Optimal Algorithms and Inapproximability Results for Every CSP?

[Extended Abstract]

Prasad Raghavendra [*][†]
Dept. of Computer Science and Eng.
University of Washington
Seattle, WA
prasad@cs.washington.edu

ABSTRACT

Semidefinite Programming(SDP) is one of the strongest algorithmic techniques used in the design of approximation algorithms. In recent years, Unique Games Conjecture(UGC) has proved to be intimately connected to the limitations of Semidefinite Programming.

Making this connection precise, we show the following result : If UGC is true, then for every constraint satisfaction problem(CSP) the best approximation ratio is given by a certain simple SDP. Specifically, we show a generic conversion from SDP integrality gaps to UGC hardness results for every CSP. This result holds both for maximization and minimization problems over arbitrary finite domains.

Using this connection between integrality gaps and hardness results we obtain a generic polynomial-time algorithm for all CSPs. Assuming the Unique Games Conjecture, this algorithm achieves the optimal approximation ratio for every CSP.

Unconditionally, for all 2-CSPs the algorithm achieves an approximation ratio equal to the integrality gap of a natural SDP used in literature. Further the algorithm achieves at least as good an approximation ratio as the best known algorithms for several problems like MaxCut, Max2Sat, MaxDiCut and Unique Games.

Categories and Subject Descriptors

F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Non Numerical Algorithms and Problems

General Terms

Algorithms, Theory.

Keywords

Semidefinite programming, Constraint Satisfaction Problem, Unique Games Conjecture, Dictatorship Tests, Rounding Schemes

1. INTRODUCTION

A Constraint Satisfaction Problem(CSP) Λ is specified by a finite domain $[q] = \{0, 1, \ldots, q - 1\}$ and a set of predicates $\mathbb{P}$. Every instance of the problem Λ consist of a set of variables $\mathcal{V}$, and a set of constraints on them. Each constraint consists of a predicate from $\mathbb{P}$ applied to a subset of variables. The objective is to find an assignment to the variables that satisfies the maximum number of constraints. The arity k of the CSP Λ is the maximum number of inputs to a predicate in $\mathbb{P}$. A large number of the fundamental optimization problems such as Max3Sat, MaxCut are examples of CSPs.

For most natural CSPs, it is NP-hard to find the optimal assignment. To cope with this intractability, the focus shifts towards approximation algorithms achieving provable guarantees. In this light, it is natural to ask the following question : *"For a given CSP Λ, what is the best approximation to the optimum that can be efficiently computed?"*

Over the past decade, this question has been answered successfully for many important CSPs like Max3Sat [10, 11] with matching approximation algorithms and NP-hardness results. Despite the success, the approximability of many interesting CSPs like MaxCut and Max2SAT still remain open. Towards obtaining tight inapproximability results, the Unique Games Conjecture was introduced by Khot [12]. An equivalent statement of the Unique Games Conjecture (UGC) is as follows:

Unique Games Conjecture[12] (equivalent version)
For any $\delta > 0$, there is a large enough number p such that : Given a set of linear equations of the form $x_i - x_j = c_{ij}$ mod p, it is NP-hard to distinguish between the following two cases:

- *There is a solution to the system that satisfies $1 - \delta$ fraction of the equations.*

- *No solution satisfies more than a δ fraction of the equations.*

The conjecture remains a notorious open problem in hardness of approximation. Assuming the Unique Games Con-

jecture, tight hardness results have been shown for many CSPs like MaxCut [13], Max2Sat [1] and Max-k-CSP[20, 3].

A salient feature of UGC hardness results is their connection to Semidefinite programming(SDP), or more precisely its limitations. On one hand, the UGC hardness results exactly match the approximation factors obtained using SDPs. Moreover, the parameters of the hardness reduction are derived from hard instances to SDPs.

The connection between UGC hardness and SDPs is most apparent in the case of 2-CSP problems over the boolean domain. A 0.878 hardness for MaxCut[13] was obtained using the hard instances for the Goemans-Williamson Max Cut algorithm. The tight hardness result for Max-2-SAT obtained in [1] crucially used SDP vectors that are hard to round. This connection to SDP gaps is also highlighted in the case of MaxCutGain[14] : a variant of MaxCut. Generalizing this line of work, Austrin [2] showed tight Unique Games hardness results for every 2-CSP over the boolean domain under certain additional conjecture.

1.1 Results

Making the connection between SDPs and Unique Games hardness precise, we obtain tight hardness results and rounding schemes for every CSP. In order to describe our results, we will need a few definitions.

A Generalized Constraint Satisfaction Problem (GCSP) is similar to a CSP, but with the predicates replaced by more general "payoff functions". A payoff function returns an arbitrary real value in $[-1, 1]$ instead of a boolean $\{0, 1\}$. The objective is to find an assignment to variables that maximizes the total payoff. By allowing the payoff functions to take negative values, even minimization problems can be formulated as GCSP problems. Some examples of GCSP include maximization problems like Max3Sat MaxCut, Max2SAT, Max-k-Cut and minimization problems like MultiwayCut and MetricLabeling.

Instead of total payoff, we will always refer to the expected payoff which is defined as : total payoff divided by the number of payoff functions. For the purpose of computing approximation ratios, the two notions are equivalent. Note that the expected payoff is always a real number in $[-1, 1]$.

For every GCSP Λ, there is a natural generic SDP relaxation $SDP(I)$ shown in Figure **1**. This relaxation is similar to the SDP used for Max3Sat in Karloff et.al. [11]. In case of 2-CSPs, the relaxation $SDP(I)$ reduces to $SDP(II)$, which is similar to ones already appeared in literature [5, 7, 16] . Further for boolean 2-CSPs, $SDP(I)$ reduces to the well known $SDP(III)$ [17, 4] also shown in Figure **1**.

For a GCSP problem Λ, define $S_\Lambda(c), U_\Lambda(c)$ as follows: $S_\Lambda(c)$: Roughly, $S_\Lambda(c)$ is the integrality gap curve of $SDP(I)$. More precisely, $S_\Lambda(c)$ is the minimum value of the integral optimum over all instances Φ with SDP value c.

$U_\Lambda(c)$: The unique games hardness curve for the problem Λ. Specifically, $U_\Lambda(c)$ is the best polynomial time computable assignment on instances with objective value c, assuming the Unique Games Conjecture is true.

Our first result is a generic conversion from SDP integrality gaps to UG hardness results for any GCSP problem. Assuming the UGC, we show that the best approximation to every GCSP problem Λ is given by $SDP(I)$.

THEOREM 1.1. *(UGC Hardness) For every constant $\eta > 0$, and every GCSP problem Λ:*

$$U_\Lambda(c) \leqslant S_\Lambda(c + \eta) + \eta \qquad \forall c \in [-1, 1)$$

Qualitatively, the result shows that if UGC is true, then $SDP(I)$ is the strongest SDP for every GCSP. Thus if UGC is true, then stronger SDPs obtained through the Lovasz-Schriver, Lasserre and Sherali-Adams hierarchies do not yield better approximation ratios for any GCSP.

Surprisingly, the soundness analysis of the above reduction yields an efficient algorithm for every GCSP. This implication was independently observed by Yi Wu.

THEOREM 1.2. *(Rounding Scheme) For every GCSP problem Λ and constant $\eta > 0$, the rounding scheme Round outputs a solution of value at least $U_\Lambda(c - \eta) - \eta$ on input an instance with SDP value $c \in (-1, 1]$.*

The running time of the algorithm Round is given by $C(\eta, \Lambda) \cdot O(n^c)$ where c is a fixed constant and $C(\eta, \Lambda)$ is a constant depending on η and the problem Λ. Roughly speaking, for every GCSP, the algorithm achieves the optimal approximation under the Unique Games Conjecture. In fact, for every CSP the algorithm achieves the best approximation ratio assuming UGC.

COROLLARY 1.3. *Assuming UGC, for every CSP Λ and every $\eta > 0$, the Round procedure obtains an approximation ratio which is within η of the optimal polynomial time algorithm.*

Most SDP rounding schemes in literature [9, 17, 4, 5, 7, 11] project the SDP vectors in to a few random directions and then use the projections to obtain an assignment to the CSP. Roughly speaking, the algorithm Round proceeds as follows : Project the vectors along constant number(depending on η) of random directions, Compute a polynomial $\mathcal{F}$ on the projections and perform simple randomized rounding on the output of the polynomial. We show that for every instance of the CSP, there exists a polynomial $\mathcal{F}$, for which the rounding performs well. The algorithm Round goes over all possible polynomials $\mathcal{F}$ and outputs the best solution.

Unconditional Results:

Using the Unique games integrality gap of Khot-Vishnoi [16], we convert the UG hardness results to integrality gaps for $SDP(II)$ and $SDP(III)$. Following the chain of implications, one obtains the following unconditional result:

THEOREM 1.4. *Unconditionally, for every GCSP problem Λ with arity 2 and every $\eta > 0$, the algorithm Round on an instance with SDP value c, outputs a solution of value at least $S_\Lambda(c - \eta) - \eta$*

COROLLARY 1.5. *Unconditionally, for every 2-CSP and $\eta > 0$, the Round scheme achieves an approximation ratio within η of the integrality gap of $SDP(II)$ and $SDP(III)$.*

To the best of our knowledge, the semidefinite programs $SDP(II)$ and $SDP(III)$ are the stronger than all SDPs used for 2-CSP algorithms in literature. Thus, the algorithm Round achieves at least as good an approximation ratio as the best known algorithm for several 2-CSPs like MaxCut [9], Max2Sat [17], Unique Games [5] and MaxDiCut [17] unconditionally.

Traditionally, SDP rounding schemes are different for each CSP with a fairly involved analysis at times. Further the proof that a rounding scheme achieves the integrality gap consists of two steps : 1) Show that the scheme achieves a ratio α. 2) Exhibit an integrality gap of α.

