Integrality Gaps for Strong SDP Relaxations of Unique Games

Prasad Raghavendra*,
Microsoft Research New England
Cambridge, MA 02142

David Steurer†,
Department of Computer Science
Princeton University, Princeton, NJ 08540-5233

Abstract— With the work of Khot and Vishnoi [18] as a starting point, we obtain integrality gaps for certain strong SDP relaxations of Unique Games. Specifically, we exhibit a Unique Games gap instance for the basic semidefinite program strengthened by all valid linear inequalities on the inner products of up to $\exp(\Omega(\log\log n)^{1/4})$ vectors. For a stronger relaxation obtained from the basic semidefinite program by R rounds of Sherali–Adams lift-and-project, we prove a Unique Games integrality gap for $R = \Omega(\log\log n)^{1/4}$.

By composing these SDP gaps with UGC-hardness reductions, the above results imply corresponding integrality gaps for every problem for which a UGC-based hardness is known. Consequently, this work implies that including any valid constraints on up to $\exp(\Omega(\log\log\ n)^{1/4})$ vectors to natural semidefinite program, does not improve the approximation ratio for any problem in the following classes: constraint satisfaction problems, ordering constraint satisfaction problems and metric labeling problems over constant-size metrics.

We obtain similar SDP integrality gaps for Balanced Separator, building on [11]. We also exhibit, for explicit constants $\gamma, \delta > 0$, an n-point negative-type metric which requires distortion $\Omega(\log\log n)^{\gamma}$ to embed into ℓ_1, although all its subsets of size $\exp(\Omega(\log\log\ n)^{\delta})$ embed isometrically into ℓ_1.

Keywords-semidefinite programming, approximation algorithms, unique games conjecture, hardness of approximation, SDP hierarchies, Sherali–Adams hierarchy, integrality gap construction

1. Introduction

Unique Games (UG) is a constraint satisfaction problem where the input consists of a constraint graph G, a label set $[q]$, and a bijection $\pi_{uv}\colon [q] \to [q]$ for each edge $e = (u, v) \in E(G)$. The objective is to find a labeling of the vertices in G so as to maximize the number of edges that are satisfied. Here an edge $e = (u, v)$ is said to be satisfied by a labeling if u is assigned a label ℓ_u and v is assigned a label ℓ_v such that $\pi_{uv}(\ell_u) = \ell_v$. The Unique Games Conjecture (UGC) of [14] asserts that for arbitrarily small constants $\eta, \delta > 0$, with a sufficiently large label set $[q]$, it is NP-hard to decide whether there is a labeling that satisfies $1 - \eta$ fraction of the edges or, no labeling satisfies more than δ fraction of the edges.

Over the last few years, the Unique Games Conjecture has fueled many of the major developments in hardness of approximation. Starting with the work of Khot [14] on Min-2SAT-Deletion, hardness of approximation results for several fundamental problems like Max Cut [15], Vertex Cover [16], non-uniform Sparsest Cut [9], [18] and Max-2-SAT [5] have been obtained assuming the Unique Games Conjecture. In more recent work [21], [13], [19], [23], assuming UGC, approximability of large classes of problems have been determined. Specifically, the work on UGC based hardness results has demonstrated the following (in a precise sense):

For every constraint satisfaction problem, ordering constraint satisfaction problem, metric labeling problem over constant size metric space, the following holds: Assuming UGC, it is NP-hard to approximate to a ratio better than the integrality gap of an simple generic semidefinite relaxation $\mathrm{SDP}_{\mathrm{gen}}$ (for its definition see e.g. [21] or [22]).

Irrespective of the truth of UGC, it is now clear that UGC precisely identifies an algorithmic barrier reached by existing work on approximation algorithms. A natural question that arises is whether stronger semidefinite programming relaxations are sufficient to breach this barrier and disprove the UGC? Or does disproving UGC warrant the use of a new technique different from semidefinite programming?

Unfortunately, progress towards answering this compelling question has been slow and difficult. In the influential paper of Khot–Vishnoi [18], the authors construct an integrality gap instance for a simple SDP relaxation of Unique Games. To the best of our knowledge, this is the sole SDP gap construction for unique games that appears in literature. On one hand, this leaves out the possibility that strong SDPs disprove UGC. More alarmingly, except in a few cases, most UGC based hardness results could possibly be falsified using a strong SDP relaxation. Except for Vertex Cover [12], and k-CSPs [25], [31], in all other cases, there are no strong SDP gaps supporting a UG hardness result.

Obtaining strong SDP gaps that support a UGC based hardness result has been a difficult exercise. In fact, the work of [18] stemmed out of an effort in this direction for the Sparsest Cut problem. Specifically, the Goemans–Linial conjecture regarding embeddability of L_2^2 metrics in to L_1 was refuted in [18] by constructing a SDP gap supporting the UGC based hardness for Sparsest Cut.

The following possibility is entirely consistent with the existing literature: Even for the Max Cut problem which is fairly well studied [15], [20], including an extra inequality on every set of 5 variables in to the standard semidefinite program yields a better approximation, thus disproving UGC.

*Work done while the author was at University of Washington, and visiting Carnegie Mellon University. Supported by NSF grant CCF-0343672.
†Research supported by NSF grants 0830673, 0832797, 528414.

1.1. Results

Our main result is an integrality gap for certain strong SDP relaxations of Unique Games. We consider two hierarchies of SDP relaxations denoted by $\{\mathsf{LH}_R\}_{R\in\mathbb{N}}$ and $\{\mathsf{SA}_R\}_{R\in\mathbb{N}}$. The R^{th} level relaxation LH_R consists of the following: 1) SDP vectors for every vertex of the unique game, 2) All valid constraints on vectors corresponding to at most R vertices. Equivalently, the LH_R relaxation consists of SDP vectors and local distributions μ_S over integral assignments to sets S of at most R variables, such that the second moments of local distributions μ_S match the corresponding inner products of SDP vectors.

The SA_R relaxation is a strengthening of LH_R with the additional constraint that for two sets S, T of size at most R, the corresponding local distribution over integral assignments μ_S, μ_T must have the same marginal distribution over $S \cap T$. The SA_R relaxation corresponds to simple SDP relaxation strengthened by R^{th} round of Sherali-Adams hierarchy [28]. Let $\mathsf{LH}_R(\Upsilon)$ and $\mathsf{SA}_R(\Upsilon)$ denote the optimum value of the corresponding SDP relaxations on the instance Υ. Further, let $\text{opt}(\Upsilon)$ denote the value of the optimum labeling for Υ. For the LH and SA hierarchies, we show:

Theorem 1.1: For all constants $\eta, \delta > 0$, there exists a Unique Games instance Υ on N vertices such that $\mathsf{LH}_R(\Upsilon) \geqslant 1 - \eta$ and $\text{opt}(\Upsilon) \leqslant \delta$ for $R = \Omega(\exp((\log\log N)^{1/4}))$.

Theorem 1.2: For all constants $\eta, \delta > 0$, there exists a Unique games instance Υ on N vertices such that $\mathsf{SA}_R(\Upsilon) \geqslant 1 - \eta$ and $\text{opt}(\Upsilon) \leqslant \delta$ for $R = \Omega((\log\log N)^{1/4})$.

Demonstrated for the first time in [18], and used in numerous later works [8], [27], [31], [21], [13], [19], it is by now well known that integrality gaps can be composed with hardness reductions. In particular, given a reduction Φ from unique games to a certain problem Λ, on starting the reduction with a integrality gap instance Υ for unique games, the resulting instance $\Phi(\Upsilon)$ is a corresponding integrality gap for Λ. Composing the integrality gap instance for LH_R or SA_R relaxation of unique games, along with UG reductions in [15], [5], [21], [13], [19], [23], one can obtain integrality gaps for LH_R and SA_R relaxations of several combinatorial optimization problems. For the sake of succinctness, we will state the following general theorem:

Theorem 1.3: Let Λ denote a problem in one of the following classes:

- Generalized Constraint Satisfaction Problem: a generalization of CSPs permitting bounded payoff functions (positive or negative), instead of predicates [21, Definition 3.1]
- Ordering Constraint Satisfaction Problem: a class of problems containing Maximum Acyclic Subgraph, Betweenness [10], with predicates/bounded payoff functions on orderings of elements.

