THEORY OF COMPUTING, Volume 10 (14), 2014, pp. 359–388
www.theoryofcomputing.org

Approximation Resistance on Satisfiable Instances for Predicates with Few Accepting Inputs

Sangxia Huang*

Received May 13, 2013; Revised January 18, 2014; Published October 15, 2014

Abstract: For every integer $k \geq 3$, we prove that there is a predicate P on k Boolean variables with $2^{\widetilde{O}(k^{1/3})}$ accepting assignments that is approximation resistant even on *satisfiable* instances. That is, given a *satisfiable* CSP instance with constraint P, we cannot achieve better approximation ratio than simply picking random assignments. This improves the best previously known result by Håstad and Khot (Theory of Computing, 2005) who showed that a predicate on k variables with $2^{O(k^{1/2})}$ accepting assignments is approximation resistant on satisfiable instances.

Our construction is inspired by several recent developments. One is the idea of using direct sums to improve soundness of PCPs, developed by Siu On Chan (STOC, 2013). We also use techniques from Cenny Wenner (Theory of Computing, 2013) to construct PCPs with perfect completeness without relying on the d-to-1 Conjecture.

ACM Classification: F.2.0

AMS Classification: 68Q17

Key words and phrases: Max-CSP, probabilistically checkable proofs, approximation resistance, satisfiable instance

A conference version of this paper appeared in the Proceedings of the 45th annual ACM Symposium on Theory of Computing, 2013 [18].

*Supported by ERC Advanced Investigator grant 226203.

 DOI: 10.4086/toc.2014.v010a014

1 Introduction

In this article we study optimization of constraint satisfaction problems with k variables in each constraint (MAX-k-CSP). A k-CSP instance contains a set of Boolean variables and constraints, where each constraint is expressed by a Boolean predicate on k literals. The goal of MAX-k-CSP is to find an assignment that maximizes the number of satisfied constraints. Given a predicate $P:\{-1,1\}^k \to \{0,1\}$ on k Boolean inputs (where for the input variables we use -1 for True and 1 for False), we can also consider the MAX-P problem, a special case of MAX-k-CSP where all constraints are expressed by the same Boolean predicate P applied to literals of k distinct variables. A MAX-P instance is *satisfiable* if there exists an assignment that satisfies all the constraints simultaneously. Let $P^{-1}(1)$ be the set of accepting inputs of P.

One naive approximation algorithm for MAX-P is to simply pick a random assignment. The expected fraction of constraints satisfied by this algorithm is $|P^{-1}(-1)|/2^k$. Somewhat surprisingly, it turns out that for some predicate P, the above naive algorithm gives the best possible performance assuming $\mathsf{P} \neq \mathsf{NP}$. We call a predicate P *approximation resistant* if it is hard to achieve better approximation ratio than simply picking random assignments. In a celebrated result, Håstad [14] showed that for MAX-k-LIN, sets of linear equations in $\mathbb{Z}_2$ on $k \geq 3$ variables, it is NP-hard to find an assignment satisfying more than a $1/2+\varepsilon$ fraction of the constraints for any $\varepsilon > 0$, even when the input has an assignment that satisfies $1-\varepsilon$ of them, while the random assignment algorithm achieves $1/2$. There has been much progress in understanding what kinds of predicates are approximation resistant, including characterization for predicates of small arity [14, 33, 12], as well as a handful of approximation resistant predicates of higher arities [14, 28, 12, 9].

The picture of approximation resistance becomes even clearer if we assume the Unique Games Conjecture (UGC) proposed by Khot [20], which states that it is NP-hard to distinguish whether certain LABEL-COVER instances are *almost* satisfiable or far from satisfiable. Austrin and Mossel [4] proved that assuming the UGC, P is approximation resistant if the set of satisfying assignments $P^{-1}(1)$ contains the support of a pairwise independent distribution. In [29], Samorodnitsky and Trevisan showed approximation resistance for the following predicate assuming the UGC: the predicate is on 2^k-1 variables, denoted as $x^{(S)}$ for $\emptyset \neq S \subseteq \{1,\ldots,k\}$, and the predicate accepts if for all $S \subseteq [k]$, $|S| \geq 2$, we have $x^{(S)} = \prod_{i\in S} x^{(\{i\})}$. Let $K = 2^k-1$. We denote the above predicate as HADAMARD$_K$. Note that HADAMARD$_K$ has only $K+1$ accepting assignments over 2^K possible assignments, giving a density of $(K+1)/2^K$. Gustav Hast [13] proved that a predicate on $K \geq 3$ variables with at most $2\lfloor K/2\rfloor+1$ accepting inputs is not approximation resistant. More generally, Charikar, Makarychev and Makarychev [6] gave a $ck/2^k$-approximation algorithm for MAX-k-CSP, for some $c > 0.44$. Thus the result of Samorodnitsky and Trevisan is optimal in terms of the sparsity of the predicate.

In a recent breakthrough [5], Siu On Chan settled the NP-hardness of MAX-HADAMARD$_K$ (and, up to a constant factor, MAX-k-CSP in general), bypassing the UGC. Chan introduced the idea of direct sums of probabilistically checkable proofs (PCPs) to improve soundness, which worked very well for predicates that are subgroups of a domain. In particular, the accepting assignments of the HADAMARD$_K$ predicate is a subgroup under elementwise product, and Chan's result implies that it is approximation resistant assuming only $\mathsf{P} \neq \mathsf{NP}$.

Let us now focus on the approximation of *satisfiable* instances. We call P approximation resistant on satisfiable instances if the best possible algorithm is still the random assignment algorithm even with the promise that there is an assignment that satisfies all constraints. In contrast to our understanding of approximation resistance as demonstrated above, approximation resistance on satisfiable instances is still largely a mystery. Most notably, if the constraints only involve linear equations, for instance k-LIN and HADAMARD$_K$, we can always find a satisfying assignment using Gaussian elimination if we are given satisfiable instances, whereas both problems are approximation resistant in general. Several other approximation algorithms for satisfiable instances were introduced [31, 33], and in particular, it is known that predicates with fewer than $(k+1)$ satisfying assignments are never approximation resistant on satisfiable instances.

On the hardness side, there have been only a handful of results: Håstad [14] proved that k-SAT is approximation resistant for satisfiable instances. The sparsest such predicate known on k variables has $2^{O(k^{1/2})}$ accepting assignments, given by Håstad and Khot [16]. This situation is not particularly surprising, as there are quite a few differences between satisfiable instances and almost satisfiable instances. Some approximation resistant predicates, such as the k-LIN predicate discussed above, are not approximation resistant on satisfiable instances. In addition to this inherent structural difference, there are challenges in techniques as well. Even in the case where we do not require perfect completeness, the best hardness result before [5] is the $2^{O(k^{1/2})}/2^k$ hardness proved by Samorodnitsky and Trevisan [28]. Engebretsen and Holmerin [9] later improved the constant in the exponents and pointed out some technical difficulties in getting better hardness than $2^{O(k^{1/2})}/2^k$ using certain kind of PCP reduction. Another challenge is that many approximation resistance results are obtained via reduction from Unique Games. This immediately introduces problem if we need perfect completeness because the Unique Games problem is solvable in polynomial time for satisfiable instances.

To address this, Khot additionally proposed the "d-to-1 Conjectures" [20]. The conjecture states that it is NP-hard to distinguish whether a "d-to-1 LABEL-COVER Instance" is satisfiable or far from satisfiable. O'Donnell and Wu proved a strong result in [26] that the Not-Two predicate (NTW)—predicate on three variables that accepts an input if the number of variables set to (-1) is not two—is approximation resistant on satisfiable instances assuming the d-to-1 conjecture for some d. Their approach was generalized by Tang [30] to MAX-3-CSP$_q$ where q is a prime greater than 3, and Huang [17] to Boolean predicates of arity $k \geq 3$ that accepts a strict superset of inputs of odd parity.

Recently, Håstad [15] and Wenner [32] proved approximation resistance for the above predicates without assuming the d-to-1 conjecture. Their proofs are based on new analytic tools as well as Khot's SMOOTH-LABEL-COVER [19]. We note that several previous results that bypassed the UGC [19, 21, 10, 11] started from SMOOTH-LABEL-COVER, although it is not needed in Chan's recent result.

An interesting question is whether we could combine these recent developments to get approximation resistance result for MAX-P on satisfiable instances for predicate P sparser than the one in Håstad and Khot [16]. From the PCP perspective, this requires PCPs that always accept correct proofs of correct statements. Not only is this a natural property to have, given the challenge of getting proofs with perfect completeness as discussed above, understanding approximability of k-CSP on satisfiable instances may also lead to new tools in both algorithms and hardness results.

An immediate proposal to achieve tight lower-bound for MAX-k-CSP on satisfiable instances would be to construct predicates as in [17, 32], that is, adding a single additional accepting assignment to the HADAMARD$_K$ predicate of arity 2^k-1. However, this simple approach does not work—the accepting inputs of HADAMARD$_K$ form a k-dimensional subspace, so if we add d new accepting inputs to it and get some other predicate P', we only need a $(k+d)$-dimensional subspace to contain all the accepting inputs of P'. Let Q be the predicate that accepts exactly all inputs from this $(k+d)$-dimensional subspace. Given any satisfiable MAX-P' instance with satisfying assignment α, we replace the predicate P' in each constraint with the predicate Q. The solution space of the instance with predicate Q is just a linear subspace satisfying the following:

- It contains the solution α to the original instance with predicate P'.
- If we project the solution space to the set of variables in each constraint, the resulting subspace has dimension at most $(k+d)$.

Therefore, if we sample a random point from this linear subspace, then for each constraint, the probability that we hit α restricted to the variables in that constraint (and hence satisfy the constraint with predicate P') is at least $1/2^{k+d}$. Thus whenever $d = o(2^k)$, the *expected* performance of the above sampling method beats simple random assignment, which only gives $(2^k+d)/2^{2^k}$.

The problem with adding more accepting assignments to HADAMARD$_K$ is that the resulting predicate does not have the group structure as in [5] any more. If we still take many rounds of direct sums as in [5], then to ensure perfect completeness, we need to accept many assignments that are products of the additional assignments we added and end up with a predicate that has more accepting assignments than we would want. On the other hand, as is demonstrated in [5], having more rounds of direct sum helps us to improve soundness dramatically and so if we are looking for sparse predicates that are approximation resistant, it would be natural to have more rounds of proofs in the direct sum. This paper is an attempt to strike a balance. We prove the following approximation resistance result.

Theorem 1.1. *There is a predicate of arity K with $2^{\widetilde{O}(K^{1/3})}$ accepting assignments that is approximation resistant on satisfiable instances.*

This improves the best previous known result of $2^{O(K^{1/2})}$ of Håstad and Khot [16].

Our result is based on many ideas developed in a number of previous works, including [9, 32, 5]. On the highest level, we use direct sum of several PCPs to get improved soundness result. However, as argued above, we also want to limit the number of PCPs involved. Therefore, we use long-code based PCP constructions that are already rather efficient, for example those used by Engebretsen and Holmerin [9]. In [32], Wenner showed how different types of noise operators behave similarly when the reduction is based on SMOOTH-LABEL-COVER. This is helpful when analyzing soundness of PCPs in that it allows us to move from correlated noise with perfect completeness to independent noise that are not perfect but easier to analyze. We also use a multivariate invariance theorem in [32], which extends methods of Mossel et al. [24, 23] to projection games. Similar techniques were developed also in other works such as [25] as well as in [5].

2 Preliminaries

In this section, we introduce some notations. In Section 2.1, we discuss variants of LABEL-COVER problems, and in particular the MULTI-LAYER-SMOOTH-LABEL-COVER that we use in the rest of the paper. We describe the general approach to proving approximation resistance via LABEL-COVER, as well as Chan's improvements in Section 2.2. In Section 2.3, we review the basics of harmonic analysis of Boolean functions.

2.1 Variants of LABEL-COVER

We first recall the definition of the LABEL-COVER problem.

Definition 2.1. A LABEL-COVER instance is defined by a tuple (U,V,E,L,R,Π). Here U and V are two sets of vertices of a bipartite multigraph, and E is the set of edges between them. L and R are label sets for vertices U and V, respectively. Π is a collection of projections, one for each edge e, $\pi_e : R \to L$. For a labeling $\sigma = (\sigma_U, \sigma_V)$ of the LABEL-COVER instance $\sigma_U : U \to L$, $\sigma_V : V \to R$, let its value be the fraction of edges $\{u,v\} \in E$ such that $\pi_{\{u,v\}}(\sigma(v)) = \sigma(u)$. The value of a LABEL-COVER instance is the maximum value of all possible assignments.