In this light, it is interesting that Round yields a uniform rounding scheme for all CSPs. Further, the proof that Round achieves the integrality gap for 2-CSPs is an indirect proof. The proof does not explicitly obtain the value of the integrality gap or the approximation ratio. Instead, one shows that if the algorithm achieves a ratio of α on a particular 2-CSP instance Φ, then it is possible to construct another instance Φ' which has an integrality gap of α.

On the flip side, the proof does not yield the exact value of the approximation ratio or the integrality gap. However, by harnessing the connection between integrality gaps and dictatorship tests we give algorithms to compute integrality gap of $\mathsf{SDP}(II)$ and $\mathsf{SDP}(III)$ for 2-CSP problems Λ.

THEOREM 1.6. *Given a 2-CSP problem Λ, the integrality gaps of $\mathsf{SDP}(II)$ and $\mathsf{SDP}(III)$ for Λ can be approximated to an additive error of η. The running time of the computation depends only on η and the domain size q.*

Note that integrality gaps are defined to be worst case ratio of SDP optimum and the integral optimum over all possible instances - an infinite set. Hence apriori, it is unclear whether integrality gaps can be computed at all.

Techniques and Limitations

The crucial ingredient in our proofs is the Invariance Principle : a generalization of the Central Limit Theorem. Specifically, the principle is a generalization of the following fact : Sum of a large number of $\{-1, 1\}$ random variables has roughly the same distribution as the sum of a large number of Gaussian random variables.

The invariance principle was first shown by V.I.Rotar [21]. More recently, it has been proved with quantitative bounds in various settings in [19, 18, 6]. Apart from several other interesting applications, it has fueled several developments in hardness of approximation [13, 8, 1, 2, 3]. In this work, we use the invariance principle for vector valued multilinear polynomials derived recently by Mossel[18].

Most of the UG hardness results [13, 1] including the recent work of Austrin [2] rely on Gaussian noise stability for their soundness analysis. While noise stability arises naturally in the case of several 2-CSPs, it is unclear how noise stability bounds can be used to obtain hardness for arbitrary CSPs. Departing from this approach, we apply the invariance principle directly, instead of using Gaussian noise stability bounds. Specifically, we use the invariance principle to argue that if an instance Φ has an integral optimum α, then a specific dictatorship test constructed using Φ has a soundness at most α.

The reductions of this paper do not apply to CSPs where either the domain size or the payoff grows with n - the input size. Further we do not obtain any UGC hardness result that grows with n. The reduction does not capture problems with hard constraints like Vertex Cover and 3-Coloring. For GCSP problems with negative payoffs, the additive error η in the approximation guarantee in Theorem 1.2 could possibly overwhelm the approximation ratio.

2. PROOF OVERVIEW

The central lemma of the paper is a conversion from integrality gaps to dictatorship tests. Towards stating the lemma, we briefly describe the problem of dictatorship testing. For the sake of exposition, we restrict our attention to boolean CSPs.

A function $\mathcal{F} : \{0, 1\}^R \to \{0, 1\}$ is said to be a *dictator* if the function is given by $\mathcal{F}(x) = x_i$ for some fixed i. The input to a dictatorship test consists of a function $\mathcal{F} : \{0, 1\}^R \to \{0, 1\}$. The objective is to query the function $\mathcal{F}$ at a few locations, and distinguish whether the function is a dictator or *far* from every dictator. The completeness of the test is the probability of success of a dictator function, while the soundness is the maximum probability of success of a function far from every dictator.

Let Φ be an instance of a GCSP Λ, with an SDP value $\mathsf{FRAC}(\Phi)$. For our applications, it is useful to work with functions $\mathcal{F}$ taking values in interval $[0, 1]$ instead of $\{0, 1\}$. In this regard, we will extend the predicates $P : \{0, 1\}^t \to \{0, 1\}$ of the problem Λ, multilinearly to obtain functions $P : [0, 1]^t \to [0, 1]$.

Using the SDP solution to Φ, we construct a dictatorship test DICT_Φ(see Figure **2**) for functions $\mathcal{F} : \{0, 1\}^R \to [0, 1]$. The test DICT_Φ has the following properties:

- All the tests made by the verifier correspond to applying a predicate/payoff P which is part of Φ. Thus if Φ was a MaxCut instance, the verifier's tests will all be of the form $\mathcal{F}(x) \neq \mathcal{F}(y)$.

- The completeness of the test is nearly equal to the SDP value. More precisely,

$$\mathsf{Completeness}(\mathsf{DICT}_\Phi) = \mathsf{FRAC}(\Phi) - o_{\epsilon,\alpha}(1)$$

  where ϵ, α are parameters independent of the size of instance Φ.

The influence of each coordinate is a measure of how *far* a function $\mathcal{F}$ is from being a dictator. In this work, we will use a slightly different notion of being *far* from a dictator which we refer to as "(γ, τ)-pseudorandom".

The soundness analysis of DICT_Φ is actually given by an efficient algorithm. Specifically, for every function $\mathcal{F} : \{0, 1\}^n \to [0, 1]$, we describe a efficient rounding scheme $\mathsf{Round}_\mathcal{F}$(Figure **3**) for the GCSP problem Λ.

Let $\mathsf{DICT}_\Phi(\mathcal{F})$ denote the performance of $\mathcal{F}$ on the test DICT_Φ. On the other hand, let $\mathsf{Round}_\mathcal{F}(\Phi)$ be the performance of function $\mathcal{F}$ in rounding the instance Φ. Using the invariance principle, we show the following main lemma.

LEMMA 2.1. *(Main Lemma : Integrality Gaps $\Rightarrow$ Dictatorship Tests) If a function $\mathcal{F} : [q]^R \to \mathsf{Conv}(\Delta_q)$ is (γ, τ) pseudorandom with respect to a GCSP instance Φ then*

$$\mathsf{Round}_\mathcal{F}(\Phi) = \mathsf{DICT}_\Phi(\mathcal{F}) \pm o_{\tau,\epsilon,\alpha,\gamma}(1)$$

For application to general CSPs, we show the above lemma with the domain $\{0, 1\}$ replaced by $[q]$ and range replaced by the q-dimensional simplex. The error term $o_{\tau,\epsilon,\alpha,\gamma}(1)$ can be made small with appropriately small choice of the parameters $\tau, \epsilon, \alpha, \gamma$. We stress here that the choice of τ, ϵ, α and γ does not depend on the size of the instance Φ, but only on the error bound required.

For every function $\mathcal{F}$, we have $\mathsf{Round}_\mathcal{F}(\Phi) \leqslant \mathsf{INT}(\Phi)$ where $\mathsf{INT}(\Phi)$ denotes the value of the integral optimum for Φ. The following corollary follows from the main lemma.

SDP(II)

Maximize $\sum_{\{u,v\}\in W} w_{\{u,v\}}\Big(\sum_{i,j\in[q]} P_{\{u,v\}}(i,j)u_i\cdot v_j\Big)$

Subject to $u_i\cdot v_j\geqslant 0,\ u_i\cdot u_j=0 \qquad \forall u,v\in\mathcal{V}, i,j\in[q]$

$$\sum_{i\in[q]}|u_i|^2=1 \qquad \forall u\in\mathcal{V}$$

$$\Big|\sum_{i\in[q]}u_i-\sum_{i\in[q]}v_i\Big|^2=0 \qquad \forall u,v\in\mathcal{V}$$

SDP(III)

Maximize $\sum_i c_i(v_0\cdot v_i)+\sum_{i,j\in[m]}A_{ij}(v_i\cdot v_j)$

Subject to $(v_0-v_i)\cdot(v_0-v_j)\geqslant 0 \qquad \forall i,j\in[m]$

$$|v_i|^2=1 \qquad \forall i\in[m]$$

SDP(I)

Maximize $\sum_{S\in W} w_S\Big(\sum_{\beta\in[q]^S} P_S(\beta(S))X_{(S,\beta)}\Big)$

Subject to
$$v_{(i,c)}\cdot v_{(i,c')}=0 \qquad\qquad \forall i\in[m], c\neq c'\in[q]$$

$$\sum_{c\in[q]}v_{(i,c)}=\mathbf{I} \qquad\qquad \forall i\in[m]$$

$$\sum_{\beta\in[q]^S,\beta(s)=c,\beta(s')=c'}X_{(S,\beta)}=v_{(s,c)}\cdot v_{(s',c')} \qquad\qquad \forall S\in W, s,s'\in S, c,c'\in[q]$$

$$|\mathbf{I}|^2=1$$

Figure 1: Semidefinite Programs

COROLLARY 2.2. *The soundness of the test* DICT_Φ *is given by* $\mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_\Phi)\leqslant\mathsf{INT}(\Phi)+o_{\tau,\epsilon,\alpha,\gamma}(1)$

In Section 5 , we sketch the proof of the main lemma for a specific example $\mathsf{GCSP}:\mathsf{MaxCut}$.
UGC Hardness: Using standard techniques, dictatorship tests can be converted in to UG hardness results. Specifically we show the following:

LEMMA 2.3. *(Dictatorship Tests $\Rightarrow$ UG Hardness) Let Φ be any instance of GCSP problem Λ. Assuming the Unique Games Conjecture, for every $\eta>0$, given an instance Φ' of the problem Λ it is NP-hard to distinguish between:*

- *The optimal assignment for Φ' has value at least* $\mathsf{Completeness}(\mathsf{DICT}_\Phi)-\eta=\mathsf{FRAC}(\Phi)-\eta$

- *Every assignment has value at most* $\mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_\Phi)+\eta$

The details of this reduction are described in Section 6. By definition of the SDP integrality gap curve $\mathcal{S}_\Lambda(c)$, there exists instances Φ^* with SDP value equal to c, while the integral optima is arbitrarily close to $\mathcal{S}_\Lambda$. The above lemmas (2.1,2.3) applied to the instance Φ^*, yield a proof of Theorem 1.1.
Algorithm : Fix an instance Φ of the problem GCSP with SDP value c. From Lemma 2.3, one can show a UGC hardness of $\mathsf{Completeness}(\mathsf{DICT}_\Phi)=c$ versus $\mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_\Phi)$. By definition, $\mathsf{U}_\Lambda(c)$ is the best possible UG hardness result on instances with value c. This implies that $\mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_\Phi)>\mathsf{U}_\Lambda(c)$. Thus there exists some (γ,τ)-pseudorandom function $\mathcal{F}$ such that $\mathsf{DICT}_\Phi(\mathcal{F})\geqslant\mathsf{U}_\Lambda(c)$. Applying Lemma 2.1, this implies $\mathsf{Round}_\mathcal{F}(\Phi)\geqslant\mathsf{U}_\Lambda(c)$. Summarizing, there exists a function $\mathcal{F}$ which yields a rounding of value at least $\mathsf{U}_\Lambda(c)$ on the instance Φ.

