# Parity is Positively Useless*

**Cenny Wenner**

**KTH – Royal Institute of Technology**
cenny@cwenner.net

**Abstract**

We give the first examples of non-trivially positively-useless predicates subject only to $\mathsf{P} \neq \mathsf{NP}$. In particular, for every constraint function $Q : \{-1, 1\}^4 \to \mathbb{R}$, we construct Contraint-Satisfaction-Problem (CSP) instances *without negations* which have value at least $1 - \varepsilon$ when evaluted for the arity-four odd-parity predicate, yet it is $\mathsf{NP}$-hard to find a solution with value significantly better than a random biased assignment when evaluated for $Q$. More generally, we show that all parities except one are positively useless.

Although we are not able to exhibit a single protocol producing hard instances when evaluated for every $Q$, we show that two protocols do the trick. The first protocol is the classical one used by Håstad with a twist. We extend the protocol to multilayered Label Cover and employ a particular distribution over layers in order to limit moments of table biases. The second protocol is a modification of Chan's multi-question protocol where queried tuples of Label Cover vertices are randomized in such a way that the tables can be seen as being independently sampled from a common distribution and in effect having identical expected biases. We believe that our techniques may prove useful in further analyzing the approximability of CSPs without negations.

**1998 ACM Subject Classification** F.2.2 Nonnumerical Algorithms and Problems

**Keywords and phrases** Approximation hardness, approximation resistance, parity, usefulness, negations, monotone, constraint satisfaction problems, smoothness, multilayer, Label Cover

**Digital Object Identifier** 10.4230/LIPIcs.APPROX-RANDOM.2014.433

## 1 Introduction

We study the usefulness of boolean maximum *Constraint Satisfaction Problems* (CSPs). The Width-3 Parity Problem (Max E3-Even) and the Width-3 Satisfiability Problem (Max E3-SAT) are two canonical CSPs. In Max E3-Even resp. Max E3-SAT, an instance consists of a collection of constraints of the form $x'_{i_1} \oplus x'_{i_2} \oplus x'_{i_3}$ resp. $x'_{i_1} \vee x'_{i_2} \vee x'_{i_3}$ where $\oplus$ denotes `exclusive or` and $x'_i$ is either a variable or its negation. A solution to an instance consists of an assignment to the variables and its value is the fraction of satisfied constraints. For the sake of analysis, varible domains is taken to be $\{-1, 1\}$ where '1' is interpreted as `false` and '$-1$' as `true`. Similarly, constraints can be seen as a function from $\{-1, 1\}^3$ to the real numbers applied to triples of variables, or their negations, and where a satisfying assignment of the variables is awarded the value 1. More generally, we define a CSP, denoted Max CSP$(P)$, by specifying the *constraint function* $P : \{-1, 1\}^k \to \mathbb{R}$ to be used instead of the 3-Even or 3-SAT constraints. The number of variables $k$ which $P$ acts on is called its width and if the range of $P$ is contained in $\{0, 1\}$, then $P$ is called a *predicate*.

It turns out that for almost all predicates, it is $\mathsf{NP}$-hard to find an assignment satisfying every constraint [19] and we turn our attention to *approximations*. We say that an algorithm

---

* Supported by ERC Advanced Grant 226203. This work was done in part while the author was visiting the Simons Institute for the Theory of Computing.


17th Int'l Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'14) / 18th Int'l Workshop on Randomization and Computation (RANDOM'14).
Editors: Klaus Jansen, José Rolim, Nikhil Devanur, and Cristopher Moore; pp. 433–448
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

$A$ is a *(factor) c-approximation (algorithm)* if, given any instance, the value of the solution produced by $A$ is within a factor $c$ of the optimal value. One of the simplest approximation algorithms one could think of is to choose a uniform at random (u.a.r) assignment. This algorithm can be derandomized and, via linearity of expectation, constitutes an $\mathbf{E}[P]$ approximation. It turns out that for Max 3-Even and Max 3-SAT, this naive algorithm achieves the optimal constant-factor approximation assuming $\mathsf{P} \neq \mathsf{NP}$. More specifically, for every $\varepsilon > 0$, it is NP-hard to distinguish whether a Max 3-Even instance has value at least $1 - \varepsilon$ or at most $1/2 + \varepsilon$. Predicates, and CSPs, with the property that it is NP-hard to beat a random assignment are said to be *approximation resistant.* In a recent result by Chan [3], combined with a result by Håstad [11], it turns out that asymptotically a fraction one of predicates as $k$ grows are in fact approximation resistant.

Inspired by stronger properties than approximability, Austrin and Håstad [1] introduced the concepts of *usefulness* and *uselessness.* A predicate $P$ is said to be *(computationally) useful* for a constraint function $Q$ if there exists a $\delta > 0$ and a polynomial-time algorithm which given instances of value at least $1-\delta$ evaluated for $P$, produces solutions of value at least $\mathbf{E}[Q]+\delta$ evaluated for $Q$. In other words, high-valued instances for $P$ permit polynomial-time non-trivial solutions for $Q$. If there is no such algorithm, then $P$ is said to be *useless for $Q$*. If $P$ is useful for some (resp. no) constraint functions $Q : \{-1, 1\}^k \to \mathbb{R}$, then $P$ is simply said to be *useful* (resp. *useless*). Note that assuming $\mathsf{P} \neq \mathsf{NP}$, uselessness can be established by showing NP-hardness. When $\mathsf{P} = \mathsf{NP}$, uselessness is instead essentially captured by the related definition called *information-theoretic uselessness.* However, information-theoretical usefulness is only of interest with respect to specific constraint functions as every constraint function is trivially information-theoretically useful for itself and hence useful given $\mathsf{P} = \mathsf{NP}$.

Assuming the Unique Games Conjecture (UGC) [14], Austrin and Håstad [1] gave a complete characterization of useless predicates: a predicate is useless if and only if there exists a pairwise-uniform distribution supported on $P^{-1}(1)$. This can be compared to the UGC-based result of Austrin and Mossel [2] which showed the same condition to be sufficient – but not necessary – for approximation resistance. This connection is not a coincidence; the general hardness of Austrin and Mossel inspired the study of Austrin and Håstad, and additionally, a predicate $P$ is approximation resistant if and only if $P$ is useless for itself. Consequently, every useless predicate is approximation resistant.

Similarly inspired to give a characterization of approximation resistance, Khot et al. [16] introduced the concept of *strong approximation resistance* where, for a predicate $P$, it is hard to not only find a solution with constant value greater than $\mathbf{E}[P]$ but it is hard to find any solutions with value significantly different than $\mathbf{E}[P]$, i.e., outside the range $\mathbf{E}\,P \pm o(1)$. Khot et al. gave sufficient and necessary condition of so called *strong approximation resistance* assuming the UGC. Curiously, if a predicate $P$ is strongly approximation resistant, then $P$ is useless for $P$ as well as its complement $1 - P$. To the best of our understanding, the converse is however not known; uselessness of $P$ for $P$ resp. $1 - P$ can be establishing using distinct instances while strong approximation resistance demands that there are instances which are *simultaneously useless* for $P$ and $1 - P$. Generalizing slightly, Austrin and Håstad named this stronger property *adaptive uselessness* for which we indeed have an equivalence.

Although we have a good understanding of the approximability of predicates assuming the UGC, until recently, little was known about approximation resistance conditioned only on $\mathsf{P} \neq \mathsf{NP}$. Using new instance constructions, Chan [3] showed the approximation resistance of all predicates $P$ such that $P^{-1}(1)$ contains a group supporting a pairwise-uniform distribution. Again, by Håstad [11], the fraction of such predicates approaches one with the width. A caveat to this and other notable results is that the study is limited to CSP instances where constraints act on both variables and their negations. This is the natural formulation for Max E3-SAT while it may be less reasonable for other problems. One well-known example of the distinction between allowing or not negations in a problem is the width-two `not-equal` predicate. If we permit negations, the predicate can encode both equality and inequality, and the resulting CSP is the Max 2-Lin-2 problem which is presently known to have a constant-approximation hardness of $\frac{11}{12}+\varepsilon$. When negations are not permitted, the problem corresponds to Max Cut which has a present constant-approximation hardness of $\frac{16}{17}+\varepsilon$ [10, 20]. Assuming the UGC, these two problems are in fact known to have the same approximability [15] but there are many other problems for which the hardness differs, such as Max E3-SAT where all constraints can be satisfied with an all-true assignment.

We denote by Max CSP$^+$(P) the restriction of Max CSP(P) where negations of variables are not automatically allowed and we call such instances *monotone* or "*without negations*". It turns out that a naive approximation algorithm for these problems can benefit from assigning variables the value 1 or $-1$ with different probabilities. The maximum expected value of such an algorithm is given by $\mathbf{E}^+P \stackrel{\text{def}}{=} \max_{b^*} \mathbf{E}_{\mathbf{x}_1,\ldots,\mathbf{x}_k \sim_{b^*} \{-1,1\}}\left[P(\mathbf{x}_1,\ldots,\mathbf{x}_k)\right]$, where $\mathbf{x} \sim_b \{-1,1\}$ denotes drawing $\mathbf{x}$ with expectation $b$. When it is NP-hard to distinguish Max CSP$^+$(P) instances which have value at least $1-\varepsilon$ from instances of value at most $\mathbf{E}^+P+\varepsilon$, we say that $P$ is *positively approximation resistant*. Similarly, we say that $P$ is *positively useful* for a constraint-function $Q$ if there exist an $\varepsilon > 0$, there exists a polynomial-time algorithm which given monotone instances with value at least $1-\varepsilon$ for $P$, produces solutions with value at least $\mathbf{E}^+Q+\varepsilon$ for $Q$. If $P$ is not *positively useful* for $Q$, then it is *positively useless* for $Q$, and $P$ is simply *positively useless* if it is so for every $Q$.

