

Explicit optimal hardness via Gaussian stability results

Anindya De*

Computer Science Division,
University of California, Berkeley

Elchanan Mossel†

Dept. of Statistics and Computer Science,
University of California, Berkeley

Abstract

The results of Raghavendra (2008) show that assuming Khot’s Unique Games Conjecture (2002), for every constraint satisfaction problem there exists a generic semi-definite program that achieves the optimal approximation factor. This result is existential as it does not provide an explicit optimal rounding procedure nor does it allow to calculate exactly the Unique Games hardness of the problem.

Obtaining an explicit optimal approximation scheme and the corresponding approximation factor is a difficult challenge for each specific approximation problem. Khot *et al.* (2004) established a general approach for determining the exact approximation factor and the corresponding optimal rounding algorithm for any given constraint satisfaction problem. However, this approach crucially relies on results explicitly proving optimal partitions in the Gaussian space. Until recently, Borell’s result (1985) was the only non-trivial Gaussian partition result known.

In this paper we derive the first explicit optimal approximation algorithm and the corresponding approximation factor using a new result on Gaussian partitions due to Isaksson and Mossel (2012). This Gaussian result allows us to determine the exact Unique Games Hardness of MAX-3-EQUAL. In particular, our results show that Zwick’s algorithm for this problem achieves the optimal approximation factor and prove that the approximation achieved by the algorithm is ≈ 0.796 as conjectured by Zwick.

We further use the previously known optimal Gaussian partitions results to obtain a new Unique Games Hardness factor for MAX-k-CSP: Using the well known fact that jointly normal pairwise independent random variables are fully independent, we show that the UGC hardness of Max-k-CSP is $\frac{\lceil (k+1)/2 \rceil}{2^{k-1}}$, improving on results of Austrin and Mossel (2009).

1 Introduction

The study of inapproximability of Constraint Satisfaction Problems (CSPs) has been an important area of research in complexity theory in the past two decades. A CSP is specified by an alphabet $[q]$ and a set of predicates \mathcal{P} such that all $P \in \mathcal{P} : [q]^k \rightarrow \{0, 1\}^1$. Here k is called the arity of the predicate. An instance of the problem (say G) is given by n variables x_1, \dots, x_n and a set of constraints \mathcal{E} such that every $e \in \mathcal{E}$ is of the form $e = (S, P)$ where $S \in [n]^k$ and $P \in \mathcal{P}$.

Now, consider any mapping $\mathcal{L} : [n] \rightarrow [q]$. A constraint $e = (S, P)$ is said to be “satisfied” if $P(\mathcal{L}(S_1), \dots, \mathcal{L}(S_k)) = 1$ where S_i is the i^{th} element of S . We also define $val_{\mathcal{L}}(G)$ as $val_{\mathcal{L}}(G) = \mathbf{E}_{v \in \mathcal{E}}[P(\mathcal{L}(S_1), \dots, \mathcal{L}(S_k))]$. The algorithmic task is to come up with the mapping \mathcal{L} such that $val_{\mathcal{L}}(G)$ is maximized. Towards this, we define $val(G) = \max_{\mathcal{L}} val_{\mathcal{L}}(G)$.

The reason for studying the very general framework of CSPs is because many specific problems of interest say MAX-CUT, MAX-3-SAT etc. fall in this framework. In the past two decades, there have been important results in the study of inapproximability of CSPs including the monumental work of Håstad [Hås01]

*anindya@cs.berkeley.edu. Research supported by Satish Rao’s NSF award CCF-1118083.

†mossel@stat.berkeley.edu. Research supported by NSF award DMS-1106999, CCF 1320105 and DOD ONR grant N000141110140

¹We are assuming a somewhat restricted form of a CSP where all the predicates have the same arity.

who obtained optimal inapproximability results for CSPs like MAX-3-SAT and MAX-3-LIN. Still, a gap continued to exist between the known algorithms and hardness results for many important CSPs like MAX-CUT and MAX-2-SAT. Towards closing this gap, Khot [Kho02] introduced the Unique Games Conjecture (UGC) which stated the following (equivalent form from [KKMO07]):

Conjecture 1. *Given any $\delta > 0$, there is a prime p such that given a set of linear equations $x_i - x_j = c_{ij} \pmod{p}$, it is NP-hard to decide which one of the following is true:*

- *There is an assignment to the x_i 's which satisfies at least $1 - \delta$ fraction of the constraints.*
- *All assignments to the x_i 's can satisfy at most δ fraction of the constraints.*

A series of (often optimal) inapproximability results were proven using the Unique Games Conjecture starting with [KR08, KKMO07] which culminated in the beautiful result of Raghavendra [Rag08] who showed that for every CSP of constant arity and alphabet size, there is a simple and generic SDP which is optimal assuming the Unique Games Conjecture. More specifically, he showed the following.

Theorem 2. *Suppose that for the generic SDP, there is an instance G such that $\text{val}(G) = s$ while the SDP objective value is c . Then, assuming the UGC, given an instance G' of the CSP such that $\text{val}(G') = c - \eta$, it is NP-hard to find a \mathcal{L} such that $\text{val}_{\mathcal{L}}(G) \geq s + \eta$ for any $\eta > 0$. Further, there is an efficient rounding algorithm such that given an instance G with value c on the instance G , it finds an assignment \mathcal{L} with value $s - \eta$ (for $\eta > 0$).*

While this result essentially settles the question of approximability of CSPs from an abstract perspective, perhaps not too surprisingly, it says nothing about the exact hardness factors for specific CSPs. This is in contrast to the situation in the case of MAX-CUT [KKMO07] or MAX-2-SAT [Aus07] where exact inapproximability factors are known. The reason is that in Raghavendra's framework (and all previous results), determining the optimal inapproximability result for a specific CSP requires knowledge of the optimal partitioning of the Gaussian space for the corresponding predicate. While the optimal partitioning is known for the predicates corresponding to MAX-CUT and MAX-2-SAT, it is not known for arbitrary predicates. In fact, it should also be mentioned that while Raghavendra's result is a generalization of the results for MAX-CUT and MAX-2-SAT, it does not imply the results for MAX-CUT or MAX-2-SAT without the knowledge of the optimal Gaussian partitioning. Likewise, even though the rounding algorithm in [Rag08] is efficient, it is a brute force search over a small space that results only in a close to optimal rounding scheme. Thus, in a sense, the result provides implicitly a sequence of rounding algorithm whose approximation factors is guaranteed to converge to the hardness factor. This again is different from the rounding algorithms in [GW95, Zwi98, LLZ02] where the rounding algorithm is far more explicit (in the first two cases, it is simply random hyperplane rounding).

We now elaborate on the reason for difficulty in establishing exact hardness factors: The exact hardness factor in the case of MAX-CUT [KKMO07] and MAX-2-SAT [Aus07] crucially rely on Gaussian Analysis. More specifically, it uses the invariance principle [MOO10] together with a result in Gaussian space specifying explicitly an *Optimal Gaussian Partition* for the particular predicate. However, only few optimal Gaussian partitions are known (or even conjectured). In fact, to the best of our knowledge, before this paper, Borell's result [Bor85] was the only non-trivial Gaussian partition result used in hardness of approximation (for e.g., [KKMO07, Aus07]).

The above issue also explains the "brute-force" search aspect of the rounding scheme in [Rag08]. The optimal rounding scheme and the optimal gaussian partitioning (for a given predicate) are known to be intimately linked to each other (see [Rag08] for a detailed explanation). In absence of knowledge of the optimal partitioning, [Rag08] uses dimension reduction ([JL84]) to reduce the dimension of the SDP solution and subsequently resorts to brute-force search in the low-dimensional space. The proof of optimality of this algorithm (assuming the UGC) uses the invariance principle.

1.1 Our contributions In this paper, we consider two maximization CSPs, namely, MAX-3-EQUAL and MAX-k-CSP. Since we are dealing with maximization problems, we set the (usual) convention that a (randomized) algorithm is said to give an α -approximation (for $\alpha \leq 1$) if (in expectation over the randomness of the algorithm), the value of the output is at least α times the optimal value.

We first start by describing our result for MAX-3-EQUAL. In MAX-3-EQUAL, the variables are boolean-valued and every constraint consists of three literals and it is satisfied if and only if all the three literals are either all zeros or all ones. We show that assuming the Unique Games Conjecture, the MAX-3-EQUAL problem is $\alpha_{EQU} \approx 0.796$ hard to approximate in polynomial time. On the complementary side, we also provide a polynomial time algorithm for this problem with the approximation ratio α_{EQU} . More formally, we prove

Theorem 3. *There is a polynomial time approximation algorithm for the MAX-3-EQUAL problem which achieves the following approximation ratio:*

$$\alpha_{EQU} := \inf_{\delta \in (0,1]} \frac{1 - \frac{3 \cos^{-1}(1-\delta)}{2\pi}}{1 - \frac{3\delta}{4}} \approx 0.796.$$

Assuming the Unique Games Conjecture, for every $\delta > 0$ there is no polynomial time that provides a better approximation ratio than $\alpha_{EQU} + \delta$.

The hardness proof uses a recent Gaussian noise stability result of Isaksson and Mossel [IM12] which does not seem to have been previously used in the literature for proving hardness of approximation results. In fact, all previous optimal hardness of approximation results with a “non-trivial” approximation ratio were dependent on the Gaussian noise stability result of Borell [Bor85] eg. MAX-CUT, MAX-2-SAT.

We also give an analytic proof of the performance of the random hyperplane rounding algorithm on the generic SDP for MAX-3-EQUAL (from [Rag08]) showing that the approximation ratio achieved by this rounding algorithm is exactly α_{EQU} .² Our proof is computer assisted but completely rigorous. We note that while Zwick [Zwi98] also considers this problem and analyzes the performance of this algorithm, the analysis is a computer based search and he notes that there is a possibility of the search having missed the worst instance for the rounding algorithm. Nevertheless, the claimed optimum in [Zwi98] is same as the optimum of our SDP³.

Remark 4. *After the publication of the preprint, David Williamson [Wil13] informed us that our analysis of the SDP is essentially identical to the analysis of MAX-DICUT SDP from [GW95]. Thus, the analysis from [GW95] can be plugged in to give a much shorter proof for the performance of our algorithm.*

While revisiting the study of the relationship between Gaussian partitions and UGC hardness, we additionally prove hardness results for MAX-k-CSPs. In particular, we investigate the hardness of the MAX-k-AND predicate, i.e., every constraint consists of k literals ℓ_1, \dots, ℓ_k and the constraint is satisfied if and only if $\ell_1 = \dots = \ell_k = 1$. Following [Mos10] and [AM09] by using the fact that in Gaussian space, pair-wise independence implies independence, we prove the following theorem:

Theorem 5. *Assuming the Unique Games Conjecture, for every $\eta > 0$, there is no polynomial time approximation algorithm that provides an approximation ratio better than $\frac{\lceil (k+1)/2 \rceil}{2^{k-1}}$ for the MAX-k-AND problem.*

²We actually do a variant of the random hyperplane rounding algorithm where we sample normal random variables with the covariance matrix given by the SDP vectors. Then each variable is assigned 0 or 1 depending on the sign of the corresponding normal random variable. Our analysis goes through even if the actual random hyperplane algorithm is used.

³We elaborate on the difference between Zwick’s SDP and our SDP in Section 6.

This improves upon [AM09] where it was shown that MAX-k-CSP is $(k + O(k^{0.525}))/2^k$ hard to approximate. Assuming the Hadamard Conjecture, they could improve it to $\lceil (k + 1)/4 \rceil / 2^{k-2}$.

It is worth mentioning that [AM09] proves the aforementioned hardness for a very general class of predicates (ones whose satisfying assignments support pairwise independent distributions) but MAX-k-AND is not included in that class of CSPs. Another important point of difference is that [AM09] shows that given a MAX-k-CSP with optimal value $1 - \eta$, it is (Unique Games) hard to find an assignment which satisfies $\frac{k + O(k^{0.525})}{2^k} + \eta$ fraction of the constraints (for any $\eta > 0$). In terms of PCPs, the PCP in [AM09] has near perfect completeness. This in fact is true even for an earlier paper on hardness of MAX-k-CSPs by Samorodnitsky and Trevisan [ST06]. In contrast, our result shows that given an instance of MAX-k-CSP with optimal value $\frac{1}{2^{\lceil (k+1)/2 \rceil}} - \eta$, it is hard to find an assignment satisfying more than $\frac{1}{2^k} + \eta$ fraction of the constraints.