The idea of the algorithm Round is to search in the space of all functions to find the best rounding function $\mathcal{F}$. Note

that the space consists of $[0,1]$ valued functions on $\{0,1\}^R$ for some constant R. Thus with appropriate discretization of the space of all functions, we finish the proof of Theorem 1.2.
Unconditional Results: Towards obtaining unconditional results, we convert UG hardness reductions back to integrality gaps along the lines of Khot-Vishnoi [16]. We show the following lemma:

LEMMA 2.4. *(Dictatorship Tests $\Rightarrow$ UG Hardness $\Rightarrow$ Integrality Gaps) Let Φ be an instance of a GCSP problem Λ with arity 2. For every $\eta>0$, there exists an instance Φ' of the problem Λ such that the*

$$\mathsf{FRAC}(\Phi')\ \geqslant\ \mathsf{Completeness}(\mathsf{DICT}_\Phi)-\eta=\mathsf{FRAC}(\Phi)-\eta$$
$$\mathsf{INT}(\Phi')\ \leqslant\ \mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_\Phi)+\eta$$

Fix an instance Φ of a GCSP Λ with SDP value c. Suppose the algorithm Round outputs a solution with value s, then for every function $\mathcal{F}$, $\mathsf{Round}_\mathcal{F}(\Phi)\leqslant s$. By Lemma 2.1, this implies that $\mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_\Phi)\leqslant s$. Using the above Lemma 2.4, we exhibit a gap instance Φ' with SDP value $c-\eta$ and integral value at most $s+\eta$.

The Lemmas 2.1 2.4 together establish an equivalence between integrality gaps and dictatorship tests for 2-CSPs. Thus the problem of finding the strongest integrality gap instance reduces to the problem of finding the dictatorship test(of a specific form) with the best soundness. The set of dictatorship tests of a specific form on functions over the finite domain $[q]^R$, can be discretized easily. This forms the outline of the proof of Theorem 1.6.
Parameters: The parameters $\eta,\gamma,\tau,\alpha,\delta,\kappa,R$ are all constants independent of the size of the instance Φ used in the reductions. For a GCSP problem Λ with arity k over a domain of size q, to achieve an error η in the reductions above : Set $\alpha=\frac{\eta}{100q^k}, \epsilon=\frac{\eta}{100k}$ and $\gamma=\frac{\eta}{100}$. The value of τ depending on ϵ,α is given by the invariance principle. Fix $\delta=\frac{1}{100}\eta\gamma\frac{\tau}{3}k^{-2}(\frac{1}{e\tau\ln 1/(1-\epsilon)})^2$ and the alphabet size R of

the unique games instance large enough to achieve a $1 - \delta$ completeness and soundness of δ.

3. PRELIMINARIES

Let $[q]$ denote the set $\{0, \ldots, q-1\}$. Let $\Delta_q = \{\mathbf{e}_0, \ldots, \mathbf{e}_{q-1}\}$ denote the standard basis in $\mathbb{R}^q$, i.e $\mathbf{e}_i$ is the vector with the $i + 1$'st coordinate equal to 1, while the remaining coordinates are zero. Let $\mathsf{Conv}(\Delta_q)$ denote the convex hull of the vectors $\{\mathbf{e}_0, \ldots, \mathbf{e}_{q-1}\}$. To deal with CSPs over the alphabet $[q]$, we shall frequently use vector valued variables. To aid in understanding, we will use bold face symbols to denote multidimensional objects.

For a set of indeterminates $\{v_1, \ldots, v_m\}$ and a subset $S \subseteq [m]$, we shall use $v_{|S}$ to denote the set $\{v_i | i \in S\}$. For a subset $S \subseteq \{1, \ldots, m\}$, we shall use $[q]^S$ to denote the set of all mappings $[q]^S = \{\beta : S \to [q]\}$

3.1 Generalized Constraint Satisfaction Problems (GCSP)

As a unifying framework for both maximization and minimization problems we define the GCSP problem.

DEFINITION 3.1. *A Generalized Constraint Satisfaction Problem Λ is specified by $\Lambda = ([q], \mathbb{P}, k)$ where $[q] = \{0, 1, \ldots, q-1\}$ is a finite domain, $\mathbb{P} = \{P : [q]^t \to [-1, 1] | t \leqslant k\}$ is a set of payoff functions. The maximum number of inputs to a payoff function $P \in \mathbb{P}$ is known as the arity of the problem Λ.*

DEFINITION 3.2. *An instance Φ of Generalized Constraint Satisfaction Problem $\Lambda = ([q], \mathbb{P}, k)$ is given by $\Phi = (\mathcal{V}, \mathbb{P}_{\mathcal{V}}, W)$ where*

- $\mathcal{V} = \{y_1, \ldots, y_m\}$: *variables taking values over $[q]$.*

- $\mathbb{P}_{\mathcal{V}}$ *consists of the payoffs applied to subsets S of variables $\mathcal{V}$ of size at most k. More precisely, for a subset $S = \{s_1, s_2, \ldots, s_t\} \subset \{1, \ldots, m\}^t$, the payoff function $P_S \in \mathbb{P}_{\mathcal{V}}$ is applied to variables $y_{|S} = (y_{s_1}, \ldots, y_{s_t})$.*

- *Positive weights $W = \{w_S\}$ satisfying $\sum_{S \subset \mathcal{V}, |S| \leqslant k} w_S = 1$. By $S \in W$, we denote a set S chosen from the probability distribution W.*

The objective is to find an assignment to the variables that maximizes the total weighted payoff/expected payoff. i.e. Maximize,

$$\mathbf{E}_{S \in W}\Big[P_S(y_{|S}) \Big] = \sum_{S \subseteq [m], |S| \leqslant k} w_S P_S(y_{|S})$$

3.2 Semidefinite Program for GCSP

Let $\Phi = (\mathcal{V}, \mathbb{P}_{\mathcal{V}}, W)$ be an instance of GCSP. The variables of the SDP(I) are given by

- For each variable $y_i \in \mathcal{V}$, introduce q variables $\mathbf{v}_i = \{v_{(i,0)}, v_{(i,1)}, \ldots, v_{(i,q-1)}\}$ taking values in $\{0, 1\}$. In the intended solution, assigning the variable $y_i = j \in [q]$ translates to $\mathbf{v}_i = \mathbf{e}_j$, i.e $v_{(i,l)} = 1$ if $l = j$ and 0 other wise.

- For every set $S = \{s_1, \ldots, s_t\} \in W$ with a payoff $P_S \in \mathbb{P}_{\mathcal{V}}$, introduce q^t variables $\{X_{(S,\beta)} | \beta \in [q]^S\}$. In the intended solution, the variable $X_{(S,\beta)} = 1$ if and only if for each $s \in S$, $y_s = \beta(s)$.

The constraints of SDP(I) ensure that for each S, the variables $X_{(S,\beta)}$ form a probability distribution $\mathsf{P}_{|S}$ over the partial assignments $\beta \in [q]^S$. Further, the inner products of the vectors $v_{(i,c)}$ are constrained to be consistent with the distribution $\mathsf{P}_{|S}$.

For the purpose of our reductions, the SDP solution need to satisfy the following special property for a constant α independent of the size of instance Φ.

DEFINITION 3.3. *For $\alpha > 0$, the SDP solution $\{v_{(i,c)}, X_{(S,\beta)}\}$ is said to be α-smooth if*

$$\min_{S \in W, \beta \in [q]^S} X_{(S,\beta)} \geqslant \alpha$$

In general, the SDP solution for an instance Φ need not be α-smooth. We describe a smoothing operation in Section 7, thus showing:

LEMMA 3.4. *For a GCSP instance Φ, let $\{v_{(i,c)}, X_{(S,\beta)}\}$ be a solution with objective value $\mathsf{FRAC}(\Phi)$ to SDP(I). For any $\alpha < \frac{1}{q^k}$, there exists an α-smooth solution $\{v'_{(i,c)}, X'_{(S,\beta)}\}$ with an objective value at least $\mathsf{FRAC}(\Phi) - 2\alpha q^k$.*

3.3 Unique Games

DEFINITION 3.5. *An instance of Unique Games represented as $\Gamma = (\mathcal{X} \cup \mathcal{Y}, E, \Pi, \langle R \rangle)$, consists of a bipartite graph over node sets $\mathcal{X}, \mathcal{Y}$ with the edges E between them. Also part of the instance is a set of labels $\langle R \rangle = \{1, \ldots, R\}$, and a set of permutations $\pi_{vw} : \langle R \rangle \to \langle R \rangle$ for each edge $e = (v, w) \in E$. An assignment A of labels to vertices is said to satisfy an edge $e = (v, w)$, if $\pi_{vw}(A(v)) = A(w)$. The objective is to find an assignment A of labels that satisfies the maximum number of edges.*

For a vertex v, we shall use $N(v)$ to denote its neighborhood. For the sake of convenience, we shall use the following version of the Unique Games Conjecture[12] which was shown to be equivalent to the original conjecture by [15].