Let SDP_{gen} denote the SDP relaxation that yields the optimal approximation ratio for Λ under UGC. Then the following holds: Given an instance Υ of the problem Λ, with $\text{SDP}_{\text{gen}}(\Upsilon) \geqslant c$ and $\text{opt}(\Upsilon) \leqslant s$, for every constant $\eta > 0$, there exists an instance Υ_η over N variables such that

- $\mathsf{LH}_R(\Upsilon_\eta) \geqslant c - \eta$ and $\text{opt}(\Upsilon_\eta) \leqslant s + \eta$ for $R = \Omega_\eta(\exp((\log\log N)^{1/4}))$,
- $\mathsf{SA}_R(\Upsilon_\eta) \geqslant c - \eta$ and $\text{opt}(\Upsilon_\eta) \leqslant s + \eta$ for $R = \Omega_\eta((\log\log N)^{1/4})$.

The classes of problems for which the above result holds include Max Cut [15], Max 2-Sat [5], Grothendieck Problem [23] k-way Cut [19] and Maximum Acyclic Subgraph [13]. Notable exceptions that do not directly fall under this framework (but are still UG-hard) are Vertex Cover and Sparsest Cut.

Reductions from Unique Games to Sparsest Cut have been exhibited in [18] and [9]. With the integrality gap for LH_R relaxation of Unique Games (Theorem 1.1), these reductions imply a corresponding LH_R integrality gap for Sparsest Cut. Viewed as a metric space, the SDP vectors of the integrality gap instance yield the following result,

Theorem 1.4: For some absolute constants $\gamma, \delta > 0$, there exists an N-point L_2^2 metric that requires distortion at least $\Omega(\log\log N)^{\delta}$ to embed in to L_1, while every set of size at most $O(\exp((\log\log N)^{\gamma}))$ embeds isometrically in to L_1.

The Uniform Sparsest Cut problem is among the many important problems for which no reduction from Unique Games is known. In [11], the techniques of [18] were extended to obtain an integrality gap for Uniform Sparsest Cut for the SDP with triangle inequalities. Roughly speaking, the SDP gap construction in [11] consists of the hypercube with its vertices identified by certain symmetries such as cyclic shift of the coordinates. Using a similar construction, we obtain the following SDP integrality gap for the Balanced Separator problem,

Theorem 1.5: For some absolute constants $\gamma, \delta > 0$, there exists an instance G on N vertices of Balanced Separator such that the ratio $\text{opt}(G)/\mathsf{LH}_R(G) \geqslant \Omega(\log\log N)^{\delta}$ for $R = O(\log\log N)^{\gamma}$.

1.2. Related Work

Considerable progress has been made in understanding the limits of linear programming hierarchies [2], [3], [29], [27], [8]. Lower bound results of this nature are fewer in the case of semidefinite programs. A $\Omega(n)$-round lower bound for proving unsatisfiability of random 3-SAT formulae was obtained for the Lovász–Schrijver SDP hierarchy (LS_+) in [6], [1]. In turn, this lower bound leads to $\Omega(n)$-round LS_+ integrality gaps for problems like Set Cover, Hypergraph Vertex Cover, where a matching NP-hardness result is known. Similarly, the $7/6$-integrality gap for $\Omega(n)$ rounds of LS_+ in [26] falls in a regime where a matching NP-hardness result is known. A significant exception is the result of Georgiou et al. [12] that exhibited a $2 - \varepsilon$-integrality gap for $\Omega(\sqrt{\log n/\log\log n})$ rounds of LS_+ hierarchy. In a beautiful work, Schoenebeck [25] exhibited integrality gaps for Random 3-SAT in the Lasserre

SDP hierarchy. Building on this result, Tulsiani [31] obtained an $\Omega(n)$-round integrality gap matching the corresponding UG hardness for k-CSP [24].

Independent of our work, Khot and Saket [17] obtained integrality gap constructions for the MAX CUT and the SPARSEST CUT problems against the SA_R hierarchy. While the integrality gap instances in [17] are nearly identical to the corresponding ones in this work, the SDP solutions are considerably different. In particular, this work develops additional technical machinery to obtain SDP gaps for UNIQUE GAMES, which in turn yields gaps for a large array of other problems.

1.3. Overview of the Technique

In this section, we will present a brief overview of our techniques and a road map for the rest of the paper.

The overall strategy in this work to construct SDP integrality gaps is along the lines of Khot–Vishnoi [18]. Let us suppose we wish to construct an SDP integrality gap for a problem Λ (say MAX CUT). Let Φ_Λ be a reduction from UNIQUE GAMES to the problem Λ. The idea is to start from an SDP integrality gap Υ for UNIQUE GAMES, and then execute the reduction Φ_Λ on the instance Υ, to obtain the SDP gap instance $\Phi_\Lambda(\Upsilon)$. Surprisingly, as demonstrated in [18], the SDP vector solution for Υ can be transformed through the reduction to obtain the SDP solution for $\Phi_\Lambda(\Upsilon)$.

Although this technique has been used extensively in numerous works [8], [27], [31], [21], [13], [19] since [18], there is a crucial distinction between [18] and later works. In all other works, starting with an SDP gap Υ for UNIQUE GAMES, one obtains an integrality gap for an SDP relaxation that is no stronger. For instance, starting with a integrality gap for 10-rounds of a SDP hierarchy, the resulting SDP gap instance satisfies at most 10 rounds of the same hierarchy.

The surprising aspect of [18], is that it harnesses the UG reduction Φ_Λ to obtain an integrality gap for a "stronger" SDP relaxation than the one which it stared with. Specifically, starting with an integrality gap Υ for a simple SDP relaxation of UNIQUE GAMES, [18] obtain an SDP gap for MAX CUT which obeys all valid constraints on 3 variables. The proof of this fact (the triangle inequality) is perhaps the most technical and least understood aspect about [18]. One of the main contributions of this work is to conceptualize and simplify this aspect of [18]. Armed with the understanding of [18], we then develop the requisite machinery to extend it to a strong SDP integrality gap for UNIQUE GAMES.

To obtain strong SDP gaps for UNIQUE GAMES, we will apply the above strategy on a reduction from UNIQUE GAMES to $\mathsf{E2Lin}_q$ due to Khot et al. [15]. Note that $\mathsf{E2Lin}_q$ is a special case of UNIQUE GAMES. In this article, an integrality gap instance for the basic semidefinite program of UNIQUE GAMES, will be referred to as a *weak gap instance* (see Definition 3.1). Formally, we show the following reduction from a weak gap instance for UNIQUE GAMES over a large alphabet to an integrality gap for a strong SDP relaxation of $\mathsf{E2Lin}_q$.

Theorem 1.6: For a prime q and $\gamma > 0$, let $\Phi_{\gamma,q}$ denote the reduction from UNIQUE GAMES to $\mathsf{E2Lin}_q$ with completeness $1-\gamma$. Let Υ be a weak $(1-\eta,\delta)$-gap instance of UNIQUE GAMES. Then, for every q of order unity, there exists an SDP solution for the $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Upsilon)$ such that

- the SDP solution is feasible for LH_R for $R = 2^{O(1/\eta^{1/4})}$,
- the SDP solution is feasible for SA_R for $R = O(\eta^{1/4})$,
- the SDP solution has value $1-\gamma-o_{\eta,\delta}(1)$ for $\Phi_{\gamma,q}(\Upsilon)$.

In particular, the $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Upsilon)$ is a $(1-\gamma-o_{\eta,\delta}(1), q^{-\eta/2}+o_{\eta,\delta}(1)))$ integrality gap instance for the relaxation LH_R for $R = 2^{\Omega(1/\eta^{1/4})}$. Further, $\Phi_{\gamma,q}(\Upsilon)$ is a $(1-\gamma-o_{\eta,\delta}(1), q^{-\eta/2}+o_{\eta,\delta}(1))$ integrality gap instance for the relaxation SA_R for $R = \Omega(1/\eta^{1/4})$.

Combining the weak gap for UNIQUE GAMES constructed in [18], with the above theorem yields Theorem 1.1 and Theorem 1.2. As already pointed out, by now it is fairly straightforward to translate an R-round integrality gap for UNIQUE GAMES to an R-round integrality gap for other problems (using known UG-hardness reductions). Hence, Theorem 1.3 is a consequence of Theorem 1.1 and Theorem 1.2.

1.3.1. Example of MAX CUT*:* For the sake of exposition, we will describe how our techniques can be used to construct an SDP integrality gap for MAX CUT. Extending the approach to obtain UNIQUE GAMES integrality gap mentioned above, requires additional ideas such as tensoring for $\mathbb{F}_q$-vectors, and a more involved construction of SDP vectors.