The following theorem combines the celebrated PCP theorem [1, 2] with Raz's parallel repetition theorem [27] and shows hardness of LABEL-COVER.

Theorem 2.2. *For every constant $\eta > 0$, there is some constant $C(\eta) < \infty$ such that for* LABEL-COVER *instances with $|R| \geq C(\eta)$, it is* NP*-hard to distinguish between those with value* 1 *and those with value no more than η.*

The MULTI-LAYERED-SMOOTH-LABEL-COVER problem is a variant of LABEL-COVER first studied in Khot [19] for showing hardness of coloring 3-colorable 3-uniform hypergraphs.

Definition 2.3 (Smoothness)**.** A LABEL-COVER instance is ξ-smooth if for any vertex $v \in V$ and any two labels $r \neq r' \in R$, over a uniformly at random neighbor u of v, we have

$$\Pr_{u \sim v}[\pi_e(r) = \pi_e(r')] \leq \xi \,. \tag{2.1}$$

Similar to LABEL-COVER, we have the following hardness result for SMOOTH-LABEL-COVER.

Theorem 2.4. *For every constant $\eta, \xi > 0$, there is some constant $D(\eta,\xi) < \infty$ such that for ξ-smooth* LABEL-COVER *instances with $|R| \geq D(\eta,\xi)$, it is* NP*-hard to distinguish between those with value 1 and those with value no more than η.*

MULTI-LAYERED-LABEL-COVER was first devised in [7] to prove strong approximation hardness result for hypergraph vertex cover, and used in [9] for improving query efficiency and hardness of approximation result for MAX-CSP. Briefly speaking, a normal LABEL-COVER instance checks consistency of labeling between a pair of vertices, while in a k-layered LABEL-COVER instance, we consider tuples of $k-1$ independently sampled edges $(\{u_1,v_1\},\{u_2,v_2\},\ldots,\{u_{k-1},v_{k-1}\})$, the k hybrid tuples of vertices $(u_1,\ldots,u_i,v_{i+1},\ldots,v_{k-1})$ for $i = 0,\ldots,k-1$ and their corresponding labelings, and we require consistency between all pairs of tuples. Formally, given a LABEL-COVER instance as defined above, the constraint between pairs of labelings on tuples is defined as follows.

Definition 2.5. Let $\vec{e} = (e_1, \ldots, e_{k-1}) \in E^{k-1}$ be a vector, and let $1 \le i < j \le k$. Define the mapping $\pi_{\vec{e}, j\to i} : L^{k-j} \times R^{j-1} \to L^{k-i} \times R^{i-1}$ as

$$(l_1, \ldots, l_{k-j}, r_{k-j+1}, \ldots, r_{k-1}) \mapsto (l_1, \ldots, l_{k-j}, \pi_{e_{k-j+1}}(r_{k-j+1}), \ldots, \pi_{e_{k-i}}(r_{k-i}), r_{k-i+1}, \ldots, r_{k-1}) .$$

It is not hard to see that the above definition preserves smoothness in the layered LABEL-COVER instances.

Lemma 2.6. *For any k-layered* LABEL-COVER *instance constructed from a ξ-smooth* LABEL-COVER *instance (U,V,E,L,R,Π), any positive integer $1 < i \le k$, vertex tuple $\vec{u} = (u_1, \ldots, u_{k-1}) \in U^{k-i} \times V^{i-1}$, two tuples of labelings $\vec{r} \ne \vec{r'} \in L^{k-i} \times R^{i-1}$, we have*

$$\Pr_{\vec{e}\sim\vec{u}}\left[\pi_{\vec{e},i\to 1}(\vec{r}) = \pi_{\vec{e},i\to 1}(\vec{r'})\right] < \xi ,$$

where we sample $\vec{e}$ by picking each $e_i \sim u_i$ independently.

Proof. If there exists $j \in \{1, \ldots, k-i\}$ such that $r_j \ne r'_j$, then we always have

$$\pi_{\vec{e},i\to 1}(\vec{r}) \ne \pi_{\vec{e},i\to 1}(\vec{r'})$$

and hence the above inequality holds for any $\xi > 0$.

We now assume that for all $j \in \{1, \ldots, k-i\}$, we have $r_j = r'_j$, and that there exists $j_0 \in \{k-i+1, \ldots, k-1\}$ such that $r_j \ne r'_j$. Observe that

$$\pi_{\vec{e},i\to 1}(\vec{r}) = \pi_{\vec{e},i\to 1}(\vec{r'})$$

implies that for all $j \in \{k-i+1, \ldots, k-1\}$, we have $\pi_{e_j}(r_j) = \pi_{e_j}(r'_j)$, and in particular

$$\pi_{e_{j_0}}(r_{j_0}) = \pi_{e_{j_0}}(r'_{j_0}) .$$

By definition of smoothness, this happens with probability less than ξ. □

In many applications, it is often easier to work with a slightly stronger notion of smoothness. We extend the definition of projection $\pi : R \to L$ to sets of labels $S \subseteq R$ by $\pi(S) := \{l \in L \mid \exists r \in S, \pi(r) = l\}$.

Definition 2.7. A LABEL-COVER instance is (J,ξ)-smooth if for any vertex $v \in V$ and any set of labels $S \subset R$, $|S| \le J$, over a uniformly at random neighbor u of v, we have

$$\Pr_{u\sim v}\left[|\pi_e(S)| < |S|\right] \le \xi . \tag{2.2}$$

Similarly, a k-layered LABEL-COVER instance is (J,ξ)-smooth if for any integer $1 < i \le k$, vertex tuple $\vec{u} = (u_1, \cdots, u_{k-1}) \in U^{k-i} \times V^{i-1}$, and set of labelings $S \subseteq L^{k-i} \times R^{i-1}$ with $|S| \le J$, we have

$$\Pr_{\vec{e}\sim\vec{u}}\left[|\pi_{\vec{e},i\to 1}(S)| < |S|\right] \le \xi .$$

Observe that $|\pi_e(S)| < |S|$ if and only if there exists $r \neq r' \in S$ such that $\pi_e(r) = \pi_e(r')$. By simple union bound over all possible pairs of labelings in S, we can show that for constant J, the above two notions of smoothness differs only by a constant factor. The same argument applies to multi-layered instances.

Lemma 2.8. *A ξ-smooth k-layered* LABEL-COVER *instance is $\left(J, \binom{J}{2}\xi\right)$-smooth.*

Combining all we have, we get the following hardness result for k-layered SMOOTH-LABEL-COVER problem.

Theorem 2.9. *For every constant $\eta, J, \xi, k > 0$, there is some constant $G(\eta, J, \xi, k) < \infty$ such that given a (J, ξ)-smooth k-layered* LABEL-COVER *instance with $|R| \geq G(\eta, J, \xi, k)$, it is* NP*-hard to distinguish between the following two cases:*

YES: *There exist assignments $\sigma_m : U^{k-m} \times V^{m-1} \to L^{k-m} \times R^{m-1} (1 \leq m \leq k)$, such that for all $\vec{e} = (e_1, \ldots, e_{k-1}) \in E^{k-1}$ and all $i, j, 1 \leq i < j \leq k$, it holds that*

$$\pi_{\vec{e}, j \to i}(\sigma_j(u_1, \ldots, u_{k-j}, v_{k-j+1}, \ldots, v_{k-1})) = \sigma_i(u_1, \ldots, u_{k-i}, v_{k-i+1}, \ldots, v_{k-1}).$$

NO: *There are no two integers l and h $(1 \leq l < h \leq k)$ such that there exist functions $P_h : U^{k-h} \times V^{h-1} \to L^{k-h} \times R^{h-1}$ and $P_l : U^{k-l} \times V^{l-1} \to L^{k-l} \times R^{l-1}$, such that for more than η fraction of $(e_1, \ldots, e_{k-1}) \in E^{k-1}$, we have*

$$\pi_{\vec{e}, l \to 1}(P_l(u_1, \ldots, u_{k-l}, v_{k-l+1}, \ldots, v_{k-1})) = \pi_{\vec{e}, h \to 1}(P_h(u_1, \ldots, u_{k-h}, v_{k-h+1}, \ldots, v_{k-1})). \quad (2.3)$$

If an edge tuple $(e_1, \ldots, e_{k-1})$ satisfies the above condition, we say that it is weakly satisfied*.*

Proof. The proof is similar to [9]. The completeness case is straightforward. For soundness, suppose there exists $1 \leq l < h \leq k$ and functions P_l, P_h, such that (2.3) holds for more than η fraction of $(e_1, \ldots, e_{k-1}) \sim E^{k-1}$. Pick any coordinate $i \in \{k-h+1, \ldots, k-l\}$. Then there is a way to fix edges $e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_{k-1}$, such that (2.3) holds for at least η fraction of the edges e_i. We conclude the proof by noting that the restriction of P_l and P_h on the i-th coordinate gives a labeling with value at least η for the original LABEL-COVER instance. □

2.2 PCP reductions

In this section, we describe a reduction from LABEL-COVER which is now the standard technique in hardness of approximation. We also discuss direct sums of PCPs introduced by Chan [5].

Consider a Boolean predicate P of arity w. The reduction typically translates labelings for $u \in U$ and $v \in V$ to $2^{|L|}$ and $2^{|R|}$ Boolean variables, respectively. These variables are viewed as functions

$$f^u : \{-1,1\}^{|L|} \to \{-1,1\} \qquad \text{and} \qquad g^v : \{-1,1\}^{|R|} \to \{-1,1\}.$$

We require that these functions be folded, that is, for any $x \in \{-1,1\}^{|L|}$, $y \in \{-1,1\}^{|R|}$,

$$f^u(-x) = -f^u(x) \qquad \text{and} \qquad g^v(-y) = -g^v(y).$$

For each pair of queries $(x, -x)$, we select one of them. If x is selected, then when $f(-x)$ is needed we return $-f(x)$ instead. Hence in the actual reduction we only use $2^{|L|-1}$ Boolean variables for each $u \in U$ and $2^{|R|-1}$ variables for each $v \in V$. This is also why we need to allow negated literals in the CSP instances. In a correct proof for a satisfiable LABEL-COVER instance, the functions are long codes for the corresponding labelings of u and v, that is, setting

$$f^u(x) = x_{\sigma_U(u)} \qquad \text{and} \qquad g^v(y) = y_{\sigma_V(v)}\,.$$

For an edge $\{u, v\}$ in the LABEL-COVER, we sample *queries*

$$\left(x^{(1)}, \ldots, x^{(m)}, y^{(m+1)}, \ldots, y^{(w)}\right)$$

according to some carefully chosen *test distribution* $\mathcal{T}$. The distribution $\mathcal{T}$ has the property that for any $l \in L$ and $r \in R$ such that $\pi_{(u,v)}(r) = l$, the predicate P accepts

$$\left(x_l^{(1)}, \ldots, x_l^{(m)}, y_r^{(m+1)}, \ldots, y_r^{(w)}\right)$$

with probability 1 (or $1 - \varepsilon$ for some small constant ε if we are considering non-perfect completeness).

Let the value of an edge be the following expectation

$$\mathop{\mathbb{E}}_{(x^{(1)},\ldots,x^{(m)},y^{(m+1)},\ldots,y^{(w)})\sim\mathcal{T}}\left[P(f^u(x^{(1)}), \ldots, f^u(x^{(m)}), g^v(y^{(m+1)}), \ldots, g^v(y^{(w)})\right]. \tag{2.4}$$

Observe that in the completeness case where the LABEL-COVER instance has an assignment that satisfies all the edges, setting f^u and g^v to the long code of the labelings would give value 1 (or close to 1) for the above expectation.

In the soundness case, of course the functions f^u and g^v are not guaranteed to be long codes. Typically, when proving approximation resistance, we start the analysis by taking the Fourier expansion of predicate P in (2.4). The constant term in the expansion is exactly the density of P. We then show that if for some non-constant terms we have that $|\mathbb{E}[\prod f^u \prod g^v]| \geq \delta$ for some small constant $\delta > 0$, then we can find a *good* labeling for the LABEL-COVER instance we started with, allowing us to distinguish between the YES case and the NO case. In some cases we can show that all possible non-constant terms—including those that do not appear in the expansion of P—are small, and this implies that predicate P is useless in the sense of [3], a stronger notion of inapproximability.