CONJECTURE 3.6. *(Unique Games Conjecture [12]) For all constants $\delta > 0$, there exists large enough constant R such that given a bipartite unique games instance $\Gamma = (\mathcal{X} \cup \mathcal{Y}, E, \Pi = \{\pi_e : \langle R \rangle \to \langle R \rangle : e \in E\}, \langle R \rangle)$ with number of labels R, it is NP-hard to distinguish between the following two cases:*

- $(1-\delta)$-satisfiable instances: *There exists an assignment A of labels such that for $1-\delta$ fraction of vertices $v \in \mathcal{X}$, all the edges incident at v are satisfied.*

- Instances that are not δ-satisfiable: *No assignment satisfies more than a δ-fraction of the edges E.*

3.4 Noise Operators and Influences

Let Ω denote a finite probability space with q different atoms $\{0, 1, \ldots, q-1\}$. Given a function $\mathcal{F} : [q]^R \to \mathsf{Conv}(\Delta_q)$, it can be thought of as a function over the probability space Ω^R.

DEFINITION 3.7. *For a function $\mathcal{F} : \Omega^R \to \mathsf{Conv}(\Delta_q)$, define*

$$\mathrm{Inf}_i(\mathcal{F}) = \mathbf{E}_{\mathbf{z}}[\mathbf{Var}_{z_i}[\mathcal{F}]]$$

Here $\mathbf{Var}_{z_i}[\mathcal{F}]$ denotes the variance of $\mathcal{F}(\mathbf{z})$ over the choice of the ith coordinate z_i.

DEFINITION 3.8. *For a function* $\mathcal{F} : \Omega^R \to \mathsf{Conv}(\Delta_q)$, *define the function* $T_{1-\epsilon}\mathcal{F}$ *as follows:*

$$T_{1-\epsilon}\mathcal{F}(\mathbf{z}^R) = \mathbf{E}[\mathcal{F}(\tilde{\mathbf{z}}^R) \mid \mathbf{z}^R]$$

where each coordinate $\tilde{z}^{(i)}$ *of* $\tilde{\mathbf{z}}^R = (\tilde{z}^{(1)}, \ldots, \tilde{z}^{(R)})$ *is equal to* $z^{(i)}$ *with probability* $1-\epsilon$ *and with the remaining probability,* $\tilde{z}^{(i)}$ *is a random element from the distribution* Ω.

Due to space constraints, we omit the proof of the following simple fact used in the hardness reduction.

LEMMA 3.9. *Given a function* $\mathcal{F} : [q]^R \to \mathsf{Conv}(\Delta_q)$, *if* $\mathcal{H} = T_{1-\epsilon}\mathcal{F}$ *then*

$$\sum_{i=1}^{R} \mathrm{Inf}_i(\mathcal{H}) \leqslant \frac{1}{e \ln 1/(1-\epsilon)}$$

4. DICTATORSHIP TESTS AND ROUNDING SCHEMES

In this section, we show a generic conversion from an instance $\Phi = (\mathcal{V}, \mathbb{P}_V, W)$ of GCSP Λ to a dictatorship test for functions on $[q]^R$. Let $\{v_{(s,c)}, X_{(S,\beta)}\}$ denote an α-smooth solution for $\mathsf{SDP}(I)$ with objective value at least $\mathsf{FRAC}(\Phi) - 2\alpha q^k$ (using Lemma 3.4).

For each set $S \in W$, define the local probability distribution $\mathsf{P}_{|S}$ as follows:

$$\mathsf{P}_{|S}(\beta) = X_{(S,\beta)} \qquad \forall \beta \in [q]^S$$

Let $\mathcal{V} = (y_1, \ldots, y_m)$ denote the variables in the GCSP instance Φ. For each $i \in [m]$, Ω_i will refer to a probability space with atoms $\{1, 2, \ldots, q\}$. In Ω_i, the probability of occurrence of an atom $c \in [q]$ is given by $|v_{(i,c)}|^2$.

Fix a function $\mathcal{F} : [q]^R \to \mathsf{Conv}(\Delta_q)$. For each i, let $\mathcal{F}_i$ denote the function $\mathcal{F}$ interpreted as a function over the product probability space Ω_i^R.

A dictatorship test distinguishes between dictator functions from functions far from every dictator. For $j \in \{1, \ldots, R\}$, the j'th dictator function is given by $\mathcal{F}(\mathbf{z}) = \mathbf{e}_{z^{(j)}}$. We will use a special definition for a function being *far from a dictator*.

DEFINITION 4.1. *A function* $\mathcal{F} : [q]^R \to \mathsf{Conv}(\Delta_q)$ *is said to* τ-*pseudorandom with respect to a subset* $S \in W$ *if for each* $s \in S$ *the following holds:*

$$\max_{j \in \{1, \ldots, R\}} \mathrm{Inf}_j(T_{1-\epsilon}\mathcal{F}_s) \leqslant \tau$$

DEFINITION 4.2. *A function* $\mathcal{F} : [q]^R \to \mathsf{Conv}(\Delta_q)$ *is said to be* (γ, τ)-*pseudorandom with respect to a* GCSP *instance* $\Phi = (\mathcal{V}, \mathbb{P}_V, W)$ *if the following holds: For a choice of subset* $S \subset \mathcal{V}$ *from the probability distribution* W, *with probability at least* $1 - \gamma$, $\mathcal{F}$ *is* τ-*pseudorandom with respect to* S.

The details of the dictatorship test DICT_Φ and the rounding scheme $\mathsf{Round}_\mathcal{F}$ are described in Figures **2** and **3** respectively. Due to space constraints, we omit the full proof of Completeness and Lemma 2.1. Instead, we will present the proof sketch for a special case(MaxCut) in the next section.

5. MAIN LEMMA

Let $\Phi = (V, E)$ denote an integrality gap instance for MaxCut. Without loss of generality, we may assume that

---

$\boxed{\begin{array}{l}\mathsf{DICT}_\Phi \text{ Test}\end{array}}$

- Pick a subset $S \subset \{1, \ldots, m\}$ at random using probability distribution $W = \{w_S\}$. Let $S = \{s_1, s_2, \ldots, s_t\}$ with $|S| = t$.

- Sample $\mathbf{z}_{|S} = \{\mathbf{z}_{s_1}, \ldots, \mathbf{z}_{s_t}\}$ from the product distribution $\mathsf{P}_{|S}^R$, i.e. For each $1 \leqslant j \leqslant R$, $z_{|S}^{(j)} = \{z_{s_1}^{(j)}, \ldots, z_{s_t}^{(j)}\}$ is sampled using the distribution $\mathsf{P}_{|S}(\beta) = X_{(S,\beta)}$.

- For each $s_i \in S$ and each $1 \leqslant j \leqslant R$, sample $\tilde{z}_{s_i}^j$ as follows: With probability $(1 - \epsilon)$, $\tilde{z}_{s_i}^{(j)} = z_{s_i}^{(j)}$, and with the remaining probability $\tilde{z}_{s_i}^{(j)}$ is a new sample from Ω_{s_i}.

- Query the function values $\mathcal{F}(\tilde{\mathbf{z}}_{s_1}), \ldots, \mathcal{F}(\tilde{\mathbf{z}}_{s_t})$.

- Return the Pay-Off : $P_S\Big(\mathcal{F}(\tilde{\mathbf{z}}_{s_1}), \ldots, \mathcal{F}(\tilde{\mathbf{z}}_{s_t})\Big)$

Figure 2: Dictatorship Test

the edges have weights w_e, and the weights form a probability distribution W. Let $\{v_1, \ldots, v_m\}$ denote the SDP vectors with objective value $\mathsf{FRAC}(\Phi)$ for the following SDP:

$$\text{Maximize } \frac{1}{2}\mathbf{E}_{e=(i,j)\in W}\Big[1 - v_i \cdot v_j\Big] \quad \text{s.t.} \quad |v_i|^2 = 1$$

The input to the dictatorship test DICT_Φ is a function $\mathcal{F} : \{-1,1\}^R \to \{-1,1\}$. Locally, for every edge $e = (i, j)$ in Φ, there exists a distribution over $\{-1,1\}$ assignments that achieve the same value as the SDP. In other words, there exists $\{-1,1\}$ valued random variables z_i, z_j such that

$$1 - v_i \cdot v_j = \mathbf{E}[1 - z_i \cdot z_j]$$

In case of MaxCut, local integral distributions exist even for the simplest SDP. For a general 2-CSP over the boolean domain, we need the triangle inequalities($\mathsf{SDP}(III)$) to ensure existence of local integral distributions. In general, the smallest SDP that guarantees existence of local integral distributions suffices to obtain a dictatorship test. For each edge $e = (i, j)$, let P_e denote the local integral distribution.

$\boxed{\begin{array}{l}\mathsf{DICT}_\Phi \text{ (MaxCut Example)}\end{array}}$

- Pick an edge $e = (i, j)$ from distribution W.

- Sample R times independently from the distribution P_e to obtain $\mathbf{z}_i^R = (z_i^{(1)}, \ldots, z_i^{(R)})$ and $\mathbf{z}_j^R = (z_j^{(1)}, \ldots, z_j^{(R)})$, both in $\{-1,1\}^R$.

- Perturb each coordinate of $\mathbf{z}_i^R$ and $\mathbf{z}_j^R$ independently with probability ϵ to obtain $\tilde{\mathbf{z}}_i^R, \tilde{\mathbf{z}}_j^R$.