We do remark that while our improvement over [AM09] might seem very minor, Makarychev and Makarychev [MM12] give a $0.62k/2^k$ approximation algorithm for MAX-k-CSP over boolean alphabet. This shows that in some sense, the scope of improvement in the existing hardness results for MAX-k-CSPs is rather limited. Of course, the question of closing the gap between our hardness result and the performance of the algorithm of Charikar *et al.* remains open.

Overview of proofs of hardness: The two main novelties in our paper are:

- Use of the new Gaussian stability result of Isaksson and Mossel [IM12] to construct a “dictatorship” test for MAX-3-EQUAL.
- Use of the “obvious” Gaussian stability result (i.e., stable partitions for independent gaussians) in a new context to construct a “dictatorship” test for MAX-k-AND.

In particular, both these dictatorship tests are constructed by a careful combination of a “good” choice of distribution (for the dictatorship test) and the relevant Gaussian stability result (along with the Invariance principle). Given the dictatorship test, getting the corresponding Unique Games hardness result is rather standard (see [KKMO07, Rag08]). For the sake of completeness, we give a complete proof of hardness of MAX-3-EQUAL using the corresponding dictatorship test. For MAX-k-AND, we do not show the conversion of the dictatorship test to a Unique Games hardness result as the proof is completely analogous to that of MAX-3-EQUAL.

To show the tightness of the UG-hardness result for MAX-3-EQUAL, we also devote a major part of the paper towards analyzing the performance of our rounding algorithm on the generic SDP from [Rag08] and showing that it indeed matches the hardness result. We would like to emphasize that while the Gaussian stability result of [IM12] applies to a set of k Gaussian variables (for any k), we do not know if this can yield a tight hardness result for MAX-k-EQUAL. In particular, while the Gaussian stability result will imply some hardness of approximation for MAX-k-EQUAL, currently, we do not have an algorithm whose approximation ratio provably matches the hardness result. We elaborate more on this in Section 7.

1.2 Organization Section 2 states all the fourier analytic and other technical preliminaries required for this paper. Section 3 describes a dictatorship test where the tester checks for equality of three literals. Section 4 describes a dictatorship test where the tester checks if all the k literals are 1. Section 5 has the two main theorems of this paper, namely a UG-hardness result for the MAX-3-EQUAL problem and a UG-hardness result for MAX-k-AND. Section 6 describes a SDP relaxation and a rounding algorithm for the MAX-3-EQUAL problem showing the tightness of the hardness result.

2 Preliminaries

2.1 Basics of Fourier analysis Our proofs are significantly dependent on fourier analysis. We start by giving several important definitions. For a more extensive reference, see lecture notes by Mossel [Mos05].

We recall that any function $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ can be written as a multi-linear polynomial.

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

where $x_S = \prod_{i \in S} x_i$. Moreover, considering the uniform measure over $\{-1, 1\}^n$, we have:

$$\mathbf{E}[f] = \hat{f}(\emptyset), \quad \text{Var}[f] = \sum_{S \neq \emptyset} \hat{f}^2(S).$$

The i 'th influence of f is given by

$$I_i(f) := \mathbf{E}_{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n} [\text{Var}[f|x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n]] = \sum_{S: i \in S} \hat{f}^2(S).$$

2.2 Noise operators and their properties We will also require the notion of noise operators. We consider a particularly important instantiation of the Bonami-Beckner operator namely that on functions over the boolean hypercube $\{-1, 1\}^n$ equipped with the uniform measure.

Definition 6. For $\rho \in [-1, 1]$, we define the Bonami-Beckner operator T_ρ on functions $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ as follows.

$$T_\rho f(x) = \mathbf{E}_{y \sim_\rho x} [f(y)],$$

where each coordinate y_i is set to be x_i independently with probability $(1 + \rho)/2$ and $-x_i$ with probability $(1 - \rho)/2$.

The effect of the Bonami-Beckner operator T_ρ can be conveniently expressed in terms of the fourier spectrum of a function. In particular, if f is as above, then

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

The following standard lemma proves a bound on the number of coordinates with high influence on a function after applying the Bonami-Beckner operator on it, see e.g. [KKMO07].

Lemma 7. Let $f : \{-1, 1\}^n \rightarrow [0, 1]$ and $\tau, \gamma > 0$. If $\mathcal{A}(f) = \{i : \text{Inf}_i(T_{1-\gamma} f) \geq \tau\}$, then $|\mathcal{A}(f)| \leq 1/(\gamma\tau)$.

The next lemma is a specialization of Lemma 6.2 from [Mos10]. It says that expected value of product of polynomials does not change by a lot when noise is added provided individual coordinates come from correlated probability spaces such that no coordinate is absolutely fixed given rest of the coordinates.

Lemma 8. For $1 \leq i \leq n$, let $(\Omega_i, \mu_i) = (\{-1, 1\}^k, \mu_i)$ where

$$\min_{x \in \{-1, 1\}^k} \mu_i(x) \geq \alpha > 0.$$

Let $(\Omega, \mu) = \prod_{i=1}^n (\Omega_i, \mu_i)$. For $1 \leq a \leq k$, let μ_i^a be the a 'th marginal of μ_i , in other words

$$\mu_i^a(x) = \mu_i(\{(x_1, \dots, x_k) : x_a = x\}).$$

Let $\mu^a = \prod_{i=1}^n \mu_i^a$. An element $x \in \Omega$ is a $k \times n$ matrix. We write x^a for the a 'th row of x which is distributed according to μ^a . For $1 \leq a \leq k$, let Q_a be a multilinear polynomial $Q_a : \{-1, 1\}^n \rightarrow [-1, 1]$. Then, for all $\epsilon > 0$, $\exists \gamma = \gamma(\epsilon, \alpha) > 0$ such that

$$\left| \mathbf{E} \left[\prod_{a=1}^k Q_a(x^a) \right] - \mathbf{E} \left[\prod_{a=1}^k T_{1-\gamma} Q_a(x^a) \right] \right| \leq \epsilon k.$$

2.3 Gaussian Stability results The following theorem from Isaksson and Mossel [IM12] is the main technical result that we use here.

Theorem 9. (Theorem 5.1, [IM12]) *Let $\Omega = \{-1, 1\}^k$, $\rho \in [0, 1]$ and let μ be a probability distribution over Ω such that*

- $\mu(x) \geq \alpha > 0$ for all x .
- For $s, t \in \{-1, 1\}$ and all $1 \leq a \neq b \leq k$:

$$\mu(x_a = s, x_b = t) = \frac{1}{2} \rho \delta(s, t) + \frac{1}{4} (1 - \rho),$$

where $\delta(s, t) = 1$ iff $s = t$. Consider the space (Ω^n, μ^n) . An element $x \in \Omega^n$ may be viewed as a $k \times n$ matrix. Write x^a for the a 'th row of this matrix for $1 \leq a \leq k$. Note that x^a is uniformly distributed in $\{-1, 1\}^n$.

Then for every $\epsilon > 0$, $\exists \tau = \tau(\epsilon, k, \alpha) > 0$ such that for any $f_1, \dots, f_k : \{-1, 1\}^n \rightarrow [0, 1]$ satisfying $\max_{i,j} \text{Inf}_i(f_j) \leq \tau$,

$$\mathbf{E} \left[\prod_{a=1}^k f_a(x^a) \right] \leq \mathbf{Pr}[\forall a \in [k] : \mathcal{Z}_a \leq t_j] + \epsilon,$$

where $\mathcal{Z}_1, \dots, \mathcal{Z}_k \sim \mathcal{N}(0, 1)$ are jointly normal and $\text{Cov}(\mathcal{Z}_a, \mathcal{Z}_{a'}) = \rho$ for all $a \neq a'$ and each t_j is chosen so that $\mathbf{Pr}[\mathcal{Z}_a \leq t_a] = \mathbf{E}[f_a]$.

To intuitively understand the above theorem, consider the case when $f_1 = \dots = f_k = f$ has range $\{0, 1\}$. Also, let $x^1, \dots, x^k \in \{-1, 1\}^n$ such that each x^i is uniform in $\{-1, 1\}^n$ and for any $j \in [n]$ and $i \neq \ell \in [k]$, the j^{th} bit of x^i and x^ℓ are ρ -correlated. Let us equip \mathbb{R}^n with the standard normal measure and define the function $\tilde{f} : \mathbb{R}^n \rightarrow \{-1, 1\}$ as follows : $\tilde{f} : x \mapsto \text{sgn}(x_1 - \theta)$ where x_1 is the first coordinate of x and θ is chosen so that $\mathbf{E}_{x \in \{-1, 1\}^n} [f(x)] = \mathbf{E}_{x \in \mathcal{N}^n(0, 1)} [\tilde{f}(x)]$. Then, for all “low-influence” function f , the probability that $\forall \ell \in [k], f(x^\ell) = 1$ is upper bounded by the probability that $\forall \ell \in [k], \tilde{f}(x^\ell) = 1$. We also consider the corollary of the above theorem when $\rho = 0$. We do remark that the following corollary can actually be obtained using the Invariance principle from Mossel [Mos10] and does not require the full strength of [IM12].

Corollary 10. *Let $\Omega = \{-1, 1\}^k$ and let μ be a probability distribution over Ω such that*

- $\mu(x) \geq \alpha > 0$ for all x .
- For $s, t \in \{-1, 1\}$ and all $1 \leq a \neq b \leq k$:

$$\mu(x_a = s, x_b = t) = \frac{1}{4}.$$

Consider the space (Ω^n, μ^n) . An element $x \in \Omega^n$ may be viewed as a $k \times n$ matrix. Write x^a for the a 'th row of this matrix for $1 \leq a \leq k$. Note that x^a is uniformly distributed in $\{-1, 1\}^n$.

Then for every $\epsilon > 0$, $\exists \tau = \tau(\epsilon, k, \alpha) > 0$ such that for any $f_1, \dots, f_k : \{-1, 1\}^n \rightarrow [0, 1]$ satisfying $\max_{i,j} \text{Inf}_i(f_j) \leq \tau$,

$$\mathbf{E} \left[\prod_{a=1}^k f_a(x^a) \right] \leq \prod_{a=1}^k \mathbf{E}[f_a] + \epsilon.$$

Proof. The corollary follows by putting $\rho = 0$ in Theorem 9 and then observing that $\mathcal{Z}_1, \dots, \mathcal{Z}_k \sim \mathcal{N}(0, 1)$ in the conclusion of Theorem 9 are simply i.i.d. $\mathcal{N}(0, 1)$ random variables. \square

2.4 Useful facts We will require the following very useful fact about Gaussians. For a reference, see [Bac63].

Fact 11. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \sim \mathcal{N}(0, 1)$ such that $(\mathcal{X}, \mathcal{Y}, \mathcal{Z})$ are jointly normal and $\text{Cov}(\mathcal{X}, \mathcal{Y}) = \rho_1$, $\text{Cov}(\mathcal{Z}, \mathcal{Y}) = \rho_2$ and $\text{Cov}(\mathcal{X}, \mathcal{Z}) = \rho_3$. Then,

$$\Pr[X, Y, Z \leq 0] = \Pr[X, Y, Z \geq 0] = \frac{1}{2} - \frac{\cos^{-1} \rho_1 + \cos^{-1} \rho_2 + \cos^{-1} \rho_3}{4\pi}.$$

We will also use the following very useful construction of pairwise independent distribution (which can be found in [BGGP12, BP89]).

Fact 12. For any $k \in \mathbb{N}$, there is a distribution D_k on $\{-1, 1\}^k$ such that the following holds:

- For any $i \in [k]$, $\mathbf{E}[x_i] = 0$.
- For any $i, j \in [k]$ and $i \neq j$, $\mathbf{E}[x_i x_j] = 0$, i.e., any two coordinates are pairwise independent.
- $\Pr_{x \in D_k}[x_1 = \dots = x_k = 1] = \frac{1}{2^{\lceil (k+1)/2 \rceil}}$.