It is not hard to adapt the above reduction to k-layered LABEL-COVERS. Instead of encoding the labelings of single vertices as long codes, we encode labelings for the hybrid vertex tuples. The rest of the analysis is similar.

In [5], Chan introduced direct sum of PCPs to get improved hardness of approximation results and proved the first general criterion for approximation resistant predicate without assuming the UGC. The main result is the following.

Theorem 2.10 (Siu On Chan). *Let $k \geq 3$ be an integer, G a finite abelian group, and C a balanced pairwise independent subgroup of G^k. It is* NP*-hard to approximate Additive-*CSP*(C) better than $|C|/|G|^k + \varepsilon$ for any constant $\varepsilon > 0$.*

In the Boolean case, we have $G = \{-1,1\}$ with product "$\cdot$" as the group operation, C consists of the accepting assignments of some predicate P, and Additive-CSP(C) is exactly MAX-P. Note that the balanced pairwise independence condition is similar to the condition of Austrin and Mossel [4].

The main idea in Chan's proof is that instead of sampling one edge, we sample c edges for some c to be determined. We also have c test distributions $\mathcal{T}_1, \ldots, \mathcal{T}_c$ corresponding to the edges we sampled, where each distribution $\mathcal{T}_i$ satisfies the same requirement as we had for $\mathcal{T}$ above. We sample queries

$$\left\{\left(x^{(1,i)},\ldots,x^{(k,i)}\right)\right\}_{i\in[c]}.$$

Note that for the same $m \in [k]$, whether the query $x^{(m,i)}$ is from $\{-1,1\}^{|L|}$ or $\{-1,1\}^{|R|}$ may vary among different $i \in [c]$ and is an important design choice. Depending on the edges we have sampled, we choose k functions to query as in the classical setting (some of those functions might be the same one). The functions now have larger domains, since the j-th function would take $\{x^{(j,i)}\}_{i\in[c]}$ as input. We also require that the functions are folded in each individual test—for any query $\{x^{(j,i)}\}_{i\in[c]}$ and $i \in [c]$ we have

$$f\left(x^{(j,1)},\ldots,-x^{(j,i)},\ldots,x^{(j,c)}\right) \;=\; -f\left(x^{(j,1)},\ldots,x^{(j,i)},\ldots,x^{(j,c)}\right).$$

The intended solution now is that each function is the product of functions for the individual PCPs, and the functions in the individual PCPs are long codes of some legitimate labelings. Similar to the classical approach, we take the answers to the queries and accept if predicate P accepts.

It is not hard to see that for the completeness to hold, we would need that the set of satisfying assignments of P has some group structure—the elementwise product of two satisfying assignments is still a satisfying assignment. Observe that the HADAMARD$_K$ predicate satisfies this property. On the soundness side, we need to bound each term in the Fourier expansion of (2.4). The important observation (Lemma 5.3 in [5]) is that the absolute value of these terms are bounded by the absolute value of these terms in each individual PCP. Hence, all we need is to show that a term in (2.4) is small in at least one of the PCPs in the direct sum unless there is good labeling.

2.3 Efron-Stein decomposition, influence and correlation

In this section, we recall basic notions from Fourier analysis, influence, the Bonami-Beckner operator, and correlation of correlated probability spaces.

Let (Ω,μ) be a finite probability space with $|\Omega| = q$. We assume that $\mu(x) > 0$ for every $x \in \Omega$. Let $\chi_0,\ldots,\chi_{q-1} : \Omega \to \mathbb{R}$ be an orthonormal basis for $L^2(\Omega,\mu)$ with respect to scalar product under μ. Let this basis be such that $\chi_0 = \mathbf{1}$ the identical one function. For $\sigma \in \mathbb{Z}_q^n$, define

$$\chi_\sigma(x_1,\ldots,x_n) = \prod_{i=1}^{n} \chi_{\sigma_i}(x_i).$$

Then $\{\chi_\sigma\}_{\sigma\in\mathbb{Z}_q^n}$ forms an orthonormal basis for $L^2(\Omega^n,\mu^{\otimes n})$, and every function $f \in L^2(\Omega^n,\mu^{\otimes n})$ can be written as

$$f(x) = \sum_{\sigma\in\mathbb{Z}_q^n} \hat{f}(\sigma)\chi_\sigma(x).$$

We also make extensive use of the following Efron-Stein decomposition [8, 23]

Theorem 2.11. *Any function $f \in L^2(\Omega^n, \mu^{\otimes n})$ can be uniquely decomposed as*

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

where

- *the function $f_S(x)$ depends only on $x_S = \{x_i \mid i \in S\}$;*
- *for every $S, T \subseteq [n]$, $S \setminus T \neq \emptyset$, $x' \in \Omega^n$, it holds that*

$$\mathbb{E}[f_S(x) \mid x_T = x'_T] = 0.$$

For $\sigma \in \mathbb{Z}_q^n$, let $\mathrm{Set}(\sigma) = \{i \mid \sigma_i > 0\}$, and let $|\sigma| = |\mathrm{Set}(\sigma)|$. It is easily verified that the Efron-Stein decomposition is related to the Fourier decomposition as follows

$$f_S(x) = \sum_{\substack{\sigma \in \mathbb{Z}_q^n \\ \mathrm{Set}(\sigma) = S}} \hat{f}(\sigma) \chi_\sigma(x).$$

A useful notion of function is the influence of a coordinate.

Definition 2.12. For $f \in L^2(\Omega^n, \mu^{\otimes n})$, $i \in [n]$, the influence of i on f is defined as

$$\mathrm{Inf}_i(f) = \mathop{\mathbb{E}}_{x_{[n] \setminus i}} [\mathop{\mathbf{Var}}_{x_i}[f(x)]].$$

Note that when we refer to influence, it is always with respect to the underlying probability space $(\Omega^n, \mu^{\otimes n})$. We have the following characterization of influence in terms of Fourier decomposition and Efron-Stein decomposition.

Proposition 2.13. *For $f \in L^2(\Omega^n, \mu^{\otimes n})$ and $i \in [n]$,*

$$\mathrm{Inf}_i(f) = \sum_{\substack{\sigma \in \mathbb{Z}_q^n \\ i \in \mathrm{Set}(\sigma)}} \hat{f}(\sigma)^2 = \sum_{S \ni i} \mathbb{E}[f_S^2].$$

Let the total influence $\mathrm{Inf}(f) = \sum_{i \in [n]} \mathrm{Inf}_i(f)$ *be the sum of influences of all coordinates on f.*

We next recall the Bonami-Beckner operator, or noise operator.

Definition 2.14. Let $0 \leq \gamma \leq 1$. The Bonami-Beckner operator T_γ is a linear operator mapping $f \in L^2(\Omega^n, \mu^{\otimes n})$ to $T_\gamma f$ as follows

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

where y is sampled by setting each bit independently to $y_i = x_i$ with probability $1 - \gamma$, and otherwise sampled according to μ with probability γ.

Again we have the following Fourier/Efron-Stein characterization of T_γ.

Proposition 2.15. *For any* $f \in L^2(\Omega^n, \mu^{\otimes n})$ *and* $0 \le \gamma \le 1$,

$$T_\gamma f = \sum_{\sigma \in \mathbb{Z}_q^n} (1-\gamma)^{|\sigma|} \hat{f}(\sigma) \chi_\sigma .$$

We define noisy influence as $\mathrm{Inf}_i^{(\gamma)}(f) = \mathrm{Inf}_i(T_\gamma f)$, and similarly $\mathrm{Inf}^{(\gamma)}(f) = \sum \mathrm{Inf}_i^{(\gamma)}(f)$. The following bound for the total noisy influence of functions with range $[-1,1]$ appears in [26, Lemma 5.9].

Proposition 2.16 (O'Donnell, Wu)**.** *For any* $f : \Omega^n \to [-1,1]$ *and* $0 < \gamma \le 1$*, we have*

$$\mathrm{Inf}^{(\gamma)}(f) = \sum_{i \in [n]} \mathrm{Inf}_i^{(\gamma)}(f) \le \gamma^{-1} .$$

The following concept of *lifted* functions is useful in the context of projection games.

Definition 2.17. Given function $f : \Omega^{nd} \to \mathbb{R}$ and d-to-1 mapping $\pi : R \to L$, define the lifted version of f as $\overline{f}^{\pi} : (\Omega^d)^n \to \mathbb{R}$ as naturally induced by π

$$\overline{f}^{\pi}(\overline{x}) = f(x) ,$$

where $\overline{x}$ satisfies $\overline{x}_{r,t} = x_{(r,t)}$ for $r \in L, t \in [d]$.

In terms of influence, we have the following relation between f and $\overline{f}$, due to Wenner [32].

Proposition 2.18 (Wenner)**.** *For any r, we have*

$$\mathrm{Inf}_r(\overline{f}) \le \sum_{r' : \pi(r') = r} \mathrm{Inf}_{r'}(f) .$$

Proof. The claim follows by applying Proposition 2.13 and comparing the terms. □

The correlation for correlated probability spaces was introduced by Mossel [23]. Given a probability measure μ defined on $\Omega \times \Psi$, we say that Ω and Ψ are correlated spaces, and we use $(\Omega \times \Psi, \mu)$ to denote correlated spaces and the corresponding measure. We use the following definition of correlation.

Definition 2.19. Let $(\Omega \times \Psi, \mu)$ be a correlated probability space, μ is a distribution on the finite product set $\Omega \times \Psi$ and that the marginals of μ on Ω and Ψ have full support. Define the correlation between Ω and Ψ to be

$$\rho(\Omega, \Psi; \mu) = \max_{\substack{f : \Omega \to \mathbb{R} \\ g : \Psi \to \mathbb{R}}} \left\{ |\mathbb{E}[fg]| \;\middle|\; \mathbb{E}[f] = 0, \mathbb{E}[f^2] \le 1, \mathbb{E}[g] = 0, \mathbb{E}[g^2] \le 1 \right\} ,$$

where the expectation $\mathbb{E}[fg]$ is under μ, and $\mathbb{E}[f]$, $\mathbb{E}[f^2]$, $\mathbb{E}[g]$ and $\mathbb{E}[g^2]$ are under marginals of μ on corresponding spaces.

A useful fact for bounding correlation of probability spaces from [23] is that the correlation of a product of correlated probability space is equal to the maximum correlation among the individual correlated spaces (excluding empty components).

Lemma 2.20. *Let $\{(\Omega_i \times \Psi, \mu_i)\}$ be a set of correlated probability spaces, then*

$$\rho\left(\prod_i \Omega_i, \prod_i \Psi_i; \prod_i \mu_i\right) \leq \max_i \rho(\Omega_i, \Psi_i; \mu_i).$$

We also need the following lemma when analyzing correlations. Intuitively, if we can decompose μ into a convex combination of two distributions and we can bound the correlation between Ω and Ψ in both sub-distributions by some constant c, then barring special cases it seems reasonable that the correlation $\rho(\Omega, \Psi; \mu)$ should also be bounded by some function of c. More formally, we have the following lemma from [32].

Lemma 2.21. *Let $(\Omega \times \Psi, \delta\nu + (1-\delta)\nu')$ be a correlated space such that the marginal distribution of at least one of Ω and Ψ is identical on both ν and ν'. Then*

$$\rho(\Omega, \Psi; \delta\nu + (1-\delta)\nu') \leq \sqrt{\delta\rho(\Omega, \Psi; \nu)^2 + (1-\delta)\rho(\Omega, \Psi; \nu')^2}.$$

Next we recall the definition of the conditional expectation operator.

Definition 2.22. Let $(\Omega \times \Psi, \mu)$ be two correlated spaces. The conditional expectation operator $\mathcal{U}$ associated with (Ω, Ψ) is the operator mapping $f \in L^2(\Psi, \mu)$ to $Uf \in L^2(\Omega, \mu)$ by

$$(\mathcal{U}f)(x) = \mathbb{E}[f(Y) \mid X = x]$$

for $x \in \Omega$ and $(X, Y) \in \Omega \times \Psi$ is distributed according to μ.

An important property we need in the analysis, due to Mossel [22], is that the Efron-Stein decomposition commutes with the conditional expectation operator.

