Soc., 2013. [doi:10.1109/CCC.2013.29, ECCC:TR12-088] 2, 3, 4, 5, 8, 11, 13, 15, 16

[9] URIEL FEIGE AND JOE KILIAN: Zero knowledge and the chromatic number. *J. Comput. System Sci.*, 57(2):187–199, 1998. Preliminary version in CCC'96. [doi:10.1006/jcss.1998.1587] 2

[10] VENKAT GURUSWAMI, PRAHLADH HARSHA, JOHAN HÅSTAD, SRIKANTH SRINIVASAN, AND GIRISH VARMA: Super-polylogarithmic hypergraph coloring hardness via low-degree long codes. *SIAM J. Comput.*, 46(1):132–159, 2017. Preliminary version in STOC'14. [doi:10.1137/140995520, arXiv:1311.7407] 7

[11] VENKATESAN GURUSWAMI, JOHAN HÅSTAD, AND MADHU SUDAN: Hardness of approximate hypergraph coloring. *SIAM J. Comput.*, 31(6):1663–1686, 2002. Preliminary version in FOCS'00. [doi:10.1137/S0097539700377165] 2, 5

[12] VENKATESAN GURUSWAMI AND EUIWOONG LEE: Strong inapproximability results on balanced rainbow-colorable hypergraphs. *Combinatorica*, 38(3):547–599, 2018. Preliminary version in SODA'15. [doi:10.1007/s00493-016-3383-0, ECCC:TR14-043] 6, 10

[13] JOHAN HÅSTAD: Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. Preliminary version in STOC'97. [doi:10.1145/502090.502098] 8, 15

[14] JONAS HOLMERIN: Vertex cover on 4-regular hyper-graphs is hard to approximate within $2-\varepsilon$. In *Proc. 34th STOC*, pp. 544–552. ACM Press, 2002. [doi:10.1145/509907.509986, ECCC:TR01-094] 5

[15] SANGXIA HUANG: $2^{(\log N)^{1/10-o(1)}}$ hardness for hypergraph coloring. 2015. [arXiv:1504.03923] 7

[16] SUBHASH KHOT AND RISHI SAKET: Hardness of coloring 2-colorable 12-uniform hypergraphs with $2^{(\log n)^{\Omega(1)}}$ colors. *SIAM J. Comput.*, 46(1):235–271, 2017. Preliminary version in FOCS'14. [doi:10.1137/15100240X, ECCC:TR14-051] 7

[17] MICHAEL KRIVELEVICH, RAM NATHANIEL, AND BENNY SUDAKOV: Approximating coloring and maximum independent sets in 3-uniform hypergraphs. *J. Discr. Algor.*, 41(1):99–113, 2001. Preliminary version in SODA'01. [doi:10.1006/jagm.2001.1173] 2

[18] ELCHANAN MOSSEL: Gaussian bounds for noise correlation of functions. *Geom. Funct. Anal.*, 19(6):1713–1756, 2010. Preliminary version in FOCS'08. [doi:10.1007/s00039-010-0047-x, arXiv:math/0703683] 6, 9, 16, 19, 20

[19] RAN RAZ: A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. Preliminary version in STOC'95. [doi:10.1137/S0097539795280895] 8

[20] RISHI SAKET: Hardness of finding independent sets in 2-colorable hypergraphs and of satisfiable CSPs. In *Proc. 29th IEEE Conf. on Comput. Complexity (CCC'14)*, pp. 78–89. IEEE Comp. Soc., 2014. [doi:10.1109/CCC.2014.16, arXiv:1312.2915] 7

[21] GIRISH VARMA: Reducing uniformity in Khot-Saket hypergraph coloring hardness reductions. *Chicago J. Theoret. Comp. Sci.*, (3):1–8, 2015. [doi:10.4086/cjtcs.2015.003, arXiv:1408.0262] 7

[22] CENNY WENNER: Circumventing *d*-to-1 for approximation resistance of satisfiable predicates strictly containing parity of width at least four. *Theory of Computing*, 9(23):703–757, 2013. Preliminary version in APPROX'12. [doi:10.4086/toc.2013.v009a023, ECCC:TR12-145] 6, 10, 18

[23] DAVID ZUCKERMAN: Linear degree extractors and the inapproximability of max clique and chromatic number. *Theory of Computing*, 3(6):103–128, 2007. Preliminary version in STOC'06. [doi:10.4086/toc.2007.v003a006, ECCC:TR05-100] 2

AUTHORS

Amey Bhangale
Assistant professor
University of California, Riverside
California, USA
ameyb@ucr.edu
https://www.cs.ucr.edu/~bhangale/

Prahladh Harsha
Associate professor
Tata Institute of Fundamental Research,
Mumbai, India
prahladh@tifr.res.in
http://www.tcs.tifr.res.in/~prahladh/

Girish Varma
Assistant professor
International Institute of Information Technology (IIIT),
Hyderabad, India
girish.varma@iiit.ac.in
https://girishvarma.in

ABOUT THE AUTHORS

AMEY BHANGALE is an Assistant Professor in the Computer Science Department at the University of California, Riverside.

Amey got his undergraduate degree from VJTI, Mumbai. In the final year of his undergraduate studies, he met Prof. Rajiv C. Gandhi, who was visiting Mumbai in 2010/11 on a Fulbright scholarship and taught a course on approximation algorithms at VJTI. This encounter influenced Amey to pursue research in the theory of computing.

Amey graduated from Rutgers University in 2017 under the supervision of Swastik Kopparty. He spent two wonderful years in Israel where he was a post-doctoral fellow at the Weizmann Institute of Science hosted by Irit Dinur. He is interested in approximation algorithms, hardness of approximations and analysis of Boolean functions. He enjoys playing and watching tennis (and hardly says no to anyone who invites him to play tennis).

PRAHLADH HARSHA is a theoretical computer scientist at the Tata Institute of Fundamental Research (TIFR). He received his B. Tech. degree in Computer Science and Engineering from the IIT Madras in 1998 and his S. M. and Ph. D. degrees in Computer Science from MIT in 2000 and 2004, respectively. Prior to joining TIFR in 2010, he was at Microsoft Research, Silicon Valley and at the Toyota Technological Institute at Chicago. His areas of interest include computational complexity, hardness of approximation, coding theory and information theory. Prahladh credits his mother for his interest in mathematics and dance. He is also deeply indebted to U Koteswara Rao, his high school mentor, for exposing him to both the beauty and rigour in mathematics.

GIRISH VARMA is a scientist at the Machine Learning Lab at IIIT Hyderabad. He received his B. Tech. degree in Computer Science and Engineering from the NIT Calicut in 2008 and his M. S. and Ph. D. degrees in Computer Science from the Tata Institute of Fundamental Research (TIFR), Mumbai in 2015. Prior to joining IIIT Hyderabad in 2016, he was at the Weizmann Institute of Science. His research focusses on using theoretical computer science concepts in solving applied problems in computer vision.