Proof. We will construct a symmetric distribution D_k with the above mentioned properties. First, we consider the case when k is odd. In this case, define D_k as follows:

$$D_k(x) = \begin{cases} \frac{1}{k+1} & \text{if } x = (1, \dots, 1), \\ \frac{k}{k+1} \cdot \frac{1}{\binom{k}{(k+1)/2}} & \text{if } \sum_{i=1}^k x_i = -1, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to verify that all the three required properties hold for this construction of D_k . We next move to the case when k is even. In this case, we define D_k as

$$D_k(x) = \begin{cases} \frac{1}{k+2} & \text{if } x = (1, \dots, 1), \\ \frac{1}{2} \cdot \frac{1}{\binom{k}{k/2}} & \text{if } \sum_{i=1}^k x_i = 0, \\ \frac{k}{2k+4} \cdot \frac{1}{\binom{k}{1+\frac{k}{2}}} & \text{if } \sum_{i=1}^k x_i = -2, \\ 0 & \text{otherwise.} \end{cases}$$

Again, it is easy to verify that all the three properties required of D_k hold for this construction. \square

3 Dictatorship test for MAX-3-EQUAL

In this section, we will construct a dictatorship test where the tester checks for equality of 3 literals. More precisely, we will prove the following theorem:

Theorem 13. *For any $0 < \delta < 1$ and $\epsilon > 0$, there is a distribution D_δ^n over $(\{-1, 1\}^n)^3$ such that if $(X, Y, Z) \sim D_\delta^n$, then for every $f : \{-1, 1\}^n \rightarrow [0, 1]$ with $\mathbf{E}[f] = 1/2$,*

- *If $f(x) = (1 + x_i)/2$ for some $i \in [n]$, then*

$$\mathbf{E}_{(X,Y,Z) \sim D_\delta^n} [f(X) \cdot f(Y) \cdot f(Z) + (1 - f(X)) \cdot (1 - f(Y)) \cdot (1 - f(Z))] = 1 - \frac{3\delta}{4}.$$

- *$\exists \tau = \tau(\delta, \epsilon) > 0$ and $\eta = \eta(\delta, \epsilon) > 0$ such that if $\max_i \text{Inf}_i(T_{1-\eta}f) \leq \tau$,*

$$\mathbf{E}_{(X,Y,Z) \sim D_\delta^n} [f(X) \cdot f(Y) \cdot f(Z) + (1 - f(X)) \cdot (1 - f(Y)) \cdot (1 - f(Z))] \leq 1 - (3 \cos^{-1}(1 - \delta))/2\pi + \epsilon.$$

Before starting the proof, we note that if f were a boolean function with range $\{0, 1\}$, then $f(X) \cdot f(Y) \cdot f(Z) + (1 - f(X)) \cdot (1 - f(Y)) \cdot (1 - f(Z))$ is 1 if and only if $f(X) = f(Y) = f(Z)$. Thus, we have a dictatorship test which checks for equality of 3 bits.

Proof. Let us define a distribution D_δ over $\{-1, 1\}^3$ as follows:

$$D_\delta(x) = \begin{cases} \frac{1}{2} - \frac{3\delta}{8} & \text{if } x = (1, 1, 1) \text{ or } x = (-1, -1, -1), \\ \frac{\delta}{8} & \text{otherwise.} \end{cases}$$

Let D_1, D_2, \dots, D_n be n i.i.d. samples of D_δ . Let $D_i(j)$ denote the j^{th} bit of D_i . With this, let us define $X, Y, Z \in \{-1, 1\}^n$ as

$$X = (D_1(1), \dots, D_n(1)) \quad Y = (D_1(2), \dots, D_n(2)) \quad Z = (D_1(3), \dots, D_n(3)).$$

We let the joint distribution (X, Y, Z) as defined here be D_δ^n . We start with the proof of the first item.

Completeness: Note for any particular $i \in [n]$, the i^{th} coordinate of D_δ has the same string with probability $1 - 3\delta/4$. Now, if $f(x) = (1 + x_i)/2$, then it means that $f(x) = 1$ if $x_i = 1$ and 0 otherwise. Hence, we have

$$\begin{aligned} \mathbf{E}_{(X,Y,Z) \in D_\delta^n} [f(X) \cdot f(Y) \cdot f(Z) + (1 - f(X)) \cdot (1 - f(Y)) \cdot (1 - f(Z))] \\ = \mathbf{E}_{(X,Y,Z) \in D_\delta^n} [\mathbf{I}(X_i = Y_i = Z_i)] = 1 - 3\delta/4, \end{aligned}$$

where $\mathbf{I}(P)$ denotes the indicator function for the predicate P . This finishes the proof of the first item. We next do the proof of the second item.

Soundness: Let \mathcal{Q} be the multilinear polynomial representation of f . Note that for any $x \in \{-1, 1\}^n$, $|\mathcal{Q}(x)| \leq 1$. Let Ω be the probability space with domain $\{-1, 1\}^3$ and probability measure D_δ on it. Note that $\forall x \in \{-1, 1\}^3$, $D_\delta(x) \geq \delta/8$. Hence, by Lemma 8, we get that $\exists \eta = \eta(\delta, \epsilon) > 0$, such that,

$$|\mathbf{E}[f(X) \cdot f(Y) \cdot f(Z) - T_{1-\eta}f(X) \cdot T_{1-\eta}f(Y) \cdot T_{1-\eta}f(Z)]| \leq \epsilon/4. \quad (1)$$

Likewise, we get that

$$|\mathbf{E}[(1 - f(X)) \cdot (1 - f(Y)) \cdot (1 - f(Z)) - (1 - T_{1-\eta}f(X)) \cdot (1 - T_{1-\eta}f(Y)) \cdot (1 - T_{1-\eta}f(Z))]| \leq \epsilon/4. \quad (2)$$

In the last two equations, $(X, Y, Z) \sim D_\delta^n$. We now apply Theorem 9. In particular, note that if $(X, Y, Z) \sim D_\delta^n$, then the variables (X_i, Y_i, Z_i) are independently and identically distributed. Also, for any $i \in [n]$, X_i , Y_i and Z_i are pairwise $\rho = (1 - \delta)$ correlated and for any $(x, y, z) \in \{-1, 1\}^3$, $\Pr[(X_i, Y_i, Z_i) = (x, y, z)] \geq \delta/8 > 0$. Finally, note that X , Y and Z are distributed as U_n . Hence

$$\mathbf{E}_X[f(X)] = \mathbf{E}_Y[f(Y)] = \mathbf{E}_Z[f(Z)] = 1/2.$$

As the Bonami Beckner operator preserves expectation of the function under the uniform distribution, we get

$$\mathbf{E}_X[T_{1-\eta}f(X)] = \mathbf{E}_Y[T_{1-\eta}f(Y)] = \mathbf{E}_Z[T_{1-\eta}f(Z)] = 1/2.$$

Thus, by Theorem 9, $\exists \tau = \tau(\delta, \epsilon)$ such that if $\max_i \text{Inf}_i(T_{1-\eta}f) \leq \tau$, then we have

$$|\mathbf{E}_{(X,Y,Z) \in D_\delta^n}[T_{1-\eta}f(X) \cdot T_{1-\eta}f(Y) \cdot T_{1-\eta}f(Z)]| \leq \Pr[\mathcal{X}, \mathcal{Y}, \mathcal{Z} \leq 0] + \epsilon/4,$$

where $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \sim \mathcal{N}(0, 1)$ and $\text{Cov}(\mathcal{X}, \mathcal{Y}) = \text{Cov}(\mathcal{Z}, \mathcal{Y}) = \text{Cov}(\mathcal{X}, \mathcal{Z}) = 1 - \delta$. Here, we again assume that τ in the hypothesis of the theorem is sufficiently small so that the hypothesis of Theorem 9 is valid. Likewise, we get that

$$|\mathbf{E}_{(X,Y,Z) \in D_\delta^n}[(1 - T_{1-\eta}f(X)) \cdot (1 - T_{1-\eta}f(Y)) \cdot (1 - T_{1-\eta}f(Z))]| \leq \Pr[\mathcal{X}, \mathcal{Y}, \mathcal{Z} \leq 0] + \epsilon/4.$$

Combining the above with (2) and (1), we get that

$$\mathbf{E}_{(X,Y,Z) \in D_\delta^n}[f(X) \cdot f(Y) \cdot f(Z) + (1 - f(X)) \cdot (1 - f(Y)) \cdot (1 - f(Z))] \leq 2\Pr[\mathcal{X}, \mathcal{Y}, \mathcal{Z} \leq 0] + \epsilon.$$

Using Fact 11, we conclude that

$$\mathbf{E}_{(X,Y,Z) \in D_\delta^n}[f(X) \cdot f(Y) \cdot f(Z) + (1 - f(X)) \cdot (1 - f(Y)) \cdot (1 - f(Z))] \leq 1 - \frac{3 \cos^{-1}(1 - \delta)}{2\pi} + \epsilon.$$

completing the proof. \square

4 Dictatorship test for MAX-k-AND

In this section, we construct a dictatorship test for MAX-k-AND i. e. the tester checks if a particular set of k literals are all set to 1. For the purposes of this section, let us assume $\rho(k) = \frac{1}{2^{\lceil (k+1)/2 \rceil}}$.

Theorem 14. *For any $k \geq 3$ and $\delta > 0$, there is a distribution D over $(\{-1, 1\}^n)^k$ such that if $(X_1, \dots, X_k) \sim D$ such that for every $f : \{-1, 1\}^n \rightarrow [0, 1]$ with $\mathbf{E}[f] = 1/2$,*

- *If $f(x) = (1 + x_i)/2$ for some $i \in [n]$, then*

$$\Pr_{(X_1, \dots, X_k) \sim D}[f(X_1) \cdot \dots \cdot f(X_k)] \geq \rho(k) - \delta.$$

- *$\exists \tau = \tau(\delta, k) > 0$ and $\eta = \eta(\delta, k) > 0$ such that if $\max_i \text{Inf}_i(T_{1-\eta}f) \leq \tau$,*

$$\Pr_{(X_1, \dots, X_k) \sim D}[f(X_1) \cdot \dots \cdot f(X_k)] \leq \frac{1}{2^k} + \delta.$$

We remark that if f were to take values in $\{0, 1\}$, then we note that $f(X_1) \cdot \dots \cdot f(X_k) = 1$ if and only if $f(X_1) \wedge \dots \wedge f(X_k) = 1$.

Proof. Let D_k be the distribution from Fact 12. We let $\xi = \delta/4$. Now, we let $D_\xi = (1 - \xi)D_k + \xi U_k$. Let D_1, \dots, D_n be n i.i.d. samples from D_ξ . Let $D_i(j)$ be the j^{th} bit of D_i . Having done this, we define X_j for $1 \leq j \leq k$ as $X_j = (D_1(j), \dots, D_n(j))$. Let D be defined as the joint distribution of (X_1, \dots, X_k) .

As before, we start with the proof of the first item.

Completeness: Since $f(x) = (1 + x_i)/2$ (for some $i \in [n]$), it means that $f(x) = 1$ if $x_i = 1$ and 0 otherwise. Hence, we have

$$\begin{aligned} \mathbf{E}_{(X_1, \dots, X_k) \in D}[f(X_1) \cdot \dots \cdot f(X_k)] &= \mathbf{E}_{(X_1, \dots, X_k) \in D}[\mathbf{I}(X_1(i) = \dots = X_k(i) = 1)] \\ &= \rho(k)(1 - \xi) + \xi 2^{-k} \geq \rho(k) - \delta. \end{aligned}$$

Soundness: Let \mathcal{Q} be the multilinear polynomial representation of f . Note that for any $x \in \{-1, 1\}^n$, $|\mathcal{Q}(x)| \leq 1$. Let Ω be the probability space with domain $\{-1, 1\}^k$ and probability measure D_ξ on it. Observe that $D_\xi(x) \geq \xi \cdot 2^{-k}$ for all $x \in \{-1, 1\}^k$. Hence, by Lemma 8, we get that $\exists \eta = \eta(\xi, k) > 0$ (note because $\xi = \delta/4$, we can also express η as a function of δ and k as required by the theorem),

$$|\mathbf{E}_{(X_1, \dots, X_k) \in D}[f(X_1) \cdot \dots \cdot f(X_k) - T_{1-\eta}f(X_1) \cdot \dots \cdot T_{1-\eta}f(X_k)]| \leq \frac{\xi}{4}. \quad (3)$$

We can now apply Corollary 10 to the function $T_{1-\eta}f$ and the random variables $(X_1, \dots, X_k) \sim D$. Much like in the proof of Theorem 13, it is easy to check that all the conditions are satisfied (In particular, note that for any $i \in [n]$, $X_1(i), X_2(i), \dots, X_k(i)$ are pairwise independent). By Corollary 10, $\exists \tau = \tau(\xi, k)$ such that if $\max_i \text{Inf}_i(f) \leq \tau$, we have

$$|\mathbf{E}_{(X_1, \dots, X_k) \in D}[T_{1-\eta}f(X_1) \cdot \dots \cdot T_{1-\eta}f(X_k)]| \leq 2^{-k} + \frac{\xi}{4}. \quad (4)$$

As before, we note that $\tau(\xi, k)$ can be expressed as $\tau(\delta, k)$. Here, we are assuming that the $\eta(\xi, k)$ and $\tau(\xi, k)$ chosen to be sufficiently small so that the hypothesis of Corollary 10 is valid. Combining (3) and (4), we get that

$$|\mathbf{E}_{(X_1, \dots, X_k) \in D}[f(X_1) \cdot \dots \cdot f(X_k)]| \leq 2^{-k} + \frac{\xi}{2}.$$

□

5 Unique games hardness from Dictatorship test

In this section, we use the dictatorship tests constructed in Section 3 and Section 4 to show the following theorems.