Proposition 2.23 (Mossel)**.** *Let $(\Omega \times \Psi, \mu) := (\prod \Omega_i \times \prod \Psi_i, \bigotimes \mu_i)$ be correlated space and let $\mathcal{U} := \bigotimes \mathcal{U}_i$ be the conditional expectation operator associated with Ω and Ψ. Suppose $f \in L^2(\Psi)$ has Efron-Stein decomposition $f(x) = \sum_{S \subseteq [n]} f_S(x_S)$. Then the Efron-Stein decomposition of $\mathcal{U}f$ satisfies $(\mathcal{U}f)_S = \mathcal{U}(f_S)$ for $S \subseteq [n]$.*

The following result, due to Mossel [22], shows that in the above setting, if the correlations between all Ω and Ψ are less than 1, then the L^2 norms of the high-degree terms of $\mathcal{U}f$ are small.

Proposition 2.24 (Mossel)**.** *Assume the setting of Proposition 2.23 and that for all i, we have*

$$\rho(\Omega_i, \Psi_i; \mu_i) \leq \rho_i.$$

Then for all f, we have

$$\|\mathcal{U}(f_S)\|_2 \leq \left(\prod_{i \in S} \rho_i\right) \|f_S\|_2.$$

3 The predicate, the PCP and outline of proof

Given the soundness parameter ε, the starting point of our reduction is a k-layered (J,ξ)-smooth LABEL-COVER, where J and ξ are constants solely dependent on ε that we will specify later.

The predicate. Fix some k, and let $[k] := \{1,2,\ldots,k\}$. Let

$$\mathscr{S}_3 := \{S \subseteq [k] \mid |S| = 3\} \qquad \text{and} \qquad \mathscr{S}_1 := \{S \subseteq [k] \mid |S| = 1\}.$$

The predicate is on variables $\{x^{(S)}\}_{S\in\mathscr{S}_1\cup\mathscr{S}_3}$ taking values from $\{-1,1\}$. We call the variables $x^{(\{i\})}$ singleton variables and the remaining ones parity check variables. The predicate accepts if there exists $\vec{w} \in \{-1,1\}^{\mathscr{S}_1\cup\mathscr{S}_3}$ such that the number of -1 entries in $\vec{w}$ is no more than k, and

$$w_S x^{(S)} \cdot \prod_{i\in S} w_{\{i\}} x^{(\{i\})} = 1$$

for all $S \in \mathscr{S}_3$.

We can view $\vec{w}$ as an error vector, and the predicate accepts inputs that are no more than Hamming distance k away from an assignment that satisfies all parity checks.

The predicate is on $k + \binom{k}{3}$ variables, and it has

$$O\left(2^k \cdot \binom{\binom{k}{3}+k}{k}\right) = 2^{O(k\log k)}$$

accepting inputs, thus the density (assuming the predicate has arity K) is $2^{\widetilde{O}(K^{1/3})}/2^K$, where the $\widetilde{O}$ hides logarithmic factors.

Outline of proof. Before going into details about the construction of our PCPs, we first give an overview of our proof and explain the intuition behind the construction.

Our PCP design is based on Chan's idea of direct sums of PCPs [5] as described in Section 2.2. We prove that all non-constant terms in the Fourier expansion of (2.4) are small.

One crucial difference between Chan's proof and ours is that we require perfect completeness. This means that sometimes there would be perfect correlation between certain queries which makes it possible for provers to find good cheating strategies. In Chan's proof as well as in many related results where perfect completeness is not required, one can usually break this correlation by applying some independent noise to each query bit. However, in the case of perfect completeness, we cannot afford perturbing each bit independently, and thus we need to take extra care when designing test distributions. That is the main reason our predicate accepts inputs that *almost* satisfy all $\binom{k}{3}$ linear constraints. In some sense, these extra accepting inputs serve as noise that breaks up perfect correlations.

Another important property that Chan uses is the "group" structure of the predicate. This makes it relatively easy to take direct sums of a large number of PCPs, each handling a small number of non-constant terms from (2.4), without worrying too much about the completeness of the resulting PCP. Our predicate, however, does not satisfy this property due to the extra noise we added. It is certainly possible that if we take the sum of two assignments that are of distance k away from assignments that satisfies all linear equations, we end up with something that is distance $2k$ away from an assignment that satisfies all linear constraints, and that would break perfect completeness. To avoid this situation, we limit both the number of PCPs in the direct sum and in each PCP the distance from an assignment that satisfies all linear constraints. More specifically, in our construction the queries to each PCP are generated such that if the provers (of each individual PCP) answer according to some consistent long code, then the answers is at most distance 1 away from an assignment that satisfies all linear equations. When taking direct sum of the k PCPs, an answer that is the direct sum of k long codes would give us an answer that would be accepted by our predicate.

It remains to find a number of suitable PCPs. If we try to generalize previous approaches, for example those in [28, 9], to larger predicates such as HADAMARD_K, one of the main adversarial strategies that we need to consider is that of inconsistent long codes. For example, consider a predicate P on variable $(x_1,\ldots,x_k)$ and a simple PCP reduction where we sample an edge $\{u,v\}$ and query functions f^u and g^v according to some test distribution $\mathcal{T}$ as described in Section 2.2. For simplicity, assume that the query to f^u corresponds to input variable x_1, and the remaining queries are on g^v. Suppose further that for a $1/2+\delta$ fraction of the accepting inputs of P, we have $x_2x_3x_4 = 1$ (both HADAMARD_K and the predicate we are studying in this work have properties similar to this.) Let g^v be long code for some arbitrary label $r \in R$. Observe that the non-constant term $g^v(x_2)g^v(x_3)g^v(x_4)$ will always have expectation roughly order of δ simply due to the requirements on $\mathcal{T}$. In this case, we get a large non-constant term but it does not help us find a consistent labeling for LABEL-COVER. A similar argument can be made for MULTI-LAYERED-LABEL-COVER. Chan's construction in [5] solves this problem by making sure that for each term, in at least one of the many PCPs in the direct sum the queries are on different functions. As discussed before, since we are aiming for fewer PCPs in the direct sum, it would be good if each PCP can carry out as many consistency checks as possible, and MULTI-LAYERED-LABEL-COVER becomes a very natural choice. We also need to decide which query should be in which layer for each PCP so that we do not miss any sets of variables that has linear relations. This is mostly done in Section 4.1.

Now we describe the PCPs in more details.

The PCPs. Let $\mathcal{C} = \{\sigma_0,\ldots,\sigma_{k-1}\}$ be the set of cyclic permutations on $[k]$. The permutation σ_i maps i to k, $i+1$ to 1, and so on. We identify 0 with k, and thus σ_0 is the identity permutation. Each permutation corresponds to a PCP for a k-layered LABEL-COVER instance, and the permutation decides which query should be in which layer in the MULTI-LAYERED-LABEL-COVER. As stated above, the final proof is the direct sum of these k PCPs.

We now describe the i-th PCP. It is based on a k-layered LABEL-COVER instance, and there are $k+\binom{k}{3}$ queries, one corresponding to an input variable. We denote the queries as $x^{(S)}$. For $S \in \mathscr{S}_1 \cup \mathscr{S}_3$, define $m_i(S) := \max \sigma_i(S)$ to be the maximum element of S under permutation σ_i. The query $x^{(S)}$ is in layer $m_i(S)$. Let

$$\mathcal{V}_i(S) := U^{k-m_i(S)} \times V^{m_i(S)-1}$$

be the set of vertex tuples in layer $m_i(S)$. The proof has a function for each vertex tuple in $\mathcal{V}_i(S)$, and the input to the functions are $\{-1,1\}$ strings indexed by the possible labelings $L^{k-m_i(S)} \times R^{m_i(S)-1}$ in layer $m_i(S)$. We denote the domain of the functions as $X_i^{(S)}$. In a correct proof of a correct labeling, the function would be a long code encoding a proper labeling for all vertices in the tuple. As described in Section 2.2, we require that all functions are folded.

The test distributions. We first define the test distributions for each individual PCP.

Fix $i \in [k]$ and consider the i-th PCP. For notational simplicity we omit i in the subscript for now. We first independently sample $k-1$ edges $\vec{e} = \{e_1,\ldots,e_{k-1}\}$. For $S \in \mathscr{S}_1$, sample $x^{(S)} \in X^{(S)}$ uniformly at random. For $S = \{s_1, s_2, s_3\} \in \mathscr{S}_3$, let $m = m(S)$ be the layer in which query $x^{(S)}$ is located, $m_j = m(s_j)$ for $j = 1, 2, 3$ be the layer query $x^{(\{s_j\})}$ is in, and set

$$x_r^{(S)} = \prod_{j=1}^{3} x^{(\{s_j\})}_{\pi_{\vec{e},m\to m_j}(r)}$$

for all possible labelings $r \in L^{k-m} \times R^{m-1}$.

We then make use of the extra inputs allowed by the predicate to add some "noise" to the distributions. As discussed above, the resulting distribution must have the property that the output obtained by applying some consistent long codes is at most distance 1 away from an assignment that satisfies all $\binom{k}{3}$ equations. The idea is to perturb one of the variables $x^{(S)}$. For each $r \in L^{k-1}$, pick a uniformly random set $N_r \in \mathscr{S}_1 \cup \mathscr{S}_3$, and for each

$$t \in \pi^{-1}_{\vec{e},m(N_r)\to 1}(r)\,,$$

set $x_t^{(N_r)}$ to a uniform random bit independently with probability $1/2$.

We denote the test distribution by $\mathcal{T}$. For each $r \in L^{k-1}$, let $\mathcal{T}_r$ be the marginal distribution of the bits that map to r under $\pi_{\vec{e},l\to 1}$ for all $l \in [k]$. Observe that we have $\mathcal{T} = \bigotimes_{r\in L^{k-1}} \mathcal{T}_r$.

Let us start by analyzing the standard completeness case.

Lemma 3.1. *For any sampling of edges, let $f^{(S)}$ be the functions we are querying, and let $x^{(S)}$ be the corresponding queries. If the k-layered* LABEL-COVER *instance has a labeling that satisfies all the edges, then we can find functions $f^{(S)}$ such that the answers*

$$\left\{f^{(S)}(x^{(S)})\right\}_{S\in\mathscr{S}_1\cup\mathscr{S}_3}$$

is at most Hamming distance 1 away from an assignment that satisfies all linear constraints on 3 *singleton variables and* 1 *parity check variable.*

Proof. The argument is similar to a standard completeness argument.

Fix a labeling that satisfies all the edges. The proof in the PCP consists of long codes encoding the labeling of all hybrid vertex tuples.

Let $r \in L^{k-1}$ be the labeling for the vertex tuple in layer 1. The answers we get from the long codes is the same as returning one bit from each query generated according to $\mathcal{T}_r$. The claim follows by observing that for each tuple of bits produced as above, either it already satisfies all linear constraints, or it would satisfy all linear constraints after we flip the N_r-th bit. □

Denote the test distribution of the i-th PCP defined above as $\mathcal{T}_i$. The distribution of the final composed PCP is simply the product of the individual test distributions $\bigotimes_{i=1}^{k} \mathcal{T}_i$. The verifier samples the edges and the inputs to the functions, queries the functions (those that correspond to the chosen vertex tuples) and accepts if the answers returned by the functions are accepted by the predicate.

It is not hard to see from above discussions that the above PCP has perfect completeness.

Lemma 3.2. *If the k-layered* LABEL-COVER *instance has a labeling that satisfies all edges, then there exists a set of functions $\{f^{(S)}\}$ such that after querying $\{f^{(S)}\}$ the verifier accepts with probability 1.*

Proof. We let our final proof be the product of proofs of the k individual PCPs given by Lemma 3.1. Since the answer for each proof is at most distance 1 away from an assignment that satisfies all linear constraints, their product is at most distance k away, which is exactly what the verifier (and our predicate) accepts. □

Remark 3.3. We briefly discuss some of the difficulties in getting hardness result better than $2^{\widetilde{O}(K^{1/3})}/2^K$. The key issue here is not unlike the one discussed in [9]. As we will see in Section 4, our construction requires us to identify a permutation cover that covers all the queries, otherwise products of a random set of dictatorship functions would be a good cheating strategy.

To be more specific, one possibility is to consider a predicate on $k+\binom{k}{4}$ variables which does "parity checks" on tuples of 4 variables and accepts everything that has less than t errors for some t, and devise a protocol which is a direct sum of t PCPs. As we can see from Section 4.1, in this case we actually need $\Omega(k^2)$ permutations to cover all queries which means that $t=\Omega(k^2)$. Such a predicate already has much more accepting inputs than the one we study here.