- Test if $\mathcal{F}(\tilde{\mathbf{z}}_i^R) \neq \mathcal{F}(\tilde{\mathbf{z}}_j^R)$

Arithmetizing the probability of success, we get

$$\mathsf{DICT}_\Phi(\mathcal{F}) = \frac{1}{2}\mathbf{E}_e \mathbf{E}_{\mathbf{z}_i^R, \mathbf{z}_j^R} \mathbf{E}_{\tilde{\mathbf{z}}_i^R, \tilde{\mathbf{z}}_j^R}\Big[1 - \mathcal{F}(\tilde{\mathbf{z}}_i^R) \cdot \mathcal{F}(\tilde{\mathbf{z}}_j^R)\Big] \quad (1)$$

Completeness : For the sake of exposition, let us assume that the verifier does not perturb its queries. In other words,

we assume $\tilde{\mathbf{z}}_i^R = \mathbf{z}_i^R$ and $\tilde{\mathbf{z}}_j^R = \mathbf{z}_j^R$. Consider the dictator function given by $\mathcal{F}(\mathbf{z}^R) = z^{(1)}$. The probability of success is given by:

$$\Pr[\text{ Success }] = \frac{1}{2}\mathbf{E}_e\mathbf{E}_{\mathbf{z}_i^R, \mathbf{z}_j^R}\left[1 - z_i^{(1)} \cdot z_j^{(1)}\right]$$

Observe that if the verifier decides to query the edge $e = (i, j)$ then it uses the distribution P_e to generate each coordinate of the query. Specifically, this means that the coordinates $z_i^{(1)}$ and $z_j^{(1)}$ satisfy:

$$\mathbf{E}_{\mathbf{z}_i^R, \mathbf{z}_j^R}\left[1 - z_i^{(1)} \cdot z_j^{(1)}\right] = 1 - v_i \cdot v_j$$

Substituting we get, $\Pr[\text{ Success }] = \frac{1}{2}\mathbf{E}_e\left[1 - v_i \cdot v_j\right]$.

Hence if the verifier does not introduce noise, the probability of success is exactly the same as the SDP value $\mathsf{FRAC}(\Phi)$. By making the noise parameter ϵ sufficiently small, the probability of success can be made arbitrarily close to $\mathsf{FRAC}(\Phi)$.
Soundness : Define the function $\mathcal{H}$ as follows:

$$\mathcal{H}(\mathbf{z}^R) = \mathbf{E}[\mathcal{F}(\tilde{\mathbf{z}}^R)|\mathbf{z}]$$

Essentially, $\mathcal{H}(\mathbf{z}^R)$ is the expected value retrieved by the verifier by querying $\mathcal{F}(\mathbf{z}^R)$ with noise added to input $\mathbf{z}^R$. In a sense, the function $\mathcal{H}$ is a smooth version of $\mathcal{F}$. The functions $\mathcal{H}, \mathcal{F}$ can be written as a multilinear polynomials in the coordinates of $\mathbf{z}^R = (z^{(1)}, \ldots, z^{(R)})$. In fact, the multilinear polynomial is just given by the Fourier expansion of $\mathcal{H}$.

$$\mathcal{F}(\mathbf{z}) = \sum_\sigma \hat{\mathcal{F}}_\sigma \prod_{i \in \sigma} z^{(i)} \qquad \mathcal{H}(\mathbf{z}) = \sum_\sigma (1-\epsilon)^{|\sigma|}\hat{\mathcal{F}}_\sigma \prod_{i \in \sigma} z^{(i)}$$

Let $F(\mathbf{x})$, $H(\mathbf{x})$ denote the multilinear polynomials in R variables $\mathbf{x} = (x^{(1)}, \ldots, x^{(R)})$ associated with $\mathcal{F}, \mathcal{H}$.

Rewriting the probability of success in Equation 1 we get

$$\mathsf{DICT}_\Phi(\mathcal{F}) = \frac{1}{2}\mathbf{E}_e\mathbf{E}_{\mathbf{z}_i^R, \mathbf{z}_j^R}\left[1 - H(\mathbf{z}_i^R) \cdot H(\mathbf{z}_j^R)\right]$$

Notice here that we are using the multilinearity of the function $P(x, y) = 1 - xy$ to move the expectation over $\tilde{\mathbf{z}}_i^R, \tilde{\mathbf{z}}_j^R$ inside.

Define the *Global Ensemble* $\mathcal{G}$ as $\mathcal{G} = \{v_1 \cdot \zeta, v_2 \cdot \zeta, \ldots, v_m \cdot \zeta\}$ where ζ is a random Gaussian vector of appropriate dimension. For convenience, let us denote $g_i = v_i \cdot \zeta$. The joint distribution of $\mathcal{G} = (g_1, g_2, \ldots, g_m)$ over $\mathbb{R}^m$ matches the first two moments with the local distributions P_e for each edge $e = (i, j)$. Specifically, for any edge $e = (i, j)$ the following hold:

$$\mathbf{E}[g_i] = \mathbf{E}[z_i] = 0 \qquad\qquad \mathbf{E}[g_i^2] = \mathbf{E}[z_i^2] = 1$$
$$\mathbf{E}[g_i g_j] = \mathbf{E}[z_i z_j] = v_i \cdot v_j$$

Consider the rounding scheme $\mathsf{Round}_\mathcal{F}$ described below for the special case of MaxCut. It is easy to check that the expected value of the cut returned by $\mathsf{Round}_\mathcal{F}$ is

$$\mathsf{Round}_\mathcal{F}(\Phi) = \frac{1}{2}\mathbf{E}_e\mathbf{E}_{\mathcal{G}^R}[1 - p_i^* \cdot p_j^*]$$

Let us suppose $\mathcal{F}$ is "pseudo-random" in the sense that it is *far* from being a dictator. More precisely, the influence of each coordinate on $\mathcal{F}$ is small. The function $\mathcal{H}$ being a noisy version of $\mathcal{F}$ is also "pseudorandom".

$\boxed{\begin{array}{l}\mathsf{Round}_\mathcal{F} \text{ (MaxCut Example)}\end{array}}$

- Generate $\mathcal{G}^R = \{\mathbf{g}_1^R, \mathbf{g}_2^R, \ldots, \mathbf{g}_m^R\}$ by sampling R independent copies of $\mathcal{G}$.

- Compute

$$p_i = H(\mathbf{g}_i^R) = \sum_\sigma \hat{\mathcal{F}}_\sigma \prod_{j \in \sigma} g_i^{(j)}$$

- Define $p_i^* = \begin{cases} 1 & \text{if } p_i > 1 \\ p_i & \text{if } p_i \in [-1, 1] \\ -1 & \text{if } p_i < -1 \end{cases}$

- Assign vertex i, the value 1 with probability $(1 + p_i^*)/2$ and -1 with the remaining probability.

We need to show that $\mathsf{DICT}_\Phi(\mathcal{F}) \approx \mathsf{Round}_\mathcal{F}(\Phi)$. Firstly, let us denote by $P : [-1, 1]^2 \to [-1, 1]$ the function given by $P(x, y) = 1 - xy$. Let us restrict our attention to a particular edge $e = (1, 2)$. For this edge, we will show that

$$\mathbf{E}_{\mathbf{z}_1^R, \mathbf{z}_2^R}\left[P(H(\mathbf{z}_1^R), H(\mathbf{z}_2^R))\right] \approx \mathbf{E}_{\mathbf{g}_1^R, \mathbf{g}_2^R}\left[P(p_1^*, p_2^*)\right] \qquad (2)$$

By using the same argument over all the edges e, the required result follows.

Here is a rough statement of the invariance principle tailored to the application at hand. We refer the reader to [19, 18] for an accurate description of the invariance principle.
(Invariance Principle) *Let $\mathbf{z} = \{z_1, z_2\}$ and $\mathbf{G} = \{g_1, g_2\}$ be two sets of random variables with matching first and second moments. Let $\mathbf{z}^R, \mathbf{G}^R$ denote R independent copies of the random variables $\mathbf{z}$ and $\mathbf{G}$.*

Let $H(\mathbf{x})$ be a low degree polynomial in $2R$ variables $\{\{x_1^{(1)}, x_2^{(1)}\}, \ldots, \{x_1^{(R)}, x_2^{(R)}\}\}$. Further H is multilinear in the following strict sense : Every monomial in H contains at most one of $\{x_1^{(i)}, x_2^{(i)}\}$ for every i.

Suppose H is far from every dictator, then the random variables $H(\mathbf{z}^R)$ and $H(\mathbf{G}^R)$ have nearly the same distribution. More precisely, for every smooth function Ψ,

$$\mathbf{E}_{\mathbf{z}^R}[\Psi(H(\mathbf{z}^R))] \approx \mathbf{E}_{\mathbf{G}^R}[\Psi(H(\mathbf{G}^R))]$$

The same holds even for a vector of multilinear polynomials given by $H = (H_1, H_2, \ldots, H_t)$.

While the function $\mathcal{F}$ could itself be of arbitrary degree, its noisy version $\mathcal{H}$ behaves like a 'low degree' function. Hence we can hope to apply the invariance principle on the polynomial H.

The proof of Equation 2 is carried out in two steps.
Step I : The predicate/payoff is currently defined as $P(x, y) = 1 - xy$ in the domain $[-1, 1]^2$. We will extend it to a smooth function over the entire space $\mathbb{R}^2$, with all its partial derivatives up to order 3 bounded uniformly throughout $\mathbb{R}^2$. We stress here that the function $P(x, y) = 1 - xy$ by itself does not have uniformly bounded derivatives in $\mathbb{R}^2$.