Let Υ be a (weak) SDP integrality gap instance for UNIQUE GAMES with alphabet $[n]$ (e.g., the one in [18]). We obtain a MAX CUT instance G from Υ by executing the UG hardness reduction for MAX CUT in [15]. Suppose the reduction has completeness c and soundness s (with $c/s = 0.878$, say). We claim that this MAX CUT instance G is essentially a (c,s)-integrality gap for a strong SDP relaxation of MAX CUT. The challenge is to exhibit a strong SDP solution of value roughly c. We start from an optimal SDP solution $\mathcal{B}$ for Υ. This SDP solution consists of "clouds" $B \in \mathcal{B}$ of n orthonormal vector. Each cloud in $\mathcal{B}$ corresponds to a vertex in Υ and each vector in the cloud corresponds to a label for that vertex. The clouds satisfy the following properties (see §3 for further justification of these properties):

1.) Matching Property: For every two clouds $A, B \in \mathcal{B}$, there exists a matching $\pi_{B\leftarrow A}\colon A \to B$ along which the inner product of vectors between A and B is maximized. Specifically, if $\rho(A,B) := \max_{a\in A, b\in B}\langle a, b\rangle$, then for each vector a in A, we have $\langle a, \pi_{B\leftarrow A}(a)\rangle = \rho(A,B)$. It is not hard to see that this matching is unique if $\rho(A,B)^2 > 1/2$ (using orthonormality).

2.) Integrality Property: The vectors in the SDP solution $\mathcal{B}$ are normalized $\{\pm 1\}$-vectors. Specifically, $\bigcup \mathcal{B} \subseteq \{1/\sqrt{d}, -1/\sqrt{d}\}^d$ for some $d \in \mathbb{N}$.

We can now describe how to construct a good SDP solution for G from the SDP solution $\mathcal{B}$ for Υ.

Global vector solution. The vectors that we assign to the vertices of G resemble (somewhat simpler in this work) the vectors in [18]. According to the reduction, the vertices of Υ are of the form (B, x) where B is a cloud and $x \in \{\pm 1\}^B$ is a $\{\pm 1\}$-assignment to the elements of the clouds. To a vertex (B, x), we can essentially assign the vector

$$\boldsymbol{V}^{B,x} := \tfrac{1}{\sqrt{n}} \textstyle\sum_{b \in B} x_b b^{\otimes t} .$$

(In the final construction we use different, more complicated vectors, because it allows us to improve the parameters significantly.) A simple calculation [18] shows that for constant $t \in \mathbb{N}$, this vector solution has SDP-objective value $\approx c$. The interesting part is to show that vector solution is feasible for the SDP. The crucial property of this SDP solution is that the inner product of two vectors $\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$ is essentially determined by terms corresponding to the matching $\pi_{B \leftarrow A}$,

Tensoring Lemma:

$$\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle = \tfrac{1}{n} \textstyle\sum_{a \in A} \langle x_a a^{\otimes t}, y_{\pi_{B \leftarrow A}(a)} \pi_{B \leftarrow A}(a)^{\otimes t} \rangle \quad \pm 2(1/2)^{t/2} .$$

In this sense, the tensoring allows the vectors to find the best matching between their clouds.

Local integral distributions. Unlike [18], instead of directly showing that the inner products of the vectors satisfy the desired inequalities (e.g., the L_2^2-triangle inequality), we exhibit a collection of distributions over $\{\pm 1\}$-assignments to subsets of vertices, such that the second-moments of the distributions match the inner products of the vectors. Specifically, we want distributions $\{\mu_S : \{\pm 1\}^S \to \mathbb{R}_+\}_{S \subseteq V(G), |S| \leqslant R}$ such that $\mathbb{E}_{Z \sim \mu_S}[Z^{A,x} Z^{B,y}] = \langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$. (Here, R is a parameter that we will choose later.) The existence of an integral distribution matching the inner products shows that the vectors satisfy all valid inequalities involving up to R vertices, including the triangle inequality (for $R \geqslant 3$).

As a first step, let us see how to construct a distribution over assignments to just two vertices $S = \{(A, x), (B, y)\}$ that roughly matches the inner product above: Pick a random vector $a \in A$. Set $b := \pi_{B \leftarrow A}(a)$. Pick a random coordinate $r \in [d^t]$. Output the assignment $Z^{A,x} = \operatorname{sign}(x_a a^{\otimes t}(r))$, $Z^{B,y} = \operatorname{sign}(y_b b^{\otimes t}(r))$. We see that the second-moment $\mathbb{E}\, Z^{A,x} Z^{B,y}$ is exactly equal to the sum that dominates the inner product $\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$ (from the Tensoring Lemma). Here, we use the integrality property of the vectors in the clouds.

For larger sets S of vertices, the challenge is to find a distribution over labelings for the clouds in S that are consistent with the matchings $\pi_{B \leftarrow A}$. Another illustrative case is when every pair A, B of clouds in S is highly correlated (say $\rho(A, B) > 0.99$). In this case, the matchings are consistent in the sense that for any three clouds A, B, C in S, we have $\pi_{A \leftarrow C} = \pi_{A \leftarrow B} \circ \pi_{B \leftarrow C}$. Hence, we can easily find a distribution over labelings $\{\ell_A\}_{A \in S}$ such that $\Pr\{\ell_A = a\} = 1/n$ for every $a \in A$ and $\Pr\{\ell_A = \pi_{A \leftarrow B}(\ell_B)\} = 1$.

Geometric partitioning schemes. In order to deal with general sets S of size up to R (where some clouds might not be highly correlated), we use geometric partitioning schemes. The goal is that every cluster of clouds in S is pairwise highly correlated. In that case, we will perform the above sampling procedure independently for each cluster.

One of our partitioning schemes consists of a distribution over partitions P of the set of clouds $\mathcal{B}$. The distribution has the property that any two clouds A, B with $\rho(A, B) < 0.99$ fall into the same cluster with probability close to zero, say δ. And, on the other hand, any two clouds A, B with $\rho(A, B)^t > \delta$ fall in the same cluster with probability $1 - \delta$ (Here, t will depend on δ). Hence for any fixed set S of size at most R, if we choose a partition according to this distribution, then all clusters of S are pairwise highly correlated with probability at least $1 - R^2\delta$. For two clouds A, B, if $\rho(A, B)^t \leqslant \delta$, then the inner product $\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle \leqslant \delta \pm 2(1/2)^{t/2}$ and an arbitrary choice of labellings would result in a correlation $\mathbb{E}\, Z^{A,x} Z^{B,y} \approx 0$, thus matching the inner product within error $O(\delta)$. On the other hand, if $\rho(A, B)^t \geqslant \delta$, with probability at least $1 - \delta - R^2\delta$, A, B fall in the same cluster and the cluster containing them is pairwise highly correlated. Therefore the resulting moment $\mathbb{E}\, Z^{A,x} Z^{B,y}$ matches the inner product within an error $O(\delta + R^2\delta)$.

Robustness lemma for strong SDP relaxations. Finally, we exhibit a procedure to modify the local distributions and the SDP vectors that are approximately consistent, to a perfectly feasible solution. In other words, we show a *robustness* property for the SDP hierarchies we consider in that, approximately feasible solutions to these hierarchies can be converted in to feasible solutions with a small loss in the objective value. To illustrate the idea, consider a set of unit vectors $\{v_i\}_{i=1}^n$ that satisfy all triangle inequalities up to an additive error of ε, i.e., $\|v_i - v_j\|^2 + \|v_j - v_k\|^2 - \|v_i - v_k\|^2 \geqslant -\varepsilon$. Let $\{w_i\}_{i=1}^n$ be a set of unit vectors that are orthogonal to all the vectors $\{v_i\}_{i=1}^n$, and to each other. Notice that the vectors $\{w_i\}$ have a *slack* on every triangle inequality, i.e, $\|w_i - w_j\|^2 + \|w_j - w_k\|^2 - \|w_i - w_k\|^2 \geqslant 2$. Define a new SDP solution $\{u_i\}_{i=1}^n$ as $u_i = \sqrt{1 - \varepsilon} \cdot v_i + \sqrt{\varepsilon} \cdot w_i$. The slack in the triangle inequality for $\{w_i\}_{i=1}^n$ would compensate for the slight infeasibility of the original vectors $\{v_i\}_{i=1}^n$ and the resulting vectors satisfy the triangle inequality.