Assuming the UGC, Austrin and Håstad [1] gave a complete characterization also of positively-useless predicates under the UGC: $P$ is positively useless if and only if $P^{-1}(1)$ supports a distribution where all biases are identical and (pairwise) correlations are non-negative and identical. For a more in-depth discussion and motivation of usefulness, we refer the reader to Austrin and Håstad [1].

Although UGC-based results are conditional, the conjecture has arguably contributed greatly in and outside the field of Approximability in the form of insights and techniques which have found applications also without the conjecture [17, 5], as well as promoting results which have subsequently been proven subject only to $\mathsf{P} \neq \mathsf{NP}$ [3, 9].

General conditional results have yet to be discovered for monotone CSPs. Despite the UGC implying that a fraction one of predicates for increasing width are positively approximation resistant, [1, 11], we are only aware of a handful of non-trivial natural predicates for which hardness is known subject only to $\mathsf{P} \neq \mathsf{NP}$. In particular, such results have been restricted to predicates where the optimal bias corresponds to balanced bits, i.e., $\mathbf{E}^+Q = \mathbf{E}\,Q$. One notable example is Max E4-Set Splitting, shown approximation resistant by Håstad [10], and generalized to greater widths by Guruswami [7]. Note that Max E$k$-Set Splitting is the same problem as Max CSP$^+$($k$-NAE) where $k$-NAE is the width-$k$ predicate accepting heterogenous assignments. While there are a few results on the approximation resistance of monotone CSPs subject to $\mathsf{P} \neq \mathsf{NP}$, to the best of our knowledge, there were no known non-trivial positively-useless predicates prior to this work.

**Organization.** Our contributions are formally stated in Section 2, while Section 3 covers the analytical preliminaries, and Section 4 gives an overview to the two protocols and their analyses. Despite a generous page limit, we unfortunately only include the multilayered protocol, in Section 5 and its analysis, in Section 6. For the multiple-questions protocol, the generalization to other parities, and a proof of the hardness of the reduced-from Label Cover variant, we refer the reader to the full version.

## 2 Our Contributions

For every $k \in \mathbb{N}^{\geq 1}$, let $k$-Even (resp. $k$-Odd) be the width-$k$ predicate satisfied by $k$-tuples containing an even (resp. odd) number of $-1$'s. Certain predicates, such as $k$-Even, are known to be trivially positively useless because their monotone instances are satisfied by an all-1 or an all-($-1$) assignment. In this work, we give the first examples of non-trivially positively-useless predicates subject only to $\mathsf{P} \neq \mathsf{NP}$.

| | $k$-Even | $k$-Odd |
|---|---|---|
| width $k = 2$ | triv. useless | pos. useful |
| $k \neq 2$ even | triv. useless | pos. useless |
| $k$ odd | triv. useless | triv. useless |

The positive usefulness of parities if $\mathsf{P} \neq \mathsf{NP}$.

▶ **Theorem 1.** *The predicate 4-Odd is positively useless iff* $\mathsf{P} \neq \mathsf{NP}$.

For presentational clarity, our analysis is chiefly concerned with $k = 4$ but the result generalizes to $k$-Odd for every even $k \geq 4$. We refer the reader to the full version for this generalization.

We note that also $k$-Odd is trivially satisfiable for odd $k$, and that for $k = 2$, Max CSP$^+$($k$-Odd) is the Max Cut problem which is not positively useless since it has a non-trivial approximation by, e.g., Williamson and Goemans [6]. Consequently, the smallest parity candidate for positive uselessness is 4-Odd and more generally we show the following complete characterization of the positive usefulness of parities.

▶ **Theorem 2.** *Let $P$ be even or odd parity of width $k \geq 1$. Then, $P$ is positively useful if and only if $P$ is* 2*-Odd or* $\mathsf{P} = \mathsf{NP}$.

A notable feature of our proof of Theorem 1 is that we exhibit two distinct protocols such that for every constraint function $Q$, one of the two protocols produce positively-useless instances for $Q$. While uselessness is implicit in many approximation-resistance proofs, to our knowledge, none of these results involve the combination of multiple protocols. This construction is somewhat non-intuitive, especially considering that a result by Austrin and Håstad [1] shows that a single protocol suffices for every positively-useless predicate assuming the UGC.

In the following sections we present a multilayered protocol and a multiple-questions protocol. The former establishes the uselessness of 4-Odd for every $Q$ with positive highest Fourier coefficient while the second establishes uselessness for $Q$'s with negative coefficient. Together they imply Theorem 1.

▶ **Lemma 3.** *The multilayered reduction $R_{\mathrm{ML},\gamma,s}$ from* LC *implies that 4-Odd is positively useless for $Q$ whenever $\hat{Q}_{[4]} \geq 0$, subject to* $\mathsf{P} \neq \mathsf{NP}$.

▶ **Lemma 4.** *The multiple-questions reduction $R_{\mathrm{MQ},\gamma,p,M}$ from* LC *implies that 4-Odd is positively useless for $Q$ whenever $\hat{Q}_{[4]} \leq 0$, subject to* $\mathsf{P} \neq \mathsf{NP}$.

## 3 Preliminaries

Random variables are for clarity denoted with bold font, as in $\mathbf{x}$, while vectors are denoted with overset arrows, as in $\vec{x}$. Vector multiplication is taken point wise. We let indexing of $n$ elements range from 0 through $n-1$ and denote by $[n]$ the integer interval $\{0, \ldots, n-1\}$. For a statement S, the indicator $1\{\mathrm{S}\}$ is 1 if S is true and otherwise 0. For a function $\pi : R \to L$, we let $\pi(T)$ denote the image of $\pi$ for the set $T \subseteq R$, and we use the notation $\pi_2(T)$ for the set of elements which $T$ maps to an odd number of times.

For sets $S$ and $T$ we may denote their union $S \uplus T$ with the added condition that $S$ and $T$ are disjoint. We denote $\mathbf{x}$ being drawn uniformly at random (u.a.r.) from $S$ with $\mathbf{x} \sim S$. When $S$ consists of two real numbers $a < b$, typically $-1$ and 1, and $\mu \in [a, b]$, we let $\mathbf{x} \sim_\mu S$ indicate the distribution where $\mathbf{x}$ is chosen from $S$ with expectation $\mu$.

With respect to an implicit graph $G$, $N(v)$ signifies the neighborhood of a vertex $v$, the notation $u \sim v$ that the vertices $u$ and $v$ are neighbors, and for a sequence of vertex sets $U_1, \ldots, U_k \subseteq V[G]$, the set of paths in $G$ contained in $U_1 \times \cdots \times U_k$ is denoted $E(U_1, \ldots, U_k)$. In particular, $E(U, W)$ is the set of edges between $U$ and $W$.

### Orthogonal Decompositions

For a function $f : \{-1, 1\}^n \to \mathbb{R}$, we define $f(\vec{x}) = \sum_{S \subseteq [n]} \hat{f}_S \prod_{i \in S} x_i$ as the *Fourier expansion* of $f$ where the quantity $\hat{f}_S$ is $\mathbf{E}_{\vec{\mathbf{x}}} \left[ f(\vec{\mathbf{x}}) \prod_{i \in S} \mathbf{x}_i \right]$. We denote $x_S = \prod_{i \in S} x_i$. For every $S \subseteq [n]$, $\min f \le \hat{f}_S \le \max f$ and Parseval's Identity is of notable interest: $\sum \hat{f}_S^2 = \mathbf{E}\left[ f^2 \right]$. For a set $\Omega$ and function $f : \Omega^n \to \mathbb{R}$, we shall also use the *Efron-Stein decomposition* $\{f_S\}_{S \subseteq [n]}$ where $f_S(\vec{x}) \stackrel{\text{def}}{=} \sum_{T \subseteq S} (-1)^{|S \setminus T|} \mathbf{E}[f(\vec{\mathbf{y}}) \mid \vec{\mathbf{y}}_T = \vec{x}_T]$, satisfying $f = \sum f_S$; $f_S(\vec{x})$ only depends on $\{x_i\}_{i \in S}$; and whenever $T \setminus S \neq \emptyset$, $\mathbf{E}[f(\vec{\mathbf{x}}) \mid \vec{x}_T] = 0$. For a motivation of this decomposition, we refer the reader to Mossel [18].

▶ **Definition 5.** For any $0 \le \gamma \le 1$ and $f : \{-1, 1\}^n \to \mathbb{R}$, the *noise operator* applied to $f$, $\mathrm{T}_{1-\gamma} f$, is defined as $\mathrm{T}_{1-\gamma} f(\vec{x}) \stackrel{\text{def}}{=} \mathbf{E}_{\vec{\mathbf{y}}}[f]$ where independently for each $i \in [n]$, $\mathbf{y}_i$ is set to $x_i$ with probability $1 - \gamma$ and otherwise sampled uniformly at random from $\{-1, 1\}$.