Let us think of $\mathbf{H} = (H(\mathbf{x}_1^R), H(\mathbf{x}_2^R))$ as a vector of multilinear polynomials over variables $\{x_1, x_2\}^R$. Apply the invariance principle with the ensembles $\mathbf{z} = \{z_1, z_2\}$ and $\mathbf{G} = \{g_1, g_2\}$, for the vector of multilinear polynomials $\mathbf{H}$ and the smooth function $\Psi = P$. This yields,

$$\mathbf{E}_{\mathbf{z}_1^R, \mathbf{z}_2^R}\left[P(H(\mathbf{z}_1^R), H(\mathbf{z}_2^R))\right] \approx \mathbf{E}_{\mathbf{g}_1^R, \mathbf{g}_2^R}\left[P(H(\mathbf{g}_1^R), H(\mathbf{g}_2^R))\right]$$

Step II : Notice that the values $H(\mathbf{z}_1^R)$ and $H(\mathbf{z}_2^R)$ are always in the range $[-1,1]$. Applying invariance principle, the values $H(\mathbf{g}_1^R), H(\mathbf{g}_2^R)$ are also nearly always in the range $[-1,1]$. Hence with high probability, we have $p_1 = p_1^*$ and $p_2 = p_2^*$ in the $\mathsf{Round}_{\mathcal{F}}$ subroutine. In fact, the invariance principle in [18] obtains precise upper bounds on $\mathbf{E}[(p_1 - p_1^*)^2]$ and $\mathbf{E}[(p_2 - p_2^*)^2]$. Since P is smooth, and p_1, p_2 are "very close" to p_1^*, p_2^* we get

$$\mathbf{E}_{\mathbf{g}_1^R, \mathbf{g}_2^R}\Big[P(p_1, p_2)\Big] \approx \mathbf{E}_{\mathbf{g}_1^R, \mathbf{g}_2^R}\Big[P(p_1^*, p_2^*)\Big]$$

Combining steps I and II, we obtain the intended result. In general, the invariance principle requires the ensembles of random variables to be hypercontractive [19, 18]. In terms of probabilities, this translates to the following constraint on all the local probability distributions $\mathsf{P}_{|S}$: Every event with non-zero probability in $\mathsf{P}_{|S}$, should occur with probability at least α for some fixed α. Hence the above reduction can only apply to SDP solutions that are α-smooth.

6. HARDNESS REDUCTION

In this section, we will make use of the dictatorship test shown in Section 4 to obtain Unique Games based hardness result, thus proving Lemma 2.3.

Let $\Gamma = (\mathcal{X} \cup \mathcal{Y}, E, \Pi, \langle R \rangle)$ be a bipartite unique games instance. Further let $\Phi = (\mathcal{V}, \mathbb{P}_{\mathcal{V}}, W)$ be an instance of a GCSP problem $\Lambda = ([q], \mathbb{P}, k)$.

Starting from the unique games instance Γ, we shall construct an instance $\Phi(\Gamma)$ of the GCSP problem Λ. For each vertex $w \in \mathcal{Y}$, we shall introduce a long code over $[q]^R$. More precisely, the instance $\Phi(\Gamma)$ is given by $\Phi(\Gamma) = (\mathcal{Y} \times [q]^R, \mathbb{P}', W')$. All the payoff functions in $\mathbb{P}'$ will belong to $\mathbb{P}_{\mathcal{V}}$, ensuring that $\Phi(\Gamma)$ is also an instance of the GCSP problem Λ. Since the set of variables of $\Phi(\Gamma)$ is given by $\mathcal{Y} \times [q]^R$, an assignment to $\Phi(\Gamma)$ consists of a set of functions,

$$\mathsf{F}^w : [q]^R \to [q] \quad \text{for each} \quad w \in \mathcal{Y}$$

For each $w \in \mathcal{Y}$, define $\mathcal{F}^w : [q]^R \to \mathsf{Conv}(\Delta_q)$ as follows:

$$\mathcal{F}^w(\mathbf{z}) = \mathbf{e}_{\mathsf{F}^w(z)}$$

For a permutation $\pi : \{1, \ldots, R\} \to \{1, \ldots, R\}$ and $\mathbf{z} \in [q]^R$, define $\pi(\mathbf{z})$ as $\big(\pi(\mathbf{z})\big)^{(j)} = z^{(\pi^{-1}(j))}$ for all $j \in \{1, \ldots, R\}$. For each $v \in \mathcal{X}$, define a function $\mathcal{F}^v : [q]^R \to \mathsf{Conv}(\Delta_q)$,

$$\mathcal{F}^v(\mathbf{z}) = \mathbf{E}_{w \in N(v)}\big[\mathcal{F}^w(\pi_{vw}(\mathbf{z}))\big]$$

The basic idea behind converting a dictatorship test to Unique Games hardness is similar to Khot et.al.[13]. Roughly speaking, the verifier performs the dictatorship test DICT_{Φ} on the functions $\mathcal{F}^v$ for $v \in \mathcal{X}$. Note that the functions $\mathcal{F}^v$ are not explicitly available to the verifier. However, this access can be simulated by accessing $\mathcal{F}^w$ for a random neighbor $w \in N(v)$.

$\mathsf{Oracle}(\mathcal{F}^v)$

- On a query $\mathcal{F}^v(\mathbf{z})$, Pick a random neighbor $w \in N(v)$, and return $\mathcal{F}^w(\pi_{vw}(\mathbf{z}))$.

$\mathsf{Verifier}(\Phi(\Gamma))$

- Pick $v \in \mathcal{X}$ at random.
- Perform the test DICT_{Φ} on $\mathcal{F}^v$, by transferring each of the queries to the $\mathsf{Oracle}(\mathcal{F}^v)$.

The queries of the $\mathsf{Verifier}(\Phi(\Gamma))$ through the oracle, translate in to tests/payoffs over the functions $\mathcal{F}^w$. In turn, this is equivalent to tests/payoffs on the values of functions $\mathsf{F}^w(z)$. Summarizing, the set of all tests of the above verifier yield a GCSP instance over the variables $\mathcal{Y} \times [q]^R$.

Let $\mathbf{z}_1, \ldots, \mathbf{z}_t \in [q]^R$ be random variables denoting the query locations of $\mathsf{Verifier}(\Phi(\Gamma))$. Further let P denote the payoff/test that the $\mathsf{Verifier}(\Phi(\Gamma))$ decides to perform on these locations. Arithmetizing the expected payoff returned by the verifier we get,

$$\mathbf{E}_{v \in \mathcal{X}}\mathbf{E}_{\mathbf{z}_i}\mathbf{E}_{w_j \in N(v)}\Big[P\Big(\mathcal{F}^{w_1}(\pi_{vw_1}(\mathbf{z}_1)), \ldots, \mathcal{F}^{w_t}(\pi_{vw_t}(\mathbf{z}_t))\Big)\Big]$$

The payoff functions P are multilinear in the region $\mathsf{Conv}(\Delta_q)$. The choices of the oracle $w_1, \ldots, w_t \in N(v)$ are independent of each other, and the verifier's query locations. In this light, we can write the expected payoff as

$$\mathbf{E}_{v \in \mathcal{X}}\mathbf{E}_{\mathbf{z}_i}\Big[P\Big(\mathbf{E}_{w_1 \in N(v)}[\mathcal{F}^{w_1}(\pi_{vw_1}(\mathbf{z}_1))], \ldots$$
$$\ldots, \mathbf{E}_{w_t \in N(v)}[\mathcal{F}^{w_t}(\pi_{vw_t}(\mathbf{z}_t))]\Big)\Big]$$

Hence the expected payoff is just equal to $\mathbf{E}_{v \in \mathcal{X}}[\mathsf{DICT}_{\Phi}(\mathcal{F}^v)]$
Completeness: Let $\mathcal{A}$ be an assignment to the Unique games instance Γ such that the following holds: For at least $1 - \delta$ fraction of the vertices $v \in \mathcal{X}$ all the edges incident at v are satisfied. Let us call such vertices *good* vertices. The assignment to the GCSP instance $\Phi(\Gamma)$ is given by the following set of functions : $\mathsf{F}^w(\mathbf{z}) = z^{(\mathcal{A}(w))}$, or equivalently $\mathcal{F}^w(\mathbf{z}) = \mathbf{e}_{z(\mathcal{A}(w))}$. For every *good* vertex $v \in \mathcal{X}$, we have:

$$\mathcal{F}^v(\mathbf{z}) = \mathbf{E}_{w \in N(v)}\big[\mathcal{F}^w(\pi_{vw}(\mathbf{z}))\big]$$
$$= \mathbf{E}_{w \in N(v)}\big[\mathbf{e}_{z(\pi_{vw}^{-1}(\mathcal{A}(w)))}\big] = \mathbf{e}_{z(\mathcal{A}(v))}$$

In other words, the functions $\mathcal{F}^v$ are dictator functions for every *good* vertex $v \in \mathcal{X}$. With at least $(1 - \delta)$ fraction of the vertices being *good*, the expected payoff is at least

$$(1 - \delta) \cdot (\mathsf{Completeness}(\mathsf{DICT}_{\Phi})) + \delta \cdot (-1)$$
$$\geqslant \mathsf{Completeness}(\mathsf{DICT}_{\Phi}) - o_{\epsilon,\delta}(1)$$

Soundness : Suppose there is an assignment to the variables $\mathcal{Y} \times [q]^R$ whose payoff is greater than $\mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_{\Phi}) + \eta$. Then we have,

$$\mathbf{E}_{v \in \mathcal{X}}[\mathsf{DICT}(\mathcal{F}^v)] > \mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_{\Phi}) + \eta$$

As all the payoff functions are bounded by 1, for at least η fraction of the vertices $v \in \mathcal{X}$, $\mathsf{DICT}_{\Phi}(\mathcal{F}^v) > \mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_{\Phi})$. Henceforth we refer to these vertices as *good* vertices. By definition of $\mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_{\Phi})$, for every *good* vertex the function $\mathcal{F}^v$ is not (γ, τ)-pseudorandom with respect to Φ.