Theorem 15. *Assuming the Unique Games Conjecture, for every $0 < \delta < 1$ and $\epsilon > 0$, it is NP-hard to distinguish an instance of MAX-3-EQUAL with value $1 - 3\delta/4 - \epsilon$ from an instance of value $1 - \frac{3 \cos^{-1}(1-\delta)}{2\pi} + \epsilon$. In other words, for every $\epsilon > 0$, MAX-3-EQUAL is $\alpha_{EQU} + \epsilon$ hard to approximate where*

$$\alpha_{EQU} = \inf_{\delta \in (0,1)} \frac{1 - \frac{3 \cos^{-1}(1-\delta)}{2\pi}}{1 - \frac{3\delta}{4}} \approx 0.796.$$

Theorem 16. *Assuming the Unique Games Conjecture, for every $\epsilon > 0$, it is NP-hard to distinguish an instance of MAX-k-AND with value $\frac{1}{2^{\lceil (k+1)/2 \rceil}} - \epsilon$ from an instance of value $2^{-k} + \epsilon$. In other words, for every $\epsilon > 0$, MAX-k-AND is $\frac{\lceil (k+1)/2 \rceil}{2^{k-1}} + \epsilon$ hard to approximate.*

Theorem 15 uses the dictatorship test in Theorem 13 to reduce Unique Games to MAX-3-EQUAL. Similarly, Theorem 16 uses the dictatorship test in Theorem 14 to reduce Unique Games to MAX-k-AND. As we said in the introduction, these reductions are by now very standard and can be found in several places. For the sake of convenience of the reader, we include the full proof of Theorem 15. The proof of Theorem 16 is exactly analogous and hence, we do not do it here.

We begin by defining the Unique Label Cover problem and then state Khot's Unique Games Conjecture (slightly differently stated than Conjecture 1).

Definition 17. *An instance of a Unique Label Cover problem (G, Σ) on alphabet size t is defined by a graph $G = (V, E)$ and a set of permutations $\Sigma = \{\sigma_{(u,v)} : [t] \rightarrow [t]\}_{(u,v) \in E}$. For any map $\mathcal{L} : V \rightarrow [t]$ and $(u, v) \in E$, $\mathcal{A}_{\mathcal{L}}(u, v) = 1$ if and only if $\mathcal{L}(v) = \sigma_{(u,v)}(\mathcal{L}(u))$, otherwise it is zero. For a map $\mathcal{L} : V \rightarrow [t]$, $\text{val}_{\mathcal{L}}(G) = \mathbf{E}_{(u,v) \sim E}[\mathcal{A}_{\mathcal{L}}(u, v)]$. The value of the unique label cover instance is (denoted by) $\text{val}(G) = \max_{\mathcal{L} : V \rightarrow [t]} \text{val}_{\mathcal{L}}(G)$.*

Conjecture 18. *[Kho02] Unique Games Conjecture: For every $\epsilon > 0$, there is a $t = t(\epsilon)$ such given a unique label cover problem (G, Σ) on alphabet size t , distinguishing whether $\text{val}(G) \leq \epsilon$ or $\text{val}(G) \geq 1 - \epsilon$ is NP-hard. We can also assume that the graph G is regular.*

Having stated the unique games conjecture, we describe a PCP verifier for the unique label cover problem which checks for equality of 3 bits. By the standard reduction between PCP verifiers and hardness of approximation, we get a hardness result for the MAX-3-EQUAL problem.

Description of the PCP verifier: Given the unique games instance (G, Σ) (on alphabet size t), we assume that $V = [n]$ and build a PCP verifier over $n \cdot 2^t$ boolean variables as follows: For every $i \in [n]$, we have a function $f_i : \{-1, 1\}^t \rightarrow \{0, 1\}$. Note that any such truth table can be described by 2^t boolean variables and hence the family of functions $\{f_i\}$ can be described in all by $n \cdot 2^t$ variables.

Remark 19. *We will also assume the functions are folded, i.e., for any x , $f(x) \neq f(-x)$. Note that this can be done without loss of generality, because whenever the verifier needs to query $f(x)$, if $x_1 = 1$, it queries $f(x)$. Else it queries $f(-x)$ and flips the output. We note that “flipping” the output can be implemented by introducing negated literals in the resulting CSP. Also, we observe that dictators satisfy this requirement.*

For a given $\delta \in (0, 1)$, let D_{δ}^t be the distribution in the hypothesis of Theorem 13. Note that the distribution D_{δ}^t is over $(\{-1, 1\}^t)^3$. Also, we use \circ to denote composition of functions. In other words, for two functions g_1 and g_2 , $g_1 \circ g_2(x)$ denotes $g_1(g_2(x))$. With this, the verifier is as follows:

- Pick $v \in V$ uniformly at random and choose three random neighbors of v , say, w_1, w_2, w_3 uniformly at random.
- Choose $(X, Y, Z) \sim D_{\delta}^t$ (described above) and accept if and only if

$$f_{w_1} \circ \sigma_{(w_1, v)}(X) = f_{w_2} \circ \sigma_{(w_2, v)}(Y) = f_{w_3} \circ \sigma_{(w_3, v)}(Z).$$

We next show the correctness of this verifier. In other words, we prove the following two lemmas.

Lemma 20. *If $\text{val}(G) \geq 1 - \epsilon$, then there is a set of functions $\{f_i : \{-1, 1\}^t \rightarrow \{0, 1\}\}_{i \in [n]}$ such that the above verifier accepts with probability at least $(1 - 3\epsilon)(1 - 3\delta/4)$.*

Lemma 21. *For any $\epsilon > 0$, if the above verifier passes with probability more than $1 - \frac{3 \cos^{-1}(1-\delta)}{2\pi} + \epsilon$, then $\exists \mathcal{L} : V \rightarrow [t]$ such that $\text{val}_{\mathcal{L}}(G) = \kappa(\epsilon, \delta) > 0$.*

Since $\text{val}_{\mathcal{L}}(G)$ in conclusion of $\kappa(\epsilon, \delta)$ does not depend on t , hence by combining Lemma 20 and Lemma 21 and the standard reduction between PCPs and hardness of CSPs, we prove Theorem 15. The proofs of Lemma 20 and Lemma 21 follow.

of Lemma 20. Since $\text{val}(G) \geq (1-\epsilon)$, $\exists \mathcal{L} : V \rightarrow [t]$ such that $\text{val}_{\mathcal{L}}(G) \geq 1-\epsilon$. Let \mathcal{L} be such a labeling of the vertices. We let $f_i : \{-1, 1\}^t \rightarrow \{0, 1\}$ be the dictator function corresponding to $\mathcal{L}(i)$. In other words, $f_i(x) = (1 + x_{\mathcal{L}(i)})/2$. Now, since $\text{val}_{\mathcal{L}}(G) \geq 1-\epsilon$ and the constraint graph G is regular, if we choose v uniformly at random and then a uniform random neighbor w_i , then $\mathcal{A}_{\mathcal{L}}(v, w_i) = 1$ with probability $1-\epsilon$. By a union bound, with probability at least $1-3\epsilon$, $\mathcal{A}_{\mathcal{L}}(v, w_1) = \mathcal{A}_{\mathcal{L}}(v, w_2) = \mathcal{A}_{\mathcal{L}}(v, w_3) = 1$. If this is indeed the case, then,

$$f_{w_1} \circ \sigma_{(w_1, v)} = f_{w_2} \circ \sigma_{(w_2, v)} = f_{w_3} \circ \sigma_{(w_3, v)}.$$

Now, applying the first part of Theorem 13, we get that in this case the test accepts with probability $1-3\delta/4$. Thus, the total probability that the test accepts is at least $(1-3\epsilon)(1-3\delta/4)$. \square

We next move to the more difficult case of soundness.

of Lemma 21. The proof follows the arguments in [KKMO07] very closely. We first describe the labeling \mathcal{L} and then describe its correctness. Our labeling is a randomized scheme. Let $\eta, \tau > 0$ be two parameters which are chosen according to the second part of the hypothesis of Theorem 13 for parameters $\epsilon/2$ and δ . First, for every $v \in V$, we define $g_v : \{-1, 1\}^t \rightarrow [0, 1]$ as

$$g_v(x) = \mathbf{E}_{(w, v) \in E} [f_w \circ \sigma_{(w, v)}(x)].$$

Again for every $v \in V$ we define $\mathcal{A}(v) \subseteq V$ as

$$\mathcal{A}(v) = \{i : \text{Inf}_i(T_{1-\eta} f_v) \geq \tau/2\} \cup \{i : \text{Inf}_i(T_{1-\eta} g_v) \geq \tau\}.$$

The randomized labeling scheme is the following: If the set $\mathcal{A}(v)$ is empty, $\mathcal{L}(v)$ is chosen arbitrarily. Else, it is chosen to be a uniformly random element from the set $\mathcal{A}(v)$. The following proposition gives us the desired result.

Proposition 22. *Over the choice of randomness for choosing \mathcal{L} , $\mathbf{E}[\text{val}_{\mathcal{L}}(G)] \geq (\epsilon\eta^2\tau^3)/64$.*

By fixing the randomness in the above proposition desirably, we get Lemma 21. So, the proof boils down to proving Proposition 22.

Proof. Let \mathcal{D} be a probability distribution over V^4 where $(v, w_1, w_2, w_3) \in \mathcal{D}$ is sampled as follows: $v \in V$ is chosen uniformly at random and w_1, w_2, w_3 are chosen to be three random neighbors of v . Further, let $(X, Y, Z) \in \mathcal{D}_{\delta}^t$. Then, the probability of acceptance of the verifier is given by

$$\begin{aligned} & \mathbf{E}[\mathbf{I}(f_{w_1} \circ \sigma_{(w_1, v)}(X) = f_{w_2} \circ \sigma_{(w_2, v)}(Y) = f_{w_3} \circ \sigma_{(w_3, v)}(Z))] \\ &= \mathbf{E}[f_{w_1} \circ \sigma_{(w_1, v)}(X) \cdot f_{w_2} \circ \sigma_{(w_2, v)}(Y) \cdot f_{w_3} \circ \sigma_{(w_3, v)}(Z)] \\ &+ \mathbf{E}[(1 - f_{w_1} \circ \sigma_{(w_1, v)}(X)) \cdot (1 - f_{w_2} \circ \sigma_{(w_2, v)}(Y)) \cdot (1 - f_{w_3} \circ \sigma_{(w_3, v)}(Z))] \\ &= \mathbf{E}[g_v(X) \cdot g_v(Y) \cdot g_v(Z) + (1 - g_v(X)) \cdot (1 - g_v(Y)) \cdot (1 - g_v(Z))]. \end{aligned}$$

Since the verifier accepts with probability at least $1 - \frac{3 \cos^{-1}(1-\delta)}{2\pi} + \epsilon$, a Markov argument gives that for at least an $\epsilon/2$ fraction of vertices $v \in V$,

$$\mathbf{E}[g_v(X) \cdot g_v(Y) \cdot g_v(Z) + (1 - g_v(X)) \cdot (1 - g_v(Y)) \cdot (1 - g_v(Z))] \geq 1 - \frac{3 \cos^{-1}(1-\delta)}{2\pi} + \epsilon/2.$$

Denote this subset (of V) by A . Note that by the second part of Theorem 13, for every $v \in A$, $\exists i \in [t]$, such that $\text{Inf}_i(T_{1-\eta} g_v) \geq \tau$. For every $v \in A$, fix an i which satisfies $\text{Inf}_i(T_{1-\eta} g_v) \geq \tau$.

$$\begin{aligned} \tau \leq \text{Inf}_i(T_{1-\eta} g_v) &= \sum_{S: i \in S} (1-\eta)^{|S|} \widehat{g}_v(S)^2 = \sum_{S: i \in S} (1-\eta)^{|S|} \left(\mathbf{E}_{(w, v) \in E} [\widehat{f_w \circ \sigma_{(w, v)}}(S)] \right)^2 \\ &= \sum_{S: i \in S} (1-\eta)^{|S|} \left(\mathbf{E}_{(w, v) \in E} [\widehat{f_w}(\sigma_{(w, v)}^{-1}(S))] \right)^2. \end{aligned}$$

