Near-Optimal UGC-hardness of Approximating Max k-CSP$_R$*

Pasin Manurangsi[1], Preetum Nakkiran[2], and Luca Trevisan[3]

1 **University of California, Berkeley, USA**
`pasin@berkeley.edu`
2 **University of California, Berkeley, USA**
`preetum@berkeley.edu`
3 **University of California, Berkeley, USA**
`luca@berkeley.edu`

Abstract

In this paper, we prove an almost-optimal hardness for Max k-CSP$_R$ based on Khot's Unique Games Conjecture (UGC). In Max k-CSP$_R$, we are given a set of predicates each of which depends on exactly k variables. Each variable can take any value from $1, 2, \dots, R$. The goal is to find an assignment to variables that maximizes the number of satisfied predicates.

Assuming the Unique Games Conjecture, we show that it is NP-hard to approximate Max k-CSP$_R$ to within factor $2^{O(k \log k)}(\log R)^{k/2}/R^{k-1}$ for any k, R. To the best of our knowledge, this result improves on all the known hardness of approximation results when $3 \leq k = o(\log R/\log\log R)$. In this case, the previous best hardness result was NP-hardness of approximating within a factor $O(k/R^{k-2})$ by Chan. When $k = 2$, our result matches the best known UGC-hardness result of Khot, Kindler, Mossel and O'Donnell.

In addition, by extending an algorithm for Max 2-CSP$_R$ by Kindler, Kolla and Trevisan, we provide an $\Omega(\log R/R^{k-1})$-approximation algorithm for Max k-CSP$_R$. This algorithm implies that our inapproximability result is tight up to a factor of $2^{O(k \log k)}(\log R)^{k/2-1}$. In comparison, when $3 \leq k$ is a constant, the previously known gap was $O(R)$, which is significantly larger than our gap of $O(\text{polylog } R)$.

Finally, we show that we can replace the Unique Games Conjecture assumption with Khot's d-to-1 Conjecture and still get asymptotically the same hardness of approximation.

1998 ACM Subject Classification F.1.3 Complexity Measures and Classes

Keywords and phrases inapproximability, unique games conjecture, constraint satisfaction problem, invariance principle

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2016.15

1 Introduction

Maximum Constraint Satisfaction Problem (Max CSP) is an optimization problem where the inputs are a set of variables, an alphabet, and a set of predicates. Each variable can be assigned any symbol from the alphabet and each predicate depends only on the assignment to a subset of variables. The goal is to find an assignment to the variables that maximizes the number of satisfied predicates.

* This material is based upon work supported by the National Science Foundation under Grant No. CCF 1540685.

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016).
Editors: Klaus Jansen, Claire Matthieu, José D. P. Rolim, and Chris Umans; Article No. 15; pp. 15:1–15:28
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Many natural optimization problems, such as Max Cut, Max k-CUT and Max k-SAT, can be formulated as Max CSP. In addition, Max CSP has been shown to help approximate other seemingly-unrelated problems such as Densest k-Subgraph [4]. Due to this, Max CSP has long been researched by the approximation algorithm community [35, 18, 6, 26, 24, 14]. Furthermore, its relation to PCPs ensures that Max CSP is also well-studied on the hardness of approximation side [32, 11, 33, 22, 2, 16, 12, 3].

The main focus of this paper is on Max k-CSP$_R$, a family of Max CSP where the alphabet is of size R and each predicate depends on only k variables. On the hardness of approximation side, most early works focused on boolean Max k-CSP. Samorodnitsky and Trevisan first showed that Max k-CSP$_2$ is NP-hard to approximate to within factor $2^{O(\sqrt{k})}/2^k$ [32]. Engebretsen and Holmerin later improved constant factors in the exponent $O(\sqrt{k})$ but still yielded hardness of a factor $2^{O(\sqrt{k})}/2^k$ [12]. To break this barrier, Samorodnitsky and Trevisan proved a hardness of approximation conditioned on Khot's Unique Games Conjecture (UGC), which will be discussed in more detail later; they achieved a ratio of $O(k/2^k)$ hardness, which is tight up to a constant for the boolean case [33]. Chan later showed that NP-hardness of approximation at this ratio can be achieved unconditionally and, thus, settled down the approximability of Max k-CSP$_2$ [3].

Unlike the boolean case, the approximability of Max k-CSP$_R$ when $R > 2$ is still not resolved. In this case, Engebretsen showed $R^{O(\sqrt{k})}/R^k$ NP-hardness of approximation [11]. Under the Unique Games Conjecture, a hardness of approximation of $O(kR/R^{k-1})$ factor was proven by Austrin and Mossel [2] and, independently, by Guruswami and Raghavendra [16]. For the case $k = 2$, results by Khot et al. [22] implicitly demonstrate UGC-hardness of approximation within $O(\log R/R)$, made explicit in [24]. Moreover, Austrin and Mossel proved UGC-hardness of approximation of $O(k/R^{k-1})$ for infinitely many ks [2], but in the regime $k \geq R$. Recently, Chan was able to remove the Unique Game Conjecture assumption from these results [3]. More specifically, Chan showed NP-hardness of approximation of factor $O(kR/R^{k-1})$ for every k, R and that of factor $O(k/R^{k-1})$ for every $k \geq R$. Due to an approximation algorithm with matching approximation ratio by Makarychev and Makarychev [26], Chan's result established tight hardness of approximation for $k \geq R$. On the other hand, when $k < R$, Chan's result gives $O(kR/R^{k-1})$ hardness of approximation whereas the best known approximation algorithm achieves only $\Omega(k/R^{k-1})$ approximation ratio [26, 14]. In an attempt to bridge this gap, we prove the following theorem.

▶ **Theorem 1** (Main Theorem). *Assuming the Unique Games Conjecture, it is NP-hard to approximate* Max k-CSP$_R$ *to within* $2^{O(k \log k)}(\log R)^{k/2}/R^{k-1}$ *factor, for any* $k \geq 2$ *and any sufficiently large* R.

When $k = o(\log R/\log\log R)$, our result improves upon the previous best known hardness of approximation result in this regime, due to Chan. In particular, when k is constant, our results are tight up to a factor of $O(\text{polylog } R)$. While Chan's results hold unconditionally, our result, similar to many of the aforementioned results (e.g. [33, 2, 16]), rely on the Unique Games Conjecture.

A *unique game* is a Max 2-CSP instance where each constraint is a permutation. The *Unique Games Conjecture (UGC)*, first proposed by Khot in his seminal paper in 2002 [20], states that, for any sufficiently small $\eta, \gamma > 0$, it is NP-hard to distinguish a unique game where at least $1-\eta$ fraction of constraints can be satisfied from a unique game where at most γ fraction of constraints can be satisfied. The UGC has since made a huge impact in hardness of approximation; numerous hardness of approximation results not known unconditionally can be derived assuming the UGC. More surprisingly, UGC-hardness of approximation for

Range of k, R	NP-Hardness	UGC-Hardness	Approximation	References
$k = 2$	$O\left(\frac{\log R}{\sqrt{R}}\right)$	$O\left(\frac{\log R}{R}\right)$	$\Omega\left(\frac{\log R}{R}\right)$	[3, 22, 24]
$3 \leq k < R$	$O\left(\frac{k}{R^{k-2}}\right)$	–	$\Omega\left(\frac{k}{R^{k-1}}\right)$	[3, 26, 14]
$R \leq k$	$O\left(\frac{k}{R^{k-1}}\right)$	–	$\Omega\left(\frac{k}{R^{k-1}}\right)$	[3, 26]
Any k, R	–	$\frac{2^{O(k \log k)}(\log R)^{k/2}}{R^{k-1}}$	$\Omega\left(\frac{\log R}{R^{k-1}}\right)$	this work

Figure 1 Comparison between our work and previous works. We list the previous best known results alongside our results. From previous works, there is either an NP-hardness or a UGC-hardness result matching the best known approximation algorithm in every case except when $3 \leq k < R$. Our hardness result improves on the best known hardness result when $k = o(\log R/\log\log R)$, and our approximation algorithm improves on the previously known algorithm when $k = o(\log R)$.

various problems, such as Max Cut [22], Vertex Cover [23] and *any* Max CSP [31][1], are known to be tight. For more details on UGC and its implications, we refer interested readers to Khot's survey [21] on the topic.

Another related conjecture from [20] is the *d-to-1 Conjecture*. In the d-to-1 Conjecture, we consider *d-to-1 games* instead of unique games. A d-to-1 game is an instance of Max 2-CSP where the constraint graph is bipartite. Moreover, each constraint must be a d-to-1 function, i.e., for each assignment to a variable on the right, there exists d assignments to the corresponding variable on the left that satisfy the constraint. The d-to-1 Conjecture states that, for any sufficiently small $\gamma > 0$, it is NP-hard to distinguish between a fully satisfiable d-to-1 game and a d-to-1 game where at most γ fraction of constraints can be satisfied. Currently, it is unknown whether the d-to-1 Conjecture implies the Unique Games Conjecture and vice versa.

While the d-to-1 Conjecture has yet to enjoy the same amount of influence as the UGC, it has been proven successful in providing a basis for hardness of graph coloring problems [9, 10, 17] and for Max 3-CSP with perfect completeness [30, 34]. Here we show that, by assuming the d-to-1 Conjecture instead of UGC, we can get a similar hardness of approximation result for Max k-CSP$_R$ as stated below.

▶ **Theorem 2.** *Assuming the d-to-1 Games Conjecture holds for some d, it is NP-hard to approximate* Max k-CSP$_R$ *to within* $2^{O(k \log k)}(\log R)^{k/2}/R^{k-1}$ *factor, for any $k \geq 2$ and any sufficiently large R.*

As mentioned earlier, there has also been a long line of works in approximation algorithms for Max k-CSP$_R$. In the boolean case, Trevisan first showed a polynomial-time approximation algorithm with approximation ratio $2/2^k$ [35]. Hast later improved the ratio to $\Omega(k/(2^k \log k))$ [18]. Charikar, Makarychev and Makarychev then came up with an $\Omega(k/2^k)$-approximation algorithm [6]. As stated when discussing hardness of approximation of Max k-CSP$_2$, this approximation ratio is tight up to a constant factor.

For the non-boolean case, Charikar, Makarychev, and Makarychev's algorithm achieve $\Omega(k \log R/R^k)$ ratio for Max k-CSP$_R$. Makarychev, and Makarychev later improved the

[1] Raghavendra showed in [31] that it is hard to approximate any Max CSP beyond what a certain type of semidefinite program can achieve. However, determining the approximation ratio of a semidefinite program is still not an easy task. Typically, one still needs to find an integrality gap for such a program in order to establish the approximation ratio.

approximation ratio to $\Omega(k/R^{k-1})$ when $k = \Omega(\log R)$ [26]. Additionally, the algorithm was extended by Goldshlager and Moshkovitz to achieve the same approximation ratio for any k, R [14]. On this front, we show the following result.

▶ **Theorem 3.** *There exists a polynomial-time* $\Omega(\log R/R^{k-1})$*-approximation algorithm for* Max k-CSP$_R$.

In comparison to the previous algorithms, our algorithm gives better approximation ratio than all the known algorithms when $k = o(\log R)$. We remark that our algorithm is just a simple extension of Kindler, Kolla and Trevisan's polynomial-time $\Omega(\log R/R)$-approximation algorithm for Max 2-CSP$_R$ [24].