2. Preliminaries

2.1. Notation

For finite sets Σ and S, we denote by Σ^S the set of functions from S to Σ. We call such a function sometimes a *Σ-assignment to S*. For a finite set X, a *distribution on X* is a function $\mu \colon X \to \mathbb{R}$ such that $\sum_{x \in X} \mu(x) = 1$ and $\mu(x) \geqslant 0$ for all $x \in X$. We let $\Delta(X)$ denote the set of distributions on X. We sometimes refer to a member of $\Delta(\Sigma^S)$ as a *distribution over Σ-assignments to S*. For a function $f \colon \Sigma^S \to \mathbb{R}$ and a distribution μ on Σ^S, we denote the *expectation of f with* *respect to* μ by $\mathrm{E}_{x\sim\mu} f(x) \overset{\text{def}}{=} \sum_{x\in\Sigma^S} \mu(x) f(x)$. For an event $\mathcal{E} \subseteq \Sigma^S$ and a distribution μ on Σ^S, we denote the *probability of* $\mathcal{E}$ *with respect to* μ by $\Pr_{x\sim\mu} \mathcal{E} \overset{\text{def}}{=} \mathrm{E}_{x\sim\mu} \mathbf{1}_{\mathcal{E}}(x) = \sum_{x\in\mathcal{E}} \mu(x)$. Here, $\mathbf{1}_{\mathcal{E}}$ denotes the 0/1-indicator function of the set $\mathcal{E}$. For a subset $T \subseteq S$, let us define the *marginal distribution* $\mathrm{margin}_T\, \mu \colon \Sigma^T \to \mathbb{R}_+$ as $\mathrm{margin}_T\, \mu(x) \overset{\text{def}}{=} \sum_{y\in\Sigma^{S\setminus T}} \mu(x, y)$. Here, (x, y) denote the Σ-assignment to S that agrees with x on T and with y on $S \setminus T$.

2.2. *SDP Hierarchies*

In this section, we present formal definitions of the LH_R and SA_R relaxations. Let Υ be a problem instance over a set of variables $\mathcal{V}$. An *SDP solution* for the instance Υ consists of the following:

1) A *collection of (local) distributions* $\{\mu_S\}_{S\subseteq\mathcal{V},|S|\leqslant R}$, where $\mu_S : \mathbb{F}_q^S \to \mathbb{R}_+$ is a distribution over $\mathbb{F}_q$-assignments to S, that is, $\mu_S \in \Delta(\mathbb{F}_q^S)$.
2) A *(global) vector solution* $\{\boldsymbol{v}_{i,a}\}_{i\in\mathcal{V},a\in\mathbb{F}_q}$, where $\boldsymbol{v}_{i,a} \in \mathbb{R}^d$ for every $i \in \mathcal{V}$ and $a \in \mathbb{F}_q$.

The intention for the local distributions $\{\mu_S\}$ is that they arise as the marginal distribution of a global distribution $\mu\colon \mathbb{F}_q^{\mathcal{V}} \to \mathbb{R}_+$ over $\mathbb{F}_q$-assignments to the variables $\mathcal{V}$. The intention for the vector solution $\{\boldsymbol{v}_{i,a}\}$ is that all vectors have only $\{0, 1\}$-coordinates and that for every i and every coordinate r, exactly one of the vectors $\boldsymbol{v}_{i,1}, \ldots, \boldsymbol{v}_{i,q}$ has a 1 in the r^{th} coordinate. We assume that the objective of the instance Υ can be expressed as a linear function in the local distributions, say $\sum_{S\subseteq\mathcal{V},|S|\leqslant R} \sum_{x\in\mathbb{F}_q^S} c_{S,x}\, \mu_S(x)$.

LH_R-Relaxation.

$$\text{maximize} \quad \sum_{S\subseteq\mathcal{V},|S|\leqslant R} \sum_{x\in\mathbb{F}_q^S} c_{S,x}\, \mu_S(x) \tag{1}$$

subject to

$$\langle \boldsymbol{v}_{i,a}, \boldsymbol{v}_{j,b}\rangle = \Pr_{x\sim\mu_S}\left\{x_i = a,\ x_j = b\right\} \quad \left(\begin{smallmatrix} S\subseteq\mathcal{V},\ |S|\leqslant R, \\ i,j\in S,\ \ a,b\in[q]\end{smallmatrix}\right), \tag{2}$$

$$\langle \boldsymbol{v}_{i,a}, \boldsymbol{v}_0\rangle = \Pr_{x\sim\mu_S}\{x_i = a\} \qquad \left(\begin{smallmatrix} S\subseteq\mathcal{V},\ |S|\leqslant R, \\ i\in S,\ \ a\in[q]\end{smallmatrix}\right). \tag{3}$$

Here, $\boldsymbol{v}_0 \in \mathbb{R}^d$ is an arbitrary fixed unit vector. We say that an SDP solution $\{\mu_S\}$, $\{\boldsymbol{v}_{i,a}\}$ for Υ is *feasible for* LH_R if it satisfies the constraints (2)–(3). We denote by $\mathsf{LH}_R(\Upsilon)$ the value of an optimal solution to this relaxation.

Remark 2.1: A feasible solution for LH_R satisfies all valid linear inequalities on the inner products of the vectors corresponding to up to R variables in $\mathcal{V}$.

The SA_R relaxation is a stronger SDP relaxation with an additional consistency requirement between the local distributions μ_S.

SA_R-Relaxation:

$$\text{maximize} \quad \sum_{S\subseteq\mathcal{V},|S|\leqslant R} \sum_{x\in\mathbb{F}_q^S} c_{S,x}\, \mu_S(x) \tag{4}$$

subject to

$$\langle \boldsymbol{v}_{i,a}, \boldsymbol{v}_{j,b}\rangle = \Pr_{x\sim\mu_S}\left\{x_i = a,\ x_j = b\right\} \quad \left(\begin{smallmatrix} S\subseteq\mathcal{V},\ |S|\leqslant R, \\ i,j\in S,\ \ a,b\in[q]\end{smallmatrix}\right), \tag{5}$$

$$\langle \boldsymbol{v}_{i,a}, \boldsymbol{v}_0\rangle = \Pr_{x\sim\mu_S}\{x_i = a\} \qquad \left(\begin{smallmatrix} S\subseteq\mathcal{V},\ |S|\leqslant R, \\ i\in S,\ \ a\in[q]\end{smallmatrix}\right), \tag{6}$$

$$\left\|\mathrm{margin}_{A\cap B}\, \mu_A - \mathrm{margin}_{A\cap B}\, \mu_B\right\|_1 = 0 \quad \left(\begin{smallmatrix} A,B\subseteq\mathcal{V}, \\ |A|,|B|\leqslant R\end{smallmatrix}\right). \tag{7}$$

We say an SDP solution $\{\mu_S\}, \{\boldsymbol{v}_{i,a}\}$ is *feasible for* SA_R if it satisfies the constraints (5)–(7). We denote by $\mathsf{SA}_R(\Upsilon)$ the value of an optimal solution to this relaxation.

2.3. *Smoothing*

Here we show that approximately feasible SDP solutions can be made into feasible SDP solutions without losing too much in the objective value. Instead of directly bounding the loss in objective value we give bounds on the "change" of the local distribution in L_1-norm. In our applications, this bound on the L_1-norm will imply a similar bound on the loss in objective value (the reason being that in our applications the objective function is Lipschitz with respect to the L_1-norm).

Definition 2.2: An SDP solution $\{\boldsymbol{v}_{i,a}\}_{i\in\mathcal{V},a\in\mathbb{F}_q}$, $\{\mu_S\}_{S\subseteq\mathcal{V},|S|\leqslant R}$ is said to be *ε-infeasible for* LH_R *(or* SA_R*)* if it satisfies all the constraints of the program up to an additive error of ε.

Theorem 2.3: Given an ε-infeasible solution $\{\boldsymbol{v}_{i,a}\}_{i\in\mathcal{V},a\in\mathbb{F}_q}$, $\{\mu_S\}_{S\subseteq\mathcal{V},|S|\leqslant R}$ to the LH_R relaxation, there exists a feasible solution $\{\boldsymbol{v}'_{i,a}\}, \{\mu'_S\}_{S\subseteq\mathcal{V},|S|\leqslant R}$ for LH_R such that for all subsets $S \subseteq \mathcal{V}, |S| \leqslant R$, $\|\mu_S - \mu'_S\|_1 \leqslant \mathrm{poly}(q) \cdot R^2 \varepsilon$.

Theorem 2.4: Given an ε-infeasible solution $\{\boldsymbol{v}_{i,a}\}_{i\in\mathcal{V},a\in\mathbb{F}_q}$, $\{\mu_S\}_{S\subseteq\mathcal{V},|S|\leqslant R}$ to the SA_R relaxation, there exists a feasible solution $\{\boldsymbol{v}'_{i,a}\}, \{\mu'_S\}_{S\subseteq\mathcal{V},|S|\leqslant R}$ for SA_R such that for all subsets $S \subseteq \mathcal{V}, |S| \leqslant R$, $\|\mu_S - \mu'_S\|_1 \leqslant \mathrm{poly}(q) \cdot \varepsilon \cdot q^R$.