4 Soundness

In this section, we analyze the soundness of our PCP. We set

$$\varepsilon_1 = \varepsilon/(7k^3+1)\,,$$
$$\xi = \varepsilon_1^2\,,$$
$$\rho_0 = 1-1/4\binom{k}{3}\,,$$
$$J = 2\lceil \log_{\rho_0}\varepsilon_1\rceil\,,$$

and γ such that $1-(1-\gamma)^{J/2}<\varepsilon_1$. Note that this gives $\rho_0^{J/2}\le\varepsilon_1$, and that all parameters depend only on k and ε. Also $\gamma<\varepsilon$.

As discussed in Section 3, we would like to prove that for all $\mathscr{S}\neq\emptyset$, the expectation

$$\mathbb{E}\left[\prod_{S\in\mathscr{S}} f^{(S)}(x^{(S)})\right] \tag{4.1}$$

is small unless there is good labeling.

Remark 4.1. Since we are able to bound (4.1) for any $\mathscr{S}\neq\emptyset$, we actually proved that our predicate is useless in the sense of [3].

Remark 4.2. The functions $f^{(S)}$ actually depend on the underlying edges we sampled. For notational convenience we suppress this dependency and save another layer of subscripts (of subscripts of subscripts).

As discussed in previous sections, we need to show that for each non-constant term, there is at least one PCP among those in the direct sum, such that if the expectation of the term under the PCP is large, we can find a good labeling for the underlying LABEL-COVER instance by looking at the functions f restricted to that PCP. Formally, we have the following lemma which is a reformulation of Lemma 5.3 in Chan [5].

Lemma 4.3. *Let $\mathcal{T} = \bigotimes_{i=1}^{k} \mathcal{T}_i$, where $\mathcal{T}_i$ is the test distribution for the i-th PCP. Suppose for some $\mathscr{S} \neq \emptyset$, we have*

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[\prod_{S \in \mathscr{S}} f^{(S)}(x^{(S)}) \right] \right| = \delta\,,$$

then for any $i \in [k]$, there exists functions $g^{(S)}$ whose inputs are query bits to the i-th PCP, such that

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}_i} \left[\prod_{S \in \mathscr{S}} g^{(S)}(x^{(S)}) \right] \right| \geq \delta\,.$$

Given $f^{(S)}$, we find $g^{(S)}$ by fixing query bits that are not in the i-th PCP in a way that does not lower the expectation.

Thus to bound each term, we need to carefully find an i, such that the test restricted to the i-th PCP has small expectation. We show how to choose such i in Section 4.1. We would be back to the traditional setting with LABEL-COVERs and dictatorship testing from then on. In Section 4.2, we show that we can instead look at the distribution where each individual bit is further perturbed independently by some random noise. Then we show in Section 4.3 how to apply an invariance-type theorem from [32] in this new setting to get our soundness result.

4.1 Permutation covering

Our k PCPs use cyclic permutations $C \in \mathcal{C}$ to decide the layer of each query and the inputs to the corresponding function. We first give a general definition of the crucial property we need from such sets of permutations.

Definition 4.4. Let $\mathcal{P}$ be a set of permutations on $[k]$. We say that $\mathcal{P}$ covers $\mathscr{S}_1 \cup \mathscr{S}_3$ if for all $\emptyset \neq \mathscr{S} \subseteq \mathscr{S}_1 \cup \mathscr{S}_3$, there exists a permutation $\sigma \in \mathcal{P}$, some $j, l_0 \in [k]$, such that

$$\Big| \{ S \in \mathscr{S} \mid j \in S, \max \sigma(S) = l_0 \} \Big| \text{ is odd}\,.$$

We now reformulate the above definition and prove a necessary and sufficient condition for general sets of permutations $\mathcal{P}$ to cover $\mathscr{S}_1 \cup \mathscr{S}_3$.

For each set $S \in \mathscr{S}_1 \cup \mathscr{S}_3$, we construct a Boolean vector $v_S^{\mathcal{P}}$ as the following: the elements in the vector are indexed by a tuple $(i,l,j) \in [|\mathcal{P}|] \times [k] \times [k]$, and $v_{S,(i,l,j)}^{\mathcal{P}} = 1$ if $\max \sigma_i(S) = l$ and $j \in S$, and $v_{S,(i,l,j)} = 0$ otherwise.

Proposition 4.5. *The set of permutations $\mathcal{P}$ covers $\mathscr{S}_1 \cup \mathscr{S}_3$ iff the vectors $\{v_S^{\mathcal{P}}\}_{S \in \mathscr{S}_1 \cup \mathscr{S}_3}$ are linearly independent over $\mathbb{F}_2$.*

Proof. If the set $\mathcal{P}$ does not cover $\mathscr{S}_1 \cup \mathscr{S}_3$, then there exists a set $\emptyset \neq \mathscr{S} \subseteq \mathscr{S}_1 \cup \mathscr{S}_3$, such that for any permutation $\sigma_i \in \mathcal{P}$ and $j, l_0 \in [k]$, we have that

$$\left| \{S \in \mathscr{S} \mid S \ni j, \max \sigma_i(S) = l_0\} \right| \text{ is even}.$$

Observe that for any $S \in \mathscr{S}_1 \cup \mathscr{S}_3$, the segment of $v_S^{\mathcal{P}}$ indexed by (i,l) for some fixed i and l is all zero if $\max \sigma_i(S) \neq l$, and otherwise it is exactly the character vector of the set S. Therefore the above is equivalent to saying that for any $i \in [|\mathcal{P}|]$ and l_0, we have

$$\sum_{S \in \mathscr{S}} v^{\mathcal{P}}_{S,(i,l_0)} = 0,$$

where the summation is modulo 2. Since the above holds for all i and l_0, we have

$$\sum_{S \in \mathscr{S}} v_S^{\mathcal{P}} = 0,$$

or the vectors $\{v_S^{\mathcal{P}}\}_{S \in \mathscr{S}}$ are linearly dependent.

Note that all the above steps are equivalent statements. Thus the other direction also holds. □

As a side note, we can see from the above argument that it is necessary to have $\Omega(k)$ permutations in order to cover $\mathscr{S}_1 \cup \mathscr{S}_3$, because otherwise we would have $\Theta(k^3)$ vectors of dimension $o(k^3)$ and thus they could not be linearly independent.

We now prove that the set of all cyclic permutations $\mathcal{C} = \{\sigma_0, \ldots, \sigma_{k-1}\}$ covers $\mathscr{S}_1 \cup \mathscr{S}_3$.

Lemma 4.6. *The set of all cyclic permutations $\mathcal{C} = \{\sigma_0, \ldots, \sigma_{k-1}\}$ covers $\mathscr{S}_1 \cup \mathscr{S}_3$.*

Proof. For any given collection of sets $\mathscr{S} \subseteq \mathscr{S}_1 \cup \mathscr{S}_3$, we show how to find the cyclic permutation σ and indices $j, l_0 \in [k]$ required in Definition 4.4.

For a set $S \in \mathscr{S}_1 \cup \mathscr{S}_3$, let

$$\operatorname{span}(S) = \min_{\sigma_i \in \mathcal{C}} \{\max \sigma_i(S) - \min \sigma_i(S)\},$$

that is, the minimum distance between the largest and the smallest element under cyclic permutations. Note that for singleton sets $S \in \mathscr{S}_1$, we have $\operatorname{span}(S) = 0$.

For a given collection of sets $\mathscr{S}$, let $S \in \mathscr{S}$ be a set with minimum span in $\mathscr{S}$ where we break ties arbitrarily. Pick i_0 such that $\sigma_{i_0}(S)$ contains 1 and $\operatorname{span}(S)+1$ as its minimum and maximum element. Let $\sigma := \sigma_{i_0}$ be the permutation we want, and let $l_0 = \operatorname{span}(S)+1$.

Now we select j. If $\operatorname{span}(S) = 0$, then let $j = \sigma^{-1}(1)$ and we are done. This is because for any non-singleton set S', $\max \sigma(S) > 1$, and for any singleton set $S'' \neq S$, clearly $\sigma(S'') \neq \sigma(S)$. Thus S would be the only set containing j with $\max \sigma(S) = 1 = l_0$.

If $\operatorname{span}(S) \neq 0$, then S has three elements, and there is no singleton set in $\mathscr{S}$. If there is any other non-singleton set $S'' \in \mathscr{S}$ with $\max \sigma(S'') = \operatorname{span}(S)+1$, then $\sigma(S'')$ and $\sigma(S)$ have the same maximum and minimum element, namely $\operatorname{span}(S)+1$ and 1. That leaves us with the middle element. But since $S \neq S''$, the middle element must be different, so each of them appear only in one set, and setting j to the inverse of any of the middle elements under σ would work. Otherwise we take $j = \max S$. □

For $\mathscr{S} \subseteq \mathscr{S}_1 \cup \mathscr{S}_3$, we consider the PCP corresponding to the cyclic permutation $\sigma_i \in \mathcal{C}$ covering $\mathscr{S}$ given by Lemma 4.6. We denote the PCP as PCP_i. As discussed before, we only need to show that if (4.1) is large even when restricted to PCP_i, we can find a good labeling for the LABEL-COVER instance we started with.

For notational simplicity, we only prove the case where $i = 0$, that is, for the identity permutation σ_0. Arguments for general cyclic permutations are entirely symmetric.

4.2 Introducing independent noise

In this section, we show that perturbing the queries does not change the expectation of the terms by too much.

Formally, let $\mathcal{T}'_r$ be the distribution where we first sample according to $\mathcal{T}_r$, and then resample each bit independently with probability γ according to its marginal distribution in $\mathcal{T}_r$—which in our case is uniform. Also define $\mathcal{T}' = \bigotimes_{r \in L^{k-1}} \mathcal{T}'_r$. We prove the following lemma which bounds the difference of expectation of (4.1) under $\mathcal{T}$ and $\mathcal{T}'$.

Lemma 4.7. *For any $\mathscr{S} \subseteq \mathscr{S}_1 \cup \mathscr{S}_3$, we have*

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[\prod_{S \in \mathscr{S}} f^{(S)}(x^{(S)}) \right] - \mathop{\mathbb{E}}_{\mathcal{T}'} \left[\prod_{S \in \mathscr{S}} f^{(S)}(x^{(S)}) \right] \right| < 7k^3 \varepsilon_1 \,, \tag{4.2}$$

where $\varepsilon_1 = \varepsilon/(7k^3+1)$ is as defined at the beginning of Section 4.

Fix some $S_0 \in \mathscr{S}_1 \cup \mathscr{S}_3$. Let $\mathcal{T}^{(S_0)}$ be the distribution where under $\mathcal{T}$, we independently resample the bits in $x^{(S_0)}$ from the uniform distribution with probability γ. We first show in Lemma 4.8 below that the expectation under $\mathcal{T}$ is close to that under $\mathcal{T}^{(S_0)}$. Lemma 4.7 follows by applying similar arguments to each $x^{(S)}$ in succession.

For $S \in \mathscr{S}_1 \cup \mathscr{S}_3$, let $m(S) = \max S$ be the maximum element in S. Recall that query $x^{(S)}$ is located in layer $m(S)$, and for $r \in L^{k-1}$, $\mathcal{T}_r$ is the distribution containing all bits in

$$\left\{ x_t^{(S)} \;\middle|\; S \in \mathscr{S}_1 \cup \mathscr{S}_3, \pi_{m(S)\to 1}(t) = r \right\},$$

that is, the query bits that map to the same r. We use $\mathcal{T}_r^{(S_0)}$ to denote the marginal distribution of $\mathcal{T}^{(S_0)}$ on bits in

$$\left\{ x_t^{(S_0)} \;\middle|\; \pi_{m(S_0)\to 1}(t) = r \right\}.$$

Let $m = m(S_0)$.

Consider the difference of expectation between $\mathcal{T}$ and $\mathcal{T}^{(S_0)}$. If $f^{(S_0)}(x^{(S_0)})$ does not appear in the product, then there would be no difference. We now assume otherwise. The following lemma shows that introducing independent noise on one query does not change the expectation by too much.

Lemma 4.8. *For any $\mathscr{S} \subseteq \mathscr{S}_1 \cup \mathscr{S}_3$, we have*

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[\prod_{S \in \mathscr{S}} f^{(S)}(x^{(S)}) \right] - \mathop{\mathbb{E}}_{\mathcal{T}^{(S_0)}} \left[\prod_{S \in \mathscr{S}} f^{(S)}(x^{(S)}) \right] \right| < 7\varepsilon_1 \,. \tag{4.3}$$

In the rest of the section, we prove Lemma 4.8. The proof follows Wenner's approach [32], especially Lemmas 3.15 through 3.17.