1.1 Summary of Techniques

Our reduction from Unique Games to Max k-CSP$_R$ follows the reduction of [22] for Max 2-CSPs. We construct a k-query PCP using a Unique-Label-Cover "outer verifier", and then design a k-query Long Code test as an "inner verifier". For simplicity, let us think of k as a constant. We essentially construct a k-query *Dictator-vs.-Quasirandom* test for functions $f : [R]^n \to [R]$, with completeness $\Omega(1/(\log R)^{k/2})$ and soundness $O(1/R^{k-1})$. Our test is structurally similar to the 2-query "noise stability" tests of [22]: first we pick a random $z \in [R]^n$, then we pick k weakly-correlated queries $x^{(1)}, \dots, x^{(k)}$ by choosing each $x^{(i)} \in [R]^n$ as a noisy copy of z, i.e., each coordinate $(x^{(i)})_j$ is chosen as z_j with some probability ρ or uniformly at random otherwise. We accept iff $f(x^{(1)}) = f(x^{(2)}) = \cdots = f(x^{(k)})$. The key technical step is our analysis of the soundness of this test. We need to show that if f is a balanced function with small low-degree influences, then the test passes with probability $O(1/R^{k-1})$. Intuitively, we would like to say that for high enough noise, the values $f(x^{(i)})$ are roughly independent and uniform, so the test passes with probability around $1/R^{k-1}$. To formalize this intuition, we utilize the *Invariance Principle* and *Hypercontractivity.*

More precisely, if we let $f^i(x) : [R]^n \to \{0, 1\}$ be the indicator function for $f(x) = i$, then the test accepts iff $f^i(x^{(1)}) = \cdots = f^i(x^{(k)}) = 1$ for some $i \in [R]$. For each $i \in [R]$, this probability can be written as the expectation of the product: $\mathbb{E}[f^i(x^{(1)})f^i(x^{(2)}) \dots f^i(x^{(k)})]$. Since $x^{(i)}$'s are chosen as noisy copies of z, this expression is related to the k-th norm of a noisy version of f^i. Thus, our problem is reduced to bounding the k-norm of a noisy function $\widetilde{f^i} : [R]^n \to [0, 1]$, which has bounded one-norm $\mathbb{E}[\widetilde{f^i}] = 1/R$ since f is balanced. To arrive at this bound, we first apply the Invariance Principle, which essentially states that a low-degree low-influence function on $[R]^n$ behaves on random input similarly to its "boolean analog" over domain $\{\pm 1\}^{nR}$. Here "boolean analog" refers to the function over $\{\pm 1\}^{nR}$ with matching Fourier coefficients.

Roughly speaking, now that we have transfered to the boolean domain, we are left to bound the k-norm of a noisy function on $\{\pm 1\}^{nR}$ based on its one-norm. This follows from Hypercontractivity in the boolean setting, which bounds a higher norm of any noisy function on boolean domain in terms of a lower norm.

The description above hides several technical complications. For example, when we pass from a function $[R]^n \to [0, 1]$ to its "boolean analog" $\{\pm 1\}^{nR} \to \mathbb{R}$, the range of the resulting function is no longer bounded to $[0, 1]$. This, along with the necessary degree truncation, means we must be careful when bounding norms. Further, Hypercontractivity only allows us to pass from k-norms to $(1 + \varepsilon)$-norms for small ε, so we cannot use the known 1-norm directly. For details on how we handle these issues, see Section 3. This allows us to prove soundness of our dictator test without passing through results on Gaussian space, as was done to prove the "Majority is Stablest" conjecture [27] at the core of the [22] 2-query dictator test.

To extend our result to work with d-to-1 Games Conjecture in place of UGC, we observe that our proof goes through even when we assume a conjecture weaker than the UGC, which we name the *One-Sided Unique Games Conjecture*. The only difference between the original UGC and the One-Sided UGC is that the completeness in UGC is allowed to be any constant smaller than one but the completeness is a fixed constant for the One-Sided UGC. The conjecture is formalized as Conjecture 13. We show that the d-to-1 Games Conjecture also implies the One-Sided UGC, which means that our inapproximability result for Max k-CSP$_R$ also holds when we change our assumption to d-to-1 Games.

Lastly, for our approximation algorithm, we simply extend the Kindler, Kolla and Trevisan's algorithm by first creating an instance of Max 2-CSP$_R$ from Max k-CSP$_R$ by projecting each constraint to all possible subsets of two variables. We then use their algorithm to approximate the instance. Finally, we set our assignment to be the same as that from KKT algorithm with some probability. Otherwise, we pick the assignment uniformly at random from $[R]$. As we shall show in Section 4, with the right probability, this technique can extend not only the KKT algorithm but any algorithm for Max k'-CSP$_R$ to an algorithm for Max k-CSP$_R$ where $k > k'$ with some loss in the advantage over the naive randomized algorithm.

1.2 Organization of the Paper

In Section 2, we define notations and list background knowledge that will be used throughout the paper. Next, we prove our hardness of approximation results, i.e., Theorem 1 and Theorem 2, in Section 3. In Section 4, we show how to extend Kindler et al.'s algorithm to Max k-CSP$_R$ and prove Theorem 3. We also explicitly present the dictator test that is implicit in our hardness proof, in Section 5. Finally, in Section 6, we discuss interesting open questions and directions for future works.

2 Preliminaries

In this section, we list notations and previous results that will be used to prove our results.

2.1 Max k-CSP$_R$

We start by giving a formal definition of Max k-CSP$_R$, which is the main focus of our paper.

▶ **Definition 4** (Max k-CSP$_R$). An instance $(\mathcal{X}, \mathcal{C})$ of (weighted) Max k-CSP$_R$ consists of

- A set $\mathcal{X} = \{x_1, \dots, x_n\}$ of variables.
- A set $\mathcal{C} = \{C_1, \dots, C_m\}$ of constraints. Each constraint C_i is a triple (W_i, S_i, P_i) of a positive weight $W_i > 0$ such that $\sum_{i=1}^m W_i = 1$, a subset of variables $S_i \subseteq \mathcal{X}$ of size k, and a predicate $P_i : [R]^{S_i} \to \{0,1\}$ that maps each assignment to variables in S_i to $\{0,1\}$. Here we use $[R]^{S_i}$ to denote the set of all functions from S_i to $[R]$, i.e., $[R]^{S_i} = \{\psi : S_i \to [R]\}$.

For each assignment of variables $\varphi : \mathcal{X} \to [R]$, we define its value to be the total weights of the predicates satisfied by this assignment, i.e., $\sum_{i=1}^m W_i P_i(\varphi\mid_{S_i})$. The goal is to find an assignment $\varphi : \mathcal{X} \to [R]$ that with maximum value. We sometimes call the optimum the value of $(\mathcal{X}, \mathcal{C})$.

Note that, while the standard definition of Max k-CSP$_R$ refers to the unweighted version where $W_1 = \cdots = W_m = 1/m$, Crescenzi, Silvestri and Trevisan showed that the approximability of these two cases are essentially the same [8].[2] Hence, it is enough for us to consider just the weight version.

2.2 Unique Games and d-to-1 Conjectures

In this subsection, we give formal definitions for unique games, d-to-1 games and Khot's conjectures about them. First, we give a formal definition of unique games.

▶ **Definition 5** (Unique Game). A unique game $(V, W, E, N, \{\pi_e\}_{e \in E})$ consists of

- A bipartite graph $G = (V, W, E)$.
- Alphabet size N.
- For each edge $e \in E$, a permutation $\pi_e : [N] \to [N]$.

For an assignment $\varphi : V \cup W \to [N]$, an edge $e = (v, w)$ is satisfied if $\pi_e(\varphi(v)) = \varphi(w)$. The goal is to find an assignment that satisfies as many edges as possible. We define the value of an instance to be the fraction of edges satisfied in the optimum solution.

The Unique Games Conjecture states that it is NP-hard to distinguish an instance of value close one from that of value almost zero. More formally, it can be stated as follows.

▶ **Conjecture 6** (Unique Games Conjecture). *For every constant $\eta, \gamma > 0$, there exists a constant $N = N(\eta, \gamma)$ such that it is NP-hard to distinguish a unique game with alphabet size N of value at least $1 - \eta$ from one of value at most γ.*

Next, we define d-to-1 games, which is similar to unique games except that each constraint is a d-to-1 function instead of a permutation.

▶ **Definition 7** (d-to-1 Game). A d-to-1 game $(V, W, E, N, \{\pi_e\}_{e \in E})$ consists of

- A bipartite graph $G = (V, W, E)$.
- Alphabet size N.
- For each edge $e \in E$, a function $\pi_e : [N] \to [N/d]$ such that $|\pi_e^{-1}(\sigma)| = d$ for every $\sigma \in [N/d]$.

Satisfiability of an edge, the goal, and an instance's value of is defined similar to that of unique games.

In contrast to the UGC, d-to-1 Conjecture requires perfect completeness, i.e., it states that we cannot distinguish even a full satisfiable d-to-1 game from one with almost zero value.

▶ **Conjecture 8** (d-to-1 Conjecture). *For every constant $\gamma > 0$, there exists a constant $N = N(\gamma)$ such that it is NP-hard to distinguish a d-to-1 game with alphabet size N of value 1 from one of value at most γ.*

---

[2] More specifically, they proved that, if the weighted version is NP-hard to approximate to within ratio r, then the unweighted version is also NP-hard to approximate to within $r - o_n(1)$ where $o_n(1)$ represents a function such that $o_n(1) \to 0$ as $n \to \infty$.

2.3 Fourier Expansions

For any function $g : [q]^n \to \mathbb{R}$ over a finite alphabet $[q]$, we define the Fourier expansion of g as follows.

Consider the space of all functions $[q] \to \mathbb{R}$, with the inner-product $\langle u, v\rangle := \mathbb{E}_{x\in[q]}[u(x)v(x)]$, where the expectation is over a uniform $x \in [q]$. Pick an orthonormal basis $l_1, \ldots, l_q$ for this space $l_i : \Sigma \to \mathbb{R}$, such that l_1 is the constant function 1. We can now write g in the tensor-product basis, as

$$g(x_1, x_2, \ldots, x_n) = \sum_{s\in[q]^n} \hat{g}(s) \cdot \prod_{i=1}^{n} l_{s(i)}(x_i). \tag{1}$$

Since we pick $l_1(x) = 1$ for all $x \in [q]$, we also have $\mathbb{E}_{x\in[q]}[l_i(x)] = \langle l_i, l_1\rangle = 0$ for every $i \neq 1$.

Throughout, we use $\hat{g}$ to refer to the Fourier coefficients of a function g.

For functions $g : [q]^n \to \mathbb{R}$, the p-norm is defined as

$$||g||_p = \mathop{\mathbb{E}}_{x\in[q]^n}[|g(x)|^p]^{1/p}. \tag{2}$$

In particular, the squared 2-norm is

$$||g||_2^2 = \mathop{\mathbb{E}}_{x\in[q]^n}[g(x)^2] = \sum_{s\in[q]^n} \hat{g}(s)^2. \tag{3}$$

2.3.1 Noise Operators

For $x \in [q]^n$, let $y \overset{\rho}{\leftarrow} x$ denote that y is a ρ-correlated copy of x. That is, each coordinate y_i is independently chosen to be equal to x_i with probability ρ, or chosen uniformly at random otherwise.

Define the noise operator T_ρ acting on any function $g : [q]^n \to \mathbb{R}$ as

$$(T_\rho g)(x) = \mathop{\mathbb{E}}_{y \overset{\rho}{\leftarrow} x}[g(y)]. \tag{4}$$

Notice that the noise operator T_ρ acts on the Fourier coefficients on this basis as follows.

$$f(x) = T_\rho g(x) = \sum_{s\in[q]^n} \hat{g}(s) \cdot \rho^{|s|} \cdot \prod_{i=1}^{n} l_{s(i)}(x_i) \tag{5}$$

where $|s|$ is defined as $|\{i \mid s(i) \neq 1\}|$.

2.3.2 Degree Truncation

For any function $g : [q]^n \to \mathbb{R}$, let $g^{\leq d}$ denote the $(\leq d)$-degree part of g, i.e.,

$$g^{\leq d}(x) = \sum_{s\in[q]^n, |s|\leq d} \hat{g}(s) \cdot \prod_{i=1}^{n} l_{s(i)}(x_i), \tag{6}$$

and similarly let $g^{>d} : [q]^n \to \mathbb{R}$ denote the $(> k)$-degree part of g, i.e.,

$$g^{>d}(x) = \sum_{s\in[q]^n, |s|>d} \hat{g}(s) \cdot \prod_{i=1}^{n} l_{s(i)}(x_i). \tag{7}$$

Notice that degree-truncation commutes with the noise-operator, so writing $T_\rho g^{\leq d}$ is unambiguous.