The proofs of the above theorems are nearly identical to the proof of Theorem 4.6 in [22], and the details are deferred to the full version.

2.4. *Integral Vectors*

Let q be a prime. In the following, we develop a natural generalization of $\{\pm 1\}$-vectors (see §1.3.1) to a q-ary alphabet and a suitable notion of tensor products of such vectors.

Definition 2.5: A $\mathbb{F}_q$-integral vector $v\colon \mathcal{R} \to \mathbb{F}_q$ is a function from a probability space $\mathcal{R}$ to $\mathbb{F}_q$. For a $\mathbb{F}_q$-integral vector $v\colon \mathcal{R} \to \mathbb{F}_q$, its symmetrization $\tilde{v}\colon \mathcal{R} \times \mathbb{F}_q^* \to \mathbb{F}_q$ is defined by $\tilde{v}(r, t) = t \cdot v(r)$.

Given a map $f\colon \mathbb{F}_q \to \mathbb{C}^d$, we denote by $f(v) := f \circ v$ the composition of functions f and v. The following function will be relevant to us:

Define $\psi\colon \mathbb{F}_q \to \mathbb{R}^{q-1}$ as $\psi(i) := \psi_i$, where $\psi_0, \psi_1, \ldots, \psi_{q-1}$ denote the corners of the q-ary simplex in $\mathbb{R}^{q-1}$, translated so that the origin is its geometric center. Thus, the function

ψ satisfies,

$$\langle \psi(a), \psi(b)\rangle = \begin{cases} 1 & \text{if } a = b\,, \\ -\frac{1}{q-1} & \text{if } a \neq b\,. \end{cases}$$

Remark 2.6: The following notions are equivalent: collection of $\mathbb{F}_q$-valued functions on some probability space $\mathcal{R} \iff$ collection of jointly-distributed, $\mathbb{F}_q$-valued random variables $\iff$ distribution over $\mathbb{F}_q$-assignments.

For the case of $\mathbb{F}_q$-integral vector, the tensor product operation is to be defined carefully, in order to mimic the properties of the traditional tensor product. We will use the following definition for the tensor operation $\otimes_q$.

Definition 2.7: Given two $\mathbb{F}_q$-valued functions $u\colon \mathcal{R} \to \mathbb{F}_q$ and $u'\colon \mathcal{R}' \to \mathbb{F}_q$, define the symmetrized tensor product $u \otimes_q u'\colon (\mathcal{R} \times \mathbb{F}_q^*) \times (\mathcal{R}' \times \mathbb{F}_q^*) \to \mathbb{F}_q$ as $(u \otimes_q u')\,(r, t, r', t') \stackrel{\text{def}}{=} t \cdot u(r) + t' \cdot u'(r')$.

Like the traditional tensor product, the inner products multiply on taking $\otimes_q$ products.

Lemma 2.8: For any $\mathbb{F}_q$-valued functions $u, v\colon \mathcal{R} \to \mathbb{F}_q$ and $u', v'\colon \mathcal{R}' \to \mathbb{F}_q$, $\langle \psi(u \otimes_q u'), \psi(v \otimes_q v')\rangle = \langle \psi(u), \psi(v)\rangle\langle \psi(u'), \psi(v')\rangle$.

We need the following simple technical observation in one of our proofs.

Observation 2.9: Let $u, v\colon \mathcal{R} \to \mathbb{F}_q$ be two "symmetric" $\mathbb{F}_q$-integral vectors,[1] that is, $\Pr_r\{u(r) - v(r) = a\} = \Pr_r\{u(r) - v(r) = b\}$ for all $a, b \in \mathbb{F}_q^*$. Then, for all $a, b \in \mathbb{F}_q$, we have $\mathrm{E}_r\langle \psi(a + u(r)), \psi(b + v(r))\rangle = \langle a \otimes_q u, b \otimes_q v\rangle$.

For notational convenience, we will abbreviate $\otimes_q$ by $\otimes$ whenever the meaning is clear from the context.

The following transformation inspired by the rounding scheme for Unique Games in Charikar et al. [7] [2] yields a way to generate $\mathbb{F}_q$-integral vectors from arbitrary vectors.

Observation 2.10: Define the function $\zeta\colon \mathcal{G}^q \to \mathbb{F}_q$ on the Gaussian domain as follows: $\zeta(x_1, \ldots, x_q) = \operatorname{argmax}_{i \in [q]} x_i$. Given a family of unit vectors $\{v_1, \ldots, v_n\} \in \mathbb{R}^d$, define the set of $\mathbb{F}_q$-valued functions $v_1^*, \ldots, v_n^*\colon \mathcal{R} \to \mathbb{F}_q$ with $\mathcal{R} = (\mathcal{G}^d)^q$ —the Gaussian space of appropriate dimension— as follows: $v_i^*(g_1, \ldots, g_q) = \zeta(\langle v_i, g_1\rangle, \ldots, \langle v_i, g_q\rangle)$ for $g_1, \ldots, g_q \in (\mathcal{G}^d)^q$. The $\mathbb{F}_q$-valued functions $\{v_i^*\}$ satisfy $\langle \psi(u^*), \psi(v^*)\rangle = 0$ whenever $\langle u, v\rangle = 0$, and $\langle \psi(u^*), \psi(v^*)\rangle = 1 - f(\varepsilon, q) = 1 - O(\sqrt{\varepsilon \log q})$ whenever $\langle u, v\rangle = 1 - \varepsilon$.

Proof: To see the first assertion, observe that if $\langle u, v\rangle = 0$, then the sets of random variables $\{\langle u, g_1\rangle, \ldots, \langle u, g_q\rangle\}$ and $\{\langle v, g_1\rangle, \ldots, \langle v, g_q\rangle\}$ are completely independent of each other. Therefore, $\langle \psi(u^*), \psi(v^*)\rangle = \mathrm{E}_{r \in \mathcal{G}^{dq}}[\psi(u^*(r))] \cdot \mathrm{E}_{r \in \mathcal{G}^{dq}}[\psi(u^*(r))] = 0$. The second assertion follows from Lemma C.8 in [7]. ■

[1] In our applications, the vectors u and v will be tensor powers. In this case, the symmetry condition is always satisfied.

[2] This observation has been communicated to us by Boaz Barak.

3. Weak Gap Instance for Unique Games

Formally, a *weak gap* instance for Unique games is defined as follows.

Definition 3.1: (Weak SDP solutions and weak gap instances) Let $\Upsilon = (V, E, \{\pi_e\colon [n] \to [n]\}_{e \in E})$. We say a collection $\mathcal{B} = \{B_u\}_{u \in V}$ is a *weak SDP solution of value* $1 - \eta$ for Υ if the following conditions hold:

1) For every vertex $u \in V$, the collection $\mathcal{B}$ contains an ordered set $B_u = \{b_{u,1}, \ldots, b_{u,n}\}$ of n orthonormal vectors in $\mathbb{R}^d$.
2) Any two vectors in $\bigcup \mathcal{B}$ have non-negative inner product and any three vectors in $\bigcup \mathcal{B}$ satisfy the ℓ_2^2-triangle inequality ($\|x - y\|^2 \leqslant \|x - z\|^2 + \|z - y\|^2$).
3) For every pair of vertices $u, v \in V$, the sets B_u and B_v satisfy the following *strong matching property*: There exists n disjoint matchings between B_u, B_v given by bijections $\pi^{(1)}, \ldots, \pi^{(n)}\colon B_u \to B_v$ such that for all $i \in [n], b, b' \in B_u$, we have $\langle b, \pi^{(i)}(b)\rangle = \langle b', \pi^{(i)}(b')\rangle$.
4) For every edge $e = (u, v) \in E$, the vector sets B_u and B_v have significant correlation under the permutation $\pi = \pi_e$. Specifically, $\langle b_{u,\ell}, b_{v,\pi(\ell)}\rangle^2 \geqslant 0.99$ for all $\ell \in [n]$.
5) The collection $\mathcal{B}$ of orthonormal sets is a good SDP solution for Υ, in the sense that

$$\mathop{\mathrm{E}}_{v \in V} \mathop{\mathrm{E}}_{\substack{w, w' \in N(v) \\ \pi = \pi_{w,v},\ \pi' = \pi_{w',v}}} \frac{1}{n} \sum_{\ell \in [n]} \langle b_{w,\pi(\ell)}, b_{w',\pi'(\ell)}\rangle \geqslant 1 - \eta.$$

We say that Υ is a *weak* $(1 - \eta, \delta)$*-gap instance* of Unique Games if Υ has a weak SDP solution of value $1 - \eta$ and no labeling for Υ satisfies more than a δ fraction of the constraints.