Here $\sigma_{(w,v)}^{-1}(S)$ is the pre-image of the set S under the map $\sigma_{(w,v)}$. Now, by Jensen's inequality we get that

$$\begin{aligned} \sum_{S:i \in S} (1-\eta)^{|S|} \left(\mathbf{E}_{(w,v) \in E} [\widehat{f_w}(\sigma_{(w,v)}^{-1}(S))] \right)^2 &\leq \mathbf{E}_{(w,v) \in E} \left[\sum_{S:i \in S} (1-\eta)^{|S|} \widehat{f_w}^2(\sigma_{(w,v)}^{-1}(S)) \right] \\ &= \mathbf{E}_{(w,v) \in E} \left[\text{Inf}_{\sigma_{(w,v)}^{-1}(i)}(T_{1-\eta} f_w) \right]. \end{aligned}$$

Using a Markov argument, this implies that for such a $v \in A$ and i such that $\text{Inf}_i(T_{1-\eta} g_v) \geq \tau$, at least a $\tau/2$ fraction of neighbors w of v satisfy,

$$\text{Inf}_{\sigma_{(w,v)}^{-1}(i)}(T_{1-\eta} f_w) \geq \tau/2.$$

We say that such a pair (v, w) of vertices is “good”. Using Lemma 7, it can be easily shown that for every $v \in V$, $|\mathcal{A}(v)| \leq 4/(\tau\eta)$. This means that for every $v \in A$, the randomized scheme \mathcal{L} assigns $\mathcal{L}(v) = i$ such that $\text{Inf}_i(T_{1-\eta} g_v) \geq \tau$ with probability at least $(\eta\tau)/4$. Observe that for any such $v \in A$, at least $\tau/2$ fraction of its neighbors w are such that (v, w) is “good”. Note that for any good pair (v, w) , if $\text{Inf}_i(T_{1-\eta} g_v) \geq \tau$, then $\text{Inf}_{\sigma_{(w,v)}^{-1}(i)}(T_{1-\eta} f_w) \geq \tau/2$. This implies that $\sigma_{(w,v)}^{-1}(i) \in \mathcal{A}(w)$. Thus, with probability at least $(\tau\eta)/4$, $\mathcal{L}(w) = \sigma_{(w,v)}^{-1}(i)$. Thus, overall the probability that $\mathcal{L}(v, w) = 1$ is at least $\epsilon\tau^3\eta^2/64$. This completes the proof of Proposition 22. \square

\square

6 Approximation algorithm for the MAX-3-EQUAL problem

In this section, we give a SDP based approximation algorithm for MAX-3-EQUAL whose performance matches the hardness result from the last section. In particular, we prove the following theorem.

Theorem 23. *There is a polynomial time approximation algorithm for the MAX-3-EQUAL problem which achieves the following approximation ration :*

$$\inf_{\delta \in (0,1)} \frac{1 - \frac{3 \cos^{-1}(1-\delta)}{2\pi}}{1 - \frac{3\delta}{4}} \approx 0.796.$$

Thus, this theorem shows that we have an approximation algorithm whose performance ratio matches the Unique Games hardness for this problem. Towards proving Theorem 23, we state a SDP relaxation for the MAX-3-EQUAL problem followed by a rounding procedure and then analyze the performance of this algorithm. The SDP formulation is essentially the generic SDP by Raghavendra [Rag08] specialized to the MAX-3-EQUAL problem. We assume that the variables are $x_1, \dots, x_n \in \{-1, 1\}$. The constraint set is $E \subseteq [n]^3 \times \{-1, 1\}^3$ such that for every $(i, j, k) \times (\eta_i, \eta_j, \eta_k) \in E$ we have a constraint that $\eta_i x_i = \eta_j x_j = \eta_k x_k$. In other words, η_i represents the polarity with which the variable x_i appears in the constraint E (likewise for η_j and η_k). The SDP relaxation is given in Figure 1.

Remark 24. *We note that Zwick [Zwi98] describes a SDP relaxation and a similar rounding procedure for the MAX-3-EQUAL problem. The paper also gives numerical evidence towards showing that the performance ratio of their algorithm is approximately 0.796. However, the paper notes that they do not have an analytical proof of this and to the best of our knowledge, no analytical proof has appeared ever since. We analyze a slightly different SDP and analytically show that the performance of it is indeed what we claim. There are a couple of differences between our SDP formulation and Zwick's SDP. The first one is that we*

SDP formulation

1. $\forall i \in [n], v_i \in \mathbb{R}^n$ and $\|v_i\|_2 = 1$.
2. $\forall i, j, k \in [n]^3, i < j < k, \alpha_{(i,j,k)}, \beta_{(i,j,k)}, \gamma_{(i,j,k)}, \delta_{(i,j,k)} \in \mathbb{R}^+ \cup \{0\}$ such that

$$\alpha_{(i,j,k)} + \beta_{(i,j,k)} + \gamma_{(i,j,k)} + \delta_{(i,j,k)} = 1.$$

3. $\forall i, j, k \in [n]^3, i < j < k,$

$$\alpha_{(i,j,k)} + \beta_{(i,j,k)} - \gamma_{(i,j,k)} - \delta_{(i,j,k)} = \langle v_i, v_j \rangle.$$

$$\alpha_{(i,j,k)} - \beta_{(i,j,k)} + \gamma_{(i,j,k)} - \delta_{(i,j,k)} = \langle v_j, v_k \rangle.$$

$$\alpha_{(i,j,k)} - \beta_{(i,j,k)} - \gamma_{(i,j,k)} + \delta_{(i,j,k)} = \langle v_i, v_k \rangle.$$

4. For $e \in E$, where $e = (i, j, k) \times (\eta_i, \eta_j, \eta_k)$, define

$$\lambda(e) = \begin{cases} \alpha_{(i,j,k)} & \text{if } \eta_i = \eta_j = \eta_k, \\ \beta_{(i,j,k)} & \text{if } \eta_i = \eta_j = -\eta_k, \\ \gamma_{(i,j,k)} & \text{if } -\eta_i = \eta_j = \eta_k, \\ \delta_{(i,j,k)} & \text{if } \eta_i = -\eta_j = \eta_k, \end{cases}$$

5. Subject to the above, Maximize $\mathbf{E}_{(i,j,k) \times (\alpha, \beta, \gamma) \in E} [\lambda(e)]$.

Figure 1: SDP relaxation for MAX-3-EQUAL problem

use some additional real variables. However, this difference is purely cosmetic as the presence of those variables does not make our relaxation any tighter than Zwick's SDP. The second difference between our SDP and Zwick's SDP is that our SDP implies an additional set of constraints, namely, for all $i, j, k \in [n]^3$, $1 + \langle v_i, v_j \rangle + \langle v_i, v_k \rangle + \langle v_j, v_k \rangle \geq 0$. We should mention that this family of constraints appears in [Zwi98] for SDP relaxations of some other CSPs but it is unclear from the paper if Zwick uses these constraints in the SDP relaxation for MAX-3-EQUAL as well. Potentially, these additional constraints make our SDP tighter than that of Zwick though the reason we use these additional constraints is that our analysis becomes simpler. As remarked earlier, after the publication of the preprint, Williamson [Wil13] pointed out to us that an analysis similar to ours had already appeared in [GW95] in the context of MAX-DICUT. Using this analysis as a black-box, we can shorten the analysis of the rounding algorithm substantially. We however keep our original analysis here so that the paper is self contained.

To see why the SDP in Figure 1 is a relaxation, consider a particular assignment to the variables x_1, \dots, x_n . Let us define $\mathbf{v}_0 \in \mathbb{R}^n$ as having 1 in the first coordinate and 0 everywhere else. If $x_i = 1$, set $v_i = \mathbf{v}_0$. Else, if $x_i = -1$, set $v_i = -\mathbf{v}_0$. The rest of the variables are set as follows. For every triple $(i, j, k), i < j < k$,

- If $x_i = x_j = x_k$, then $\alpha_{(i,j,k)} = 1, \beta_{(i,j,k)} = \gamma_{(i,j,k)} = \delta_{(i,j,k)} = 0$.
- If $x_i = x_j = -x_k$, then $\beta_{(i,j,k)} = 1, \alpha_{(i,j,k)} = \gamma_{(i,j,k)} = \delta_{(i,j,k)} = 0$.
- If $-x_i = x_j = x_k$, then $\gamma_{(i,j,k)} = 1, \alpha_{(i,j,k)} = \beta_{(i,j,k)} = \delta_{(i,j,k)} = 0$.

- If $x_i = -x_j = x_k$, then $\delta_{(i,j,k)} = 1$, $\alpha_{(i,j,k)} = \beta_{(i,j,k)} = \gamma_{(i,j,k)} = 0$.

It is easy to verify that with these assignments of $\alpha_{(i,j,k)}$, $\beta_{(i,j,k)}$, $\gamma_{(i,j,k)}$, $\delta_{(i,j,k)}$ and v_i , constraints 1, 2 and 3 are indeed satisfied. Further, for this assignment, if a constraint $e \in E$ is satisfied, then it is easy to see that $\lambda(e) = 1$. Also, if a constraint e is not satisfied, then $\lambda(e) = 0$. Thus, the objective value of the program for this assignment is exactly the fraction of constraints $e \in E$ which are satisfied and hence its a relaxation.

6.1 Rounding algorithm Our rounding algorithm is as follows: Let $\Sigma \in \mathbb{R}^{n \times n}$ be the matrix such that $\Sigma_{i,j} = \langle v_i, v_j \rangle$. Note that Σ is positive semidefinite. So, we let $\mathcal{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$, i.e., \mathcal{X} be a jointly normal distribution in \mathbb{R}^n with mean at the origin and the covariance matrix Σ . The rounding algorithm gets a sample \mathcal{X} and assigns $x_i = 1$ if $\mathcal{X}_i \geq 0$ and -1 otherwise. Here \mathcal{X}_i denotes the i^{th} coordinate of \mathcal{X} . We will call this rounding as the “random gaussian” rounding. We now prove Theorem 23 by analyzing the performance of this rounding algorithm.

We would also like to remark that (perhaps not too surprisingly), if instead of the “random gaussian” rounding, we would have used “random hyperplane” rounding, the performance of the algorithm would have been the same and our analysis would have also gone through without any changes.

Proof of Theorem 23. We start by considering a particular constraint $e \in E$. Without loss of generality, assume that $e = (i, j, k) \times (\eta_i, \eta_j, \eta_k)$, where $\eta_i = \eta_j = \eta_k = 1$. We note that if the triple (η_i, η_j, η_k) were to take some other value in $\{-1, 1\}^3$, our analysis would remain unchanged.

Now, for the particular edge e , its contribution to the SDP objective is $\lambda(e) = \alpha_{(i,j,k)}$. On the other hand, let the expected contribution to the true objective from this edge be $\kappa(e)$. Note that

$$\kappa(e) = \Pr[(\mathcal{X}_i, \mathcal{X}_j, \mathcal{X}_k \geq 0) \cup (\mathcal{X}_i, \mathcal{X}_j, \mathcal{X}_k < 0)]. \quad (5)$$

It is obvious that the performance ratio of the algorithm is lower bounded by $\inf \kappa(e)/\lambda(e)$. Hence, we will simply aim to prove a lower bound on $\inf \kappa(e)/\lambda(e)$. Observe that for any (i, j, k) , $\alpha_{(i,j,k)} + \beta_{(i,j,k)} + \gamma_{(i,j,k)} + \delta_{(i,j,k)} = 1$. Now, using this and plugging Fact 11 into (5), we get (below, we use α as a shorthand for $\alpha_{(i,j,k)}$ and likewise for β, γ and δ),

$$\begin{aligned} \kappa(e) &= 1 - \frac{\cos^{-1}(\langle v_i, v_j \rangle) + \cos^{-1}(\langle v_j, v_k \rangle) + \cos^{-1}(\langle v_i, v_k \rangle)}{2\pi} \\ &= 1 - \frac{\cos^{-1}(2(\alpha + \beta) - 1) + \cos^{-1}(2(\alpha + \gamma) - 1) + \cos^{-1}(2(\alpha + \delta) - 1)}{2\pi}. \end{aligned}$$

Thus, for $a, b, c, d \in \mathbb{R}^+ \cup \{0\}$, if we define

$$g(a, b, c, d) = \frac{1 - \frac{\cos^{-1}(2(a+b)-1) + \cos^{-1}(2(a+c)-1) + \cos^{-1}(2(a+d)-1)}{2\pi}}{a},$$

$$\text{then, } \frac{\kappa(e)}{\lambda(e)} \geq \inf_{a,b,c,d} g(a, b, c, d) \text{ subjected to } a + b + c + d = 1 \text{ and } a, b, c, d \geq 0.$$

For the purposes of the analysis, it is helpful to fix the value of a , and then find the optimum choice of b, c, d for that value of a to minimize $g(a, b, c, d)$. Subsequently, one optimizes over the choice of a . In other words, let us define $h_a(b, c, d)$ as

$$h_a(b, c, d) = \cos^{-1}(2(a+b)-1) + \cos^{-1}(2(a+c)-1) + \cos^{-1}(2(a+d)-1).$$

$$\Psi(a) = \sup_{b,c,d} h_a(b, c, d) \text{ subjected to } b + c + d = 1 - a \text{ and } b, c, d \geq 0 \text{ where } a > 0. \quad (6)$$

Hence, we now get that

$$\frac{\kappa(e)}{\lambda(e)} \geq \inf_{0 < a \leq 1} \frac{1 - \frac{\Psi(a)}{2\pi}}{a}. \quad (7)$$

Thus, we now focus on finding $\Psi(a)$ for every $a \in (0, 1]$. In order to find out $\Psi(a)$, we find out the local minima by evaluating the partial derivatives of the function $h_a(b, c, d)$ and also investigate the value of $h_a(b, c, d)$ at the boundaries of the domain. Williamson [Wil13] noted to us that the expression $\inf_{0 < a \leq 1} \frac{1 - (\Psi(a)/2\pi)}{a}$ had already been analyzed in [GW95] (see Lemma 7.3.2). However, we keep our original analysis here.