Also, notice that truncating does not increase 2-norm:

$$||g^{\leq d}||_2^2 = \sum_{s \in [q]^n, |s| \leq d} \hat{g}(s)^2 \leq \sum_{s \in [q]^n} \hat{g}(s)^2 = ||g||_2^2. \tag{8}$$

We frequently use the fact that noisy functions have small high-degree contributions. That is, for any function $g : [q]^n \to [0, 1]$, we have

$$||T_\rho g^{>d}||_2^2 = \sum_{s \in [q]^n, |s| > d} \rho^{2|s|} \hat{g}(s)^2 \leq \rho^{2d} \sum_{s \in [q]^n} \hat{g}(s)^2 = \rho^{2d} ||g||_2^2 \leq \rho^{2d}. \tag{9}$$

2.3.3 Influences

For any function $g : [q]^n \to \mathbb{R}$, the influence of coordinate $i \in [n]$ is defined as

$$Inf_i[g] = \mathop{\mathbb{E}}_{x \in [q]^n} [Var_{x_i \in [q]}[g(x) \mid \{x_j\}_{j \neq i}]]. \tag{10}$$

This can be expressed in terms of the Fourier coefficients of g as follows:

$$Inf_i[g] = \sum_{s \in [q]^n :\ s(i) \neq 1} \hat{g}(s)^2. \tag{11}$$

Further, the degree-d influences are defined as

$$Inf_i^{\leq d}[g] = Inf_i[g^{\leq d}] = \sum_{\substack{s \in [q]^n : \\ |s| \leq d,\ s(i) \neq 1}} \hat{g}(s)^2. \tag{12}$$

2.3.4 Binary Functions

The previous discussion of Fourier analysis can be applied to boolean functions, by specializing to $q = 2$. However, in this case the Fourier expansion can be written in a more convenient form. Let $G : \{+1, -1\}^n \to \mathbb{R}$ be a boolean function over n bits. We can choose orthonormal basis functions $l_1(y) = 1$ and $l_2(y) = y$, so G can be written as

$$G(y) = \sum_{S \subseteq [n]} \hat{G}(S) \prod_{i \in S} y_i \tag{13}$$

for some coefficients $\hat{G}(S)$.

Degree-truncation then results in

$$G^{\leq d}(y) = \sum_{S \subseteq [n] : |S| \leq d} \hat{G}(S) \prod_{i \in S} y_i, \tag{14}$$

and the noise-operator acts as follows:

$$T_\rho G(y) = \sum_{S \subseteq [n]} \hat{G}(S) \rho^{|S|} \prod_{i \in S} y_i. \tag{15}$$

2.3.5 Boolean Analogs

To analyze k-CSP$_R$, we will want to translate between functions over $[R]^n$ to functions over $\{\pm 1\}^{nR}$. The following notion of *boolean analogs* will be useful.

For any function $g : [R]^n \to \mathbb{R}$ with Fourier coefficients $\hat{g}(s)$, define the boolean analog of g to be a function $\{g\} : \{\pm 1\}^{n \times R} \to \mathbb{R}$ such that

$$\{g\}(z) = \sum_{s \in [R]^n} \hat{g}(s) \cdot \prod_{i \in [n], s(i) \neq 1} z_{i,s(i)}. \tag{16}$$

Note that

$$||g||_2^2 = \sum_{s \in [R]^n} \hat{g}(s)^2 = ||\{g\}||_2^2, \tag{17}$$

and that

$$\{g^{\leq d}\} = \{g\}^{\leq d}. \tag{18}$$

Moreover, the noise operator acts nicely on $\{g\}$ as follows:

$$T_\rho \{g\} = \{T_\rho g\}. \tag{19}$$

For simplicity, we use T_ρ to refer to both boolean and non-boolean noise operators with correlation ρ.

2.4 Invariance Principle and Mollification Lemma

We start with the Invariance Principle in the form of Theorem 3.18 in [27]:

▶ **Theorem 9** (General Invariance Principle [27]). *Let $f : [R]^n \to \mathbb{R}$ be any function such that $Var[f] \leq 1$, $Inf_i[f] \leq \delta$, and $deg(f) \leq d$. Let $F : \{\pm 1\}^{nR} \to \mathbb{R}$ be its boolean analog: $F = \{f\}$. Consider any "test-function" $\psi : \mathbb{R} \to \mathbb{R}$ that is $\mathcal{C}^3$, with bounded 3rd-derivative $|\psi'''| \leq C$ everywhere. Then,*

$$\left| \mathop{\mathbb{E}}_{y \in \{\pm 1\}^{nR}} [\psi(F(y))] - \mathop{\mathbb{E}}_{x \in [R]^n} [\psi(f(x))] \right| \leq C 10^d R^{d/2} \sqrt{\delta}. \tag{20}$$

Note that the above version follows directly from Theorem 3.18 and Hypothesis 3 of [27], and the fact that uniform ± 1 bits are $(2, 3, 1/\sqrt{2})$-hypercontractive as described in [27].

As we shall see later, we will want to apply the Invariance Principle for some functions ψ that are not in $\mathcal{C}^3$. However, these functions will be Lipschitz-continuous with some constant $c \in \mathbb{R}$ (or "c-Lipschitz"), meaning that

$$|\psi(x + \Delta) - \psi(x)| \leq c|\Delta| \quad \text{for all } x, \Delta \in \mathbb{R}. \tag{21}$$

Therefore, similar to Lemma 3.21 in [27], we can "smooth" it to get a function $\widetilde{\psi}$ that is that is $\mathcal{C}^3$, and has arbitrarily small pointwise difference to ψ.

▶ **Lemma 10** (Mollification Lemma [27]). *Let $\psi : \mathbb{R} \to \mathbb{R}$ be any c-Lipschitz function. Then for any $\zeta > 0$, there exists a function $\widetilde{\psi} : \mathbb{R} \to \mathbb{R}$ such that*
- *$\widetilde{\psi} \in \mathcal{C}^3$,*
- *$||\widetilde{\psi} - \psi||_\infty \leq \zeta$, and,*
- *$||\widetilde{\psi}'''||_\infty \leq \widetilde{C} c^3/\zeta^2$.*

For some universal constant $\widetilde{C}$, not depending on ζ, c.

For completeness, the full proof of the lemma can be found in Appendix A.1.

Now we state the following version of the Invariance Principle, which will be more convenient to invoke. It can be proved simply by just combining the two previous lemmas. We list a full proof in Appendix A.2.

▶ **Lemma 11** (Invariance Principle). *Let $\psi : \mathbb{R} \to \mathbb{R}$ be one of the following functions:*

1. $\psi_1(t) := |t|$,
2. $\psi_k(t) := \begin{cases} t^k & \text{if } t \in [0,1], \\ 0 & \text{if } t < 0, \\ 1 & \text{if } t \geq 1. \end{cases}$

Let $f : [R]^n \to [0,1]$ be any function with all $Inf_i^{\leq d}[f] \leq \delta$. Let $F : \{\pm 1\}^{nR} \to \mathbb{R}$ be its boolean analog: $F = \{f\}$. Let $f^{\leq d} : [R]^n \to \mathbb{R}$ denote f truncated to degree d, and similarly for $F^{\leq d} : \{\pm 1\}^{nR} \to \mathbb{R}$.

Then, for parameters $d = 10k \log R$ and $\delta = 1/(R^{10+100k\log(R)})$, we have

$$\left| \mathop{\mathbb{E}}_{y \in \{\pm 1\}^{nR}} [\psi(F^{\leq d}(y))] - \mathop{\mathbb{E}}_{x \in [R]^n} [\psi(f^{\leq d}(x))] \right| \leq O(1/R^k). \tag{22}$$

2.5 Hypercontractivity Theorem

Another crucial ingredient in our proof is the hypercontractivity lemma, which says that, on $\{\pm 1\}^n$ domain, the operator T_ρ smooths any function so well that the higher norm can be bound by the lower norm of the original (unsmoothed) function. Here we use the version of the theorem as stated in [28].

▶ **Theorem 12** (Hypercontractivity Theorem [28]). *Let $1 \leq p \leq q \leq \infty$. For any $\rho \leq \sqrt{\frac{p-1}{q-1}}$ and for any function $h : \{\pm 1\}^n \to \mathbb{R}$, the following inequality holds:*

$$||T_\rho h||_q \leq ||h||_p. \tag{23}$$

In particular, for choice of parameter $\rho = 1/\sqrt{(k-1)\log R}$, we have

$$||T_{2\rho} h||_k \leq ||h||_{1+\varepsilon}. \tag{24}$$

where $\varepsilon = 4/\log(R)$.

3 Inapproximability of Max k-CSP$_R$

In this section, we prove Theorem 1 and Theorem 2. Before we do so, we first introduce a conjecture, which we name *One-Sided Unique Games Conjecture*. The conjecture is similar to UGC except that the completeness parameter ζ is fixed in contrast to UGC where the completeness can be any close to one.

▶ **Conjecture 13** (One-Sided Unique Games Conjecture). *There exists a constant $\zeta > 0$ such that, for every constant $\gamma > 0$, there exists a constant $N = N(\gamma)$ such that it is NP-hard to distinguish a unique game with alphabet size N of value at least ζ from one of value at most γ.*

It is obvious that the UGC implies One-Sided UGC with $\zeta = 1 - \eta$ for any sufficiently small η. It is also not hard to see that, by repeating each alphabet on the right d times and spreading each d-to-1 constraint out to be a permutation, d-to-1 Games Conjecture implies One-Sided UGC with $\zeta = 1/d$. A full proof of this can be found in Appendix B.

Since both UGC and d-to-1 Games Conjecture imply One-Sided UGC, it is enough for us to show the following theorem, which implies both Theorem 1 and Theorem 2.

▶ **Theorem 14.** *Unless the One-Sided Unique Games Conjecture is false, for any $k \geq 2$ and any sufficiently large R, it is NP-hard to approximate* Max k-CSP$_R$ *to within $2^{O(k \log k)}(\log R)^{k/2}/R^{k-1}$ factor.*

The theorem will be proved in Subsection 3.3. Before that, we first prove an inequality that is the heart of our soundness analysis in Subsection 3.2.

3.1 Parameters

We use the following parameters throughout, which we list for convenience here:
- Correlation[3]: $\rho = 1/\sqrt{(k-1)\log R}$
- Degree: $d = 10k \log R$
- Low-degree influences: $\delta = 1/(R^{10+100k\log(R)})$

3.2 Hypercontractivity for Noisy Low-Influence Functions

Here we show a version of hypercontractivity for noisy low-influence functions over large domains. Although hypercontractivity does not hold in general for noisy functions over large domains, it turns out to hold in our setting of high-noise and low-influences. The main technical idea is to use the Invariance Principle to reduce functions over larger domains to boolean functions, then apply boolean hypercontractivity.

▶ **Lemma 15** (Main Lemma)**.** *Let $g : [R]^n \to [0,1]$ be any function with $\mathbb{E}_{x\in[R]^n}[g(x)] = 1/R$. Then, for our choice of parameters ρ, d, δ: If $Inf_i^{\leq d}[g] \leq \delta$ for all i, then*

$$\mathop{\mathbb{E}}_{x\in[R]^n}[(T_\rho g(x))^k] \leq 2^{O(k)}/R^k.$$

Before we present the full proof, we outline the high-level steps below. Let $f = T_\rho g$, and define boolean analogs $G = \{g\}$, and $F = \{f\}$. Let $\psi_k : \mathbb{R} \to \mathbb{R}$ be defined as in Lemma 11. Then,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x\in[R]^n}[f(x)^k] \approx \mathbb{E}[\psi_k(f^{\leq d}(x))] && (25)\\
\text{(Lemma 11: Invariance Principle)} \quad &\approx \mathop{\mathbb{E}}_{y\in\{\pm 1\}^{nR}}[\psi_k(F^{\leq d}(y))] && (26)\\
\text{(Definition of } \psi_k) \quad &\leq ||F^{\leq d}||_k^k && (27)\\
\text{(Definition of } F) \quad &= ||T_\rho G^{\leq d}||_k^k && (28)\\
&= ||T_{2\rho} T_{1/2} G^{\leq d}||_k^k && (29)\\
\text{(Hypercontractivity, for small } \varepsilon) \quad &\leq ||T_{1/2} G^{\leq d}||_{1+\varepsilon}^k && (30)\\
\text{(Invariance, etc.)} \quad &\approx 2^{O(k)}||g||_1^k && (31)\\
\text{(Since } \mathbb{E}[|g|] = 1/R) \quad &= 2^{O(k)}/R^k. && (32)
\end{aligned}$$

[3] Note that for $k = 2$, this correlation yields a stability of $\approx 1/R$ for the plurality. That is, $\Pr_{z,x,y}[\text{plur}(x_1,\ldots,x_n) = \text{plur}(y_1,\ldots,y_n)] \approx 1/R$ where each $z_i \in [R]$ is iid uniform, and x_i, y_i are ρ-correlated copies of z_i.

Proof. To establish line (25), first notice that

$$\psi_k(f(x)) = \psi_k(f^{\leq d}(x) + f^{>d}(x)) \leq \psi_k(f^{\leq d}(x)) + k|f^{>d}(x)| \tag{33}$$

where the last inequality is because the function ψ_k is k-Lipschitz.