For notational simplicity let F' be the product of all terms but $f^{(S_0)}(x^{(S_0)})$ and we abbreviate $f^{(S_0)}$ as f. We use $\overline{X}^{(S_0)}$ to abbreviate

$$\prod_{S \in \mathscr{S}_1 \cup \mathscr{S}_3, S \neq S_0} X^{(S)} .$$

Similarly we define $\overline{X}_r^{(S_0)}$ for $r \in L^{k-1}$. The first step is to use Lemma 2.21 to bound the correlation

$$\rho(X^{(S_0)}, \overline{X}^{(S_0)}; \mathcal{T}) \qquad \text{and} \qquad \rho(X^{(S_0)}, \overline{X}^{(S_0)}; \mathcal{T}^{(S_0)}) .$$

Since $\mathcal{T}$ is simply a product of $\mathcal{T}_r$ with different values r, by Lemma 2.20, we only need to bound

$$\rho(X_r^{(S_0)}, \overline{X}_r^{(S_0)}; \mathcal{T}_r) \qquad \text{and} \qquad \rho(X_r^{(S_0)}, \overline{X}_r^{(S_0)}; \mathcal{T}_r^{(S_0)}) .$$

Claim 4.9. *For any $S_0 \in \mathscr{S}_3$, the correlation $\rho(X_r^{(S_0)}, \overline{X}_r^{(S_0)}; \mathcal{T}_r)$ is upper-bounded by*

$$1 - \frac{1}{4\binom{k}{3}} \triangleq \rho_0 .$$

The same bound holds for $\rho(X_r^{(S_0)}, \overline{X}_r^{(S_0)}; \mathcal{T}_r^{(S_0)})$.

Proof. We divide $\mathcal{T}_r$ into two parts: (i) the set S_0 is chosen as N_r; or (ii) some set other than S_0 is chosen. It is not hard to verify that the marginal of $X_r^{(S_0)}$ after conditioning on either of them remains uniform and thus we can apply Lemma 2.21. Let μ be the conditional distribution assuming (i) happens, and ν be the one assuming (ii) happens. We have that

$$\rho(X_r^{(S_0)}, \overline{X}_r^{(S_0)}; \nu) = 1 .$$

For the correlation of the other part, we have

$$\rho(X_r^{(S_0)}, \overline{X}_r^{(S_0)}; \mu) = 1 ,$$

achieved by dictatorship functions. Therefore, the overall correlation is upper-bounded by

$$\sqrt{(1 - 1/\binom{k}{3}) + 1/\binom{k}{3} \cdot (1/2)^2} \leq \sqrt{1 - 1/2\binom{k}{3}} < 1 - 1/4\binom{k}{3} .$$

Intuitively, the correlation under $\mathcal{T}_r^{(S_0)}$ could not exceed that under $\mathcal{T}_r$ since the noise we added are all independent. In particular, the part corresponding to

$$\rho(X_r^{(S_0)}, \overline{X}_r^{(S_0)}; \nu)$$

becomes less than 1 due to lack of perfect correlation, and the part corresponding to

$$\rho(X_r^{(S_0)}, \overline{X}_r^{(S_0)}; \mu)$$

remains the same. Thus the result follows by similar calculations as in $\mathcal{T}_r$. □

Take the Efron-Stein decomposition $f = \sum_{T \subseteq L^{k-1}} f_T$. More specifically, for $T \subseteq L^{k-1}$, we have that

$$f_T = \sum_{\substack{U \subseteq L^{k-m} \times R^{m-1} \\ \pi_{m\to 1}(U)=T}} \hat{f}_U \chi_U \,.$$

Again for notational simplicity, we temporarily drop the subscript and write $\pi_{m\to 1}$ as π. We decompose the terms in the expectation in (4.3) as following

$$fF' \;=\; F' \sum_{T \subseteq L^{k-1}} f_T \;=\; F' \sum_{\substack{T \subseteq L^{k-1} \\ |T| \le J/2}} f_T + F' \sum_{\substack{T \subseteq L^{k-1} \\ |T| > J/2}} f_T \,. \tag{4.4}$$

We first bound the expectation of the high degree parts under both $\mathcal{T}$ and $\mathcal{T}^{(S_0)}$.

This is a standard correlation argument. We first consider the expectation under $\mathcal{T}$. Let $\mathcal{U}_{\mathcal{T}}$ be the conditional expectation operator mapping a function in $L_2(X^{(S_0)})$ to a function in $L_2(\overline{X}^{(S_0)})$ with respect to distribution $\mathcal{T}$. We have

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[F' \sum_{\substack{T \subseteq L^{k-1} \\ |T| > J/2}} f_T \right] \right| = \left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[F' \sum_{\substack{T \subseteq L^{k-1} \\ |T| > J/2}} \mathcal{U}_{\mathcal{T}} f_T \right] \right| . \tag{4.5}$$

Note that the expectation on the right hand side is in fact taken under the marginals of $\mathcal{T}$ on $\overline{X}^{(S_0)}$. Applying Cauchy-Schwarz, we have

$$\begin{aligned} \left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[F' \sum_{\substack{T \subseteq L^{k-1} \\ |T| > J/2}} f_T^{(S_0)} \right] \right| &\le \sqrt{ \mathop{\mathbb{E}}_{\mathcal{T}} \left[\sum_{\substack{T \subseteq L^{k-1} \\ |T| > J/2}} \mathcal{U}_{\mathcal{T}} (f_T^{(S_0)})^2 \right] } \sqrt{\mathop{\mathbb{E}}_{\mathcal{T}}[F'^2]} & (4.6) \\ &\le \sqrt{ \sum_{\substack{T \subseteq L^{k-1} \\ |T| > J/2}} \left\| \mathcal{U}_{\mathcal{T}} f_T^{(S_0)} \right\|^2 } & (4.7) \\ &\le \sqrt{ \sum_{\substack{T \subseteq L^{k-1} \\ |T| > J/2}} \rho_0^{|T|} \left\| f_T^{(S_0)} \right\|^2 } \le \rho_0^{J/2} \le \varepsilon_1 \,, & (4.8) \end{aligned}$$

where the inequality in (4.8) follows from Proposition 2.24 and that the norm in (4.8) is with respect to the marginal of $\mathcal{T}$ on $X^{(S_0)}$, which is uniform. The analysis for expectation under $\mathcal{T}^{(S_0)}$ is identical as it only involves correlation. Therefore

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[fF' \right] - \mathop{\mathbb{E}}_{\mathcal{T}^{(S_0)}} \left[fF' \right] \right| \le \left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[F' \sum_{\substack{T \subseteq L^{k-1} \\ |T| \le J/2}} f_T \right] - \mathop{\mathbb{E}}_{\mathcal{T}^{(S_0)}} \left[F' \sum_{\substack{T \subseteq L^{k-1} \\ |T| \le J/2}} f_T \right] \right| + 2\varepsilon_1 \,. \tag{4.9}$$

Now we turn to the low degree parts. Further unraveling the Efron-Stein decomposition, we have

$$F' \sum_{\substack{T\subseteq L^{k-1}\\ |T|\le J/2}} f_T \tag{4.10}$$

$$= F' \sum_{\substack{U\subseteq L^{k-m}\times R^{m-1}\\ |\pi(U)|\le J/2}} \hat{f}_U \chi_U \tag{4.11}$$

$$= F' \sum_{\substack{U\subseteq L^{k-m}\times R^{m-1}\\ |U|=|\pi(U)|\\ |\pi(U)|\le J/2}} \hat{f}_U \chi_U + F' \sum_{\substack{U\subseteq L^{k-m}\times R^{m-1}\\ |U|>|\pi(U)|\\ |\pi(U)|\le J/2}} \hat{f}_U \chi_U \,. \tag{4.12}$$

Following the terminology in [15, 32], we refer to the first term as *shattered* term, and the second as non-shattered term. We study these two terms separately. From (4.9), we have

$$\left|\mathop{\mathbb{E}}_{\mathcal{T}}[fF'] - \mathop{\mathbb{E}}_{\mathcal{T}^{(S_0)}}[fF']\right| \le 2\varepsilon_1 + \left|\mathop{\mathbb{E}}_{\mathcal{T}}\left[F' \sum_{\substack{U\subseteq L^{k-m}\times R^{m-1}\\ |U|=|\pi(U)|\\ |\pi(U)|\le J/2}} \hat{f}_U \chi_U\right] - \mathop{\mathbb{E}}_{\mathcal{T}^{(S_0)}}\left[F' \sum_{\substack{U\subseteq L^{k-m}\times R^{m-1}\\ |U|=|\pi(U)|\\ |\pi(U)|\le J/2}} \hat{f}_U \chi_U\right]\right| \tag{4.13}$$

$$+ \left|\mathop{\mathbb{E}}_{\mathcal{T}}\left[F' \sum_{\substack{U\subseteq L^{k-m}\times R^{m-1}\\ |U|>|\pi(U)|\\ |\pi(U)|\le J/2}} \hat{f}_U \chi_U\right] - \mathop{\mathbb{E}}_{\mathcal{T}^{(S_0)}}\left[F' \sum_{\substack{U\subseteq L^{k-m}\times R^{m-1}\\ |U|>|\pi(U)|\\ |\pi(U)|\le J/2}} \hat{f}_U \chi_U\right]\right| . \tag{4.14}$$

We first use smoothness to bound the non-shattered terms. The process is very similar to that in [32], and we get

$$\left|\mathop{\mathbb{E}}_{\mathcal{T}}\left[F' \sum_{\substack{U\subseteq L^{k-m}\times R^{m-1}\\ |U|>|\pi(U)|\\ |\pi(U)|\le J/2}} \hat{f}_U \chi_U\right]\right| \le 2\varepsilon_1 \,. \tag{4.15}$$

The same argument holds under distribution $\mathcal{T}^{(S_0)}$. For the difference involving the shattered terms, we have

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[F' \sum_{\substack{U \subseteq L^{k-m} \times R^{m-1} \\ |U| = |\pi(U)| \\ |\pi(U)| \leq J/2}} \hat{f}_U \chi_U \right] - \mathop{\mathbb{E}}_{\mathcal{T}^{(S_0)}} \left[F' \sum_{\substack{U \subseteq L^{k-m} \times R^{m-1} \\ |U| = |\pi(U)| \\ |\pi(U)| \leq J/2}} \hat{f}_U \chi_U \right] \right| \tag{4.16}$$

$$= \left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[F' \sum_{\substack{U \subseteq L^{k-m} \times R^{m-1} \\ |U| = |\pi(U)| \\ |\pi(U)| \leq J/2}} \hat{f}_U \chi_U \right] - \mathop{\mathbb{E}}_{\mathcal{T}} \left[F' \sum_{\substack{U \subseteq L^{k-m} \times R^{m-1} \\ |U| = |\pi(U)| \\ |\pi(U)| \leq J/2}} (1-\gamma)^{|U|} \hat{f}_U \chi_U \right] \right| \tag{4.17}$$

$$= \left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[F' \sum_{\substack{U \subseteq L^{k-m} \times R^{m-1} \\ |U| = |\pi(U)| \leq J/2}} (1-(1-\gamma)^{|U|}) \hat{f}_U \chi_U \right] \right| \tag{4.18}$$

$$\leq 1-(1-\gamma)^{J/2} \leq \varepsilon_1 \,. \tag{4.19}$$

The key step is (4.17) where we switch the distribution of the second term from $\mathcal{T}^{(S_0)}$ to $\mathcal{T}$. We rely crucially on the fact that $|U| = |\pi(U)|$. To see why this holds, denote the query to f as x (just for the current argument). Observe that the variables x_t are independent for $t \in U$ with different $\pi(t)$, so we first focus on those values t that map to the same $r \in U$. Looking at each $r \in \pi(U)$, $|\pi(U)| = |U|$ implies that there is a unique $t \in U$ such that $\pi(t) = r$, and thus perturbing those x_t satisfying $\pi(t) = r$ with probability γ would give exactly a multiplicative factor of $(1-\gamma)$ to the expectation. Since each $r \in \pi(U)$ contributes a factor of $(1-\gamma)$, the final factor thus becomes $(1-\gamma)^{|\pi(U)|} = (1-\gamma)^{|U|}$.

Summing up the above, we have

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}}[fF'] - \mathop{\mathbb{E}}_{\mathcal{T}^{(S_0)}}[fF'] \right| \leq 7\varepsilon_1 \,. \tag{4.20}$$

This completes the proof.