Starting with the integrality gap instance Υ for Unique Games constructed in Khot–Vishnoi [18], delete all edges of Υ that contribute less than $\sqrt{3/4}$ to the SDP objective. It is easy to see that the resulting instance Υ' is a *weak gap* instance. Thus, the following is a direct consequence of Theorem 9.2 and Theorem 9.3 in [18].

Observation 3.2: For all $\eta, \delta > 0$, there exists a weak $(1 - \eta, \delta)$-gap instance with $2^{2^{O(\log(1/\delta)/\eta)}}$ vertices.

3.1. $\mathbb{F}_q$-integrality

Here we use $\langle \cdot, \cdot\rangle_\psi := \langle \psi(\cdot), \psi(\cdot)\rangle$ as inner product for $\mathbb{F}_q$-integral vectors. Observation 2.10 implies that without much loss we can assume that a weak SDP solution is $\mathbb{F}_q$*-integral*, that is, all vectors are $\mathbb{F}_q$-integral. Formally,

Lemma 3.3: Let $\Upsilon = (V, E, \{\pi_e\}_{e \in E})$ be a weak $(1 - \eta, \delta)$-gap instance. Then, for every $q \in \mathbb{N}$, we can find a weak $\mathbb{F}_q$-integral SDP solution of value $1 - O(\sqrt{\eta \log q})$ for a Unique Games instance Υ' which is obtained from Υ by deleting $O(\sqrt{\eta \log q})$ edges.

(The proof of the lemma is deferred to the full version.)

3.2. Tensoring

Definition 3.4: For $A, B \in \mathcal{B}$, we denote $\rho(A,B) \stackrel{\text{def}}{=} \max_{a\in A, b\in B} |\langle a, b\rangle|$. We define $\pi_{B\leftarrow A} \colon A \to B$ to be any[3] bijection from A to B such that $|\langle a, \pi_{B\leftarrow A}(a)\rangle| = \rho(A,B)$ for all $a \in A$.

As a direct consequence of the orthogonality of the clouds in $\mathcal{B}$, we have the following fact about the uniqueness of $\pi_{B\leftarrow A}$ for highly correlated clouds $A, B \in \mathcal{B}$.

Fact 3.5: Let $A, B \in \mathcal{B}$. If $\rho(A,B)^2 > 3/4$, then there exists exactly one bijection $\pi \colon A \to B$ such that $|\langle a, \pi(a)\rangle| = \rho(A,B)$ for all $a \in A$. (The constant $3/4$ is not optimal.)

Notice that for an edge $e = (u, v)$ in a *weak gap* instance, and the corresponding clouds B_u, B_v, the above matching is nothing but the permutation $\pi_{u\to v}$. We observe the following consequence of Fact 3.5 and item 4 of Definition 3.1.

Observation 3.6: If $\mathcal{B} = \{B_u\}_{u\in V}$ is a weak SDP solution for $\Upsilon = (V, E, \{\pi_e\}_{e\in E})$, then for any two edges $(w, v), (w', v) \in E$, the two bijections $\pi = \pi^{-1}_{(w',v)} \circ \pi_{(w,v)}$ and $\pi_{B_{w'}\leftarrow B_w}$ (see Def. 3.4) give rise to the same matching between the vector sets B_w and $B_{w'}$, $\pi(i) = j \iff \pi_{B_{w'}\leftarrow B_w}(b_{w,i}) = b_{w',j}$.
Now we state the crucial tensoring lemma, that underlies a large part of the construction.

Lemma 3.7 (Tensoring Lemma): For $t \in \mathbb{N}$ and every pair of clouds $A, B \in \mathcal{B}$,

$$\frac{1}{n} \sum_{\substack{a\in A, b\in B\\ a\neq \pi_{A\leftarrow B}(b)}} |\langle a, b\rangle|^t \leqslant 2 \cdot (3/4)^{t/2} .$$

Proof: By orthogonality, $\sum_{a\in B} \langle a, b\rangle^2 \leqslant 1$ for every $b \in B$. Hence, $\langle a, b\rangle^2 \leqslant 1/2$ for all $a \neq \pi_{A\leftarrow B}(b)$. Thus,

$$\frac{1}{n} \sum_{\substack{a\in A, b\in B\\ a\neq \pi_{A\leftarrow B}(b)}} |\langle a, b\rangle|^t \leqslant (1/2)^{\frac{t-2}{2}} \cdot \frac{1}{n} \sum_{a\in A, b\in B} |\langle a, b\rangle|^2 \leqslant (1/2)^{\frac{t-2}{2}} .$$

(In the lemma, we state the suboptimal bound $2\cdot(3/4)^t$ because it also holds in a more general setting where the vectors are not orthogonal but only near-orthogonal.) ∎
The notation $X = Y \pm Z$ means that $|X - Y| \leqslant Z$.

Corollary 3.8: For an even number $t \in \mathbb{N}$ and every pair of clouds $A, B \in \mathcal{B}$,

$$\langle \frac{1}{\sqrt{n}} \sum_{a\in A} a^{\otimes t}, \frac{1}{\sqrt{n}} \sum_{b\in B} b^{\otimes t}\rangle = \rho(A,B)^t \quad \pm 2 \cdot (3/4)^{t/2} .$$

This fact that the functional $\rho(A,B)^t$ is closely approximated by inner products of averaged-tensored vectors has implicitly been used in [18] and was explicitly noted in [4, Lemma 2.2].

3.3. Local Distributions for Unique Games

First, we recall a few facts that are direct consequences of the (symmetrized) ℓ_2^2-triangle inequality.

Fact 3.9: Let $a, b, c \in \bigcup \mathcal{B}$ with $|\langle a, b\rangle| = 1 - \eta_{ab}$ and $|\langle b, c\rangle| = 1 - \eta_{bc}$. Then, $|\langle a, c\rangle| \geqslant 1 - \eta_{ab} - \eta_{bc}$.

Fact 3.10: Let $A, B, C \in \mathcal{B}$ with $\rho(A,B) = 1 - \eta_{AB}$ and $\rho(B,C) = 1 - \eta_{BC}$. Then, $\rho(A,C) \geqslant 1 - \eta_{AB} - \eta_{BC}$.

The construction in the proof of the next lemma is closely related to propagation-style UG algorithms [30], [4].

Definition 3.11: A set $\mathcal{S} \subseteq \mathcal{B}$ is *consistent* if for all $A, B \in \mathcal{S}$, $\rho(A,B) \geqslant 1 - 1/16$.

Lemma 3.12: If $\mathcal{S} \subseteq \mathcal{B}$ is consistent, there exists bijections $\{\pi_A \colon [n] \to A\}_{A\in\mathcal{S}}$ such that for all $A, B \in \mathcal{S}$, $\pi_B = \pi_{B\leftarrow A} \circ \pi_A$.

Proof: We can construct the bijections in a greedy fashion: Start with an arbitrary cloud $C \in \mathcal{S}$ and choose an arbitrary bijection $\pi_C \colon [n] \to C$. For all other clouds $B \in \mathcal{S}$, choose $\pi_B := \pi_{B\leftarrow C} \circ \pi_C$.

Let A, B be two arbitrary clouds in $\mathcal{S}$. Let $\sigma_{A\leftarrow B} := \pi_A \circ \pi_B^{-1}$. To prove the lemma, we have to verify that $\sigma_{A\leftarrow B} = \pi_{A\leftarrow B}$. By construction, $\sigma_{A\leftarrow B} = \pi_{A\leftarrow C} \circ \pi_{C\leftarrow B}$. Let $\eta = 1/16$. Since $\rho(A,C) \geqslant 1-\eta$ and $\rho(B,C) \geqslant 1-\eta$, we have $|\langle b, \sigma_{A\leftarrow B}(b)\rangle| \geqslant 1 - 2\eta$ for all $b \in B$ (using Fact 3.9). Since $(1-2\eta)^2 > 1 - 4\eta = 3/4$, Fact 3.5 (uniqueness of bijection) implies that $\sigma_{A\leftarrow B} = \pi_{A\leftarrow B}$. ∎

Hence, for a consistent set of clouds $\mathcal{S}$, the distribution over local unique games labelings $\mu_{\mathcal{S}}$ can be defined easily as follows:

> Sample $\ell \in [n]$ uniformly at random, and for every cloud $A \in \mathcal{S}$, assign $\pi_A(\ell)$ as label.