Moreover, since $g(x) \in [0, 1]$, we have $f(x) \in [0, 1]$, so

$$f(x)^k = \psi_k(f(x)). \tag{34}$$

Thus,

$$\begin{aligned}
\mathbb{E}[f(x)^k] &= \mathbb{E}[\psi_k(f(x))] && (35)\\
&\leq \mathbb{E}[\psi_k(f^{\leq d}(x))] + k\,\mathbb{E}[|f^{>d}(x)|] && (36)\\
&= \mathbb{E}[\psi_k(f^{\leq d}(x))] + k||f^{>d}||_1 && (37)\\
&\leq \mathbb{E}[\psi_k(f^{\leq d}(x))] + k||f^{>d}||_2. && (38)\\
& && (39)
\end{aligned}$$

And we can bound the 2-norm of $f^{>d}$, since f is noisy, we have

$$||f^{>d}||_2^2 = ||T_\rho g^{>d}||_2^2 \leq \rho^{2d} \leq O(1/R^{2k}). \tag{40}$$

The last inequality comes from our choice of ρ and d.

So line (25) is established:

$$\mathbb{E}[f(x)^k] \leq \mathbb{E}[\psi_k(f^{\leq d}(x))] + O(k/R^k). \tag{41}$$

Line (26) follows directly from our version of the Invariance Principle (Lemma 11), for the function ψ_k:

$$\mathop{\mathbb{E}}_{x\in[R]^n}[\psi_k(f^{\leq d}(x))] \leq \mathop{\mathbb{E}}_{y\in\{\pm1\}^{nR}}[\psi_k(F^{\leq d}(y))] + O(1/R^k). \tag{42}$$

We can now rewrite $\mathbb{E}_{y\in\{\pm1\}^{nR}}[\psi_k(F^{\leq d}(y))]$ as

$$\begin{aligned}
\mathop{\mathbb{E}}_{y\in\{\pm1\}^{nR}}[\psi_k(F^{\leq d}(y))] &\leq \mathop{\mathbb{E}}_{y\in\{\pm1\}^{nR}}[|F^{\leq d}(y)|^k] && (43)\\
&= ||F^{\leq d}||_k^k && (44)\\
&= ||T_\rho G^{\leq d}||_k^k && (45)\\
&= ||T_{2\rho} T_{1/2} G^{\leq d}||_k^k. && (46)\\
& && (47)
\end{aligned}$$

Now, from the Hypercontractivity Theorem, Equation (24), we have

$$||T_{2\rho} T_{1/2} G^{\leq d}||_k \leq ||T_{1/2} G^{\leq d}||_{1+\varepsilon} \tag{48}$$

for $\varepsilon = 4/\log R$. This establishes line (30):

$$||T_{2\rho} T_{1/2} G^{\leq d}||_k^k \leq ||T_{1/2} G^{\leq d}||_{1+\varepsilon}^k = \mathbb{E}[|T_{1/2} G^{\leq d}(y)|^{1+\varepsilon}]^{k/(1+\varepsilon)}. \tag{49}$$

To show the remaining steps, we will apply the Invariance Principle once more. Notice that for all $t \in \mathbb{R} : |t|^{1+\varepsilon} \leq |t| + t^2$. Hence, we can derive the following bound:

$$\begin{aligned}
& \mathbb{E}[|T_{1/2} G^{\leq d}(y)|^{1+\varepsilon}] \leq \mathbb{E}[|T_{1/2} G^{\leq d}(y)|] + \mathbb{E}[(T_{1/2} G^{\leq d}(y))^2] && (50)\\
\text{(Matching Fourier expansion)} \quad & = \mathbb{E}[|T_{1/2} G^{\leq d}(y)|] + \mathbb{E}[(T_{1/2} g^{\leq d}(y))^2] && (51)\\
\text{(Lemma 11, Invariance Principle)} \quad & \leq \mathbb{E}[|T_{1/2} g^{\leq d}(x)|] + \mathbb{E}[(T_{1/2} g^{\leq d}(x))^2] + O(1/R^k). && (52)
\end{aligned}$$

Here we applied our Invariance Principle (Lemma 11) for the function ψ_1 as defined in Lemma 11. We will bound each of the expectations on the RHS, using the fact that g is balanced, and $T_{1/2}g$ is noisy.

First,

$$\begin{aligned}
\mathbb{E}[|T_{1/2}g^{\leq d}(x)|] &= \mathbb{E}[|T_{1/2}g(x) - T_{1/2}g^{>d}(x)|] && (53)\\
\text{(Triangle Inequality)} \quad &\leq \mathbb{E}[|T_{1/2}g(x)|] + \mathbb{E}[|T_{1/2}g^{>d}(x)|] && (54)\\
&= ||g||_1 + ||T_{1/2}g^{>d}||_1 && (55)\\
&\leq ||g||_1 + ||T_{1/2}g^{>d}||_2 && (56)\\
&\leq 1/R + (1/2)^d && (57)\\
\text{(By our choice of } d) \quad &= O(1/R). && (58)
\end{aligned}$$

Second,

$$\begin{aligned}
\mathbb{E}[(T_{1/2}g^{\leq d}(x))^2] &= \sum_{s\in[R]^n, |s|\leq d} (1/2)^{2|s|}\hat{g}(s)^2 && (59)\\
&\leq \sum_{s\in[R]^n} (1/2)^{2|s|}\hat{g}(s)^2 && (60)\\
&= \mathbb{E}[(T_{1/2}g(x))^2] && (61)\\
\text{(Since } g \in [0,1]) \quad &\leq \mathbb{E}[T_{1/2}g(x)] && (62)\\
&= \mathbb{E}[g(x)] = 1/R. && (63)
\end{aligned}$$

Finally, plugging these bounds into (52), we find:

$$\begin{aligned}
||T_{1/2}G^{\leq d}||_{1+\varepsilon}^k &= \mathbb{E}[|T_{1/2}G^{\leq d}(y)|^{1+\varepsilon}]^{k/(1+\varepsilon)} && (64)\\
&\leq (O(1/R))^{k/(1+\varepsilon)} && (65)\\
&= 2^{O(k)}/R^{k/(1+\varepsilon)} && (66)\\
&\leq 2^{O(k)}/R^{k(1-\varepsilon)} && (67)\\
\text{(Recall } \varepsilon = 4/\log R) \quad &= 2^{O(k)}/R^k. && (68)
\end{aligned}$$

This completes the proof of the main lemma. ◀

3.3 Reducing Unique Label Cover to Max k-CSP$_R$

Here we reduce unique games to Max k-CSP$_R$. We will construct a PCP verifier that reads k symbols of the proof (with an alphabet of size R) with the following properties:

- **Completeness.** If the unique game has value at least ζ, then the verifier accepts an honest proof with probability at least $c = \zeta^k/((\log R)^{k/2}2^{O(k\log k)})$.
- **Soundness.** If the unique game has value at most $\gamma = 2^{O(k)}\delta^2/(4dR^k)$, then the verifier accepts any (potentially cheating) proof with probability at most $s = 2^{O(k)}/R^{k-1}$.

Since each symbol in the proof can be viewed as a variable and each accepting predicate of the verifier can be viewed as a constraint of Max k-CSP$_R$, assuming the One-sided UGC, this PCP implies NP-hardness of approximating Max k-CSP$_R$ of factor $s/c = 2^{O(k\log k)}(\log R)^{k/2}/R^{k-1}$ and, hence, establishes our Theorem 14.

3.3.1 The PCP

Given a unique game $(V, W, E, n, \{\pi_e\}_{e\in E})$, the proof is the truth-table of a function $h_w : [R]^n \to [R]$ for each vertex $w \in W$. By folding, we can assume h_w is balanced, i.e. h_w takes on all elements of its range with equal probability: $\Pr_{x\in[R]^n}[h_w(x) = i] = 1/R$ for all $i \in [R]$.[4]

Notationally, for $x \in [R]^n$, let $(x \circ \pi)$ denote permuting the coordinates of x as: $(x \circ \pi)_i = x_{\pi(i)}$. Also, for an edge $e = (v, w)$, we write $\pi_e = \pi_{v,w}$, and define $\pi_{w,v} = \pi^{-1}_{v,w}$.

The verifier picks a uniformly random vertex $v \in V$, and k independent uniformly random neighbors of v: $w_1, w_2, \dots, w_k \in W$. Then pick $z \in [R]^n$ uniformly at random, and let $x^{(1)}, x^{(2)}, \dots, x^{(k)}$ be independent ρ-correlated noisy copies of z (each coordinate x_i chosen as equal to z_i w.p. ρ, or uniformly at random otherwise). The verifier accepts if and only if

$$h_{w_1}(x^{(1)} \circ \pi_{w_1,v}) = h_{w_2}(x^{(2)} \circ \pi_{w_2,v}) = \dots = h_{w_k}(x^{(k)} \circ \pi_{w_k,v}). \tag{69}$$

To achieve the desired hardness result, we pick $\rho = 1/\sqrt{(k-1)\log R}$.

3.3.2 Completeness Analysis

First, note that that we can assume without loss of generality that the graph is regular on V side.[5] Let the degree of each vertex in V be Δ.

Suppose that the original unique game has an assignment of value at least ζ. Let us call this assignment φ. The honest proof defines h_w at each vertex $w \in W$ as the long code encoding of this assignment, i.e., $h_w(x) = x_{\varphi(w)}$. We can written the verifier acceptance condition as follows:

$$\begin{aligned} \text{The verifier accepts} &\Leftrightarrow h_{w_1}(x^{(1)} \circ \pi_{w_1,v}) = \dots = h_{w_k}(x^{(k)} \circ \pi_{w_k,v}) && (70)\\ &\Leftrightarrow (x^{(1)} \circ \pi_{w_1,v})_{\varphi(w_1)} = \dots = (x^{(k)} \circ \pi_{w_k,v})_{\varphi(w_k)} && (71)\\ &\Leftrightarrow (x^{(1)})_{\pi_{w_1,v}(\varphi(w_1))} = \dots = (x^{(k)})_{\pi_{w_k,v}(\varphi(w_k))}. && (72)\end{aligned}$$

Observe that, if the edges $(v, w_1), \dots, (v, w_k)$ are satisfied by φ, then $\pi_{w_1,v}(\varphi(w_1)) = \dots = \pi_{w_k,v}(\varphi(w_k)) = \varphi(v)$. Hence, if the aforementioned edges are satisfied and $x^{(1)}, \dots, x^{(k)}$ are not perturbed at coordinate $\varphi(v)$, then $(x^{(1)})_{\pi_{w_1,v}(\varphi(w_1))} = \dots = (x^{(k)})_{\pi_{w_k,v}(\varphi(w_k))}$.

For each $u \in V$, let s_u be the number of satisfied edges touching u. Since $w_1, \dots, w_k$ are chosen from the neighbors of v independently from each other, the probability that the edges $(v, w_1), (v, w_2), \dots, (v, w_k)$ are satisfied can be bounded as follows:

$$\begin{aligned} &\Pr_{v,w_1,\dots,w_k}[(v, w_1), \dots, (v, w_k) \text{ are satisfied}] && (73)\\ &= \sum_{u\in V} \Pr_{w_1,\dots,w_k}[(v, w_1), \dots, (v, w_k) \text{ are satisfied} \mid v = u] \Pr[v = u] && (74)\\ &= \sum_{u\in V} (s_u/\Delta)^k \Pr[v = u] && (75)\\ &= \mathop{\mathbb{E}}_{u\in V}\left[(s_u/\Delta)^k\right] && (76)\\ &\geq \mathop{\mathbb{E}}_{u\in V}[s_u/\Delta]^k. && (77)\end{aligned}$$

[4] More precisely, if the truth-table provided is of some function $\widetilde{h}_w : [R]^n \to [R]$, we define the "folded" function h_w as $h_w(x_1, x_2, x_3, \dots x_n) := \widetilde{h}_w(x-(x_1, x_1, \dots, x_1))+x_1$, where the $\pm$ is over mod R. Notice that the folded h_w is balanced, and also that folding does not affect dictator functions. Thus we define our PCP in terms of h_w, but simulate queries to h_w using the actual proof $\widetilde{h}_w$.

[5] See, for instance, Lemma 3.4 in [23].