6.2 Supremum of $h_a(b, c, d)$ at the boundary of the domain: The next claim gets the supremum of $h_a(b, c, d)$ when b, c, d lie on the boundary of the domain defined in Equation 6.

Claim 25. *The supremum of $h_a(b, c, d)$ when b, c and d lie on the boundary of the domain defined in (6) is $\cos^{-1}(2a - 1) + 2\cos^{-1}(a)$.*

Proof. Note that because $b, c, d \geq 0$ and $b + c + d = 1 - a$, the boundary of the domain is defined by at least one of b, c and d being 0. Without loss of generality, we assume $b = 0$. Note that because a is fixed, we are viewing the domain as a two dimensional object. In that case,

$$h_a(0, c, d) = \cos^{-1}(2a - 1) + \cos^{-1}(2(a + c) - 1) + \cos^{-1}(2(a + d) - 1),$$

with $c + d = 1 - a$ and $c, d \geq 0$. Performing the substitution $d = 1 - a - c$, we get

$$h_a(0, c, d) = \cos^{-1}(2a - 1) + \cos^{-1}(2(a + c) - 1) + \cos^{-1}(1 - 2c), \quad (8)$$

where $0 \leq c \leq 1 - a$. Now, note that since a is fixed, $h_a(0, c, d)$ is solely a function of c . Hence, to find out the supremum of $h_a(0, c, d)$, we evaluate it at the end points of the domain, i.e., at $c = 0$, $c = 1 - a$ and at its critical points.

- If $c = 0$, then $d = 1 - a$. Hence, at this point, $h_a(b, c, d) = h_a(0, 0, 1 - a) = \cos^{-1}(2a - 1) + \cos^{-1}(2a - 1) + \cos^{-1}(1) = 2\cos^{-1}(2a - 1)$.
- If $c = 1 - a$, then $d = 0$. Hence, at this point, $h_a(b, c, d) = h_a(0, 1 - a, 0) = \cos^{-1}(2a - 1) + \cos^{-1}(1) + \cos^{-1}(2a - 1) = 2\cos^{-1}(2a - 1)$.

Having evaluated $h_a(0, c, d)$ at the boundary points, we now find out the critical points of this function. Differentiating the expression in (8), we get

$$\frac{\partial h_a(0, c, d)}{\partial c} = \frac{-2}{\sqrt{1 - (2(a + c) - 1)^2}} + \frac{2}{\sqrt{1 - (1 - 2c)^2}} = 0.$$

This implies that

$$\begin{aligned} 1 - (2(a + c) - 1)^2 &= 1 - (1 - 2c)^2 \\ \Rightarrow (2(a + c) - 1) &= \pm(1 - 2c). \end{aligned}$$

This means that either $a = 0$ or $a + 2c = 1$. Since $a > 0$, we can neglect the first condition. Thus, the only condition we need to consider is $a + 2c = 1$. Because $a + c + d = 1$, this means that $c = d = (1 - a)/2$. Thus, $h_a(0, c, d) = \cos^{-1}(2a - 1) + 2\cos^{-1}(a)$. Thus, we get that

$$\sup_{c, d} h_a(0, c, d) = \sup\{\cos^{-1}(2a - 1) + 2\cos^{-1}(a), 2\cos^{-1}(2a - 1)\} = \cos^{-1}(2a - 1) + 2\cos^{-1}(a). \quad (9)$$

The last equality uses Fact 27. □

6.3 Evaluation of $h_a(b, c, d)$ at the critical points: The next claim evaluates the supremum of $h_a(b, c, d)$ at the critical points of the domain.

Claim 26. *The supremum of $h_a(b, c, d)$ at the critical points inside the domain defined in (6) is given by*

$$\sup h_a(b, c, d) = \begin{cases} \pi + \cos^{-1}(4a - 1) & \text{if } 0 \leq a \leq 1/4, \\ 3 \cos^{-1}((4a - 1)/3) & \text{if } 1/4 < a \leq 1. \end{cases}$$

Proof. Note that $b + c + d = 1 - a$. Thus, we get

$$h_a(b, c, d) = \cos^{-1}(1 - 2c - 2d) + \cos^{-1}(2(a + c) - 1) + \cos^{-1}(2(a + d) - 1).$$

As a is fixed, $h_a(b, c, d)$ is a function of c and d alone. At the critical point,

$$\begin{aligned} \frac{\partial h_a(b, c, d)}{\partial c} &= \frac{2}{\sqrt{1 - (1 - 2c - 2d)^2}} - \frac{2}{\sqrt{1 - (1 - 2a - 2c)^2}} = 0, \\ \frac{\partial h_a(b, c, d)}{\partial d} &= \frac{2}{\sqrt{1 - (1 - 2c - 2d)^2}} - \frac{2}{\sqrt{1 - (1 - 2a - 2d)^2}} = 0. \end{aligned}$$

Thus, at the critical point,

$$\begin{aligned} (1 - 2c - 2d)^2 &= (1 - 2a - 2c)^2 = (1 - 2a - 2d)^2 \\ \Rightarrow \pm(1 - 2c - 2d) &= \pm(1 - 2a - 2c) = \pm(1 - 2a - 2d). \end{aligned}$$

We now solve for c, d for the various possibilities listed above.

- $1 - 2c - 2d = 1 - 2a - 2c = 1 - 2a - 2d$. In this case, we get $a = c = d$ and hence $b = 1 - 3a$. Since $b \geq 0$, this possibility occurs only when $0 \leq a \leq (1/3)$. If this indeed holds,

$$h_a(b, c, d) = \cos^{-1}(1 - 4a) + \cos^{-1}(4a - 1) + \cos^{-1}(4a - 1) = \pi + \cos^{-1}(4a - 1).$$

- $1 - 2c - 2d = -(1 - 2a - 2c) = 1 - 2a - 2d$. In this case, we get $a = c = b$ and $d = 1 - 3a$. Again as $d \geq 0$, this possibility occurs only when $0 \leq a \leq (1/3)$. As before,

$$h_a(b, c, d) = \cos^{-1}(1 - 4a) + \cos^{-1}(4a - 1) + \cos^{-1}(4a - 1) = \pi + \cos^{-1}(4a - 1).$$

- $1 - 2c - 2d = 1 - 2a - 2c = -(1 - 2a - 2d)$. This goes exactly the same way as in the previous case. Here again, we have

$$h_a(b, c, d) = \cos^{-1}(1 - 4a) + \cos^{-1}(4a - 1) + \cos^{-1}(4a - 1) = \pi + \cos^{-1}(4a - 1).$$

- $-(1 - 2c - 2d) = 1 - 2a - 2c = 1 - 2a - 2d$. In this case, $b = c = d = (1 - a)/3$. Now, we get

$$h_a(b, c, d) = \cos^{-1}((4a - 1)/3) + \cos^{-1}((4a - 1)/3) + \cos^{-1}((4a - 1)/3) = 3 \cos^{-1}((4a - 1)/3).$$

Hence at the critical points, we have

$$\sup h_a(b, c, d) = \begin{cases} \sup\{\pi + \cos^{-1}(4a - 1), 3 \cos^{-1}((4a - 1)/3)\} & \text{if } 0 < a \leq 1/3, \\ 3 \cos^{-1}((4a - 1)/3) & \text{if } a > 1/3. \end{cases}$$

However, using Fact 28, the above simplifies to saying that at the critical points,

$$\sup h_a(b, c, d) = \begin{cases} \pi + \cos^{-1}(4a - 1) & \text{if } 0 \leq a \leq 1/4, \\ 3 \cos^{-1}((4a - 1)/3) & \text{if } 1/4 < a \leq 1. \end{cases}$$

□

Define ζ to be the smallest of the following three quantities:

$$\left\{ \inf_{a \in (0, 1/4]} \frac{1 - \frac{\pi + \cos^{-1}(4a-1)}{2\pi}}{a}, \inf_{a \in (1/4, 1]} \frac{1 - \frac{3 \cos^{-1}((4a-1)/3)}{2\pi}}{a}, \inf_{a \in (0, 1]} \frac{1 - \frac{2 \cos^{-1}(a) + \cos^{-1}(2a-1)}{2\pi}}{a} \right\}.$$

Combining Claims 25 and 26 along with (7), we get that $\frac{\kappa(e)}{\lambda(e)} \geq \zeta$. By Fact 29, the first quantity inside the definition of ζ simplifies to 1 as follows:

$$\inf_{a \in (0, 1/4]} \frac{1 - \frac{\pi + \cos^{-1}(4a-1)}{2\pi}}{a} = \frac{1 - \frac{\pi + \cos^{-1}(4 \cdot (1/4) - 1)}{2\pi}}{(1/4)} = 1. \quad (10)$$

At this point, we are left with the task of finding the following quantities:

$$\inf_{a \in (0, 1]} \frac{1 - \frac{2 \cos^{-1}(a) + \cos^{-1}(2a-1)}{2\pi}}{a} \quad \inf_{a \in (1/4, 1]} \frac{1 - \frac{3 \cos^{-1}((4a-1)/3)}{2\pi}}{a}.$$

Thus, we are now left with the task of finding the infimum of two single-variable functions and then taking the minima of these two quantities. We do this computation by evaluating these two functions at sufficiently many points and then taking the infimum of these. For a mathematical justification, see Appendix A.2. Doing the numerical computation, we get,

$$\inf_{a \in (0, 1]} \frac{1 - \frac{2 \cos^{-1}(a) + \cos^{-1}(2a-1)}{2\pi}}{a} = [0.803125, 0.803325]. \quad (11)$$

$$\inf_{a \in (1/4, 1]} \frac{1 - \frac{3 \cos^{-1}((4a-1)/3)}{2\pi}}{a} = [0.795970, 0.796170]. \quad (12)$$

Further, the value of a achieving the infimum in (12) is $a = 0.700296 \pm 0.000001$. Hence, we have that

$$\frac{\kappa(e)}{\lambda(e)} \geq \inf_{a \in (1/4, 1]} \frac{1 - \frac{3 \cos^{-1}((4a-1)/3)}{2\pi}}{a} = \inf_{a \in (0, 1]} \frac{1 - \frac{3 \cos^{-1}((4a-1)/3)}{2\pi}}{a}.$$

The second equality (i.e., making the domain $(0, 1]$ instead of $(1/4, 1]$) follows because Fact 28 and (10) can be combined as:

$$\forall 0 < a \leq 1/4 \quad \frac{1 - \frac{3 \cos^{-1}((4a-1)/3)}{2\pi}}{a} \geq \frac{1 - \frac{\pi + \cos^{-1}(4a-1)}{2\pi}}{a} \geq 1.$$

Put $\delta = 4(1 - a)/3$. Then, we get that

$$\inf_{a \in (0, 1]} \frac{1 - \frac{3 \cos^{-1}((4a-1)/3)}{2\pi}}{a} = \inf_{0 \leq \delta < 4/3} \frac{1 - \frac{3 \cos^{-1}(1-\delta)}{2\pi}}{1 - \frac{3\delta}{4}} = \inf_{0 \leq \delta \leq 1} \frac{1 - \frac{3 \cos^{-1}(1-\delta)}{2\pi}}{1 - \frac{3\delta}{4}}.$$

Here the last equality is true because we have earlier observed that the infimum of the expression in (12) is obtained when $a \approx 0.700$. This means the corresponding value of $\delta \approx 0.400 < 1$. Thus, making the domain of δ to be $(0, 1]$ instead of $(0, 4/3]$ does not affect the value of the infimum. This also conclude the proof of the theorem. \square

7 Difficulty in getting optimal results for MAX-k-EQUAL

Given our results on MAX-3-EQUAL, a very obvious question is whether or not our results can be extended to MAX-k-EQUAL for $k > 3$. More concretely, since it is known that assuming the UGC, Raghavendra's SDP achieves the optimal approximation ratio for every CSP, it is natural to ask if the “random gaussian” rounding algorithm described in subsection 6.1 also achieves this ratio. We now explain the difficulty in proving such a result in a nutshell.