For a function $f : \{-1, 1\}^n \to \mathbb{R}$, the Fourier expansion of $\mathrm{T}_{1-\gamma} f$ is particularly simple: $\sum_S (1 - \gamma)^{|S|} \hat{f}_S x_S$. Additionally, mean values are invariant of noise: $\mathbf{E}[\mathrm{T}_{1-\gamma} f] = \mathbf{E}[f]$.

## 3.1 Label Cover and Inapproximability

The Label Cover (LC) problem is a common starting point of strong inapproximability results. In particular, we will concern ourselves with a smooth multilayered variant which has an additional property which we call *path samplability*. Multilayered LC, or PCPs, were first introduced by Dinur et al. [4] for studying the approximability of Hypergraph Vertex Cover.

▶ **Definition 6.** An instance $(G, \{L_i\}_{[m_{\mathrm{LC}}]}, \{\pi^e\}_{E(G)})$ of the *$m_{LC}$-multilayered maximization problem* Label Cover consists of $m_{\mathrm{LC}}$ label sets $L_0, \ldots, L_{m_{\mathrm{LC}}-1}$, and an $m_{\mathrm{LC}}$-partite graph $G = (V_0 \uplus \cdots \uplus V_{m_{\mathrm{LC}}-1}, E)$ associating for every pair of vertex sets $V_a, V_b, a < b$; and edge $(\mathbf{u}, \mathbf{v}) \in E(V_a, V_b)$, a *projection* $\pi^{\mathbf{v},\mathbf{u}} : L_b \to L_a$. A solution to the instance consists of a labeling $\lambda : \biguplus V_a \to \biguplus L_a$ and its value is the maximum fraction of edges between two distinct vertex sets for which the labeling satisfies the associated projection:

$$\max_{a<b} \mathop{\mathbf{P}}_{(\mathbf{u},\mathbf{v}) \sim E(V_a, V_b)} \{\pi^{\mathbf{v},\mathbf{u}}(\lambda(\mathbf{v})) = \lambda(\mathbf{u})\}.$$

For a vertex $v \in \biguplus V_i$, we will also denote by $L(v)$ the label set $L_i$ where $V_i$ is the unique layer containing $v$. When restricting ourselves to the more common bipartite case, we denote $U = V_0; L = L_0; V = V_1$, and $R = L_1$. Since the label set $L$ is typically smaller than $R$, we shall also refer to the vertices $U$ resp. $V$ as small- resp. large-side vertices. *Smoothness* of projections is a condition akin to but weaker than unique projections introduced by Khot [13] for analyzing approximate coloring of 3-colorable hypergraphs. We use the following definition of smoothness which is equivalent up to constants.

▶ **Definition 7.** An $m_{\text{LC}}$-multilayered LC instance is $(J,\xi)$-*smooth* when for every $0 \le a < b < m_{\text{LC}}$, vertex $v \in V_b$, and set of at most $J$ labels $S \subseteq L_b$, over a u.a.r. neighbor $\mathbf{u} \in V_a$ of $v$,

$$\mathop{\mathbf{P}}_{\mathbf{u} \sim V_a \cap N(v)} \{|\pi^{v,\mathbf{u}}(S)| \neq |S|\} \le \xi. \tag{1}$$

The most well-used bipartite LC constructions are known to be biregular and has the following sampling property: choosing a vertex $v$ u.a.r. from either layer and a u.a.r. neighbor of $v$ yields a u.a.r. edge in the graph. Our analysis requires a slightly stronger property: for every two LC layers $a < b$, choosing a u.a.r. path $\vec{p}$ from the first to the last layer and considering the vertices chosen in $V_a$ and $V_b$ yields a u.a.r. edge between $V_a$ and $V_b$. We call this property path samplability and note that it is relatively easy to verify that, e.g., Khot's multilayered construction [13] satisfies this property. For a proof, we refer the reader to the full version.

▶ **Theorem 8.** *For every $J, \xi, \varepsilon_{LC} > 0$ and $m_{LC} \ge 2$, there exist path-samplable $(J,\xi)$-smooth $m_{LC}$-multilayered* LC *instances for which it is* NP*-hard to distinguish instances of value 1 from instances of value at most $\varepsilon_{LC}$.*

## 4 Protocols for Useless Instances

We have not been able to find a single protocol producing instances useless for every $Q$. However, by combining a new protocol with a multilayered-variant of a classical protocol, we argue that for every $Q$, at least one of these two protocols produces instances which are positively useless for $Q$.

Håstad's classical protocol [10] establishing the approximation resistance of 4-Odd with negations, samples u.a.r. an edge $(\mathbf{u}, \mathbf{v})$ from a bipartite LC instance, issues one query to an associated *table* $f^{\mathbf{u}}$ – a collection of variables seen as a function – and three to an associated table $f^{\mathbf{v}}$. With negations, each table can be assumed to have a balanced assignment of 1's and $-1$'s via a folding trick. From this, one can argue that instances in the `no` case have value at most $\mathbf{E}[\text{4-Odd}] + \varepsilon$, implying approximation resistance, i.e., the uselessness of 4-Odd for 4-Odd. Without negations, the analysis does not carry through for the 4-Odd predicate for a simple reason. Since the protocol always issues one query to the first layer and three to the second, we could simply let the first layer consist entirely of $-1$'s and the second entirely of 1's. Every constraint would then include one variable assigned $-1$ and three assigned 1, satisfying the 4-Odd predicate.[1]

Notably, Håstad's protocol does show the positive uselessness of 4-Odd for some constraint functions such as 4-Even. In turns out that a slight modification of the protocol suffices to extend this to every $Q$ such that $\hat{Q}_{[4]} \ge 0$. The modification is to reduce from multilayered LC instances and sample tables from pairs of layers according to a particular distribution. The distribution ensures that the value analytically approximately corresponds to issuing randomly the three queries to either the lower or higher of the two layers, rather than always querying the higher layer thrice. This property ensures that whenever $\hat{Q}_{[4]} \ge 0$, the optimal table choices are essentially equally unbalanced. We note that a positive coefficient of $\hat{Q}_{[4]}$

[1] Note that one can show the positive approximation resistance of 4-Odd using a protocol which chooses a random $u \in U$, two neighbors $v_1, v_2 \sim u$ and issues two queries each to $v_1$ and $v_2$. A third possibility is the protocol which on a multilayered instance issues one query each to two distinct layers and two queries to a third. However, it turns out that there are still contraint functions $Q$ for which one can deduce non-trivial solutions to all three of these protocols.

essentially means that even-parity assignments, such as zero, two, or four 1's; have a positive effect on the value while odd-parity assignments have a negative effect. This explanation agrees with the intuition of the protocol – if we can use the hardness of Label Cover to decouple the table queries, then we randomly query some table once and another thrice. We should then expect biased tables to only be able to achieve greater probabilities of odd-parity outcomes which is not beneficial.

Our second protocol is similar to Chan's protocol [3] for groups supporting pairwise independence. However, asking each table about several LC vertices serves a slightly different purpose in our construction than for Chan's hardness amplification. We construct constraints by sampling some large number of edges from a bipartite LC instance and independently for each edge and table, we ask the table about one of the two edge endpoints. This ensures that for any fixed set of edges, the four tables are independently drawn with replacement from a common distribution, implying that the four tables *have the same expected bias*. For technical reasons, asking each table about a single random endpoint is not sufficient. We pick many edges to ensure that each table is asked about both small-side and large-side endpoints, and almost always choose the large-side endpoint to avoid certain deleterious cases. Typically in the analysis of protocols showing approximation resistance, one argues that when reducing

from low-valued LC instances, the value of the protocol roughly corresponds to issuing independent queries to the respective chosen tables. Our construction does not have this property and the arguments can be highly correlated even for low-valued LC instances. However, we show that a certain kind of previously-known folding trick between arguments to a table still works and is enough to argue that the tables cannot coordinate better than independent biased assignments whenever $\hat{Q}_{[4]} \leq 0$.

## 5 The Multilayered Classical Protocol

We introduce the distribution over LC layers alluded to in Section 4. The distribution and lemmas are in particular taken from Guruswami et al., where the same was used to generate arguments for ordering problems [8]. In particular, we are using slight reformulations of their general distributions restricted to a domain of size $k = 2$.

For an integer $s > 0$, the distribution $\mathcal{D}_s$ could be seen as generating bit vectors by choosing a random suffix-length $0 \leq r < s$, a prefix $\vec{p}$ of length $s - r - 1$, and producing two random length-$s$ strings $(\mathbf{a}, \mathbf{b})$, one with prefix $p \cdot (0)$ and one with prefix $p \cdot (1)$. This ensures that we always produce pairs for which $\mathbf{a} < \mathbf{b}$.

▶ **Definition 9.** For two integers $r, p \geq 0$, define $\mathcal{D}_{r,p}$ as the uniform distribution over $\{2^r \cdot p, \ldots, 2^r \cdot p + 2^r - 1\}$.

▶ **Definition 10** (Special case $k = 2$ of Definition 11.2, [8])**.** The distribution $\mathcal{D}_s$ is a distribution over pairs from $[2^s]$ defined as follows.

1. Pick a random $\mathbf{r}$ uniformly in $[s]$.
2. Pick a random $\mathbf{p}$ uniformly in $\left[2^{s-\mathbf{r}-1}\right]$.
3. Output $(\mathbf{a}, \mathbf{b})$ where $\mathbf{a}$ is sampled from $\mathcal{D}_{\mathbf{r}-1,2\mathbf{p}}$ and $\mathbf{b}$ from $\mathcal{D}_{\mathbf{r}-1,2\mathbf{p}+1}$.