Notice that $\mathbb{E}_{u\in V}[s_u/\Delta]$ is exactly the value of φ, which is at least ζ. As a result,

$$\Pr_{v,w_1,\dots,w_k}[(v,w_1),\dots,(v,w_k) \text{ are satisfied}] \geq \zeta^k.$$

Furthermore, it is obvious that the probability that $x_1,\dots,x_k$ are not perturbed at the coordinate $\varphi(v)$ is ρ^k. As a result, the PCP accepts with probability at least $\zeta^k\rho^k$. When $\rho = 1/\sqrt{(k-1)\log R}$ and ζ is a constant not depending on k and R, the completeness is $1/((\log R)^{k/2}2^{O(k\log k)})$.

3.3.3 Soundness Analysis

Suppose that the unique game has value at most $\gamma = 2^{O(k)}\delta^2/(4dR^k)$. We will show that the soundness is $2^{O(k)}/R^{k-1}$.

Suppose for the sake of contradiction that the probability that the verifier accepts $\Pr[accept] > t = 2^{\Omega(k)}/R^{k-1}$ where $\Omega(\cdot)$ hides some large enough constant.

Let $h^i_w(x) : [R]^n \to \{0,1\}$ be the indicator function for $h_w(x) = i$ and let $x \overset{\rho}{\leftarrow} z$ denote that x is a ρ-correlated copy of z. We have

$$\begin{aligned}
\Pr[accept] &= \Pr[h_{w_1}(x^{(1)}\circ\pi_{w_1,v}) = \cdots = h_{w_k}(x^{(k)}\circ\pi_{w_k,v})] && (78)\\
&= \sum_{i\in[R]} \Pr[i = h_{w_1}(x^{(1)}\circ\pi_{w_1,v}) = \cdots = h_{w_k}(x^{(k)}\circ\pi_{w_k,v})] && (79)\\
&= \sum_{i\in[R]} \mathbb{E}[h^i_{w_1}(x^{(1)}\circ\pi_{w_1,v})\cdots h^i_{w_k}(x^{(k)}\circ\pi_{w_k,v})] && (80)\\
&= \sum_{i\in[R]} \mathbb{E}\left[\mathbb{E}_{w_1}[h^i_{w_1}(x^{(1)}\circ\pi_{w_1,v})]\cdots \mathbb{E}_{w_k}[h^i_{w_k}(x^{(k)}\circ\pi_{w_k,v})]\right]. && (81)
\end{aligned}$$

Where the last equality follows since the w_i's are independent, given v.

Now define $g^i_v : [R]^n \to [0,1]$ as

$$g^i_v(x) = \mathbb{E}_{w\sim v}[h^i_w(x\circ\pi_{w,v})] \quad (82)$$

where $w \sim v$ denotes neighbors w of v.

We can rewrite $\Pr[accept]$ as follows:

$$\begin{aligned}
\Pr[accept] &= \sum_{i\in[R]} \mathbb{E}[g^i_v(x^{(1)})g^i_v(x^{(2)})\cdots g^i_v(x^{(k)})] && (83)\\
\text{(Since } x^{(j)}\text{'s are independent given } z) &= \sum_{i\in[R]} \mathbb{E}\left[\mathbb{E}_{x\overset{\rho}{\leftarrow} z}[g^i_v(x)]^k\right] && (84)\\
&= \sum_{i\in[R]} \mathbb{E}_{v,z}[(T_\rho g^i_v(z))^k] && (85)\\
&= \mathbb{E}_v\left[\sum_{i\in[R]} \mathbb{E}_z[(T_\rho g^i_v(z))^k\right]. && (86)
\end{aligned}$$

Next, notice that $\sum_{i\in[R]}\mathbb{E}_z[(T_\rho g^i_v(z))^k]$ is simply the probability the verifier accepts given it picks vertex v, and thus this sum is bounded above by 1.

Therefore, since $\Pr[accept] > t$, by (86), at least $t/2$ fraction of vertices $v \in V$ have

$$\sum_{i\in[R]} \mathbb{E}_z[(T_\rho g^i_v(z))^k] \geq t/2. \quad (87)$$

For these "good" vertices, there must exist some $i \in [R]$ for which

$$\mathop{\mathbb{E}}_{z}[(T_\rho g_v^i(z))^k] \geq t/(2R). \tag{88}$$

Then for "good" v and i as above,

$$\mathop{\mathbb{E}}_{z}[(T_\rho g_v^i(z))^k] > 2^{\Omega(k)}/R^k. \tag{89}$$

By Lemma 15 (Main Lemma), this means g_v^i has some coordinate j for which

$$Inf_j^{\leq d}[g_v^i] > \delta \tag{90}$$

for our choice of d, δ as defined in Subsection 3.1. Pick this j as the label of vertex $v \in V$.

Now to pick the label of a vertex $w \in W$, define the candidate labels as

$$Cand[w] = \{j \in [n] : \exists\ i \in [R] \text{ s.t. } Inf_j^{\leq d}[h_w^i] \geq \delta/2\}. \tag{91}$$

Notice that

$$\sum_{j\in[n]} Inf_j^{\leq d}[h_w^i] = \sum_{s\in[R]^n:\ |s|\leq d} |s|\hat{h}_w^i(s)^2 \leq d \sum_{s:|s|>0} \hat{h}_w^i(s)^2 = d\, Var[h_w^i] \leq d. \tag{92}$$

So for each $i \in [R]$, the projection h_w^i can have at most $2d/\delta$ coordinates with influence $\geq \delta/2$. Therefore the number of candidate labels is bounded:

$$|Cand[w]| \leq 2dR/\delta. \tag{93}$$

Now we argue that picking a random label in $Cand[w]$ for $w \in W$ is in expectation a good decoding. We will show that if we assigned label j to a "good" $v \in V$, then $\pi_{v,w}(j) \in Cand[w]$ for a constant fraction of neighbors $w \sim v$. Note here that $\pi_{v,w} = \pi_{w,v}^{-1}$.

First, since $g_v^i(x) = \mathbb{E}_{w\sim v}[h_w^i(x \circ \pi_{w,v})]$, the Fourier transform of g_v^i is related to the Fourier transform of the long code labels h_w^i as

$$\hat{g}_v^i(s) = \mathop{\mathbb{E}}_{w\sim v}[\hat{h}_w^i(s \circ \pi_{w,v})]. \tag{94}$$

Hence, the influence $Inf_j^{\leq d}[g_v^i]$ of being large implies the expected influence $Inf_{\pi_{v,w}^{-1}(j)}^{\leq d}[h_w^i]$ of its neighbor labels $w \sim v$ is also large as formalized below.

$$\delta < Inf_j^{\leq k}[g_v^i] = \sum_{\substack{s\in[R]^n \\ |s|\leq k, s_j\neq 1}} \hat{g}_v^i(s)^2 \tag{95}$$

$$= \sum \mathop{\mathbb{E}}_{w\sim v}[\hat{h}_w^i(s \circ \pi_{w,v})]^2 \tag{96}$$

$$\leq \sum \mathop{\mathbb{E}}_{w\sim v}[\hat{h}_w^i(s \circ \pi_{w,v})^2] \tag{97}$$

$$= \mathop{\mathbb{E}}_{w\sim v}[\sum_{\substack{s\in[R]^n \\ |s|\leq k, s_j\neq 1}} \hat{h}_w^i(s \circ \pi_{w,v})^2] \tag{98}$$

$$= \mathop{\mathbb{E}}_{w\sim v}[\sum_{\substack{s\in[R]^n \\ |s|\leq k, s_{\pi_{w,v}^{-1}(j)}\neq 1}} \hat{h}_w^i(s)^2] \tag{99}$$

$$(\text{Since } \pi_{v,w} = \pi_{w,v}^{-1}) = \mathop{\mathbb{E}}_{w\sim v}[\sum_{\substack{s\in[R]^n \\ |s|\leq k, s_{\pi_{v,w}(j)}\neq 1}} \hat{h}_w^i(s)^2] \tag{100}$$

$$= \mathop{\mathbb{E}}_{w\sim v}[Inf_{\pi_{v,w}(j)}^{\leq d}[h_w^i]] \tag{101}$$

Therefore, at least $\delta/2$ fraction of neighbors $w \sim v$ must have $Inf^{\leq d}_{\pi_{v,w}(j)}[h^i_w] \geq \delta/2$, and so $\pi_{v,w}(j) \in Cand[w]$ for at least $\delta/2$ fraction of neighbors of "good" vertices v.

Finally, recall that at least $(t/2)$ fraction of vertices $v \in V$ are "good". These vertices have at least $(\delta/2)$ fraction of neighbors $w \in W$ with high-influence labels and the matching label $w \in W$ is picked with probability at least $\delta/(2dR)$. Moreover, as stated earlier, we can assume that the graph is regular on V side. Hence, the expected fraction of edges satisfied by this decoding is at least

$$(t/2)(\delta/2)(\delta/2dR) = t\delta^2/(4dR) = 2^{\Omega(k)}\delta^2/(4dR^k) > \gamma, \tag{102}$$

which contradicts our assumption that the unique game has value at most γ. Hence, we can conclude that the soundness is at most $2^{O(k)}/R^{k-1}$ as desired.

4 $\Omega(\log R/R^{k-1})$-Approximation Algorithm for Max k-CSP$_R$

Instead of just extending the KKT algorithm to work with Max k-CSPs, we will show a more general statement that *any* algorithm that approximates Max CSPs with small arity can be extended to approximate Max CSPs with larger arities. In particular, we show how to extend any $f(R)/R^{k'}$-approximation algorithm for Max k'-CSP$_R$ to an $(f(R)/2^{O(\min\{k',k-k'\})})/R^k$-approximation algorithm for Max k-CSP$_R$ where $k > k'$.

Since the naive algorithm that assigns every variable randomly has an approximation ratio of $1/R^k$, we think of $f(R)$ as the advantage of algorithm A over the randomized algorithm. From this perspective, our extension lemma preserves the advantage up to a factor of $1/2^{O(\min\{k',k-k'\})}$.

The extension lemma and its proof are stated formally below.

▶ **Lemma 16.** *Suppose that there exists a polynomial-time approximation algorithm A for* Max *k'-CSP$_R$ that outputs an assignment with expected value at least $f(R)/R^{k'}$ times the optimum. For any $k > k'$, we can construct a polynomial-time approximation algorithm B for* Max *k-CSP$_R$ that outputs an assignment with expected value at least $(f(R)/2^{O(\min\{k',k-k'\})})/R^k$ times the optimum.*

Proof. The main idea of the proof is simple. We turn an instance of Max k-CSP$_R$ to an instance of Max k'-CSP$_R$ by constructing $\binom{k}{k'}R^{k-k'}$ new constraints for each original constraint; each new constraint is a projection of the original constraint to a subset of variables of size k'. We then use A to solve the newly constructed instance. Finally, B simply assigns each variable with the assignment from A with a certain probability and assigns it randomly otherwise.

For convenience, let α be $\frac{k-k'}{k}$. We define B on input $(\mathcal{X}, \mathcal{C})$ as follows:

1. Create an instance $(\mathcal{X}, \mathcal{C}')$ of Max k'-CSP$_R$ with the same variables and, for each $C = (W, S, P) \in \mathcal{C}$ and for every subset S' of S with $|S'| = k'$ and every $\tau \in [R]^{S-S'}$, create a constraint $C^{S',\tau} = (W', S', P')$ in $\mathcal{C}'$ where $W' = \frac{W}{\binom{k}{k'}R^{k-k'}}$ and $P' : [R]^{S'} \to \{0,1\}$ is defined by

$$P'(\psi) = P(\psi \circ \tau).$$

Here $\psi \circ \tau$ is defined as follows:

$$\psi \circ \tau(x) = \begin{cases} \psi(x) & \text{if } x \in S', \\ \tau(x) & \text{otherwise.} \end{cases}$$

2. Run A on input $(\mathcal{X}, \mathcal{C}')$. Denote the output of A by φ_A.
3. For each $x \in \mathcal{X}$, with probability α, pick $\varphi_B(x)$ randomly from $[R]$. Otherwise, let $\varphi_B(x)$ be $\varphi_A(x)$.
4. Output φ_B.

We now show that φ_B has expected value at least $(f(R)/2^{O(\min\{k', k-k'\})})/R^k$ times the optimum.