3.3.1. Local Distributions via Geometric Decomposition: To construct a local distribution for a set $\mathcal{S}$ which is not consistent, we partition the set $\mathcal{S}$ into consistent clusters. To this end, we make the following definition:

Definition 3.13: A set $\mathcal{S} \subseteq \mathcal{B}$ is *consistent* with respect to a partition P of $\mathcal{B}$ (denoted $\mathrm{Cons}(\mathcal{S}, P)$) if

$$\forall C \in P. \quad \forall A, B \in C \cap \mathcal{S}. \qquad \rho(A,B) \geqslant 1 - 1/16 .$$

We use $\mathrm{Incons}(\mathcal{S}, P)$ to denote the event that $\mathcal{S}$ is not consistent with P. The following is a corollary of Lemma 3.12.

Corollary 3.14: Let P be a partition of $\mathcal{B}$ and let $\mathcal{S} \subseteq \mathcal{B}$. If $\mathrm{Cons}(\mathcal{S}, P)$, then there exists bijections $\{\pi_A \colon [n] \to A \mid A \in \mathcal{S}\}$ such that for all $C \in P$, and $A, B \in C \cap \mathcal{S}$, $\pi_B = \pi_{B\leftarrow A} \circ \pi_A$.

Lemma 3.15: For every $t \in \mathbb{N}$, there exists a distribution over partitions P of $\mathcal{B}$ such that

$$\rho(A,B) \geqslant 1 - \varepsilon \Longrightarrow \Pr\{P(A) = P(B)\} \geqslant 1 - O(t\sqrt{\varepsilon}) ,$$

$$\rho(A,B) \leqslant 1 - 1/16 \Longrightarrow \Pr\{P(A) = P(B)\} \leqslant (3/4)^t .$$

Proof: Let $s \in \mathbb{N}$ be even and large enough (we will determine the value of s later). For every set $B \in \mathcal{B}$, define a vector $\boldsymbol{v}_B \in \mathbb{R}^D$ with $D := d^s$ as $\boldsymbol{v}_B := \frac{1}{\sqrt{n}} \sum_{v\in B} v^{\otimes s}$. We consider the following distribution over partitions P of $\mathcal{B}$: Choose t random hyperplanes $H_1, \ldots, H_t$ through the origin in $\mathbb{R}^D$. Consider the partition of $\mathbb{R}^D$ formed by these hyperplanes. Output the induced partition P of $\mathcal{B}$ (two sets $A, B \in \mathcal{B}$ are in the same cluster of P if and only if $\boldsymbol{v}_A$ and $\boldsymbol{v}_B$ are not separated by any of the hyperplanes $H_1, \ldots, H_t$).

[3]The matching property asserts that such a matching exists. If it is not unique, we pick an arbitrary one. We will assume $\pi_{A\to B} = \pi^{-1}_{B\to A}$.

Since s is even, Corollary 3.8 shows that for any two sets $A, B \in \mathcal{B}$, $\langle \boldsymbol{v}_A, \boldsymbol{v}_B\rangle = \rho(A,B)^s \pm 2\cdot(3/4)^{-s/2}$.

Furthermore, if $\rho(A,B) = 1-\varepsilon$, then $\langle \boldsymbol{v}_A, \boldsymbol{v}_B\rangle \geqslant (1-\varepsilon)^s \geqslant 1-s\varepsilon$. Let $\eta = 1/16$. We choose s minimally such that $(1-\eta)^s + 2\cdot(3/4)^{-s/2} \leqslant 1/\sqrt{2}$. (So s is an absolute constant.) Then for any two sets $A, B \in \mathcal{B}$ with $\rho(A,B) \leqslant 1-\eta$, their vectors have inner product $\langle \boldsymbol{v}_A, \boldsymbol{v}_B\rangle \leqslant 1/\sqrt{2}$. Thus, a random hyperplane through the origin separates $\boldsymbol{v}_A$ and $\boldsymbol{v}_B$ with probability at least $1/4$. Therefore, $\Pr\{P(A) = P(B)\} \leqslant (3/4)^t$. On the other hand, if $\rho(A,B) = 1-\varepsilon$, then the vectors of A and B have inner product $\langle \boldsymbol{v}_A, \boldsymbol{v}_B\rangle \geqslant 1-s\varepsilon$. Thus, a random hyperplane through the origins separates the vectors with probability at most $O(\sqrt{\varepsilon})$. Hence, $\Pr\{P(A)=P(B)\} \geqslant \left(1-O(\sqrt{\varepsilon})\right)^t \geqslant 1-O(t\sqrt{\varepsilon})$. ∎

The previous lemma together with a simple union bound imply the next corollary.

Corollary 3.16: The distribution over partitions from Lemma 3.15 satisfies the following property: For every set $\mathcal{S} \subseteq \mathcal{B}$, $\Pr\left\{\mathrm{Incons}(\mathcal{S}, P)\right\} \leqslant |\mathcal{S}|^2 \cdot (3/4)^t$

4. Integrality Gap Instance for Unique Games

Khot et al. [15] show a UGC-based hardness result for the $\mathsf{E2Lin}_q$ problem. Specifically, they exhibit a reduction $\Phi_{\gamma,q}$ (see Figure 1) that maps a Unique Games instance Υ to an $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Upsilon)$ such that the following holds: For every $\gamma > 0$ and all $q \geqslant q_0(\gamma)$,

- Completeness: If Υ is $1-\eta$-satisfiable then $\Phi_{\gamma,q}(\Upsilon)$ is $1-\gamma-o_{\eta,\delta}(1)$ satisfiable.
- Soundness: If Υ has no labeling satisfying more than δ-fraction of the constraints, then no assignment satisfies more than $q^{-\eta/2}+o_{\eta,\delta}(1)$-fraction of equations in $\Phi_{\gamma,q}(\Upsilon)$.

Here the notation $o_{\eta,\delta}(1)$ refers to any function that tends to 0 whenever η and δ go to naught.

5. Construction of SDP Solutions for E2LIN(q)

For a vertex $(B,x) \in \mathcal{B}\times\mathbb{F}_q^n$, we index the coordinates of x by the elements of B. Specifically, we have $x = (x_b)_{b\in B} \in \mathbb{F}_q^B$.

Geometric Partitioning: Apply Lemma 3.15 to the collection of sets of vectors $\mathcal{B}$. We obtain a distribution $\mathcal{P}$ over partitions P of $\mathcal{B}$ into T disjoint subsets $\{P_\alpha\}_{\alpha=1}^T$. For a subset $\mathcal{S} \subset \mathcal{B}$, let $\mathcal{S} = \{\mathcal{S}_\alpha\}_{\alpha=1}^T$ denote the partition induced on the set $\mathcal{S}$, that is, $\mathcal{S}_\alpha := P_\alpha \cap \mathcal{S}$. For a family $B \in \mathcal{B}$, let α_B denote the index of the set P_{α_B} in the partition P that contains B.

5.1. Vector Solution

Let $\mathcal{R}$ is the measure space over which the tensored vectors $b^{\otimes t}$ are defined. The notation $P_\alpha(B)$ denotes the 0/1-indicator for the event $B \in P_\alpha$.

For a vertex $(B,x) \in \mathcal{B}\times\mathbb{F}_q^n$, the corresponding SDP vectors are given by functions $\boldsymbol{V}_j^{B,x} \colon \mathcal{P}\times[T]\times\mathcal{R} \to \mathbb{R}^q$ defined as follows:

$$\boldsymbol{W}_j^{B,x}(r) = \tfrac{1}{\sqrt{n}}\textstyle\sum_{b\in B}\psi(x_b - j + b^{\otimes t}(r)) \qquad (8)$$

$$\boldsymbol{U}_j^{B,x}(P,\alpha,r) = P_\alpha(B)\cdot \boldsymbol{W}_j^{B,x}(r) \qquad (9)$$

Input A Unique Games instance Υ with vertex set V, edge set $E \subseteq V\times V$ (we assume the graph (V,E) to be regular), and permutations $\{\pi_e\colon [n]\to[n]\}_{e\in E}$.

Output An $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Upsilon)$ with vertex set $\mathcal{V} = V\times\mathbb{F}_q^n$. Let $\{\mathcal{F}_v\colon \mathbb{F}_q^n \to \mathbb{F}_q\}_{v\in V}$ denote an $\mathbb{F}_q$-assignment to $\mathcal{V}$. The constraints of $\Phi_{\gamma,q}(\Upsilon)$ are given by the tests performed by the following probabilistic verifier:

- Pick a random vertex $v \in V$. Choose two random neighbors $w, w' \in N(v) \subseteq V$. Let π, π' denote the permutations on the edges (w,v) and (w',v).
- Sample $x \in \mathbb{F}_q^n$ uniformly at random. Generate $y \in \mathbb{F}_q^n$ as follows:

$$y_i = \begin{cases} x_i & \text{with probability } 1-\gamma \\ \text{uniformly random in } \mathbb{F}_q & \text{with probability } \gamma\end{cases}$$

- Generate a uniform random element $c \in \mathbb{F}_q$.
- Test if $\mathcal{F}_w(y\circ\pi + c\cdot\mathbb{1}) = \mathcal{F}_{w'}(x\circ\pi') + c$. (Here, $x\circ\pi$ denotes the vector $(x_{\pi(i)})_{i\in[n]}$.)