The crucial property that we use this distribution for is that in spite of the distribution always generating pairs $(\mathbf{a}, \mathbf{b})$ for which $\mathbf{a} < \mathbf{b}$, when evaluated for discretized functions $f$ and $g$, their expected product w.r.t. the distribution is close to an expectation over distributions where $\mathbf{a}$ and $\mathbf{b}$ are sampled independently.

In our analysis, we in particular use the following lemma which follows from discretizing the domain $[-1,1]$ and applying properties of $\mathcal{D}_s$ separately to $f$ and $g$.

▶ **Lemma 11.** *Let $s \geq 0$ and consider $f, g : [2^s] \to [-1,1]$. Then for any integers $q > 0; k_1, k_2 \geq 0$,*

$$\left| \mathop{\mathbf{E}}_{(\mathbf{a},\mathbf{b})\sim\mathcal{D}_s} \left[f(\mathbf{a})^{k_1} g(\mathbf{b})^{k_2}\right] - \mathop{\mathbf{E}}_{\mathbf{r},\mathbf{p}} \left[ \mathop{\mathbf{E}}_{\mathcal{D}_{\mathbf{r},\mathbf{p}}} \left[f(\mathbf{a})^{k_1}\right] \mathop{\mathbf{E}}_{\mathcal{D}_{\mathbf{r},\mathbf{p}}} \left[g(\mathbf{a})^{k_2}\right]\right]\right| \leq 2\frac{k_1+k_2}{q} + 8\sqrt{\frac{q}{s}}.$$

**Proof.** Deferred to Section 6. ◀

## 5.1 The Reduction

As is typical for LC reductions, we define a reduction to Max CSP$^+$ instances by specifying a probabilistic protocol where the weight of variable tuple in the produced instance corresponds to the probability that it is generated by the protocol.

▶ Procedure 12 (The Multi-Layered Protocol reduction $R_{\mathrm{ML},\gamma,s}$). Let $\gamma > 0; s \geq 2$ be arbitrary. The reduction $R_{\mathrm{ML},\gamma,s}$ from path-samplable $M = 2^s$-multilayered LC instances $(G, \{L_i\}_{[m_{\mathrm{LC}}]}, \{\pi^e\}_{E(G)})$ is defined with the following protocol.

1. Sample a pair of layer indices $(\mathbf{a},\mathbf{b})$ from $\mathcal{D}_s$ and an edge $\mathbf{e} = (\mathbf{w_a}, \mathbf{w_b})$ from $E(V_\mathbf{a}, V_\mathbf{b})$.
2. Draw $\vec{\mathbf{x}}$ u.a.r. from $\{-1,1\}^{L_\mathbf{a}}$; $\vec{\mathbf{y}}^{(1)}$ and $\vec{\mathbf{y}}^{(2)}$ from $\{-1,1\}^{L_\mathbf{b}}$; and for each $j \in L_\mathbf{b}$, set $\mathbf{y}_j^{(3)} = -\mathbf{x}_{\pi^\mathbf{e}(j)} \mathbf{y}_j^{(1)} \mathbf{y}_j^{(2)}$.
3. For each $i \in L_\mathbf{a}$, resp. $j \in L_\mathbf{b}$, sample $\zeta_i^{(0)}, \zeta_j^{(1)}, \zeta_j^{(2)}$, and $\zeta_j^{(3)} \sim_{1-\gamma} \{-1,1\}$.
4. Output a random permutation of the tuple

$$\left(f^{(\mathbf{w_a})}(\vec{\zeta}^{(0)}\vec{\mathbf{x}}), f^{(\mathbf{w_b})}(\vec{\zeta}^{(1)}\vec{\mathbf{y}}^{(1)}), f^{(\mathbf{w_b})}(\vec{\zeta}^{(2)}\vec{\mathbf{y}}^{(2)}), f^{(\mathbf{w_b})}(\vec{\zeta}^{(3)}\vec{\mathbf{y}}^{(3)})\right).$$

The distribution over pairs of vertices draws the pair $(\mathbf{a},\mathbf{b}) \in [2^s]^2$ from $\mathcal{D}_s$ and outputs a random edge $(\mathbf{w_a}, \mathbf{w_b}) \in E(V_\mathbf{a}, V_\mathbf{b})$. Note that path samplability of LC instances implies that this distribution is equivalent to choosing a u.a.r. path $\vec{W} = (\mathbf{w}_0, \ldots, \mathbf{w}_{2^s-1}) \sim E(V_0, \ldots, V_{2^s-1})$ and a pair $(\mathbf{a},\mathbf{b}) \sim \mathcal{D}_s$, generating the tuple $(\mathbf{w_a}, \mathbf{w_b})$.

▶ **Lemma 13** (Completeness). *Reducing from a satisfiable* LC *instance $\mathcal{I}$, the instance $R_{\mathrm{ML},\gamma,s}(\mathcal{I})$ produced by the Multilayered Protocol has value at least $1-4\gamma$ when evaluated for 4-Odd.*

This proof is standard. Consider a dictatorship assignment of an arbitrary satisfying labeling. If none of the coordinates used by the four dictators are noised, which happens with probability at least $1-4\gamma$, then the tuple of values of the queried tables equals $(x_1, x_2, x_3, -x_1x_2x_3)$ for some $x_1, x_2, x_3 \in \{-1,1\}$. For the 4-Odd predicate, such tuples are awarded value 1 and in effect, the expected value of the protocol is at least $1-4\gamma$.

▶ **Lemma 14** (Soundness). *Let $\xi, J, \varepsilon_{LC} > 0; s \geq 2$. Whenever $Q$ satisfies $\hat{Q}_{[4]} \geq 0$ and the reduced-from $2^s$-layered* LC *instance is path-samplable, $(J,\xi)$-smooth, and of value at most $\varepsilon_{LC}$, then the instance produced by $R_{\mathrm{ML},\gamma,s}$ has value at most*

$$\mathbf{E}^+ Q + 2\xi + 2(1-\gamma)^{3J} + \frac{2^9}{\sqrt[3]{s}} + \frac{\sqrt{\varepsilon_{LC}}}{\gamma}.$$

**Proof.** Deferred to Section 6. ◀

For appropriate choices of constants, the completeness and soundness of the reduction implies Theorem 3 – the uselessness for every $Q$ such that $\hat{Q}_{[4]} \geq 0$.

# 6 Analysis of the Multilayered Protocol

In this section we complete the analysis of the multilayered protocol $R_{\mathrm{ML},\gamma,s}$. The analysis is split into five parts. First, we argue that the claimed completeness and soundness of $R_{\mathrm{ML},\gamma,s}$ indeed implies Theorem 3 – the positive uselessness of constraint-functions $Q$ satisfying $\hat{Q}_{[4]} \geq 0$. Second, we give an outline of the soundness analysis of $Q$, split its proof into three lemmas, and argue that they imply the desired soundness. These three lemmas are subsequently proved, and respectively analyze the decoupling of table arguments, properties of the layer-distribution $\mathcal{D}_s$, and the bounding the value of decoupled evaluations of $Q$ with averaged evaluations of $Q$.

## 6.1 Uselessness from the Protocol $R_{\mathrm{ML},\gamma,s}$

We argue that Theorem 3 in Section 2 follows from the supposed completeness and soundness of the multilayered protocol.

**Proof of Theorem 3.** The lemma follows if we can argue that, assuming $\mathsf{P} \neq \mathsf{NP}$, for arbitrary $\varepsilon' > 0$ and $Q' : \{-1,1\}^4 \to \mathbb{R}$ such that $\hat{Q'}_{[4]} \geq 0$, given a MAX CSP$^+$ instance with value at least $1-\varepsilon$ when evaluated for 4-Odd, there is no polynomial-time algorithm to determine if the value is at least $\mathbf{E}^+ Q' + \varepsilon'$ evaluated for $Q'$. If $\min Q' = \max Q'$, the claim is trivial. Otherwise, we introduce the constraint function $Q = \frac{Q' - \min Q'}{\max Q' - \min Q'}$ and show the statement for $\varepsilon = \varepsilon'(\max Q' - \min Q')$. Note in particular that the codomain of $Q'$ is $[0,1]$.

For the sake of contradiction, suppose that we had a polynomial-time algorithm $A$ as above for some $\varepsilon > 0$. Let $\gamma = \varepsilon/4; \xi = \varepsilon/8; J \geq \ln_{1-\gamma}(\varepsilon/8); q \geq 2^{10}\varepsilon^{-1}; s \geq 2^6 q \varepsilon^{-2}$; and $\varepsilon_{\mathrm{LC}} = \gamma^2\varepsilon^2/16$. From Theorem 8, it is $\mathsf{NP}$-hard to distinguish $2^s$-multilayered LC instances of value 1 from path-samplable $(J,\xi)$-smooth instances of value at most $\varepsilon_{\mathrm{LC}}$. Given such an instance $\mathcal{I}$, consider running the supposed algorithm $A$ on the $R_{\mathrm{ML},\gamma,s}(\mathcal{I})$. From Theorem 13, $R_{\mathrm{ML},\gamma,s}(\mathcal{I})$ has value at least $1-\varepsilon$ evaluated for 4-Odd. Consequently, $A$ produces in polynomial time an instance of value at least $\mathbf{E}^+ Q + \varepsilon$ for $Q$ by the choice of parameters, contradicting the assumption that $\mathsf{P} \neq \mathsf{NP}$. ◀

## 6.2 Properties of the Layer Distribution $\mathcal{D}_s$

We prove the property Theorem 11 of the layer-distribution $\mathcal{D}_s$.