4.3 Influence based decoding

Suppose we have that for some $\mathscr{S} \subseteq \mathscr{S}_1 \cup \mathscr{S}_3$, the following term is large

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}'} \left[\prod_{S \in \mathscr{S}} f^{(S)}(x^{(S)}) \right] \right| > \varepsilon_1 \,, \tag{4.21}$$

then for at least an $\varepsilon_1/2$ fraction of all possible edge samplings, we have

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}'} \left[\prod_{S \in \mathscr{S}} f^{(S)}(x^{(S)}) \right] \right| > \varepsilon_1/2 \,. \tag{4.22}$$

In the rest of the proof, we focus on samplings of edges where (4.22) holds. We show how to extract a labeling for these edges.

Observe that after we fixed the edges, which function we query only depends on the layer of the query, so for the rest of this section, let f_l be the function on layer l. Also recall that $m(S) = \max S$ is the layer query $x^{(S)}$ is in, and thus in the PCP query $x^{(S)}$ goes to function $f_{m(S)}$. Let $l_m = \max_{S\in\mathscr{S}} m(S)$ be the maximum layer among queries that appears in $\mathscr{S}$.

For $l \in [k]$, denote the queries that appear on layer l as $\mathcal{L}_l := \{S \in \mathscr{S}_1 \cup \mathscr{S}_3 \mid \max S = l\}$, and let $\mathcal{L}_{\leq l} := \cup_{l'\leq l}\mathcal{L}_{l'}$, and similarly define $\mathcal{L}_{<l}$. We need the following observation on independence between queries.

Claim 4.10. *For any $l \in [k]$ and $S_0 \in \mathcal{L}_l$, $x^{(S_0)}$ and $\{x^{(S)}\}_{S\in\mathcal{L}_{<l}}$ are independent under both $\mathcal{T}$ and $\mathcal{T}'$.*

Proof. We first consider $\mathcal{T}$. We can write $x^{(S_0)} = x_e \cdot x^{(\{l\})}$, where $x^{(\{l\})}$ is a uniform random string, "$\cdot$" denotes the elementwise product, and x_e depends on: (1) S_0, (2) $\{x^{(S)}\}_{S\in\mathcal{L}_{<l}}$, (3) the choice of the locations N_r for $r \in L^{k-1}$, and (4) the decision whether the bits in query $x^{(N_r)}$ are resampled. Observe that $x^{(\{l\})}$ is independent of $\{x^{(S)}\}_{S\in\mathcal{L}_{<l}}$, the N_r, and whether the bits are resampled, thus its marginal is still uniform no matter how we fix everything else, and so is the marginal of $x^{(S_0)}$. This implies that $x^{(S_0)}$ is independent of everything else and in particular $\{x^{(S)}\}_{S\in\mathcal{L}_{<l}}$.

For $\mathcal{T}'$, note that the additional noise is applied independently to each bit, and we can use a similar argument as above to show that the marginal of $x^{(S_0)}$ is always uniform, no matter how we fix the other parameters. □

We rewrite the left hand side of (4.22) as

$$\mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{S\in\mathscr{S}} f_{m(S)}(x^{(S)})\right] = \mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{l\in[k]}\prod_{S\in\mathcal{L}_l\cap\mathscr{S}} f_l(x^{(S)})\right].$$

By our choice of permutation and Lemma 4.6, there exists l_0 and j_0 such that

$$|\{S \in \mathcal{L}_{l_0} \cap \mathscr{S} \mid S \ni j_0\}| \text{ is odd}.$$

Then flipping $x^{(\{j_0\})}$ while leaving all other $x^{(\{j'\})}$ unchanged changes the sign of the following

$$\prod_{S\in\mathcal{L}_{l_0}\cap\mathscr{S}} f_{l_0}(x^{(S)}),$$

and since the marginal of $x^{(\{j_0\})}$ is uniform and all functions are folded, we have

$$\mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{S\in\mathcal{L}_{l_0}\cap\mathscr{S}} f_{l_0}(x^{(S)})\right] = 0.$$

To complete the proof of soundness, we show that if

$$\left|\mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{S\in\mathscr{S}} f_{m(S)}(x^{(S)})\right]\right| = \left|\mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{l\in[k]}\prod_{S\in\mathcal{L}_l\cap\mathscr{S}} f_l(x^{(S)})\right]\right|$$
$$= \left|\mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{l\in[k]}\prod_{S\in\mathcal{L}_l\cap\mathscr{S}} f_l(x^{(S)})\right] - \prod_{l\in[k]}\mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{S\in\mathcal{L}_l\cap\mathscr{S}} f_l(x^{(S)})\right]\right| > \varepsilon_1/2, \tag{4.23}$$

then there exists two layers $1 \le l < l_m \le k$ such that

$$\sum_{\substack{r_l \in L^{k-l} \times R^{l-1} \\ r_m \in L^{k-l_m} \times R^{l_m-1} \\ \pi_{l_m \to 1}(r_m) = \pi_{l \to 1}(r_l)}} \mathrm{Inf}_{r_l}^{(\gamma)}(f_l)\,\mathrm{Inf}_{r_m}^{(\gamma)}(f_{l_m}) > \frac{\varepsilon_1^2}{4Z}, \tag{4.24}$$

where $Z = Z(k,\gamma) := 2^{4k^3} k^9 \gamma^{-1}$. This enables us to define a good labeling as the following: choose r_l with probability $\mathrm{Inf}_{r_l}^{(\gamma)}(f_l)/\mathrm{Inf}^{(\gamma)}(f_l)$, and similarly choose r_m with probability $\mathrm{Inf}_{r_m}^{(\gamma)}(f_{l_m})/\mathrm{Inf}^{(\gamma)}(f_{l_m})$, then the probability that the labeling weakly satisfies the edge is

$$\sum_{\substack{r_l \in L^{k-l} \times R^{l-1} \\ r_m \in L^{k-l_m} \times R^{l_m-1} \\ \pi_{l_m \to 1}(r_m) = \pi_{l \to 1}(r_l)}} \frac{\mathrm{Inf}_{r_l}^{(\gamma)}(f_l)}{\mathrm{Inf}^{(\gamma)}(f_l)} \frac{\mathrm{Inf}_{r_m}^{(\gamma)}(f_{l_m})}{\mathrm{Inf}^{(\gamma)}(f_{l_m})} > \gamma^2 \sum_{\substack{r_l \in L^{k-l} \times R^{l-1} \\ r_m \in L^{k-l_m} \times R^{l_m-1} \\ \pi_{l_m \to 1}(r_m) = \pi_{l \to 1}(r_l)}} \mathrm{Inf}_{r_l}^{(\gamma)}(f_l)\,\mathrm{Inf}_{r_m}^{(\gamma)}(f_{l_m}) \ge \frac{\gamma^2 \varepsilon_1^2}{4Z}.$$

This holds for at least $\varepsilon_1/2$ fraction of choices of edges, thus the expected value achieved by the above random labeling procedure is at least $\gamma^2\varepsilon_1^3/8Z$, a value depending only on k and ε.

The key step to proving (4.23) is to bound the following difference

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{S \in \mathscr{S}} f_{m(S)}(x^{(S)})\right] - \mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{l < l_m} \prod_{S \in \mathcal{L}_l \cap \mathscr{S}} f_l(x^{(S)})\right] \mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{S \in \mathcal{L}_{l_m} \cap \mathscr{S}} f_{l_m}(x^{(S)})\right] \right|, \tag{4.25}$$

where we recall that l_m is the highest layer of queries involved in $\mathscr{S}$. We can iteratively apply the bound on (4.25) to get (4.23). In order to establish (4.25), we use an invariance-type result from [32].

Theorem 4.11 (Wenner)**.** *Consider functions*

$$\left\{f^{(t)} \in L^\infty(\Omega_t^n)\right\}_{t \in [d]} \qquad \text{on a probability space} \qquad \mathcal{P} = \left(\prod_{t=1}^d \Omega_t, P\right)^{\otimes n}$$

and a set $M \subsetneq [d]$. Furthermore, let $\mathcal{C}$ be the collection of minimal sets $C \subseteq [d]$, $C \not\subseteq M$, such that the spaces $\{\Omega_t\}_{t \in C}$ are dependent. Then

$$\left| \mathbb{E}\left[\prod f^{(t)}\right] - \prod_{t \notin M} \mathbb{E}[f^{(t)}]\,\mathbb{E}\left[\prod_{t \in M} f^{(t)}\right] \right| \le 2^{2d} \max_{C \in \mathcal{C}} \sqrt{\min_{r \in C} \mathrm{Inf}(f^{(r)}) \sum_i \prod_{t \in C \setminus \{r\}} \mathrm{Inf}_i(f^{(t)})} \prod_{t \notin C} \left\| f^{(t)} \right\|_\infty.$$

To apply the above theorem, we first combine all functions that are not in the highest layer. Let

$$Q = \prod_{S \in \mathscr{S} \cap \mathcal{L}_{<l_m}} X^{(S)},$$

and $q \in Q$ simply be concatenation of $\{x^{(S)}\}_{S \in \mathscr{S} \cap \mathcal{L}_{<l_m}}$. Define the combined function

$$F = \prod_{S \in \mathscr{S} \cap \mathcal{L}_{<l_m}} f_{m(S)},$$

and the noisy version

$$F' = \prod_{S \in \mathscr{S} \cap \mathcal{L}_{<l_m}} T_\gamma f_{m(S)} \,.$$

We still have by Claim 4.10 that Q and $X^{(S_0)}$ are independent for all $S_0 \in \mathcal{L}_{l_m}$. Then the first term in (4.25) becomes

$$\mathop{\mathbb{E}}_{\mathcal{T}'}\left[\prod_{S \in \mathscr{S}} f_{m(S)}(x^{(S)})\right] = \mathop{\mathbb{E}}_{\mathcal{T}}\left[F'(q) \prod_{S \in \mathscr{S} \cap \mathcal{L}_{l_m}} T_\gamma f_{l_m}(x^{(S)})\right] .$$

Let us set $M = \mathscr{S} \cap \mathcal{L}_{l_m}$. Consider the sets C in Theorem 4.11. Since Theorem 4.11 requires that $C \nsubseteq M$, we have that C must include variable q. Due to the independence in Claim 4.10, C must also include at least two variables from $\mathscr{S} \cap \mathcal{L}_{l_m}$. Applying Theorem 4.11, we have

$$\begin{aligned}
&\left|\mathop{\mathbb{E}}_{\mathcal{T}}\left[F'(q) \prod_{S \in \mathscr{S} \cap \mathcal{L}_{l_m}} T_\gamma f_{l_m}(x^{(S)})\right] - \mathop{\mathbb{E}}_{\mathcal{T}}\left[F'(q)\right] \mathop{\mathbb{E}}_{\mathcal{T}}\left[\prod_{S \in \mathscr{S} \cap \mathcal{L}_{l_m}} T_\gamma f_{l_m}(x^{(S)})\right]\right| \\
&\leq 2^{2k^3} \sqrt{\mathrm{Inf}(\overline{T_\gamma f_{l_m}}) \sum_{r \in L^{k-1}} \mathrm{Inf}_r(\overline{F'})\, \mathrm{Inf}_r(\overline{T_\gamma f_{l_m}})} \,,
\end{aligned}$$

where $\overline{F'}$ and $\overline{f}_{l_m}$ are lifted versions of F' and f_{l_m} as defined in Definition 2.17.

Using Proposition 2.18, we have

$$\mathrm{Inf}(\overline{T_\gamma f_{l_m}}) \leq \mathrm{Inf}(T_\gamma f_{l_m}) = \mathrm{Inf}^{(\gamma)}(f_{l_m}) \leq \gamma^{-1} ,$$

and similarly

$$\mathrm{Inf}_r(\overline{T_\gamma f_{l_m}}) \leq \sum_{\substack{r_m \in L^{k-l_m} \times R^{l_m - 1} \\ \pi_{l_m \to 1}(r_m) = r}} \mathrm{Inf}_{r_m}(T_\gamma f_{l_m}) = \sum_{\substack{r_m \in L^{k-l_m} \times R^{l_m - 1} \\ \pi_{l_m \to 1}(r_m) = r}} \mathrm{Inf}^{(\gamma)}_{r_m}(f_{l_m}) \,.$$

Now we need to relate $\mathrm{Inf}_r(\overline{F'})$ with $\mathrm{Inf}^{(\gamma)}_r(f_{m(S)})$. We use the following generalization of Lemma 6.5 from [23].