Consider the case of MAX-k-EQUAL. Let (g_1, \dots, g_k) be jointly normally distributed random variables such that each $g_i \sim \mathcal{N}^n(0, 1)$ (the value of n is immaterial as long as $n \geq k$). Assume that for all $1 \leq a \leq n$, the covariance matrix of $g_1(a), \dots, g_k(a)$ is given by $\rho \in \mathbb{R}^{k \times k}$ (and ρ is independent of a). Here $g_i(a)$ represents the i^{th} coordinate of g_i .

Let us define a family of distributions $\mathcal{D}(\rho)$ over $\{-1, 1\}^k$ in the following way: $\mathcal{A} \in \mathcal{D}(\rho)$ if and only if

- $\forall 1 \leq i \leq k$, $\mathcal{A}(i)$ is a uniformly random bit.
- $\forall 1 \leq i < j \leq k$, $\mathbf{E}[\mathcal{A}(i) \cdot \mathcal{A}(j)] = \rho_{ij}$.

We next define the following two quantities:

$$\begin{aligned} h_s(\rho) &= \mathbf{Pr}_{g \in \mathcal{N}^n(0,1)}[\forall i \in [k] \ g \cdot g_i \geq 0] + \mathbf{Pr}_{g \in \mathcal{N}^n(0,1)}[\forall i \in [k] \ g \cdot g_i < 0]. \\ h_c(\rho) &= \max_{\mathcal{A} \in \mathcal{D}(\rho)} [\mathcal{A}(1) = \dots = \mathcal{A}(k)]. \end{aligned}$$

It is easy to show that the approximation ratio achieved by the “random gaussian” rounding algorithm on Raghavendra's SDP is (lower)-bounded by $\inf_{\rho} h_s(\rho)/h_c(\rho)$.

If we want to show that the “random gaussian” rounding algorithm on Raghavendra's SDP indeed achieves the optimal approximation ratio (assuming the UGC), then the task essentially boils down to constructing a dictatorship test for MAX-k-EQUAL whose ratio of soundness to completeness is $\inf_{\rho} h_s(\rho)/h_c(\rho)$. To do this, let us assume that $\arg \inf_{\rho} h_s(\rho)/h_c(\rho) = \rho'$.

First of all, we construct a dictatorship test for MAX-k-EQUAL whose completeness is $h_c(\rho')$. To do this, let us assume that the distribution (in $\mathcal{D}(\rho')$) which achieves the maximum in the definition of $h_c(\rho')$ is \mathcal{A} . The dictatorship test is as follows: Given a function $f : \{-1, 1\}^n \rightarrow \{0, 1\}$, we sample $(X_1, \dots, X_k) \in \mathcal{A}^n$ and accept if and only if $f(X_1) = \dots = f(X_k)$.

It is easy to see that if f is a dictator, then the probability that $f(X_1) = \dots = f(X_k)$ is exactly $h_c(\rho')$. Thus, the completeness of the dictatorship test is exactly $h_c(\rho')$. The hard part is to bound the soundness of the dictatorship test. In other words, assuming that f is a balanced function where every coordinate has a low-influence, we need to bound the probability that $f(X_1) = \dots = f(X_k)$. An application of the invariance principle [Mos10] says that it suffices to bound the following quantity: Let $f' : \mathbb{R}^n \rightarrow \{0, 1\}$ be a function on the gaussian space such that $\mathbf{E}[f'(x)] = 1/2$. Let $g_1, \dots, g_k \sim \mathcal{N}^n(0, 1)$ be jointly normally distributed random variables where for all $1 \leq a \leq n$, the covariance matrix of $g_1(a), \dots, g_k(a)$ is given by ρ' . We need to upper bound the probability that $f'(g_1) = \dots = f'(g_k)$. The result in [IM12] says that as long as all the off-diagonal entries of ρ'^{-1} are non-positive, the probability is maximum when f' is a halfspace. However, if f' is indeed a halfspace, then $\mathbf{Pr}[f'(g_1) = \dots = f'(g_k)] = h_s(\rho')$. Thus, if all the off-diagonal entries of ρ'^{-1} are non-positive, then the soundness of the dictatorship test is $h_s(\rho')$.

For the case of $k = 3$, by a direct analysis of the rounding algorithm, we showed that ρ' is a matrix whose diagonal entries are all 1 and all the off-diagonal entries are the same positive quantity. From this, it is easy to check that all the off diagonal entries of ρ'^{-1} are non-positive and hence the results of [IM12] are applicable here. On the other hand, for $k > 3$, it seems difficult to compute ρ' exactly or even prove that all the off diagonal entries of ρ'^{-1} are non-positive. This makes it impossible to apply the results of [IM12]

here. One might consider the possibility of doing computer simulations to find $\inf_{\rho} h_s(\rho)/h_c(\rho)$ for $k > 3$ (or to make a reasonable conjecture about this quantity). However, note that for $k > 3$, $h_c(\rho)$ is not even completely determined by ρ . As a result, even doing computer simulations for $k > 3$ is rather complicated. This summarizes the difficulty in extending our results to MAX-k-EQUAL for $k > 3$.

8 Conclusion

Our results illustrate the importance of Gaussian partition results in establishing exact optimal UGC hardness and rounding schemes. Not only did we show that a new Gaussian partition result allows to obtain exact UGC hardness of MAX-3-EQUAL, we also showed how the trivial Gaussian partition gives near optimal hardness for MAX-k-CSPs.

There are many interesting open problems that emerge from our work and previous work. Perhaps the most natural open problem is regarding the hardness of MAX-k-EQUAL. In particular, is it true that the generic SDP from [Rag08] followed by the random gaussian / hyperplane rounding is optimal for MAX-k-EQUAL (assuming the Unique Games Conjecture)?

A more general challenge is to obtain further optimal Gaussian partition results. In particular we recall the Standard Simplex Conjecture from [IM12] which says that if (X, Y) are jointly normal random variables in \mathbb{R}^n such that $X, Y \sim \mathcal{N}^n(0, 1)$ and $\text{Cov}(X, Y) = \rho I_n$ where $\rho > 0$, then a partitioning of the gaussian space into k parts of equal measure such that (X, Y) fall in the same partition is maximized when the partition corresponds to a k -simplex centered at the origin. Proving this, will have consequences for hardness of MAX-k-CUT.

Acknowledgements

We are grateful to Ori Gurel-Gurevich and Ron Peled for answering several questions related to [BGGP12]. We are grateful to Per Austrin and Jelani Nelson for helpful comments on an earlier draft. We also thank Anand Bhaskar and Piyush Srivastava for help with using Mathematica. We are grateful to David Williamson for letting us know that our analysis of the MAX-3-EQUAL SDP is essentially identical to the analysis of the MAX-DICUT SDP from [GW95].

AD is grateful to Luca Trevisan and Madhur Tulsiani for numerous discussions and Satish Rao for financially supporting him during the period when this work was done.

References

- [AM09] P. Austrin and E. Mossel. Approximation Resistant Predicates from Pairwise Independence. *Computational Complexity*, 18(2):249--271, 2009. Conference version in CCC 2009.
- [Aus07] P. Austrin. Balanced MAX-2-SAT might not be the hardest. In *Proc. 39th Annual ACM Symposium on Theory of Computing (STOC)*, pages 189--197. ACM, 2007.
- [Bac63] R. Bacon. Approximations to Multivariate Normal Orthant probabilities. *Annals of Mathematical Statistics*, 34:191--198, 1963.
- [BGGP12] I. Benjamini, O. Gurel-Gurevich, and R. Peled. On K-wise Independent Distributions and Boolean Functions. Available at <http://arxiv.org/abs/1201.3261>, 2012.
- [Bor85] C. Borell. Geometric bounds on the Ornstein-Uhlenbeck velocity process. *Probability Theory and Related fields*, 70:1--13, 1985.

- [BP89] E. Boros and A. Prekopa. Closed Form Two-Sided Bounds for Probabilities that at Least r and Exactly r out of n Events Occur. *Mathematics of Operations Research*, 14:317--342, 1989.
- [GW95] M.X. Goemans and D.P. Williamson. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. *Journal of the ACM*, 42(6):1115--1145, 1995.
- [Hås01] J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798--859, 2001.
- [IM12] M. Isaksson and E. Mossel. Maximally stable Gaussian partitions with discrete applications. *Israel Journal of Mathematics*, 189:347--396, 2012.
- [JL84] W. Johnson and J. Lindenstrauss. Extensions of Lipschitz mapping into Hilbert space. *Contemporary Mathematics*, 26:189--206, 1984.
- [Kho02] S. Khot. On the power of unique 2-prover 1-round games. In *Proc. 34th Annual ACM Symposium on the Theory of Computing*, pages 767--775, 2002.
- [KKMO07] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for Max-Cut and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319--357, 2007.
- [KR08] S. Khot and O. Regev. Vertex cover might be hard to approximate to within $2 - \epsilon$. *Journal of Computer & System Sciences*, 74(3):335--349, 2008.
- [LLZ02] M. Lewin, D. Livnat, and U. Zwick. Improved Rounding Techniques for MAX-DI-CUT and MAX-2-SAT. In *Integer Programming and Combinatorial Optimization*, pages 67--82, 2002.
- [MM12] K. Makarychev and Y. Makarychev. Approximation Algorithm for Non-boolean MAX k-CSP. In *International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX*, pages 254--265, 2012.
- [MOO10] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. *Annals of Mathematics*, 171:295--341, 2010.
- [Mos05] E. Mossel. Lecture notes on Fourier Analysis. <http://www.stat.berkeley.edu/~mossel/teach/206af05/>, 2005.
- [Mos10] E. Mossel. Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis*, 19(6):1713--1756, 2010.
- [Rag08] P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proc. 40th Annual ACM Symposium on Theory of Computing (STOC)*, pages 245--254, 2008.
- [ST06] A. Samorodnitsky and L. Trevisan. Gowers uniformity, influence of variables, and PCPs. In *Proc. 38th Annual ACM Symposium on Theory of Computing (STOC)*, pages 11--20, 2006.
- [Wil13] D. Williamson. . Personal communication, 2013.
- [Zwi98] U. Zwick. Approximation Algorithms for Constraint Satisfaction Problems Involving at Most Three Variables per Constraint. In *Proc. of the Ninth Annual ACM/SIGACT-SIAM Symposium on Discrete Algorithms --- (SODA 1998)*, pages 201--210, 1998.

APPENDIX

A.1 Useful Trigonometric facts

Fact 27. For every $0 \leq a \leq 1$, $2 \cos^{-1}(a) - \cos^{-1}(2a - 1) \geq 0$.

Proof. Note that

$$\cos(2 \cos^{-1}(a)) = 2a^2 - 1 \leq 2a - 1 = \cos(\cos^{-1}(2a - 1)).$$

Now recall that if $0 \leq \theta, \phi \leq \pi$, then $\cos \theta \leq \cos \phi$ if and only if $\theta \geq \phi$. Clearly, as $a \geq 0$, $0 \leq 2 \cos^{-1}(a) \leq \pi$. Also, $0 \leq \cos^{-1}(2a - 1) \leq \pi$. This concludes the proof. \square

Fact 28. Let $-1 \leq x \leq 1$. Then, if $x \geq 0$, then $\pi + \cos^{-1}(x) \leq 3 \cos^{-1}(x/3)$. Else if, $x \leq 0$, then $\pi + \cos^{-1}(x) \geq 3 \cos^{-1}(x/3)$.