▶ **Lemma 15** (Special case $k=2$ of Lemma 11.3, [8]). *Let $f$ be an arbitrary function from $[2^s]$ to a set $S$. When $\mathbf{r},\mathbf{p},\mathbf{a},\mathbf{b}$ are chosen as in Theorem 10, for a random $\mathbf{j} \in \{0,1\}$,*

$$\sum_{\sigma \in S} \mathop{\mathbf{E}}_{\mathbf{r},\mathbf{p},\mathbf{j}} \left[ \left| \mathop{\mathbf{P}}_{\mathbf{a} \sim \mathcal{D}_{\mathbf{r},\mathbf{p}}} \{f(\mathbf{a}) = \sigma\} - \mathop{\mathbf{P}}_{\mathbf{a} \sim \mathcal{D}_{\mathbf{r}-1,2\mathbf{p}+\mathbf{j}}} \{f(\mathbf{a}) = \sigma\} \right| \right] \leq \sqrt{\frac{|S|}{s}}.$$

Given two functions $f, g : [2^s] \to S$, Theorem 15 implies that $(f(\mathbf{a}), g(\mathbf{b}))$ where $(\mathbf{a},\mathbf{b}) \sim \mathcal{D}_s$ is close in distribution to $(f(\mathbf{a}), g(\mathbf{b}))$ where $\mathbf{a},\mathbf{b}$ are independently drawn from $\mathcal{D}_{\mathbf{r},\mathbf{p}}$ where $\mathbf{r} \sim [s], \mathbf{p} \sim \left[2^{s-\mathbf{r}-1}\right]$. In particular, although $\mathbf{a} < \mathbf{b}$ when drawn from $\mathcal{D}_s$, the same event only occurs with probability roughly 1/2 in the latter case.

The decoupling Theorem 11 stated in Section 5 is a corollary of Theorem 15. We refer the interested reader to the full version for the proof.

## 6.3 Soundness of $R_{\mathrm{ML},\gamma,s}$

In this section, we prove the claimed soundness Theorem 14 of the $R_{\mathrm{ML},\gamma,s}$.

### 6.3.1 Notation

Let $\rho \stackrel{\text{def}}{=} 1-\gamma$ and for natural numbers $x, y$, let $x \nmid y$ denote that $x$ does not divide $y$. Parametrized by an edge $e \in E$, we define the test distribution $\mathcal{T}^e = \mathcal{T}^{(u,v)}$ which independently samples $\{\zeta^{(t)}\}_{t=0}^{3}$ as random $\rho$-biased strings; draws $\vec{\mathbf{x}}, \vec{\mathbf{y}}^{(1)}, \vec{\mathbf{y}}^{(2)}$ uniformly at random; and for every $j \in L_b$, sets $\mathbf{y}_j^{(3)} = -\mathbf{x}_{\pi^{(e)}(j)}\mathbf{y}_j^{(1)}\mathbf{y}_j^{(2)}$. For notational clarity, we shall let the vertices $\mathbf{w_a}$, and $\mathbf{w_b}$ be implicit and denote $\mathbf{N} = N(\mathbf{w_b}) \cap V_\mathbf{a}$, $\pi = \pi^{(\mathbf{w_b},\mathbf{w_a})}$, $\mathbf{f} = f^{\mathbf{w_a}}$, $\mathbf{g} = f^{(\mathbf{w_b})}$, and $\mathcal{T} = \mathcal{T}^{(\mathbf{w_a},\mathbf{w_b})}$. For a sequence of vertices $\vec{w} \in V^{(0)} \times \cdots \times V^{(2^s-1)}$, introduce the function $\delta^{\vec{w}} : [2^s] \to [-1,1]$ as a shorthand for the bias of the table in Layer $a$, i.e., $\delta^{\vec{w}}(a) \stackrel{\text{def}}{=} \mathbf{E}_{\vec{\mathbf{x}}}\left[f^{(w_a)}(\vec{\mathbf{x}})\right]$. Finally, introduce the value of the protocol as

$$\mathrm{Val} \stackrel{\text{def}}{=} \mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e}=(\mathbf{w_a},\mathbf{w_b}),\mathcal{T}}\left[Q\left(\mathbf{f}(\zeta^{(0)}\vec{\mathbf{x}}), \mathbf{g}(\zeta^{(1)}\vec{\mathbf{y}}^{(1)}), \mathbf{g}(\zeta^{(2)}\vec{\mathbf{y}}^{(2)}), \mathbf{g}(\zeta^{(3)}\vec{\mathbf{y}}^{(3)})\right)\right],$$

and the *argument-decoupled value*,

$$\mathrm{Val}_\perp \stackrel{\text{def}}{=} \mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e}=(\mathbf{w_a},\mathbf{w_b}),\mathcal{T}}\left[Q\left(\mathbf{E}[\mathbf{f}], \mathbf{E}[\mathbf{g}], \mathbf{E}[\mathbf{g}], \mathbf{E}[\mathbf{g}]\right)\right].$$

### 6.3.2 Outline

The proof outline is as follows. We argue that for low-valued LC instances, the value of the protocol is not significantly altered by drawing the arguments to the queried tables independently and uniformly at random. That is, the value of the protocol is approximately that of $Q\left(\delta^{\vec{\mathbf{w}}}(\mathbf{a}), \delta^{\vec{\mathbf{w}}}(\mathbf{b}), \delta^{\vec{\mathbf{w}}}(\mathbf{b}), \delta^{\vec{\mathbf{w}}}(\mathbf{b})\right)$ over random choices of the path $\vec{\mathbf{w}}$ and layer indices $\mathbf{a} < \mathbf{b}$. Exploiting the the distribution $\mathcal{D}_s$ of $(\mathbf{a}, \mathbf{b})$, the value is roughly equal when choosing $\mathbf{a}$ and $\mathbf{b}$ independently from a random distribution $\mathcal{D}_{\mathbf{r},\mathbf{p}}$. When $\mathbf{a}$ and $\mathbf{b}$ are drawn independently, we are able to compare the value to $Q\left(\frac{\delta^{\vec{\mathbf{w}}}(\mathbf{a})+\delta^{\vec{\mathbf{w}}}(\mathbf{b})}{2}, \ldots, \frac{\delta^{\vec{\mathbf{w}}}(\mathbf{a})+\delta^{\vec{\mathbf{w}}}(\mathbf{b})}{2}\right)$; which can be seen as drawing two tables and for each query picking a random point from one of the two tables with equal probability. In fact, this value and $Q\left(\delta^{\vec{\mathbf{w}}}(\mathbf{a}), \delta^{\vec{\mathbf{w}}}(\mathbf{b}), \delta^{\vec{\mathbf{w}}}(\mathbf{b}), \delta^{\vec{\mathbf{w}}}(\mathbf{b})\right)$ agree up to and including third moments. More carefully, the difference between the value of the protocol and the value of the random queries is given by $-\hat{Q}_{[4]} \cdot \left(\frac{\delta^{\vec{\mathbf{w}}}(\mathbf{a})-\delta^{\vec{\mathbf{w}}}(\mathbf{b})}{2}\right)^4$ which is non-positive for $\hat{Q}_{[4]} \geq 0$, in favor of the random queries. Consequently, the value of the protocol is approximately bounded from above by $Q\left(\frac{\delta^{\vec{\mathbf{w}}}(\mathbf{a})+\delta^{\vec{\mathbf{w}}}(\mathbf{b})}{2}, \ldots, \frac{\delta^{\vec{\mathbf{w}}}(\mathbf{a})+\delta^{\vec{\mathbf{w}}}(\mathbf{b})}{2}\right)$, which in turn is bounded by the maximum over independent biased bits: $\mathbf{E}^+Q = \max_{b^*} Q(b^*, b^*, b^*, b^*)$.

### 6.3.3 Steps of the Soundness Analysis

The formal proof is divided into three steps with their respective lemmas. The first lemma argues that we can decouple table arguments when reducing from low-valued LC instances. The methods used to prove this lemma are standard and should come as no surprise to those familiar with the analysis of LC reductions.

▶ **Lemma 16.** *When the reduced-from* LC *instance is* $(J, \xi)$*-smooth and of value at most* $\varepsilon_{LC}$*, the value produced by the reduction* $R_{\mathrm{ML},\gamma,s}$ *satisfies*

$$|\mathrm{Val} - \mathrm{Val}_\perp| \leq 2\rho^{3J} + 2\xi + \frac{\sqrt{\varepsilon_{LC}}}{\gamma}.$$

In the second step, we prove properties of the layer-distribution $\mathcal{D}_s$, Theorem 11, and that decoupled arguments drawn according to the $\mathcal{D}_s$ distribution has a simple expression in terms of the constraint-function $Q$.

▶ **Lemma 17.** *Using properties of the layer distribution $\mathcal{D}_s$, over some distribution of pairs $(\mathbf{x}, \mathbf{y}) \in [-1, 1]$, for every $s \geq 2$,*

$$\left| \mathrm{Val}_\perp - \mathbf{E}\left[\frac{Q(\mathbf{x}, \mathbf{y}, \mathbf{y}, \mathbf{y}) + Q(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{y})}{2}\right]\right| \leq \frac{2^9}{\sqrt[3]{s}}.$$

Finally, we argue that regardless of the queried pair, when $\hat{Q}_{[4]} \geq 0$, evaluating $Q$ using one random element of the pair once and the other thrice, is never beneficial in comparison to evaluating $Q$ only with their average.