Lemma 4.12. *Let $(\prod_{i=1}^m \Omega_i^n, \mu)$ be correlated probability space, and $f_i : \Omega_i^n \to [-1,1]$ for $i = 1, \ldots, m$. Then for all r:*

$$\mathrm{Inf}_r\left(\prod_{i=1}^m f_i\right) \leq m \sum_{i=1}^m \mathrm{Inf}_r(f_i) \,.$$

The argument goes exactly the same so we omit the proof here.

Applying Lemma 4.12, we can upper-bound $\mathrm{Inf}_r(\overline{F'})$ by the following

$$\mathrm{Inf}_r(\overline{F'}) \;\leq\; k^3 \sum_{S \in \mathscr{S} \cap \mathcal{L}_{<l_m}} \mathrm{Inf}_r(\overline{T_\gamma f_{m(S)}}) \;\leq\; k^6 \sum_{l < l_m} \mathrm{Inf}_r(\overline{T_\gamma f_l}) \;\leq\; k^6 \sum_{l < l_m} \sum_{\substack{r_l \in L^{k-l} \times R^{l-1} \\ \pi_{l \to 1}(r_l) = r}} \mathrm{Inf}^{(\gamma)}_{r_l}(f_l) \,,$$

where we used Proposition 2.18 to obtain the last inequality.

Summing up, we have

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}} \left[T_\gamma F(q) \prod_{S \in \mathscr{S} \cap \mathcal{L}_{l_m}} T_\gamma f_{l_m}(x^{(S)}) \right] - \mathop{\mathbb{E}}_{\mathcal{T}} \left[T_\gamma F(q) \right] \mathop{\mathbb{E}}_{\mathcal{T}} \left[\prod_{S \in \mathscr{S} \cap \mathcal{L}_{l_m}} T_\gamma f_{l_m}(x^{(S)}) \right] \right| \tag{4.26}$$

$$\leq 2^{2k^3} \sqrt{k^6 \gamma^{-1} \sum_{\substack{1 \leq l < l_m \\ r_l \in L^{k-l} \times R^{l-1} \\ r_m \in L^{k-l_m} \times R^{l_m - 1} \\ \pi_{l_m \to 1}(r_m) = \pi_{l \to 1}(r_l)}} \mathrm{Inf}^{(\gamma)}_{r_l}(f_l) \, \mathrm{Inf}^{(\gamma)}_{r_m}(f_{l_m})} \,. \tag{4.27}$$

Let $Z' = 2^{2k^3} \sqrt{k^6 \gamma^{-1}}$, applying (4.27) to all layers, we get

$$\left| \mathop{\mathbb{E}}_{\mathcal{T}'} \left[\prod_{S \in \mathscr{S}} f^{(S)}(x^{(S)}) \right] \right| < Z' \sum_{2 \leq l_m < k} \sqrt{ \sum_{\substack{1 \leq l < l_m \\ r_l \in L^{k-l} \times R^{l-1} \\ r_m \in L^{k-l_m} \times R^{l_m - 1} \\ \pi_{l_m \to l}(r_m) = r_l}} \mathrm{Inf}^{(\gamma)}_{r_l}(f_l) \, \mathrm{Inf}^{(\gamma)}_{r_m}(f_{l_m})} \,.$$

Thus if the left hand side of the above is larger than $\varepsilon_1/2$, then there exists $1 \leq l < l_m \leq k$ such that

$$\sum_{\substack{r_l \in L^{k-l} \times R^{l-1} \\ r_m \in L^{k-l_m} \times R^{l_m - 1} \\ \pi_{l_m \to l}(r_m) = r_l}} \mathrm{Inf}^{(\gamma)}_{r_l}(f_l) \, \mathrm{Inf}^{(\gamma)}_{r_m}(f_{l_m}) > \left(\frac{\varepsilon_1/2}{kZ'} \right)^2 \cdot \frac{1}{k} = \frac{\varepsilon_1^2}{4Z} \,.$$

This completes the proof. □

5 Acknowledgments

The author would like to thank Johan Håstad for inspiring discussions and invaluable advice; and Cenny Wenner for explaining many of his ideas, as well as numerous comments on improving this manuscript. The author would also like to thank the referees for a careful reading of the paper and many suggestions for improvements.

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. See also at ECCC. [doi:10.1145/278298.278306] 363

[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] 363

[3] PER AUSTRIN AND JOHAN HÅSTAD: On the usefulness of predicates. *ACM Trans. Computation Theory*, 5(1):1, 2013. Preliminary version in CCC'12. [doi:10.1145/2462896.2462897] 366, 374

[4] PER AUSTRIN AND ELCHANAN MOSSEL: Approximation resistant predicates from pairwise independence. *Comput. Complexity*, 18(2):249–271, 2009. Preliminary version in CCC'08. See also at ECCC. [doi:10.1007/s00037-009-0272-6] 360, 367

[5] SIU ON CHAN: Approximation resistance from pairwise independent subgroups. In *Proc. 45th STOC*, pp. 447–456. ACM Press, 2013. See also at ECCC. [doi:10.1145/2488608.2488665] 360, 361, 362, 365, 366, 367, 371, 372, 375

[6] MOSES CHARIKAR, KONSTANTIN MAKARYCHEV, AND YURY MAKARYCHEV: Near-optimal algorithms for maximum constraint satisfaction problems. *ACM Trans. Algorithms*, 5(3):32:1–32:14, 2009. Preliminary version in SODA'07. [doi:10.1145/1541885.1541893] 360

[7] IRIT DINUR, VENKATESAN GURUSWAMI, SUBHASH KHOT, AND ODED REGEV: A new multilayered PCP and the hardness of hypergraph vertex cover. *SIAM J. Comput.*, 34(5):1129–1146, 2005. Preliminary version in STOC'03. [doi:10.1137/S0097539704443057] 363

[8] BRADLEY EFRON AND CHARLES STEIN: The jackknife estimate of variance. *Ann. Stat.*, 9(3):586–596, 1981. [doi:10.1214/aos/1176345462] 367

[9] LARS ENGEBRETSEN AND JONAS HOLMERIN: More efficient queries in PCPs for NP and improved approximation hardness of maximum CSP. *Random Structures & Algorithms*, 33(4):497–514, 2008. Preliminary version in STACS'05. [doi:10.1002/rsa.20226] 360, 361, 362, 363, 365, 372, 374

[10] VITALY FELDMAN, VENKATESAN GURUSWAMI, PRASAD RAGHAVENDRA, AND YI WU: Agnostic learning of monomials by halfspaces is hard. *SIAM J. Comput.*, 41(6):1558–1590, 2012. Preliminary version in FOCS'09. [doi:10.1137/120865094] 361

[11] VENKATESAN GURUSWAMI, PRASAD RAGHAVENDRA, RISHI SAKET, AND YI WU: Bypassing UGC from some optimal geometric inapproximability results. In *Proc. 23rd Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'12)*, pp. 699–717. ACM Press, 2012. See also at ECCC. [ACM:2095174] 361

[12] GUSTAV HAST: Beating a random assignment. In *Proc. 9th Internat. Workshop on Randomization and Computation (RANDOM'05)*, volume 3624 of *LNCS*, pp. 134–145. Springer, 2005. [doi:10.1007/11538462_12] 360

[13] GUSTAV HAST: *Beating a Random Assignment: Approximating Constraint Satisfaction Problems*. Ph. D. thesis, KTH Royal Institute of Technology, 2005. DiVA. 360

[14] JOHAN HÅSTAD: Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. Preliminary version in STOC'97. [doi:10.1145/502090.502098] 360, 361

[15] JOHAN HÅSTAD: On the NP-hardness of Max-Not-2. *SIAM J. Comput.*, 43(1):179–193, 2014. Preliminary version in APPROX'12. [doi:10.1137/120882718] 361, 380

[16] JOHAN HÅSTAD AND SUBHASH KHOT: Query efficient PCPs with perfect completeness. *Theory of Computing*, 1(7):119–148, 2005. Preliminary version in FOCS'01. [doi:10.4086/toc.2005.v001a007] 361, 362

[17] SANGXIA HUANG: Approximation resistance on satisfiable instances for predicates strictly dominating parity. *Electron. Colloq. on Comput. Complexity (ECCC)*, 19:40, 2012. ECCC. 361

[18] SANGXIA HUANG: Approximation resistance on satisfiable instances for predicates with few accepting inputs. In *Proc. 45th STOC*, pp. 457–466. ACM Press, 2013. [doi:10.1145/2488608.2488666] 359

[19] SUBHASH KHOT: Hardness results for coloring 3-colorable 3-uniform hypergraphs. In *Proc. 43rd FOCS*, pp. 23–32. IEEE Comp. Soc. Press, 2002. [doi:10.1109/SFCS.2002.1181879] 361, 363

[20] SUBHASH KHOT: On the power of unique 2-prover 1-round games. In *Proc. 34th STOC*, pp. 767–775. ACM Press, 2002. [doi:10.1145/509907.510017] 360, 361

[21] SUBHASH KHOT AND DANA MOSHKOVITZ: NP-hardness of approximately solving linear equations over reals. *SIAM J. Comput.*, 42(3):752–791, 2013. Preliminary version in STOC'11. See also at ECCC. [doi:10.1137/110846415] 361

[22] ELCHANAN MOSSEL: Gaussian bounds for noise correlation of functions and tight analysis of long codes. In *Proc. 49th FOCS*, pp. 156–165. IEEE Comp. Soc. Press, 2008. [doi:10.1109/FOCS.2008.44] 370

[23] 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] 362, 367, 369, 384

[24] ELCHANAN MOSSEL, RYAN O'DONNELL, AND KRZYSZTOF OLESZKIEWICZ: Noise stability of functions with low influences: invariance and optimality. *Ann. of Math.*, 171(1):295–341, 2010. Preliminary version in FOCS'05. [doi:10.4007/annals.2010.171.295] 362

[25] RYAN O'DONNELL AND JOHN WRIGHT: A new point of NP-hardness for unique games. In *Proc. 44th STOC*, pp. 289–306. ACM Press, 2012. [doi:10.1145/2213977.2214005] 362

[26] RYAN O'DONNELL AND YI WU: Conditional hardness for satisfiable 3-CSPs. In *Proc. 41st STOC*, pp. 493–502. ACM Press, 2009. [doi:10.1145/1536414.1536482] 361, 369

[27] RAN RAZ: A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. Preliminary version in STOC'95. [doi:10.1137/S0097539795280895] 363

[28] ALEX SAMORODNITSKY AND LUCA TREVISAN: A PCP characterization of NP with optimal amortized query complexity. In *Proc. 32nd STOC*, pp. 191–199. ACM Press, 2000. [doi:10.1145/335305.335329] 360, 361, 372

[29] ALEX SAMORODNITSKY AND LUCA TREVISAN: Gowers uniformity, influence of variables, and PCPs. *SIAM J. Comput.*, 39(1):323–360, 2009. Preliminary version in STOC'06. See also at ECCC. [doi:10.1137/070681612] 360

[30] LINQING TANG: Conditional hardness of approximating satisfiable Max 3CSP-q. In *Internat. Symp. Algorithms and Computation (ISAAC'09)*, pp. 923–932. Springer, 2009. [doi:10.1007/978-3-642-10631-6_93] 361

[31] LUCA TREVISAN: Approximating satisfiable satisfiability problems. *Algorithmica*, 28(1):145–172, 2000. Preliminary version in ESA'97. [doi:10.1007/s004530010035] 361

[32] 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. Conference version in APPROX'12. [doi:10.4086/toc.2013.v009a023] 361, 362, 369, 370, 375, 378, 380, 383

[33] URI ZWICK: Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint. In *Proc. 9th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'98)*, pp. 201–210. ACM Press, 1998. [ACM:314701] 360, 361

AUTHOR

Sangxia Huang
Ph. D. Student
KTH Royal Institute of Technology
Stockholm, Sweden
huang.sangxia@gmail.com
http://www.csc.kth.se/~sangxia/

ABOUT THE AUTHOR

SANGXIA HUANG is a Ph. D. student at KTH Royal Institute of Technology, Stockholm, Sweden, advised by Johan Håstad. His research interest is computational complexity, and in particular, approximation algorithms and hardness of approximation for Constraint Satisfaction Problems (CSPs) and other combinatorial optimization problems. He grew up in Shanghai, China, and spent his undergraduate days at Shanghai Jiao Tong University. During his spare time, he enjoys reading, playing billiards, and traveling. He is also an avid Arsenal fan.