First, observe that the optimum of $(\mathcal{X}, \mathcal{C}')$ is at least $1/R^{k-k'}$ times that of $(\mathcal{X}, \mathcal{C})$. To see that this is true, consider any assignment $\varphi : \mathcal{X} \to [R]$ and any constraint $C = (W, S, P)$. Its weighted contribution in $(\mathcal{X}, \mathcal{C})$ is $WP(\varphi|_S)$. On the other hand, $\frac{W}{\binom{k}{k'}R^{k-k'}}P(\varphi|_S)$ appears $\binom{k}{k'}$ times in $(\mathcal{X}, \mathcal{C}')$, once for each subset $S' \subseteq S$ of size k'. Hence, the value of φ with respect to $(\mathcal{X}, \mathcal{C}')$ is at least $1/R^{k-k'}$ times its value with respect to $(\mathcal{X}, \mathcal{C})$

Recall that the algorithm A gives an assignment of expected value at least $f(R)/R^{k'}$ times the optimum of $(\mathcal{X}, \mathcal{C}')$. Hence, the expected value of φ_A is at least $f(R)/R^k$ times the optimum of $(\mathcal{X}, \mathcal{C})$.

Next, we will compute the expected value of φ_B (with respect to $(\mathcal{X}, \mathcal{C})$). We start by computing the expected value of φ_B with respect to a fixed constraint $C = (W, S, P) \in \mathcal{C}$, i.e., $\mathbb{E}_{\varphi_B}[WP(\varphi_B|_S)]$. For each $S' \subseteq S$ of size k, we define $D_{S'}$ as the event where, in step 3, $\varphi_B(x)$ is assigned to be $\varphi_A(x)$ for all $x \in S'$ and $\varphi_B(x)$ is assigned randomly for all $x \in S - S'$.

Since $D_{S'}$ is disjoint for all $S' \subseteq S$ of size k', we have the following inequality.

$$\mathop{\mathbb{E}}_{\varphi_B}[WP(\varphi_B|_S)] \geq \sum_{\substack{S' \subseteq S \\ |S'|=k'}} \Pr[D_{S'}] \mathop{\mathbb{E}}_{\varphi_B}[WP(\varphi_B|_S) \mid D_{S'}] \tag{103}$$

$$(\text{Since } \Pr[D_{S'}] = \alpha^{k-k'}(1-\alpha)^{k'}) = \alpha^{k-k'}(1-\alpha)^{k'} \sum_{\substack{S' \subseteq S \\ |S'|=k'}} W \mathop{\mathbb{E}}_{\varphi_B}[P(\varphi_B|_S) \mid D_{S'}] \tag{104}$$

Moreover, since every vertex in $S - S'$ is randomly assigned when $D_{S'}$ occurs, $\mathbb{E}[P(\varphi_B|_S) \mid D_{S'}]$ can be view as the average value of $P((\varphi_A|_{S'}) \circ \tau)$ over all $\tau \in [R]^{S-S'}$. Hence, we can derive the following:

$$\mathop{\mathbb{E}}_{\varphi_B}[P(\varphi_B|_S) \mid D_{S'}] = \frac{1}{R^{k-k'}} \mathop{\mathbb{E}}_{\varphi_A}\left[\sum_{\tau \in [R]^{S-S'}} P((\varphi_A|_{S'}) \circ \tau)\right]. \tag{105}$$

As a result, we have

$$\mathop{\mathbb{E}}_{\varphi_B}[WP(\varphi_B|_S)] \geq \frac{\alpha^{k-k'}(1-\alpha)^{k'}}{R^{k-k'}}\left(\mathop{\mathbb{E}}_{\varphi_A}\left[\sum_{\substack{S' \subseteq S \\ |S'|=k'}} \sum_{\tau \in [R]^{S-S'}} WP((\varphi_A|_{S'}) \circ \tau)\right]\right). \tag{106}$$

By summing (106) over all constraints $C \in \mathcal{C}$, we arrive at the following inequality.

$$\mathop{\mathbb{E}}_{\varphi_B}\left[\sum_{C=(W,S,P)\in\mathcal{C}} WP(\varphi_B|_S)\right] \tag{107}$$

$$\geq \frac{\alpha^{k-k'}(1-\alpha)^{k'}}{R^{k-k'}} \mathop{\mathbb{E}}_{\varphi_A}\left[\sum_{C=(W,S,P)\in\mathcal{C}}\left(\sum_{\substack{S'\subseteq S\\|S'|=k'}}\sum_{\tau\in[R]^{S-S'}} WP((\varphi_A|_{S'})\circ\tau)\right)\right] \tag{108}$$

$$= \binom{k}{k'}\alpha^{k-k'}(1-\alpha)^{k'} \mathop{\mathbb{E}}_{\varphi_A}\left[\sum_{C=(W,S,P)\in\mathcal{C}}\left(\sum_{\substack{S'\subseteq S\\|S'|=k'}}\sum_{\tau\in[R]^{S-S'}} \frac{W}{\binom{k}{k'}R^{k-k'}}P((\varphi_A|_{S'})\circ\tau)\right)\right] \tag{109}$$

$$= \binom{k}{k'}\alpha^{k-k'}(1-\alpha)^{k'} \mathop{\mathbb{E}}_{\varphi_A}\left[\sum_{C'=(W',S',P')\in\mathcal{C}} W'P'(\varphi_A|_{S'})\right] \tag{110}$$

The first expression is the expected value of φ_B whereas the last is $\binom{k}{k'}\alpha^{k-k'}(1-\alpha)^{k'}$ times the expected value of φ_A. Since the expected value of φ_A is at least $f(R)/R^k$ times the optimum of $(\mathcal{X},\mathcal{C})$, the expected value of φ_B is at least $(\binom{k}{k'}\alpha^{k-k'}(1-\alpha)^{k'})(f(R)/R^k)$ times the optimum of $(\mathcal{X},\mathcal{C})$.

Finally, we substitute $\alpha = \frac{k-k'}{k}$ in to get

$$\binom{k}{k'}\alpha^{k-k'}(1-\alpha)^{k'} = \binom{k}{k'}\left(\frac{k-k'}{k}\right)^{k-k'}\left(\frac{k'}{k}\right)^{k'}. \tag{111}$$

Let $l = \min\{k', k-k'\}$. We then have

$$\binom{k}{k'}\left(\frac{k-k'}{k}\right)^{k-k'}\left(\frac{k'}{k}\right)^{k'} = \binom{k}{l}\left(\frac{k-l}{k}\right)^{k-l}\left(\frac{l}{k}\right)^{l} \tag{112}$$

$$\geq \left(\frac{k}{l}\right)^{l}\left(\frac{k-l}{k}\right)^{k-l}\left(\frac{l}{k}\right)^{l} \tag{113}$$

$$\geq \left(\frac{k-l}{k}\right)^{k} \tag{114}$$

$$= \left((1-l/k)^{2k/l}\right)^{2l} \tag{115}$$

$$(\text{From Bernoulli's inequality and from } l \leq k/2) \geq 1/2^{2l}. \tag{116}$$

Hence, φ_B has expected value at least $(f(R)/2^{O(l)})/R^k$ times the optimum of $(\mathcal{X},\mathcal{C})$, which completes the proof of this lemma. ◀

Finally, Theorem 3 is an immediate consequence of applying Lemma 16 to the algorithm from [24] with $k' = 2$ and $f(R) = \Omega(R \log R)$.

5 k-Query Large Alphabet Dictator Test

We remark that the results of Section 3 also implicitly yield a k-query nonadaptive *Dictator-vs.-Quasirandom* test for functions over large alphabet. A Dictator-vs.-Quasirandom test aims to distinguish dictator functions from functions with small low-degree influences ("quasirandom").

This concept was essentially introduced in [19], and we borrow the "quasirandom" terminology from [29] (adapted here for functions over non-binary alphabets). Specifically, we have the following test:

▶ **Theorem 17.** *For any function $f : [R]^n \to [R]$, and any $i \in [R]$, let $f^i : [R]^n \to \{0,1\}$ denote the indicator function for $f(x) = i$. For any $k, R \geq 2$, set parameters $\rho = 1/\sqrt{(k-1)\log R}$, $d = 10k \log R$, and $\delta = 1/(R^{10+100k\log(R)})$. Then there exists a k-query nonadaptive Dictator-vs. Quasirandom test with the following guarantees:*

Completeness: *If f is a dictator, i.e. $f(x) = x_j$ for some coordinate $j \in [n]$, then the test passes with probability at least*

$$\rho^k = 1/((\log R)^{k/2} 2^{O(k \log k)}) .$$

Soundness: *If f has $Inf_j^{\leq d}[f^i] \leq \delta$ for all coordinates $j \in [n]$ and all projections $i \in [R]$, then the test passes with probability at most*

$$2^{O(k)}/R^{k-1} .$$

Notice that if we assume f is balanced, then this theorem is immediately implied by the techniques of Section 3. However, to extend this to general functions via "folding", we must technically show that the operation of folding keeps low-influence functions as low-influence. The full proof can be found in Appendix C.

6 Conclusions and Open Questions

We conclude by posting interesting open questions regarding the approximability of Max k-CSP$_R$ and providing our opinions on each question. First, as stated earlier, even with our results, current inapproximability results do not match the best known approximation ratio achievable in polynomial time when $3 \leq k < R$. Hence, it is intriguing to ask what the right ratio that Max k-CSP$_R$ becomes NP-hard to approximate is. Since our hardness factor $2^{O(k \log k)}(\log R)^{k/2}/R^{k-1}$ does not match Chan's hardness factor $O(k/R^{k-2})$ when $k = R$, it is likely that there is a k between 3 and $R-1$ such that a drastic change in the hardness factor, and technique that yields that factor, occurs.

Moreover, since our PCP has completeness of $1/(2^{O(k)}(\log R)^{k/2})$, even if one cannot improve on the inapproximability factor, it is still interesting if one can come up with a hardness result with almost perfect completeness. In fact, even for $k = 2$, there is no known hardness of approximation of factor better than $O(\log R/\sqrt{R})$ with near perfect completeness whereas the best UGC-hardness known is $O(\log R/R)$.

It is also interesting to try to relax assumptions for other known inapproximability results from UGC to the One-Sided UGC. Since the One-Sided UGC is implied by d-to-1 Games Conjecture, doing so will imply inapproximability results based on the d-to-1 Games Conjecture. Moreover, without going into too much detail, we remark that most attempts to refute the UGC and the d-to-1 Conjecture need the value of the game to be high [1, 5, 7, 15, 20, 25, 36]. Hence, these algorithms are not candidates to refute the One-Sided UGC. In addition, Arora, Barak and Steurer's [1] subexponential time algorithm for unique games suggest that unique games have *intermediate complexity*, meaning that, even if the UGC is true, the UGC-hardness would not imply exponential time lower bounds. On the other hand, to the best of the authors' knowledge, the ABS algorithm does not run in subexponential time when the completeness is small. Hence, the One-Sided UGC may require exponential time to solve, which could give similar running time lower bounds for the resulting hardness of approximation results. Finally, there are evidences suggesting that relaxing completeness or soundness conditions of a conjecture can make it easier; the most relevant such result is that from Feige and Reichman who proved that, if one only cares about the approximation ratio and not completeness and soundness, then unique game is hard to approximate to within factor ε for any $\varepsilon > 0$ [13].

References

1 Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for unique games and related problems. *J. ACM*, 62(5):42, 2015. doi:10.1145/2775105.

2 P. Austrin and E. Mossel. Approximation resistant predicates from pairwise independence. In *Computational Complexity, 2008. CCC'08. 23rd Annual IEEE Conference on*, pages 249–258, June 2008. doi:10.1109/CCC.2008.20.

3 Siu On Chan. Approximation resistance from pairwise independent subgroups. In *Proceedings of the Forty-fifth Annual ACM Symposium on Theory of Computing*, STOC'13, pages 447–456, New York, NY, USA, 2013. ACM. doi:10.1145/2488608.2488665.

4 Moses Charikar, MohammadTaghi Hajiaghayi, and Howard Karloff. Improved approximation algorithms for label cover problems. *Algorithmica*, 61(1):190–206, 2011. doi:10.1007/s00453-010-9464-3.

5 Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for unique games. In *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing*, STOC'06, pages 205–214, New York, NY, USA, 2006. ACM. doi:10.1145/1132516.1132547.

6 Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for maximum constraint satisfaction problems. *ACM Transactions on Algorithms*, 5(3), 2009. doi:10.1145/1541885.1541893.

7 Eden Chlamtac, Konstantin Makarychev, and Yury Makarychev. How to play unique games using embeddings. In *47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), 21-24 October 2006, Berkeley, California, USA, Proceedings*, pages 687–696, 2006. doi:10.1109/FOCS.2006.36.