▶ **Lemma 18.** *For any $x, y \in \mathbb{R}$ and symmetric constraint-function $Q$ s.t. $\hat{Q}_{[4]} \geq 0$,*

$$\frac{Q(x, y, y, y) + Q(x, x, x, y)}{2} \leq Q\left(\frac{x+y}{2}, \frac{x+y}{2}, \frac{x+y}{2}, \frac{x+y}{2}\right).$$

More generally, and recalling that $p \nmid q$ denotes that $p$ does not divide $q$, Theorem 18 is a corollary of the following lemma.

▶ **Lemma 19.** *For any $x, y \in \mathbb{R}$, even $k$, and (multilinear extension of a) symmetric constraint function $Q : \{-1, 1\}^k \to \mathbb{R}$ s.t. $\hat{Q}_{[k]} \geq 0$,*

$$\mathop{\mathbf{E}}_{\mathbf{k}_0 \sim \mathrm{Bin}(k, \frac{1}{2}) \,:\, 2 \nmid \mathbf{k}_0}\Big[Q(\underbrace{x, \ldots, x}_{\mathbf{k}_0}, \underbrace{y, \ldots, y}_{k - \mathbf{k}_0})\Big] \leq Q\Big(\underbrace{\frac{x+y}{2}, \ldots, \frac{x+y}{2}}_{k}\Big).$$

## 6.3.4 Soundness from Step Lemmas

Theorems 16 to 18 indeed implies the soundness of the reduction.

**Proof of Theorem 14.** Applying the three lemmas, over some distribution of pairs $(\mathbf{x}, \mathbf{y}) \in [-1, 1]$,

$$\mathrm{Val} \leq \mathbf{E}\left[Q\left(\frac{\mathbf{x}+\mathbf{y}}{2}, \frac{\mathbf{x}+\mathbf{y}}{2}, \frac{\mathbf{x}+\mathbf{y}}{2}, \frac{\mathbf{x}+\mathbf{y}}{2}\right)\right] + 2\xi + 2\rho^{3J} + \frac{\sqrt{\varepsilon_{\mathrm{LC}}}}{\gamma} + \frac{2^9}{\sqrt[3]{s}}.$$

Since the expectation of random variable is always bounded by its maximum, the first term is at most $\max_{b^*} Q(b^*, b^*, b^*, b^*) = \mathbf{E}^+ Q$ and hence corresponds to the stated bound. ◀

## 6.3.5 Decoupling Table Arguments

**Proof of Theorem 16.** The value of the protocol is given by

$$\mathrm{Val} \stackrel{\mathrm{def}}{=} \mathop{\mathbf{E}}_{\mathcal{D}_s, \mathbf{e}=(\mathbf{w_a}, \mathbf{w_b}), \mathcal{T}}\left[Q\left(\mathbf{f}(\zeta^{(0)}\vec{\mathbf{x}}), \mathbf{g}(\zeta^{(1)}\vec{\mathbf{y}}^{(1)}), \mathbf{g}(\zeta^{(2)}\vec{\mathbf{y}}^{(2)}), \mathbf{g}(\zeta^{(3)}\vec{\mathbf{y}}^{(3)})\right)\right]. \tag{2}$$

Using the Fourier decomposition of $Q$ and the definition of the noise operator,

$$(2) = \sum_{\Gamma \subseteq [4]} \hat{Q}_\Gamma \mathop{\mathbf{E}}_{\mathcal{D}_s, \mathcal{T}}\left[\mathrm{T}_\rho f(\vec{\mathbf{x}})^{1\{0 \in \Gamma\}} \prod_{t \in \Gamma \setminus \{0\}} \mathrm{T}_\rho g(\vec{\mathbf{y}}^{(t)})\right]. \tag{3}$$

We would like to compare this value to that of decoupled table arguments,

$$\mathrm{Val}_\perp \stackrel{\mathrm{def}}{=} \mathop{\mathbf{E}}_{\vec{\mathbf{w}}, \mathcal{D}_s}\left[Q\left(\delta^{\vec{\mathbf{w}}}(\mathbf{a}), \delta^{\vec{\mathbf{w}}}(\mathbf{b}), \delta^{\vec{\mathbf{w}}}(\mathbf{b}), \delta^{\vec{\mathbf{w}}}(\mathbf{b})\right)\right] = \mathop{\mathbf{E}}_{\mathcal{D}_s, \mathbf{e}}[Q(\mathbf{E}\,\mathbf{f}, \mathbf{E}\,\mathbf{g}, \mathbf{E}\,\mathbf{g}, \mathbf{E}\,\mathbf{g})]$$

$$= \sum_{\Gamma \subseteq [4]} \hat{Q}_\Gamma \mathop{\mathbf{E}}_{\mathcal{D}_s, \mathbf{e}}\left[(\mathbf{E}\,\mathbf{f})^{1\{0 \in \Gamma\}} \prod_{t \in \Gamma \setminus \{0\}} \mathbf{E}\,\mathbf{g}\right]. \tag{4}$$

Due to independence in the test distribution $\mathcal{T}$, all terms in the expansion of $Q$ agree in Eq. (3) and Eq. (4), with exception of $\Gamma = \{1,2,3\}$ and $\Gamma = \{0,1,2,3\}$. Since we do not necessarily have balanced tables, our analysis cannot entirely follow Håstad's analysis [10] and instead we additionally employ smoothness similar to, e.g., Holmerin and Khot [12].

Note that if we can provide an absolute bound on the case $\Gamma = [4]$ for every pair of layers $a < b$ and every assignment to tables $\{f^{(w)}\}_{w \in V_a \cup V_b}$, then we also have a bound on the case $\Gamma = \{1,2,3\}$ by setting $f^{(w)} \equiv 1$ for all $w \in V_a$.

Hence for arbitrary fixed $a < b$ we proceed to bound the error for $\Gamma = [4]$:

$$\mathop{\mathbf{E}}_{\mathcal{D}_s, \mathbf{e}=(\mathbf{w_a},\mathbf{w_b}), \mathcal{T}} \left[ \mathbf{f}(\zeta^{(0)}\vec{\mathbf{x}}) \prod_{t=1}^{3} \mathbf{g}(\zeta^{(t)}\vec{\mathbf{y}}^{(t)}) \right]. \tag{5}$$

The following section shows that for LC instances of small value, the expectation Eq. (5) does not change much when the arguments are drawn independently. In particular, we bound

$$\left| \mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e}} \left[ \mathop{\mathbf{E}}_{\mathcal{T}} \left[ \mathbf{f}(\zeta^{(0)}\vec{\mathbf{x}}) \prod_{t=1}^{3} \mathbf{g}(\zeta^{(t)}\vec{\mathbf{y}}^{(t)}) \right] - \mathbf{E}\,\mathbf{f} \prod_{t=1}^{3} \mathbf{E}\,\mathbf{g} \right] \right|. \tag{6}$$

Taking the Fourier expansions of the functions, Eq. (6) equals

$$\left| \mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e},\mathcal{T}} \left[ \mathrm{T}_\rho \mathbf{f}(\vec{\mathbf{x}}) \prod_{t=1}^{3} \mathrm{T}_\rho \mathbf{g}(\vec{\mathbf{y}}^{(t)}) - \mathbf{E}\,\mathbf{f} \prod_{t=1}^{3} \mathbf{E}\,\mathbf{g} \right] \right| \tag{7}$$

$$= \left| \mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e}} \left[ \sum_{\substack{S \subseteq L_\mathbf{a} \\ T_1,T_2,T_3 \subseteq L_b \\ S \cup T_1 \cup T_2 \cup T_3 \neq \emptyset}} \rho^{|S|} \hat{\mathbf{f}}_S \cdot \left( \prod_{t=1}^{3} \rho^{|T|} \hat{\mathbf{g}}_{T_t} \right) \mathop{\mathbf{E}}_{\mathcal{T}} \left[ \left( \prod_{i \in S} \mathbf{x}_i \right) \left( \prod_{t=1}^{3} \prod_{j \in T_t} \mathbf{y}_j^{(t)} \right) \right] \right] \right|, \tag{8}$$

where we recall that $\mathbf{E}\,\mathbf{f} \prod_{t=1}^{3} \mathbf{E}\,\mathbf{g}$ expands to $\hat{\mathbf{f}}_\emptyset \prod_{t=1}^{3} \hat{\mathbf{g}}_\emptyset$, explaining the condition in the summation.

Using the definition of the test distribution $\mathcal{T}$, the inner expectation evaluates to 1 precisely when $S = \pi_2(T_3); T_1 = T_2 = T_3$ and otherwise to 0:

$$(8) = \left| \mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e}} \left[ \sum_{\emptyset \neq T \subseteq L_\mathbf{b}} \rho^{|\pi_2(T)|+3|T|} \hat{\mathbf{f}}_{\pi_2(T)} \hat{\mathbf{g}}_T^3 \right] \right| \tag{9}$$

$$\leq \left| \mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e}} \left[ \sum_{\substack{T \neq \emptyset \\ \pi_2(T) = \emptyset}} \rho^{3|T|} \hat{\mathbf{f}}_\emptyset \hat{\mathbf{g}}_T^3 \right] \right| + \left| \mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e}} \left[ \sum_{\substack{T \neq \emptyset \\ \pi_2(T) \neq \emptyset}} \rho^{|\pi_2(T)|+3|T|} \hat{\mathbf{f}}_{\pi_2(T)} \hat{\mathbf{g}}_T^3 \right] \right|. \tag{10}$$

We bound separately the terms with $\pi_2(T)$ empty resp. non-empty.