Consider a *good* vertex $v \in \mathcal{X}$. For a random choice of subset $S \subset [m]$, from the probability distribution W, with probability at least γ, the function $\mathcal{F}^v$ is not τ-pseudorandom with respect to S. By an averaging argument, there exists a subset S such that for at least γ-fraction of the *good* vertices, the function $\mathcal{F}^v$ is not τ-pseudorandom with respect to S. Fix such a set $S = \{s_1, \ldots, s_t\}$. For convenience, let us denote $\mathcal{H}_s^v = T_{1-\epsilon}\mathcal{F}_s^v$ for each $v \in \mathcal{X} \cup \mathcal{Y}$ and $s \in S$.
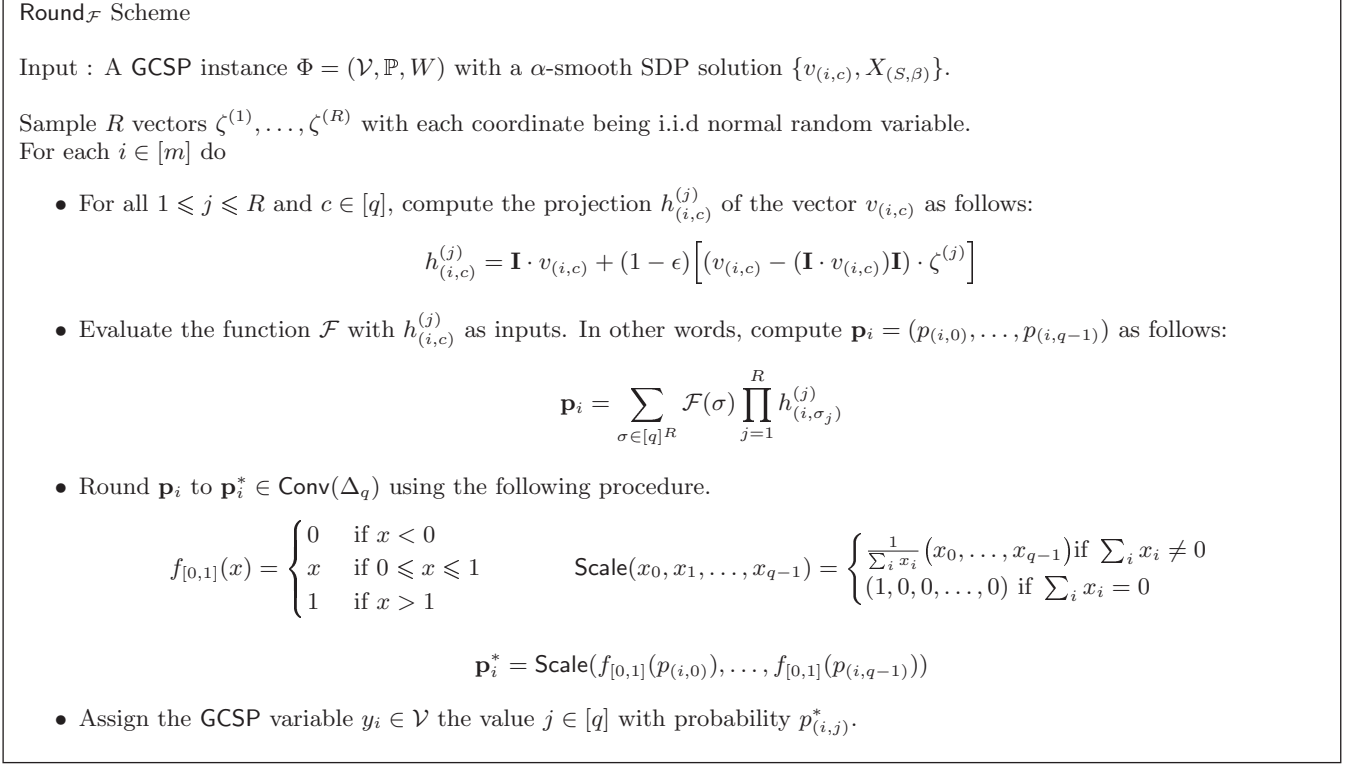
For each vertex $v \in \mathcal{X}$ define the set of labels $\mathsf{L}(v)$ as

$$\mathsf{L}(v) = \{j | \exists s \in S, \mathrm{Inf}_j(\mathcal{H}_s^v) > \tau\}$$

Similarly, for each $w \in \mathcal{Y}$ define,

$$\mathsf{L}(w) = \{j | \exists s \in S, \mathrm{Inf}_j(\mathcal{H}_s^w) > \tau/3\}$$

Round$_\mathcal{F}$ Scheme

Input : A GCSP instance $\Phi = (\mathcal{V}, \mathbb{P}, W)$ with a α-smooth SDP solution $\{v_{(i,c)}, X_{(S,\beta)}\}$.

Sample R vectors $\zeta^{(1)}, \ldots, \zeta^{(R)}$ with each coordinate being i.i.d normal random variable.
For each $i \in [m]$ do

- For all $1 \leqslant j \leqslant R$ and $c \in [q]$, compute the projection $h_{(i,c)}^{(j)}$ of the vector $v_{(i,c)}$ as follows:

$$h_{(i,c)}^{(j)} = \mathbf{I} \cdot v_{(i,c)} + (1 - \epsilon)\Big[(v_{(i,c)} - (\mathbf{I} \cdot v_{(i,c)})\mathbf{I}) \cdot \zeta^{(j)}\Big]$$

- Evaluate the function $\mathcal{F}$ with $h_{(i,c)}^{(j)}$ as inputs. In other words, compute $\mathbf{p}_i = (p_{(i,0)}, \ldots, p_{(i,q-1)})$ as follows:

$$\mathbf{p}_i = \sum_{\sigma \in [q]^R} \mathcal{F}(\sigma) \prod_{j=1}^{R} h_{(i,\sigma_j)}^{(j)}$$

- Round $\mathbf{p}_i$ to $\mathbf{p}_i^* \in \mathsf{Conv}(\Delta_q)$ using the following procedure.

$$f_{[0,1]}(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leqslant x \leqslant 1 \\ 1 & \text{if } x > 1 \end{cases} \qquad \mathsf{Scale}(x_0, x_1, \ldots, x_{q-1}) = \begin{cases} \frac{1}{\sum_i x_i}(x_0, \ldots, x_{q-1}) & \text{if } \sum_i x_i \neq 0 \\ (1, 0, 0, \ldots, 0) & \text{if } \sum_i x_i = 0 \end{cases}$$

$$\mathbf{p}_i^* = \mathsf{Scale}(f_{[0,1]}(p_{(i,0)}), \ldots, f_{[0,1]}(p_{(i,q-1)}))$$

- Assign the GCSP variable $y_i \in \mathcal{V}$ the value $j \in [q]$ with probability $p_{(i,j)}^*$.

Figure 3: Rounding Scheme

Consider the following Labeling for the unique games instance Γ : For each vertex $v \in \mathcal{X} \cup \mathcal{Y}$, assign a random label from $\mathsf{L}(v)$ if it is nonempty, else assign a uniformly random label.

At least η fraction of the vertices $v \in \mathcal{X}$ are *good* vertices. By the choice of S, at least γ fraction of the *good* vertices $v \in \mathcal{X}$ have a non-empty label set $\mathsf{L}(v)$. Fix a *good* vertex v with a nonempty label set $\mathsf{L}(v)$. Consider a label $j \in \mathcal{L}(v)$. By definition of $\mathsf{L}(v)$, we have $\mathrm{Inf}_j(\mathcal{H}_i^v) \geqslant \tau$ for some $s \in S$. The function $\mathcal{H}_s^v$ is given by $\mathcal{H}_s^v(\mathbf{z}) = \mathbf{E}_{w \in N(v)}[\mathcal{H}_s^w(\pi_{vw}(\mathbf{z}))]$. By convexity of influences, if $\mathrm{Inf}_j(\mathcal{H}_s^v) \geqslant \tau$ then
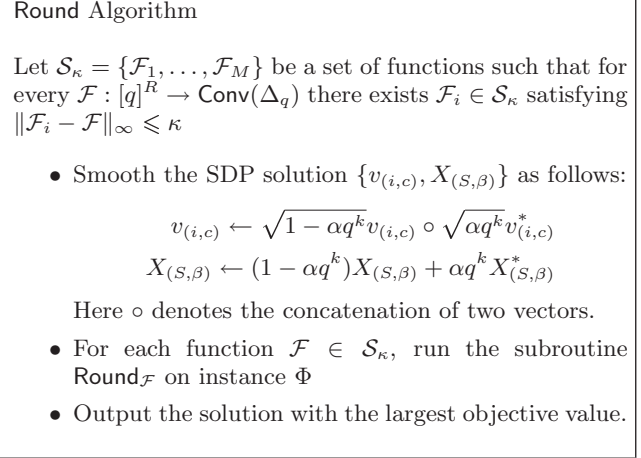
$$\mathbf{E}_{w \in N(v)}[\mathrm{Inf}_{\pi_{vw}(j)}(\mathcal{H}_s^w)] \geqslant \tau$$

Since the range of the function $\mathcal{H}_s^w$ is Δ_q, we have $\mathrm{Inf}_j(\mathcal{H}_s^w) \leqslant 2$. Hence for at least $\tau/3$ fraction of neighbors $w \in N(v)$, $\mathrm{Inf}_{\pi_{vw}(j)}(\mathcal{H}_s^w) \geqslant \frac{\tau}{3}$. Thus for at least $\tau/3$ fraction of the neighbors $w \in N(v)$, there exists $j \in [R]$ such that $j \in \mathsf{L}(v)$ and $\pi_{vw}(j) \in \mathsf{L}(w)$. For every such neighbor w, the edge constraint π_{vw} is satisfied with probability at least $\frac{1}{|\mathsf{L}(v)||\mathsf{L}(w)|}$.