8 Pierluigi Crescenzi, Riccardo Silvestri, and Luca Trevisan. On weighted vs unweighted versions of combinatorial optimization problems. *Information and Computation*, 167(1):10–26, 2001.

9 Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. *SIAM Journal on Computing*, 39(3):843–873, 2009. doi:10.1137/07068062X.

10 Irit Dinur and Igor Shinkar. On the conditional hardness of coloring a 4-colorable graph with super-constant number of colors. In *Proceedings of the 13th International Conference on Approximation, and 14 the International Conference on Randomization, and Combinatorial Optimization: Algorithms and Techniques*, APPROX/RANDOM'10, pages 138–151, Berlin, Heidelberg, 2010. Springer-Verlag. URL: http://dl.acm.org/citation.cfm?id=1886521.1886533.

11 Lars Engebretsen. The nonapproximability of non-boolean predicates. *SIAM J. Discret. Math.*, 18(1):114–129, January 2005. doi:10.1137/S0895480100380458.

12 Lars Engebretsen and Jonas Holmerin. More efficient queries in pcps for np and improved approximation hardness of maximum csp. *Random Struct. Algorithms*, 33(4):497–514, December 2008. doi:10.1002/rsa.v33:4.

13 Uriel Feige and Daniel Reichman. On systems of linear equations with two variables per equation. In Klaus Jansen, Sanjeev Khanna, JoséD.P. Rolim, and Dana Ron, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*,

volume 3122 of *Lecture Notes in Computer Science*, pages 117–127. Springer Berlin Heidelberg, 2004. doi:10.1007/978-3-540-27821-4_11.

14 Gil Goldshlager and Dana Moshkovitz. Approximating kCSP for large alphabets. https://people.csail.mit.edu/dmoshkov/papers/Approximating%20MAX%20kCSP.pdf, 2015.

15 Anupam Gupta and Kunal Talwar. Approximating unique games. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2006, Miami, Florida, USA, January 22-26, 2006*, pages 99–106, 2006. URL: http://dl.acm.org/citation.cfm?id=1109557.1109569.

16 Venkatesan Guruswami and Prasad Raghavendra. Constraint satisfaction over a non-boolean domain: Approximation algorithms and unique-games hardness. In Ashish Goel, Klaus Jansen, JoséD.P. Rolim, and Ronitt Rubinfeld, editors, *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, volume 5171 of *Lecture Notes in Computer Science*, pages 77–90. Springer Berlin Heidelberg, 2008. doi:10.1007/978-3-540-85363-3_7.

17 Venkatesan Guruswami and Ali Kemal Sinop. The complexity of finding independent sets in bounded degree (hyper)graphs of low chromatic number. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 1615–1626, 2011. doi:10.1137/1.9781611973082.125.

18 Gustav Hast. Approximating Max kCSP – outperforming a random assignment with almost a linear factor. In *Proceedings of the 32Nd International Conference on Automata, Languages and Programming*, ICALP'05, pages 956–968, Berlin, Heidelberg, 2005. Springer-Verlag. doi:10.1007/11523468_77.

19 Johan Håstad. Clique is hard to approximate within $n^{1-\varepsilon}$. In *Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on*, pages 627–636. IEEE, 1996.

20 Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thirty-fourth Annual ACM Symposium on Theory of Computing*, STOC'02, pages 767–775, New York, NY, USA, 2002. ACM. doi:10.1145/509907.510017.

21 Subhash Khot. On the unique games conjecture (invited survey). In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity, CCC 2010, Cambridge, Massachusetts, June 9-12, 2010*, pages 99–121, 2010. doi:10.1109/CCC.2010.19.

22 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):319–357, 2007. doi:10.1137/S0097539705447372.

23 Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within 2-ε. *J. Comput. Syst. Sci.*, 74(3):335–349, May 2008. doi:10.1016/j.jcss.2007.06.019.

24 Guy Kindler, Alexandra Kolla, and Luca Trevisan. Approximation of non-boolean 2csp. *CoRR*, abs/1504.00681, 2015. URL: http://arxiv.org/abs/1504.00681.

25 Alexandra Kolla. Spectral algorithms for unique games. *Computational Complexity*, 20(2):177–206, 2011. doi:10.1007/s00037-011-0011-7.

26 Konstantin Makarychev and Yury Makarychev. Approximation algorithm for non-boolean max-k-csp. *Theory of Computing*, 10:341–358, 2014. doi:10.4086/toc.2014.v010a013.

27 Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. *Ann. Math. (2)*, 171(1):295–341, 2010. doi:10.4007/annals.2010.171.295.

28 Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014. URL: http://www.cambridge.org/de/academic/subjects/computer-science/algorithmics-complexity-computer-algebra-and-computational-g/analysis-boolean-functions.

29 Ryan O'Donnell and Yi Wu. 3-bit dictator testing: 1 vs. 5/8. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2009, New York, NY, USA, January 4-6, 2009*, pages 365–373, 2009. URL: `http://dl.acm.org/citation.cfm?id=1496770.1496811`.

30 Ryan O'Donnell and Yi Wu. Conditional hardness for satisfiable 3-csps. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 – June 2, 2009*, pages 493–502, 2009. `doi:10.1145/1536414.1536482`.

31 Prasad Raghavendra. Optimal algorithms and inapproximability results for every csp? In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC'08, pages 245–254, New York, NY, USA, 2008. ACM. `doi:10.1145/1374376.1374414`.

32 Alex Samorodnitsky and Luca Trevisan. A PCP characterization of NP with optimal amortized query complexity. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, May 21-23, 2000, Portland, OR, USA*, pages 191–199, 2000. `doi:10.1145/335305.335329`.

33 Alex Samorodnitsky and Luca Trevisan. Gowers uniformity, influence of variables, and pcps. In *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing*, STOC'06, pages 11–20, New York, NY, USA, 2006. ACM. `doi:10.1145/1132516.1132519`.

34 Linqing Tang. Conditional hardness of approximating satisfiable max 3csp-q. In Yingfei Dong, Ding-Zhu Du, and Oscar Ibarra, editors, *Algorithms and Computation*, volume 5878 of *Lecture Notes in Computer Science*, pages 923–932. Springer Berlin Heidelberg, 2009. `doi:10.1007/978-3-642-10631-6_93`.

35 Luca Trevisan. Parallel approximation algorithms by positive linear programming. *Algorithmica*, 21(1):72–88, 1998. `doi:10.1007/PL00009209`.

36 Luca Trevisan. Approximation algorithms for unique games. *Theory of Computing*, 4(1):111–128, 2008. `doi:10.4086/toc.2008.v004a005`.

A Proofs of Preliminary Results

For completeness, we prove some of the preliminary results, whose formal proofs were not found in the literature by the authors.

A.1 Mollification Lemma

Below is the proof of the Mollification Lemma. We remark that, while its main idea is explained in [28], the full proof is not shown there. Hence, we provide the proof here for completeness.

Proof of Lemma 10. Let $p : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^4$ function supported only on $[-1, +1]$, such that $p(y)$ forms a probability distribution. (For example, an appropriately normalized version of $e^{-1/(1+y^2)}$ for $|y| \leq 1$). Define $p_\lambda(y)$ to be re-scaled to have support $[-\lambda, +\lambda]$ for some $\lambda > 0$:

$$p_\lambda(y) := (1/\lambda)p(y/\lambda). \tag{117}$$

Let Y_λ be a random variable with distribution $p_\lambda(y)$, supported on $[-\lambda, +\lambda]$. We will set $\lambda = \zeta/c$.

Now, define

$$\widetilde{\psi} := \mathop{\mathbb{E}}_{Y_\lambda}[\psi(x + Y_\lambda)]. \tag{118}$$

This is pointwise close to ψ, since ψ is c-Lipschitz:

$$|\widetilde{\psi}(x)-\psi(x)| = |\mathop{\mathbb{E}}_{Y_\lambda}[\psi(x+Y_\lambda)-\psi(x)]| \leq \mathop{\mathbb{E}}_{Y_\lambda}[|\psi(x+Y_\lambda)-\psi(x)|] \leq \mathop{\mathbb{E}}_{Y_\lambda}[c|Y_\lambda|] \leq c\lambda = \zeta. \quad (119)$$

Further, $\widetilde{\psi}$ is $\mathcal{C}^3$, because $\widetilde{\psi}(x)$ can be written as a convolution:

$$\widetilde{\psi}(x) = (\psi * p_\lambda)(x) \implies \widetilde{\psi}''' = (\psi * p_\lambda)''' = (\psi * p_\lambda'''). \quad (120)$$

To see that $\widetilde{\psi}'''$ is bounded, for a fixed $x \in \mathbb{R}$, define the constant $z := \psi(x)$. Then,

$$\begin{aligned}
|\widetilde{\psi}'''(x)| &= |(\psi * p_\lambda''')(x)| && (121)\\
(z \text{ is constant, so } z'=0) \quad &= |(\psi * p_\lambda''' - z' * p_\lambda'')(x)| && (122)\\
&= |(\psi * p_\lambda''' - z * p_\lambda''')(x)| && (123)\\
&= |((\psi - z) * p_\lambda''')(x)| && (124)\\
&= \left|\int_{-\infty}^{+\infty} p_\lambda'''(y)(\psi(x-y)-z)dy\right| && (125)\\
&= \left|\int_{-\infty}^{+\infty} p_\lambda'''(y)(\psi(x-y)-\psi(x))dy\right| && (126)\\
&\leq \int_{-\lambda}^{+\lambda} |p_\lambda'''(y)||\psi(x-y)-\psi(x)|dy && (127)\\
(c\text{-Lipschitz}) \quad &\leq ||p_\lambda'''||_\infty \int_{-\lambda}^{+\lambda} |cy|dy && (128)\\
&= ||p_\lambda'''||_\infty c\lambda^2. && (129)
\end{aligned}$$

Define the universal constant $\widetilde{C} := ||p'''||_\infty$. We have

$$p_\lambda'''(y) = (1/\lambda^4)p'''(y/\lambda) \implies ||p_\lambda'''||_\infty \leq (1/\lambda^4)\widetilde{C}. \quad (130)$$

With our choice of $\lambda = \zeta/c$, this yields $|\widetilde{\psi}'''(x)| \leq \widetilde{C}c^3/\zeta^2$, which completes the proof of Lemma 10. ◂

A.2 Proof of Lemma 11

Below we show the proof of Lemma 11.

Proof. First, we "mollify" the function ψ to construct a $\mathcal{C}^3$ function $\widetilde{\psi}$, by applying Lemma 10 for $\zeta = 1/R^k$. Notice that both choices of ψ are k-Lipschitz. Therefore the Mollification Lemma guarantees that $|\widetilde{\psi}'''(x)| \leq \widetilde{C}k^3R^{2k}$ for some universal constant $\widetilde{C}$.

Since $\widetilde{\psi}$ is pointwise close to ψ, with deviation at most $1/R^k$, we have

$$\begin{aligned}
&\left|\mathop{\mathbb{E}}_{y\in\{\pm1\}^{nR}}[\psi(F^{\leq d}(y))] - \mathop{\mathbb{E}}_{x\in[R]^n}[\psi(f^{\leq d}(x))]\right|\\
\leq &\left|\mathop{\mathbb{E}}_{y\in\{\pm1\}^{nR}}[\widetilde{\psi}(F^{\leq d}(y))] - \mathop{\mathbb{E}}_{x\in[R]^n}[\widetilde{\psi}(f^{\leq d}(x))]\right| + O(1/R^k). \quad (131)
\end{aligned}$$

Applying the General Invariance Principle (Theorem 9) with the function $\widetilde{\psi}$, we have

$$\left|\mathop{\mathbb{E}}_{y\in\{\pm1\}^{nR}}[\widetilde{\psi}(F^{\leq d}(y))] - \mathop{\mathbb{E}}_{x\in[R]^n}[\widetilde{\psi}(f^{\leq d}(x))]\right| \leq \widetilde{C}k^3R^{2k}10^dR^{d/2}\sqrt{\delta}. \quad (132)$$

By our choice of parameters d, δ, this is $O(1/R^k)$. ◂

B d-to-1 Games Conjecture implies One-Sided Unique Games Conjecture

In this section, we prove that if d-to-1 Games Conjecture is true, then so is One-Sided Unique Games Conjecture.