## Case $\pi_2(T)$ empty

Consider first the sum of terms with $\pi_2(T) = \emptyset$ and rewrite the expression as

$$\left| \mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{w_a},\mathbf{w_b}} \left[ \sum_{T \neq \emptyset} (\hat{\mathbf{g}}_T^2) \hat{\mathbf{f}}_\emptyset \hat{\mathbf{g}}_T \rho^{3|T|} 1 \left\{ \pi_2^{(\mathbf{w_b},\mathbf{w_a})}(T) = \emptyset \right\} \right] \right|. \tag{11}$$

Using $\left|\hat{\mathbf{f}}_{\emptyset}\hat{\mathbf{g}}_T\right| \le 1$,

$$(11) \le \left|\mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{w_b}}\left[\sum_{T\neq\emptyset}\hat{\mathbf{g}}_T^2\rho^{3|T|}\mathop{\mathbf{P}}_{\mathbf{w_a}\sim\mathbf{N}}\left\{\pi_2^{(\mathbf{w_b},\mathbf{w_a})}(T)=\emptyset\right\}\right]\right|. \tag{12}$$

Since $\sum\hat{\mathbf{g}}_T^2 \le 1$ via Parseval's identity, Eq. (12) is bounded by

$$\mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{w_b}}\left[\max_{T\neq\emptyset}\rho^{3|T|}\mathop{\mathbf{P}}_{\mathbf{w_a}\sim\mathbf{N}}\{\pi_2^{\mathbf{w_b},\mathbf{w_a}}(T)=\emptyset\}\right]. \tag{13}$$

When the reduced-from LC instance is $(J,\xi)$-smooth, the probability in the expression is by definition at most $\xi$ whenever $|T| \le J$. For larger sets, the factor $\rho^{3|T|}$ is at most $\rho^{3J}$. Consequently, we have a bound on the terms with $\pi_2(T)=\emptyset$,

$$(11) \le (13) \le \rho^{3J}+\xi. \tag{14}$$

### Case $\pi_2(T)$ non-empty

We proceed to bound terms satisfying $\pi_2(T)\neq\emptyset$. Using the Cauchy-Schwarz Inequality and that $\sum\mathbf{g}_T^4 \le 1$ via Parseval's identity, the second term in Eq. (9) is bounded by

$$\mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e}}\left[\left(\sum_{T:\pi_2(T)\neq\emptyset}\rho^{2|\pi_2(T)|+6|T|}\hat{\mathbf{f}}_{\pi_2(T)}^2\hat{\mathbf{g}}_T^2\right)^{\frac{1}{2}}\left(\sum_T\mathbf{g}_T^4\right)^{\frac{1}{2}}\right], \tag{15}$$

$$\le \max_{k>0}\{k\rho^k\}\mathop{\mathbf{E}}_{\mathcal{D}_s,\mathbf{e}}\left[\left(\sum_{T:\pi_2(T)\neq\emptyset}\frac{1}{|\pi_2(T)||T|}\hat{\mathbf{f}}_{\pi_2(T)}^2\hat{\mathbf{g}}_T^2\right)^{\frac{1}{2}}\right], \tag{16}$$

where the factor $\max_{k>0}\{k\rho^k\}$ is an upper bound on the ratio between $\left(\rho^{|\pi_2(T)|+|T|}\right)^{\frac{1}{2}}$ and $\left(\frac{1}{|\pi_2(T)||T|}\right)^{\frac{1}{2}}$. Note that an upper bound on $k\rho^k$ is in turn $\rho^1+\cdots+\rho^k \le (1-\rho)^{-1}=\gamma^{-1}$.

When Eq. (16) is significant, one expects to be able to derive a good labeling of the reduced-from LC instance. The labeling strategy we consider in showing this is the natural generalization of Håstad classical decoding. Given an assignment of the tables $\{f^{(w)}\}_{w\in\biguplus V_r}$, for every vertex $w\in\biguplus V_r$, choose the label $i$ with probability $\sum_{S\ni i}\frac{1}{|S|}\widehat{f^{(w)}}_S^2$, and with the remaining probability choose an arbitrary label. This indeed defines probability distributions since, via Parseval's Identity, $\sum_i\sum_{S\ni i}\frac{1}{|S|}\widehat{f^{(w)}}_S^2 \le \sum_S\widehat{f^{(w)}}_S^2 = 1$.

Between any two layers $a<b$, the labeling satisfies at least a fraction of constraints,

$$\mathop{\mathbf{E}}_{\mathbf{e}=(\mathbf{w}_a,\mathbf{w}_b)\sim E(V_a,V_b)}\left[\sum_{(i,j)\in\pi^{(\mathbf{e})}}\left(\sum_{S\ni i}\frac{1}{|S|}\widehat{\mathbf{f}}_S^2\right)\left(\sum_{T\ni j}\frac{1}{|T|}\widehat{\mathbf{g}}_T^2\right)\right]$$

$$=\mathop{\mathbf{E}}_{\mathbf{e}}\left[\sum_{S,T}\#\{(i,j)\in\pi\ :\ i\in S, j\in T\}\frac{\hat{\mathbf{f}}_S^2\hat{\mathbf{g}}_T^2}{|S||T|}\right] \ge \mathop{\mathbf{E}}_{\mathbf{e}}\left[\sum_{T:\pi_2(T)\neq\emptyset}\frac{1}{|\pi_2(T)||T|}\hat{\mathbf{f}}_{\pi_2(T)}^2\hat{\mathbf{g}}_T^2\right]. \tag{17}$$

Since the value of a multilayered LC instance was defined as the maximum fraction of satisfied edges between two distinct layers, Eq. (17) is bounded from above by the value of the reduced-from LC instance, which in turn by assumption is at most $\varepsilon_{\mathrm{LC}}$.

Returning to the value of considered term,

$$(16) \leq \gamma^{-1} \mathop{\mathbf{E}}_{\mathcal{D}_s, \mathbf{e}} \left[ \left( \sum_{T: \pi_2(T) \neq \emptyset} \frac{1}{|\pi_2(T)||T|} \hat{\mathbf{f}}^2_{\pi_2(T)} \hat{\mathbf{g}}^2_T \right)^{\frac{1}{2}} \right] \leq \frac{\sqrt{\varepsilon_{\mathrm{LC}}}}{\gamma}. \tag{18}$$

### Completing the Argument Decoupling

Combining the bounds Eqs. (14) and (18), the error introduced by sampling the arguments independently for the term $\Gamma = [4]$ is $(5) \leq \rho^{3J} + \xi + \frac{\sqrt{\varepsilon_{\mathrm{LC}}}}{\gamma}$. From this, Eq. (14), and that $Q \to [0,1]$, we conclude

$$|\mathrm{Val} - \mathrm{Val}_{\perp}| \leq 2\rho^{3J} + 2\xi + \frac{\sqrt{\varepsilon_{\mathrm{LC}}}}{\gamma}. \tag{19}$$

◀

## 6.3.6 Decoupled Value to Symmetric Evaluation

We proceed to prove Theorem 17 – relating the decoupled value to an expectation of $Q$ evaluated symmetrically for a random pair.

**Proof of Theorem 17.** Taking the Fourier expansion of the constraint function $Q$ and recalling the definition of $\delta^{\vec{w}}$,

$$\mathrm{Val}_{\perp} = \sum \hat{Q}_\Gamma \mathop{\mathbf{E}}_{\mathcal{D}_s, \mathbf{e}} \left[ \mathbf{E}\left[\mathbf{f}\right]^{1\{0 \in \Gamma\}} \mathbf{E}\left[\mathbf{g}\right]^{|\Gamma \setminus \{0\}|} \right] = \sum \hat{Q}_\Gamma \mathop{\mathbf{E}}_{\vec{\mathbf{w}}, \mathcal{D}_s} \left[ \delta^{\vec{\mathbf{w}}}(\mathbf{a})^{1\{0 \in \Gamma\}} \delta^{\vec{\mathbf{w}}}(\mathbf{b})^{|\Gamma \setminus \{0\}|} \right].$$

It is an issue for our analysis that the pair $(\mathbf{a}, \mathbf{b}) \sim \mathcal{D}_s$ is always chosen so that $\mathbf{a} < \mathbf{b}$. However, due to the choice of the distribution $\mathcal{D}_s$, the value of two functions evaluated for $(\mathbf{a}, \mathbf{b})$ is roughly unchanged when $\mathbf{a}$ and $\mathbf{b}$ are drawn independently from a random distribution. More specifically, using Theorem 11, for every choice of $\vec{w}$,

$$\left| \mathop{\mathbf{E}}_{\mathcal{D}_s} \left[ \delta^{\vec{w}}(\mathbf{a})^{k_1} \delta^{\vec{w}}(\mathbf{b})^{k_2} \right] - \mathop{\mathbf{E}}_{\mathbf{r}, \mathbf{p}} \left[ \mathop{\mathbf{E}}_{\mathcal{D}_{\mathbf{r},\mathbf{p}}} \left[ \delta^{\vec{w}}(\mathbf{a})^{k_1} \right] \mathop{\mathbf{E}}_{\mathcal{D}_{\mathbf{r},\mathbf{p}}} \left[ \delta^{\vec{w}}(\mathbf{b})^{k_2} \right] \right] \right| \leq 2\frac{k_1 + k_2}{q} + 8\sqrt{\frac{q}{s}}.$$