Proof. Consider $f(x) = 3 \cos^{-1}(x/3) - \pi - \cos^{-1}(x)$. Then, note that within the domain $(-1, 1)$, the function is differentiable and hence

$$\frac{df(x)}{dx} = \frac{-1}{\sqrt{1 - x^2/9}} + \frac{1}{\sqrt{1 - x^2}}.$$

It is easy to see that for all $x \in (-1, 1)$, $df(x)/dx \geq 0$. As $f(0) = 0$, we can conclude that for all $x \in [-1, 0]$, $f(x) \leq 0$ and for all $x \in [0, 1]$, $f(x) \geq 0$. This concludes the proof. \square

Fact 29. Let $f : (0, 1/4] \rightarrow \mathbb{R}$ be defined as

$$f(x) = \frac{1 - \frac{\pi + \cos^{-1}(4x-1)}{2\pi}}{x}.$$

Then, $f(x)$ is decreasing in the interval $(0, 1/4]$.

Proof. We do a change of variables. Put $\cos \theta = 4x - 1$. Thus proving the claim is equivalent to showing that for $\pi/2 \leq \theta \leq \pi$, $g(\theta)$ (defined below) is an increasing function in the said interval.

$$g(\theta) = 4 \cdot \frac{\frac{1}{2} - \frac{\theta}{2\pi}}{1 + \cos \theta}.$$

Next, we evaluate $g'(\theta)$.

$$g'(\theta) = 4 \cdot \frac{(1 + \cos \theta) \cdot \frac{-1}{2\pi} + \sin \theta \cdot \left(\frac{1}{2} - \frac{\theta}{2\pi}\right)}{(1 + \cos \theta)^2}.$$

Note that if we show $g'(\theta) \geq 0$ in the interval $\theta \in [\pi/2, \pi]$, then it implies that $g(\theta)$ is an increasing function in the same interval. Thus, we need to show that for $\theta \in [\pi/2, \pi]$

$$(1 + \cos \theta) \cdot \frac{-1}{2\pi} + \sin \theta \cdot \left(\frac{1}{2} - \frac{\theta}{2\pi}\right) \geq 0.$$

Using the identities $1 + \cos \theta = 2 \cos^2(\theta/2)$ and $\sin \theta = 2 \cos(\theta/2) \cdot \sin(\theta/2)$, we get

$$(\pi - \theta) \sin(\theta/2) \geq \cos(\theta/2) \iff \pi - \theta - \cot(\theta/2) \geq 0.$$

So, we finally need to show that $h(\theta) = \pi - \theta - \cot(\theta/2)$ is non-negative in the interval $\theta \in [\pi/2, \pi]$. But $h'(\theta) = -\cot^2(\theta/2) < 0$. This means that $h(\theta) \geq h(\pi) = 0$ proving our claim. \square

Fact 30. For $0 \leq x \leq 1$, $\cos^{-1}(x) \leq \pi/2 - x$.

Proof.

$$\sin x \leq x \quad \Rightarrow \quad \cos(\pi/2 - x) \leq x \quad \Rightarrow \quad \pi/2 - x \geq \cos^{-1}(x).$$

□

Fact 31. For $0 \leq x \leq 1$, $\cos^{-1}(x - 1) \leq \pi - \sqrt{x}$.

Proof. Let $g(x) = \cos(\sqrt{x}) - 1 + x$. Observe that $g(0) = 0$. Also,

$$g'(x) = -\frac{\sin \sqrt{x}}{2\sqrt{x}} + 1 > 0.$$

This implies that $g(x) \geq 0$ for all $0 \leq x \leq 1$. This implies

$$\begin{aligned} \cos(\sqrt{x}) - 1 + x \geq 0 & \Rightarrow x - 1 \geq -\cos(\sqrt{x}) = \cos(\pi - \sqrt{x}) \\ & \Rightarrow \cos^{-1}(x - 1) \leq \pi - \sqrt{x}. \end{aligned}$$

□

Fact 32. For $0 \leq x \leq 1$, $\cos^{-1}(x) \leq 3\sqrt{1 - x}$.

Proof. Put $x = 1 - \epsilon$. Then, the claim is equivalent to proving that for $0 \leq \epsilon \leq 1$, $\cos^{-1}(1 - \epsilon) \leq 3\sqrt{\epsilon}$. Towards this, define $g(\epsilon) = 3\sqrt{\epsilon} - \cos^{-1}(1 - \epsilon)$. Clearly, $g(0) = 0$. Next, we note that

$$g'(\epsilon) = \frac{3}{2\sqrt{\epsilon}} - \frac{1}{\sqrt{1 - (1 - \epsilon)^2}} = \frac{3}{2\sqrt{\epsilon}} - \frac{1}{\sqrt{2\epsilon - \epsilon^2}} = \frac{3}{\sqrt{4\epsilon}} - \frac{1}{\sqrt{2\epsilon - \epsilon^2}}.$$

It is easy to see that for $\epsilon \in [0, 1]$, $g'(\epsilon) \geq 0$. Hence, for $\epsilon \in [0, 1]$, $g(\epsilon) \geq 0$ finishing the proof. □

Fact 33. For $0.9 \leq x \leq 1$, $\cos^{-1}(2x - 1) \leq 5\sqrt{1 - x}$.

Proof. Note that putting $x = 1 - \epsilon$, this is equivalent to proving that for $0 \leq \epsilon \leq 0.1$, $\cos^{-1}(1 - 2\epsilon) \leq 5\sqrt{\epsilon}$. To prove this, consider the function $g(\epsilon) = 5\sqrt{\epsilon} - \cos^{-1}(1 - 2\epsilon)$. Clearly, $g(0) = 0$. Also,

$$g'(\epsilon) = \frac{5}{2\sqrt{\epsilon}} - \frac{2}{\sqrt{1 - (1 - 2\epsilon)^2}} = \frac{5}{2\sqrt{\epsilon}} - \frac{1}{\sqrt{\epsilon - \epsilon^2}}.$$

Now, note that for $\epsilon \in [0, 0.1]$, $g'(\epsilon) \geq 0$. Hence, for $\epsilon \in [0, 0.1]$, $g(\epsilon) \geq 0$ finishing the proof. □

A.2 Justification for numerically finding the minima In Section 6, we numerically evaluate the minimum of two single variable functions using the software “Mathematica”. We now give a detailed explanation of how we find the minima of these functions to the desired error and the mathematical soundness of this computer-assisted procedure.

A.2.1 Infimum of $h_1(a)$ Given the function $h_1 : (0, 1] \rightarrow \mathbb{R}$ from Section 6 (which is defined as)

$$h_1(a) = \frac{1 - \frac{2 \cos^{-1}(a) + \cos^{-1}(2a-1)}{2\pi}}{a}.$$

To find $\inf_{a \in (0,1]} h_1(a)$, we do the following:

- Show that for the interval $A_1 = (0, x_s]$ and $A_2 = [x_t, 1]$ (where $x_s = 0.179$ and $x_t = 0.99$), $\inf_{x \in A_1} h_1(x) \geq 0.85$ and $\inf_{x \in A_2} h_1(x) \geq 0.83$.
- Show that for the interval $A_3 = (x_s, x_t)$, and $x \in A_3$, $|h'_1(x)| \leq \Delta$ where $\Delta = 500$.
- Divide the interval A_3 into Δ/η (with $\eta = 10^{-4}$) intervals of equal length and evaluate h_1 at each of these points where $h_1(a)$ is evaluated at each point with an error of $\epsilon = 10^{-6}$. Subsequently, take the minimum of all these numbers.

It is clear that the above procedure returns the infimum of h_1 in the interval $(0, 1]$ to within error $\epsilon + \eta/2 \leq 10^{-4}$. Following this procedure, $\inf_{a \in (0,1]} h_1(a)$ was obtained to be 0.803225. Since, we note that the error can be at most 10^{-4} , hence $\inf_{a \in (0,1]} h_1(a) \in [0.803125, 0.803325]$.

We now give proofs for the first and the second item in the above procedure.

Proposition 34. *Let $h_1 : [0, 1] \rightarrow \mathbb{R}$ be defined as*

$$h_1(a) = \frac{1 - \frac{2 \cos^{-1}(a) + \cos^{-1}(2a-1)}{2\pi}}{a}.$$

Then, for $0 \leq a \leq 0.179$, $h(a) \geq 0.85$.

Proof. Using Fact 30 and Fact 31, we have

$$h_1(a) = \frac{1 - \frac{2 \cos^{-1}(a) + \cos^{-1}(2a-1)}{2\pi}}{a} \geq \frac{2a + \sqrt{2a}}{2\pi a} = \frac{1}{\pi} + \frac{1}{\pi\sqrt{2a}}.$$

Plugging in the values, this implies that as long as $a \leq 0.179$, $h_1(a) \geq 0.85$. □

Proposition 35. *Let $h_1 : [0, 1] \rightarrow \mathbb{R}$ be defined as*

$$h_1(a) = \frac{1 - \frac{2 \cos^{-1}(a) + \cos^{-1}(2a-1)}{2\pi}}{a}.$$

Then, for $0.99 \leq a \leq 1$, $h(a) \geq 0.83$.

Proof. Using Fact 32 and Fact 33, we have

$$h_1(a) = \frac{1 - \frac{2 \cos^{-1}(a) + \cos^{-1}(2a-1)}{2\pi}}{a} \geq \frac{1 - \frac{6\sqrt{1-a} + 5\sqrt{1-a}}{2\pi}}{a}.$$

Plugging in the values, this implies that as long as $0.99 \leq a \leq 1$, $h_1(a) \geq 0.83$. □

Proposition 34 and Proposition 35 imply the proof of the first item. The next proposition implies the correctness of the third item.

Proposition 36. *For every $a \in [0.179, 0.99]$, $|h'_1(a)| \leq 500$.*

Proof.

$$h'_1(a) = \frac{\frac{a}{\pi\sqrt{1-a^2}} + \frac{a}{\pi\sqrt{1-(1-2a)^2}} - 1 + \frac{\cos^{-1}(a)}{\pi} + \frac{\cos^{-1}(2a-1)}{2\pi}}{a^2}.$$

This implies that

$$|h'_1(a)| \leq \frac{\frac{a}{\pi\sqrt{1-a^2}} + \frac{a}{2\pi\sqrt{a-a^2}} + 3}{a^2} \leq \frac{3}{a^2} + \frac{\frac{1}{\pi\sqrt{1-a^2}} + \frac{1}{2\pi\sqrt{a-a^2}}}{a}.$$

To bound the value of $|h'_1(a)|$, we consider the two cases: when $0.179 \leq a \leq 0.5$ and when $0.99 \geq a > 0.5$. Splitting into these two cases, it is easy to show

$$|h'_1(a)| \leq 500.$$

□

A.2.2 Infimum of $h_2(a)$ Recall that we need to find the following quantity:

$$\inf_{a \in (1/4, 1]} h_2(a) \quad \text{where} \quad h_2(a) = \frac{1 - \frac{3 \cos^{-1}((4a-1)/3)}{2\pi}}{a}.$$

We do the following change of variables: We put $(4a - 1)/3 = \cos x$. Then, the problem becomes finding the quantity

$$\inf_{x \in [0, \pi/2)} g(x) \quad \text{where} \quad g(x) = 4 \cdot \frac{1 - \frac{3x}{2\pi}}{1 + 3 \cos x}.$$

To find $\inf_{x \in [0, \pi/2)} g(x)$, we do the following:

- Show that for $x \in [0, \pi/2)$, $|g'(x)| \leq \Delta$ where $\Delta = 50$.
- Divide the interval $[0, \pi/2)$ into Δ/η (with $\eta = 10^{-4}$) intervals of equal length and evaluate $g(x)$ at each of these points where $g(x)$ is evaluated at each point with an error of $\epsilon = 10^{-6}$. Subsequently, take the minimum of all these numbers.

It is clear that the above procedure returns the infimum of h_2 in the interval $(0, 1]$ to within error $\epsilon + \eta \cdot (\pi/4) \leq 10^{-4}$. Following this procedure, $\inf_{a \in (0, 1]} h_2(a)$ was obtained to be 0.796070. Since the error is bounded by 10^{-3} , we know $\inf_{a \in (0, 1]} h_2(a) \in [0.795970, 0.796170]$. We now give proof for the first item in the above procedure.

Proposition 37. *Let $g : [0, \pi/2) \rightarrow \mathbb{R}$ be defined as above. Then, for $x \in [0, \pi/2)$, $|g'(x)| \leq 50$.*

Proof.

$$g'(x) = 12 \cdot \frac{\sin x - \frac{3x \sin x}{2\pi} - \frac{1}{2\pi} - \frac{3 \cos x}{2\pi}}{(1 + 3 \cos x)^2}.$$

It is now trivial to see that the absolute value of $g'(x)$ is bounded by 50 at all points in $[0, \pi/2)$. □