▶ **Lemma 18.** *For every $d \in \mathbb{N}$, d-to-1 Games Conjecture implies One-Sided UGC.*

Proof. Suppose that d-to-1 Games Conjecture is true for some $d \in \mathbb{N}$. We will prove One-Sided UGC; more specifically, ζ in the One-Sided UGC is $1/d$. The reduction from a d-to-1 game $(V, W, E, N, \{\pi_e\}_{e \in E})$ to a unique game $(V', W', E', N', \{\pi'_e\}_{e \in E})$ can be described as follows:

- Let $V' = V, W' = W, E' = E$, and $N' = N$
- We define π'_e as follows. For each $\theta \in [N/d]$, let the elements of $\pi_e^{-1}(\theta)$ be $\sigma_1, \sigma_2, \dots, \sigma_d \in [N]$. We then define $\pi'_e(\sigma_i) = d(\theta - 1) + i$.

Now, we will prove the soundness and completeness of this reduction.

Completeness. Suppose that the d-to-1 game is satisfiable. Let $\varphi : V \cup W \to [N]$ be the assignment that satisfies every constraint in the d-to-1 game. We define $\varphi' : V' \cup W' \to [N']$ by first assign $\varphi'(v) = \varphi(v)$ for every $v \in V$. Then, for each $w \in W$, pick $\varphi'(w)$ to be an assignment that satisfies as many edges touching w in the unique game as possible, i.e., for a fixed w, $\varphi'(w)$ is select to maximize $|\{v \in N(w) \mid \pi'_{(v,w)}(\varphi(v)) = \varphi'(w)\}|$ where $N(w)$ is the set of neighbors of w. From how $\varphi'(w)$ is picked, we have

$$\begin{aligned}&|\{v \in N(w) \mid \pi'_{(v,w)}(\varphi(v)) = \varphi'(w)\}| \\ \geq &\frac{1}{d}\sum_{i=1}^{d} |\{v \in N(w) \mid \pi'_{(v,w)}(\varphi(v)) = d(\varphi(w) - 1) + i\}|. \end{aligned} \tag{133}$$

Let $1[\pi'_{(v,w)}(\varphi(v)) = d(\varphi(w) - 1) + i]$ be the indicating variable whether $\pi'_{(v,w)}(\varphi(v)) = d(\varphi(w) - 1) + i$, we can rewrite the right hand side as follows:

$$\frac{1}{d}\sum_{i=1}^{d} |\{v \in N(w) \mid \pi'_{(v,w)}(\varphi(v)) = d(\varphi(w) - 1) + i\}| \tag{134}$$

$$= \frac{1}{d}\sum_{i=1}^{d} \sum_{v \in N(w)} 1[\pi'_{(v,w)}(\varphi(v)) = d(\varphi(w) - 1) + i] \tag{135}$$

$$= \frac{1}{d}\sum_{v \in N(w)} \sum_{i=1}^{d} 1[\pi'_{(v,w)}(\varphi(v)) = d(\varphi(w) - 1) + i]. \tag{136}$$

From how $\pi'_{(v,w)}$ is defined and since $\pi_{(v,w)}(\varphi(v)) = \varphi(w)$, there exists $i \in [d]$ such that $\pi'_{(v,w)}(\varphi(v)) = d(\varphi(w) - 1) + i$. As a result, we have

$$\frac{1}{d}\sum_{v \in N(w)} \sum_{i=1}^{d} 1[\pi'_{(v,w)}(\varphi(v)) = d(\varphi(w) - 1) + i] \geq \frac{1}{d}\sum_{v \in N(w)} 1 = \frac{|N(w)|}{d}. \tag{137}$$

In other words, at least $1/d$ fraction of edges touching w is satisfied in the unique game for every $w \in W$. Hence, φ' has value at least $1/d$, which means that the unique game also has value at least $1/d$.

Soundness. Suppose that the value of the d-to-1 game is at most γ. For any assignment $\varphi' : V' \cup W' \to [N']$ to the unique game, we can define an assignment $\varphi : V \cup W \to [N]$ by

$$\varphi(u) = \begin{cases} \varphi'(u) & \text{if } u \in V, \\ \lfloor (\varphi'(u) - 1)/d \rfloor + 1 & \text{if } u \in W. \end{cases} \tag{138}$$

From how π'_e is defined, it is easy to see that, if $\pi'_e(\varphi'(v)) = \varphi'(w)$, then $\pi_e(\varphi(v)) = \varphi(w)$. In other words, the value of φ' with respect to the unique game is no more than the value of φ with respect to the d-to-1 game. As a result, the value of the unique game is at most ε.

As a result, if it is NP-hard to distinguish a satisfiable d-to-1 game from one with value at most γ, then it is also NP-hard to distinguish a unique game of value at least $\zeta = 1/d$ from that with value at most γ, which concludes the proof of this lemma. ◀

C Proof of Dictator Test

Here we prove our result for the Dictator-vs.-Quasirandom test (Theorem 17).

Proof of Theorem 17. For $c \in [R]$, define the function

$$f_c(x_1, x_2, \ldots, x_n) := f(x_1 + c, x_2 + c, \ldots, x_n + c) - c. \tag{139}$$

Note that $\pm c$ is performed modulo R.

The test works as follows: Pick $z \in [R]^n$ uniformly at random, and let $x^{(1)}, x^{(2)}, \ldots, x^{(k)}$ be independent ρ-correlated noisy copies of z. Then, pick $c_1, c_2, \ldots, c_k$ independently uniformly at random, where each $c_i \in [R]$. Accept iff

$$f_{c_1}(x^{(1)}) = f_{c_2}(x^{(2)}) = \cdots = f_{c_k}(x^{(k)}). \tag{140}$$

For completeness, notice that if f is a dictator, then $f_c = f$ for all $c \in [R]$. Say f is a dictator on the j-th coordinate: $f(x) = x_j$. Then the test clearly accepts with probability at least ρ^k (if none of the coordinates j were perturbed in all the noisy copies $x^{(i)} \overset{\rho}{\leftarrow} z$).

For soundness: For any $i \in [R]$, let $f^i : [R]^n \to \{0,1\}$ denote the indicator function for $f(x) = i$, and similarly for $f^i_c : [R]^n \to \{0,1\}$. Notice that

$$f^i_c(x) = f^{i+c}(x + (c, c, \ldots, c)) \tag{141}$$

Then, write the acceptance probability as

$$\Pr[accept] = \Pr_{c_i, z, x^{(j)} \overset{\rho}{\leftarrow} z}[f_{c_1}(x^{(1)}) = f_{c_2}(x^{(2)}) = \cdots = f_{c_k}(x^{(k)})] \tag{142}$$

$$= \sum_{i \in [R]} \Pr_{c_i, z, x^{(j)} \overset{\rho}{\leftarrow} z}[i = f_{c_1}(x^{(1)}) = f_{c_2}(x^{(2)}) = \cdots = f_{c_k}(x^{(k)})] \tag{143}$$

$$= \sum_{i \in [R]} \mathop{\mathbb{E}}_{c_i, z, x^{(j)} \overset{\rho}{\leftarrow} z}[f^i_{c_1}(x^{(1)}) f^i_{c_2}(x^{(2)}) \ldots f^i_{c_k}(x^{(k)})] \tag{144}$$

$$(\text{Independence of } c_i) \quad = \sum_{i \in [R]} \mathop{\mathbb{E}}_{z, x^{(j)} \overset{\rho}{\leftarrow} z}[\mathop{\mathbb{E}}_{c_1}[f^i_{c_1}(x^{(1)})] \mathop{\mathbb{E}}_{c_2}[f^i_{c_2}(x^{(2)})] \ldots \mathop{\mathbb{E}}_{c_2}[f^i_{c_k}(x^{(k)})]]. \tag{145}$$

$$\tag{146}$$

If we define the function $g^i : [R]^n \to [0,1]$ as

$$g^i(x) := \mathop{\mathbb{E}}_{c}[f^i_c(x)]. \tag{147}$$

Then this acceptance probability is

$$\Pr[accept] = \sum_{i\in[R]} \mathop{\mathbb{E}}_{z, x_i \overset{\rho}{\leftarrow} z} [g^i(x^{(1)}) g^i(x^{(2)}) \dots g^i(x^{(k)})] \tag{148}$$

$$= \sum_{i\in[R]} \mathop{\mathbb{E}}_{z}[(T_\rho g^i(z))^k]. \tag{149}$$

Notice that $\mathbb{E}_x[g^i(x)] = 1/R$, because

$$\mathop{\mathbb{E}}_{x}[g^i(x)] = \mathop{\mathbb{E}}_{x,c}[f^i_c(x)] = \mathop{\mathbb{E}}_{x,c}[f^{i+c}(x + (c, c, \dots, c))] \tag{150}$$

$$(c, x \text{ same joint distribution as } i + c, x + c) \quad = \mathop{\mathbb{E}}_{x,c}[f^c(x)] \tag{151}$$

$$= \mathop{\mathbb{E}}_{x}[\mathop{\mathbb{E}}_{c}[f^c(x)]] = \mathop{\mathbb{E}}_{x}[1/R] = 1/R. \tag{152}$$

Thus, if the function g^i has small low-degree influences, then Lemma 15 (Main Lemma) applied to g^i in line (149) directly implies that this acceptance probability is $2^{O(k)}/R^{k-1}$. We will now formally show that the influences of the "expected folded function" g^i are bounded by the influences of the original f^i.

First, the Fourier coefficients of g^i are

$$\hat{g}^i(s) = \mathop{\mathbb{E}}_{c}[\hat{f}^i_c(s)]. \tag{153}$$

Thus the low-degree influences of g^i are bounded as

$$Inf_j^{\leq d}[g^i] = \sum_{\substack{s\in[R]^n \\ s(j)\neq 1, |s|\leq d}} \hat{g}^i(s)^2 \tag{154}$$

$$= \sum_{\substack{s\in[R]^n \\ s(j)\neq 1, |s|\leq d}} \mathop{\mathbb{E}}_{c}[\hat{f}^i_c(s)]^2 \tag{155}$$

$$\leq \sum_{\substack{s\in[R]^n \\ s(j)\neq 1, |s|\leq d}} \mathop{\mathbb{E}}_{c}[\hat{f}^i_c(s)^2] \tag{156}$$

$$= \mathop{\mathbb{E}}_{c}[\sum_{\substack{s\in[R]^n \\ s(j)\neq 1, |s|\leq d}} \hat{f}^i_c(s)^2] \tag{157}$$

$$= \mathop{\mathbb{E}}_{c}[Inf_j^{\leq d}[f^i_c]]. \tag{158}$$

Finally, we must relate the influences of f^i_c to the influences of f^i. For a fixed $c \in [R]$, we have

$$Inf_j^{\leq d}[f^i_c] = Inf_j[(f^i_c)^{\leq d}] \tag{159}$$

$$= \mathop{\mathbb{E}}_{x\in[R]^n}[Var_{x_j\in[R]}[(f^i_c)^{\leq d}]] \tag{160}$$

$$= \mathop{\mathbb{E}}_{x\in[R]^n}[Var_{x_j\in[R]}[(f^{i+c})^{\leq d}(x_1 + c, x_2 + c, \dots, x_n + c)]] \tag{161}$$

$$= \mathop{\mathbb{E}}_{x\in[R]^n}[Var_{x_j\in[R]}[(f^{i+c})^{\leq d}(x_1, x_2, \dots, x_n)]] \tag{162}$$

$$= Inf_j[(f^{i+c})^{\leq d}] \tag{163}$$

$$= Inf_j^{\leq d}[f^{i+c}]. \tag{164}$$

Therefore, if $Inf_j^{\leq d}[f^i] \leq \delta$ for all coordinates $j \in [n]$ and all projections $i \in [R]$ (as we assume for soundness), then from (158) and (164) we have

$$Inf_j^{\leq d}[g^i] \leq \mathop{\mathbb{E}}_c[Inf_j^{\leq d}[f_c^i]] = \mathop{\mathbb{E}}_c[Inf_j^{\leq d}[f^{i+c}]] \leq \delta. \tag{165}$$

Thus the function g^i has small low-degree influences as well.

So we can complete the proof, continuing from line (149) and applying our Main Lemma to g^i:

$$\begin{aligned} \Pr[accept] &= \sum_{i\in[R]} \mathop{\mathbb{E}}_z[(T_\rho g^i(z))^k] && (166)\\ \text{(Lemma 15)} \quad &\leq \sum_{i\in[R]} 2^{O(k)}/R^k && (167)\\ &= 2^{O(k)}/R^{k-1}. && (168) \end{aligned}$$

◀