Applying the bound once for every $\Gamma \subseteq [4]$,

$$\left| \mathrm{Val}_{\perp} - \sum \hat{Q}_\Gamma \mathop{\mathbf{E}}_{\vec{\mathbf{w}}, \mathbf{r}, \mathbf{p}} \left[ \mathop{\mathbf{E}}_{\mathcal{D}_{\mathbf{r},\mathbf{p}}} \left[ \delta^{\vec{\mathbf{w}}}(\mathbf{a})^{1\{0 \in \Gamma\}} \right] \mathop{\mathbf{E}}_{\mathcal{D}_{\mathbf{r},\mathbf{p}}} \left[ \delta^{\vec{\mathbf{w}}}(\mathbf{b})^{|\Gamma \setminus \{0\}|} \right] \right] \right| \leq \min_q \left( \frac{2^7}{q} + 2^7 \sqrt{\frac{q}{s}} \right).$$

Undoing the expansion of $Q$,

$$\left| \mathrm{Val}_{\perp} - \mathop{\mathbf{E}}_{\vec{\mathbf{w}}, \mathbf{r}, \mathbf{p}} \left[ \mathop{\mathbf{E}}_{\mathbf{a}, \mathbf{b} \sim \mathcal{D}_{\mathbf{r},\mathbf{p}}} \left[ Q(\mathbf{x}, \mathbf{y}, \mathbf{y}, \mathbf{y}) \right] \right] \right| \leq \min_q \left( \frac{2^7}{q} + 2^7 \sqrt{\frac{q}{s}} \right),$$

$\mathbf{x} = \delta^{\mathbf{w}}(\mathbf{a})$ and $\mathbf{y} = \delta^{\mathbf{w}}(\mathbf{b})$. Since $\mathbf{a}$ and $\mathbf{b}$ are drawn independently from the same distribution, this indeed yields the desired bound for $s \geq 2$:

$$\left| \mathrm{Val}_{\perp} - \mathop{\mathbf{E}}_{\vec{\mathbf{w}}, \mathbf{a}, \mathbf{b}} \left[ \frac{Q(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{y}) + Q(\mathbf{x}, \mathbf{y}, \mathbf{y}, \mathbf{y})}{2} \right] \right| \leq \min_{q \in \mathbb{Z}^{\geq 1}} \left( \frac{2^7}{q} + 2^7 \sqrt{\frac{q}{s}} \right) \leq \frac{2^9}{\sqrt[3]{s}}.$$

◀

### 6.3.7 Symmetric Evaluation to the Value of Querying Random Tables

We prove Theorem 18 with the more general Theorem 19.

**Proof of Theorem 19.** Fix arbitrary $x, y \geq \mathbb{R}$, $k$ even, and symmetric $Q : \{-1,1\}^k \to \mathbb{R}$ such that $\hat{Q}_{[k]} \geq 0$. Since $Q$ is symmetric, for every $m \in [k]$, we can introduce constants $\{\hat{Q}_m\}_{m\in[k]}$ such that for all $\Gamma \subseteq [k]$, $\hat{Q}_\Gamma = \hat{Q}_{|\Gamma|}$. We wish to upper-bound the expression

$$\mathop{\mathbf{E}}_{\mathbf{k}_0 \sim \mathrm{Bin}(k,\frac{1}{2}) \ | \ 2 \nmid \mathbf{k}_0}\left[Q(\underbrace{x,\ldots,x}_{\mathbf{k}_0},\underbrace{y,\ldots,y}_{k-\mathbf{k}_0})\right] \tag{20}$$

$$= \sum_{m=0}^{k} \hat{Q}_m \sum_{\Gamma : |\Gamma|=m} \mathop{\mathbf{E}}_{\mathbf{k}_0 \sim \mathrm{Bin}(k,\frac{1}{2}) \ | \ 2 \nmid \mathbf{k}_0}\left[x^{|\Gamma\cap[\mathbf{k}_0]|} y^{|\Gamma \cap \{\mathbf{k}_0,\ldots,k-1\}|}\right]$$

with

$$Q\Big(\underbrace{\frac{x+y}{2},\ldots,\frac{x+y}{2}}_{k}\Big) = \sum_{m=0}^{k} \hat{Q}_m \binom{k}{m}\left(\frac{x+y}{2}\right)^m. \tag{21}$$

By symmetry,

$$(20) = \sum_{m=0}^{k} \hat{Q}_m \binom{k}{m} \mathop{\mathbf{E}}_{\mathbf{S}\subseteq[k] \, : \, 2\nmid|S|}\left[x^{|\mathbf{S}\cap[m]|} y^{|\bar{\mathbf{S}}\cap[m]|}\right]. \tag{22}$$

Since the elements in $S$ are drawn $(k-1)$-wise independently, for $m < k$, the respective terms simply evaluate to $\hat{Q}_m\binom{k}{m}\left(\frac{x+y}{2}\right)^m$ which agree with the corresponding terms in Eq. (21).

We consider the term $m = k$ and use that $k$ is even,

$$\hat{Q}_k \sum_{\Gamma : |\Gamma|=k} \mathop{\mathbf{E}}_{\mathbf{k}_0 \sim \mathrm{Bin}(k,\frac{1}{2}) \ | \ 2 \nmid \mathbf{k}_0}\left[x^{|\Gamma\cap[\mathbf{k}_0]|} y^{|\Gamma \cap \{\mathbf{k}_0,\ldots,k-1\}|}\right]$$

$$= \hat{Q}_k \mathop{\mathbf{P}}_{\mathbf{S}\subseteq[k]}\{2 \nmid |S|\}^{-1} \mathop{\mathbf{E}}_{\mathbf{S}\subseteq[k]}\left[x^{|\mathbf{S}|}y^{|\bar{\mathbf{S}}|}1\{2\nmid|S|\}\right]$$

$$= \hat{Q}_k \mathop{\mathbf{E}}_{\mathbf{S}\subseteq[k]}\left[x^{|\mathbf{S}|}y^{k-|\mathbf{S}|}\left(1-(-1)^{k-|\mathbf{S}|}\right)\right]$$

$$= \hat{Q}_k\left(\left(\frac{x+y}{2}\right)^m - \left(\frac{x-y}{2}\right)^m\right) \leq \hat{Q}_k\left(\frac{x+y}{2}\right)^m,$$

where the final step uses that $\hat{Q}_k \geq 0$. The upper bound agrees with the term $m = k$ in Eq. (21), establishing the lemma. ◀

**Acknowledgments.** The author is grateful to Johan Håstad and Lukàš Polàček for many fruitful discussions and suggestions on this project, to anonymous referees for constructive comments, and to the hospitable Simons Institute for the Theory of Computing where part of this work was done.

## References

1 Per Austrin and Johan Håstad. On the usefulness of predicates. *TOCT*, 5(1):1, 2013.
2 Per Austrin and Elchanan Mossel. Approximation resistant predicates from pairwise independence. *Computational Complexity*, 18(2):249–271, 2009.

3 Siu On Chan. Approximation resistance from pairwise independent subgroups. In *STOC*, pages 447–456, 2013.
4 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), 2005.
5 Uriel Feige, Guy Kindler, and Ryan O'Donnell. Understanding parallel repetition requires understanding foams. In *IEEE Conference on Computational Complexity*, 2007.
6 Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 1995.
7 Venkatesan Guruswami. Inapproximability results for set splitting and satisfiability problems with no mixed clauses. *Algorithmica*, 38(3):451–469, 2003.
8 Venkatesan Guruswami, Johan Håstad, Rajsekar Manokaran, Prasad Raghavendra, and Moses Charikar. Beating the random ordering is hard: Every ordering csp is approximation resistant. *SIAM J. Comput.*, 40(3):878–914, 2011.
9 Venkatesan Guruswami, Prasad Raghavendra, Rishi Saket, and Yi Wu. Bypassing ugc from some optimal geometric inapproximability results. In *SODA*, 2012.
10 Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.
11 Johan Håstad. On the approximation resistance of a random predicate. *Computational Complexity*, 18(3):413–434, 2009.
12 Jonas Holmerin and Subhash Khot. A new pcp outer verifier with applications to homogeneous linear equations and max-bisection. In *STOC*, 2004.
13 Subhash Khot. Hardness results for coloring 3 -colorable 3 -uniform hypergraphs. In *FOCS*, pages 23–32, 2002.
14 Subhash Khot. On the power of unique 2-prover 1-round games. In *STOC*, 2002.
15 Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM J. Comput.*, 37(1), 2007.
16 Subhash Khot, Madhur Tulsiani, and Pratik Worah. A characterization of strong approximation resistance. *CoRR*, abs/1305.5500, 2013.
17 Subhash Khot and Nisheeth K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into $l_1$. In *FOCS*, 2005.
18 Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis*, 19, 2010.
19 Thomas J. Schaefer. The complexity of satisfiability problems. In *STOC*, 1978.
20 Luca Trevisan, Gregory B. Sorkin, Madhu Sudan, and David P. Williamson. Gadgets, approximation, and linear programming. *SIAM J. Comput.*, 29(6):2074–2097, 2000.