From Lemma 3.9, each function $\mathcal{H}_s^v$ can have at most $C(\tau, \epsilon) = \frac{1}{e\tau \ln 1/(1-\epsilon)}$ influential coordinates. Thus the maximum size of the label set $\mathsf{L}(v)$ is $kC(\tau, \epsilon)$. In conclusion, the expected fraction of unique games constraints satisfied is at least $\eta \times \gamma \times \frac{\tau}{3} \times k^{-2}C(\tau, \epsilon)^{-2}$. For small enough choice of unique games soundness δ, the expected fraction of satisfied edges exceeds δ.

7. ALGORITHM

For every instance Φ, there is a canonical SDP solution $\{v_{(i,c)}^*, X_{(S,\beta)}^*\}$ which corresponds to a uniform distribution on all the integral solutions. More precisely, call a SDP solution *integral* if all the variables take $\{0, 1\}$ values. Since $\mathrm{SDP}(I)$ is a convex program, there exists a feasible solution

---

Round Algorithm

Let $\mathcal{S}_\kappa = \{\mathcal{F}_1, \ldots, \mathcal{F}_M\}$ be a set of functions such that for every $\mathcal{F} : [q]^R \to \mathsf{Conv}(\Delta_q)$ there exists $\mathcal{F}_i \in \mathcal{S}_\kappa$ satisfying $\|\mathcal{F}_i - \mathcal{F}\|_\infty \leqslant \kappa$

- Smooth the SDP solution $\{v_{(i,c)}, X_{(S,\beta)}\}$ as follows:

$$v_{(i,c)} \leftarrow \sqrt{1 - \alpha q^k} v_{(i,c)} \circ \sqrt{\alpha q^k} v_{(i,c)}^*$$
$$X_{(S,\beta)} \leftarrow (1 - \alpha q^k) X_{(S,\beta)} + \alpha q^k X_{(S,\beta)}^*$$

Here $\circ$ denotes the concatenation of two vectors.

- For each function $\mathcal{F} \in \mathcal{S}_\kappa$, run the subroutine Round$_\mathcal{F}$ on instance Φ
- Output the solution with the largest objective value.

Figure 4: Algorithm

$\{v_{(i,c)}^*, X_{(S,\beta)}^*\}$ that corresponds to the average of all the integral solutions.

To apply Lemma 2.1, we need the SDP solution to be α-smooth for some fixed constant α. In general, the optimal SDP solution need not be α-smooth for any fixed constant α. In order to obtain an α-smooth solution, take a convex combination of the SDP solution $\{v_{(i,c)}, X_{(S,\beta)}\}$ with the canonical SDP solution.

Using the fact that payoff functions are smooth, the following lemma can be shown(proof in full version):

LEMMA 7.1. *For two functions* $\mathcal{F}, \mathcal{F}' : [q]^R \to \mathsf{Conv}(\Delta_q)$ *and a* GCSP *instance* Φ, *we have*

$$\|\mathsf{Round}_\mathcal{F}(\Phi) - \mathsf{Round}_{\mathcal{F}'}(\Phi)\| \leqslant O(\|\mathcal{F}' - \mathcal{F}\|_\infty)$$

where the constant in O depends on the GCSP *problem* Λ.

Among (γ, τ)-pseudorandom functions, let $\mathcal{F}^*$ be the function that achieves the optimal expected payoff. i.e. $\mathsf{DICT}_\Phi(\mathcal{F}^*) = \mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_\Phi)$. There exists certain $\mathcal{F}_i \in \mathcal{S}_\kappa$ which is κ-close to $\mathcal{F}^*$. By the above lemma, for sufficiently small choice of κ, the output of Round algorithm is at least $\mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_\Phi) - \eta$. In conjunction with Lemmas 2.1 and 2.4, this completes the proofs of Theorems 1.2 and 1.4.

8. CONSTRUCTING INTEGRALITY GAPS

Except for some minor modifications, the following theorem is a direct consequence of [16].

THEOREM 8.1. *For every $\delta > 0$, there exists a UG instance, $\Gamma = (\mathcal{X} \cup \mathcal{Y}, E, \Pi = \{\pi_e : [R] \to [R] \mid e \in E\}, [R])$ and vectors $\{w^i\}$ for every $w \in \mathcal{Y}$, $i \in [R]$ such that the following conditions hold :*

- *Every assignment satisfies $< \delta$ fraction of constraints Π.*

- *For all $w, w_1, w_2 \in \mathcal{Y}, i, j \in [R]$,*

$$w_1^i \cdot w_2^j \geqslant 0 \qquad\quad w^i \cdot w^j = 0.$$
$$\sum_{j \in [R]} w^j = \mathbf{I} \qquad\quad |\mathbf{I}|^2 = 1.$$

- *The SDP value is at least $1 - \delta$:*

$$\mathbf{E}_{v \in \mathcal{X}, w_1, w_2 \in N(v)}\Big[\sum_{r \in R} w_1^{\pi_1(i)} \cdot w_2^{\pi_2(i)} \Big] \geqslant 1 - \delta$$

Let Φ be an instance of GCSP problem Λ of arity 2. Apply Theorem 8.1, with a sufficiently small δ to obtain a UGC instance Γ and SDP vectors $\{w^j | w \in \mathcal{Y}, j \in [q]\} \cup \{\mathbf{I}\}$. Consider the instance $\Phi(\Gamma)$ constructed by running the hardness reduction in Section 6 on the integrality gap instance Γ. The variables of the instance $\Phi(\Gamma)$ are given by $\mathcal{Y} \times [q]^R$.

The program $\mathsf{SDP}(II)$ on the instance $\Phi(\Gamma)$ contains q vectors $\{\mathbf{V}_i^{(w,\mathbf{z})} | i \in [q]\}$ for each vertex $(w, \mathbf{z}) \in \mathcal{Y} \times [q]^R$ and a special vector $\mathbf{I}$ denoting the constant 1. Define a solution to $\mathsf{SDP}(II)$ as follows: Set the vector $\mathbf{I}$ to be the corresponding vector in the instance Γ. For each $(w, \mathbf{z}) \in \mathcal{Y} \times [q]^R$ and $i \in [q]$

$$\mathbf{V}_i^{(w,\mathbf{z})} = \sum_{z^{(j)} = i} w^j$$

It is easy to check that the vectors $\mathbf{V}_i^{(w,\mathbf{z})}$ satisfy the constraints of $\mathsf{SDP}(II)$ and have an SDP value close to $\mathsf{FRAC}(\Phi)$. On the other hand, the soundness analysis in Section 6 implies that the integral optimum for $\Phi(\Gamma)$ is at most $\mathsf{Soundness}_{\gamma,\tau}(\Phi) + \eta$. Details of proof of Lemma 2.4 will appear in the full version.

9. REFERENCES

[1] Per Austrin. Balanced max 2-sat might not be the hardest. In *ACM STOC*, pages 189–197. 2007.

[2] Per Austrin. Towards sharp inapproximability for any 2-csp. In *IEEE FOCS*, pages 307–317. 2007.

[3] Per Austrin, Elchannan Mossel. Approximation Resistant Predicates from Pairwise Independence. In *IEEE CCC*, 2008.

[4] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for maximum constraint satisfaction problems. In *ACM-SIAM SODA 2007*, pages 62–68. 2007.

[5] Moses Charikar, Yury Makarychev, and Konstantin Makarychev. Near-optimal algorithms for unique games. In *ACM STOC*, pages 205–214. 2006.

[6] Sourav Chatterjee. A simple invariance theorem, arxiv:math/0508213v1., 2005.

[7] Eden Chlamtac, Yury Makarychev, and Konstantin Makarychev. How to play unique games using embeddings. In *IEEE FOCS*, pages 687–696. 2006.

[8] Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. In *ACM STOC*, pages 344–353. 2006.

[9] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *JACM*, 42(6):1115–1145, 1995.

[10] Johann Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.

[11] Howard Karloff and Uri Zwick. A 7/8-approximation algorithm for MAX 3SAT? In *IEEE FOCS*, pages 406–415. 1997.

[12] Subhash Khot. On the power of unique 2-prover 1-round games. In *ACM STOC*, pages 767–775. 2002.

[13] Subhash Khot, Guy Kindler, Elchannan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable CSPs? In *IEEE FOCS*, pages 146–154. 2004.

[14] Subhash Khot and Ryan O'Donnell. SDP gaps and UGC-hardness for MAXCUTGAIN. In *IEEE FOCS*, pages 217–226. 2006.

[15] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within /spl epsi/. In *IEEE CCC*, volume 18, pages 379–386. 2003.

[16] Subhash Khot and Nisheeth K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into l_1. In *IEEE FOCS*, pages 53–62. 2005.

[17] Michael Lewin, Dror Livnat, and Uri Zwick. Improved rounding techniques for the MAX 2-SAT and MAX DI-CUT problems. In *IPCO* , pages 67–82, 2002.

[18] Elchanan Mossel. Gaussian bounds for noise correlation of functions, arxiv:math/0703683. 2007.

[19] Elchannan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. In *IEEE FOCS*, pages 21–30. 2005.

[20] Alex Samorodnitsky and Luca Trevisan. Gowers uniformity, influence of variables, and PCPs. In *ACM STOC*, pages 11–20. 2006.

[21] V.I.Rotar'. Limit theorems for polylinear forms. *Journal of Multivariate Analysis*, 9(4):511–530, 1979.