Figure 1. Hardness reduction from Unique Games to $\mathsf{E2Lin}_q$ [15]

Further, $\boldsymbol{V}_0$ is a unit vector orthogonal to all the vectors $\boldsymbol{U}_j^{B,x}$. Define, $\boldsymbol{V}_j^{B,x} = \frac{1}{q}\boldsymbol{V}_0 + \frac{\sqrt{q-1}}{q}\boldsymbol{U}_j^{B,x}$. Let us evaluate the inner product between two vectors $\boldsymbol{V}_i^{A,x}$ and $\boldsymbol{V}_j^{B,y}$, (in this way, we also clarify the intended measure on the coordinate set)

$$\begin{aligned}\langle \boldsymbol{V}_i^{A,x}, \boldsymbol{V}_j^{B,y}\rangle &= \tfrac{1}{q^2} + \tfrac{q-1}{q^2}\langle \boldsymbol{U}_i^{A,x}, \boldsymbol{U}_j^{B,y}\rangle \\ &= \tfrac{1}{q^2} + \tfrac{q-1}{q^2}\mathop{\mathrm{E}}_{P\sim\mathcal{P}}\textstyle\sum_{\alpha=1}^T P_\alpha(A)P_\alpha(B)\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle \\ &= \tfrac{1}{q^2} + \tfrac{q-1}{q^2}\mathop{\Pr}_{P\sim\mathcal{P}}\{P(A)=P(B)\}\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle \qquad (10)\end{aligned}$$

Let us also compute the inner product of $\boldsymbol{W}_i^{A,x}$ and $\boldsymbol{W}_j^{B,y}$. Recall the notation $\langle u,v\rangle_\psi := \langle\psi(u),\psi(v)\rangle$. The inner product $\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle$ is given by,

$$\begin{aligned}&= \tfrac{1}{n}\textstyle\sum_{a\in A,b\in B}\mathrm{E}_{r\sim\mathcal{R}}\langle x_a - i + a^{\otimes t}(r), y_b - j + b^{\otimes t}(r)\rangle_\psi \\ &= \tfrac{1}{n}\sum_{a\in A,b\in B}\langle (x_a-i)\otimes a^{\otimes t}, (y_b-j)\otimes b^{\otimes t}\rangle_\psi \quad \text{(by Obs. 2.9)} \\ &= \tfrac{1}{n}\sum_{a\in A,b\in B}\langle\psi(x_a-i),\psi(y_b-j)\rangle\langle a,b\rangle_\psi^t . \quad \text{(by Lem. 2.8)} \qquad (11)\end{aligned}$$

5.2. Local Distributions

Fix a subset $\mathcal{S}\subset\mathcal{B}$ of size at most R. In this section, we will construct a local distribution over $\mathbb{F}_q$-assignments for the vertex set $S = \mathcal{S}\times\mathbb{F}_q^n$ (see Figure 2). Clearly, the same construction also yields a distribution for a general set of vertices $S' \subset \mathcal{B}\times\mathbb{F}_q^n$ of size at most R.

We need the following two simple observations.

Observation 5.1: For all $a,b \in \mathbb{F}_q$, we have

$$\Pr_{\kappa\in\mathbb{F}_q}\{a+\kappa = i,\ b+\kappa = j\} = \tfrac{1}{q^2} + \tfrac{q-1}{q^2}\langle\psi(a-i),\psi(b-j)\rangle .$$

Observation 5.2: Fix $a,b\in\mathbb{F}_q$, over a random choice of $h_1,h_2\in\mathbb{F}_q$, we have $\mathrm{E}_{h_1,h_2\in\mathbb{F}_q}[\langle\psi(a+h_1),\psi(b+h_2)\rangle] = 0$.

The next lemma shows that the second-order correlations of the distribution μ_S approximately match the inner products of the vector solution $\{V_i^{A,x}\}$.

Lemma 5.3: For any two vertices $(A,x),(B,y) \in S$,

$$\Pr_{Z\sim\mu_S}\left\{Z^{A,x}=i,\ Z^{B,y}=j\right\} = \langle V_i^{A,x}, V_j^{B,y}\rangle \ \pm 10|\mathcal{S}|^2(3/4)^{t/2}\,.$$

Proof: Firstly, since $\Pr[\mathrm{Cons}(\mathcal{S},P)] \geqslant 1-|S|^2(3/4)^t$ (by Corollary 3.16), the probability $\Pr_{\mu_S}\{Z^{A,x}{=}i,\ Z^{B,y}{=}j\}$ equals

$$\Pr_{\mu_S}\left\{Z^{A,x}{=}i,\ Z^{B,y}{=}j \;\middle|\; \mathrm{Cons}(\mathcal{S},P)\right\} \quad \pm |S|^2(3/4)^t\,. \qquad (12)$$

Using Observation 5.1, and the definition of $Z^{A,x}$ and $Z^{B,y}$ we can write the probability $\Pr_{\mu_S}\left\{Z^{A,x}{=}i,\ Z^{B,y}{=}j \;\middle|\; \mathrm{Cons}(\mathcal{S},P)\right\}$ as

$$\frac{1}{q^2}+\frac{q-1}{q^2}\mathop{\mathbb{E}}_{P,H,L,r}\left[\langle\psi(F^{A,x}{+}h_A{-}i),\psi(F^{B,y}{+}h_B{-}j)\rangle \;\middle|\; \mathrm{Cons}(\mathcal{S},P)\right]. \qquad (13)$$

If A,B fall in the same set in the partition P (that is $\alpha_A=\alpha_B$), then we have $h_A=h_B$. If A,B fall in different sets (that is $\alpha_A\neq\alpha_B$), then h_A,h_B are independent random variables uniformly distributed over $\mathbb{F}_q$. Using Observation 5.2, we can write

$$\begin{aligned}&\mathop{\mathbb{E}}_{P,H,L,r}[\langle\psi(F^{A,x}+h_A-i),\psi(F^{B,y}+h_B-j)\rangle \mid \mathrm{Cons}(\mathcal{S},P)]\\ &= \mathop{\mathbb{E}}_{P,L,r}[\mathbf{1}(\alpha_A=\alpha_B)\langle\psi(F^{A,x}-i),\psi(F^{B,y}-j)\rangle \mid \mathrm{Cons}(\mathcal{S},P)].\quad (14)\end{aligned}$$

Let P be a partition such that $\mathrm{Cons}(\mathcal{S},P)$ and $\alpha_A=\alpha_B=\alpha$. The bijections π_A,π_B (see step 4 Figure 2) satisfy $\pi_A=\pi_{A\leftarrow B}\circ\pi_B$. Note that therefore $a=\pi_{A\leftarrow B}(b)$ whenever $a=\pi_A(\ell)$ and $b=\pi_B(\ell)$ for some $\ell\in[n]$. Hence, over a random choice of $\ell_\alpha\in[n]$, the pair $(\pi_A(\ell_\alpha),\pi_A(\ell_\alpha))$ has the same distribution as a uniformly random pair (a,b) with $a=\pi_{A\to B}(b)$. Thus,

$$\begin{aligned}&\mathop{\mathbb{E}}_{L}\mathop{\mathbb{E}}_{r}[\langle\psi(F^{A,x}(P,L,r)-i),\psi(F^{B,y}(P,L,r)-j)\rangle]\\ &=\frac{1}{n}\sum_{\substack{a\in A,b\in B\\ a=\pi_{A\leftarrow B}(b)}}\mathop{\mathbb{E}}_{r}\langle\psi(x_a-i+a^{\otimes t}(r)),\psi(y_b-j+b^{\otimes t}(r))\rangle\\ &=\frac{1}{n}\sum_{\substack{a\in A,b\in B\\ a=\pi_{A\leftarrow B}(b)}}\langle\psi(x_a-i),\psi(y_b-j)\rangle\cdot\langle a,b\rangle_\psi^t \quad \text{(by Obs. 2.9 and Lem. 2.8)}\\ &=\langle W_i^{A,x},W_j^{B,y}\rangle\pm 2\cdot(3/4)^{t/2} \quad \text{(by Eq. (11) and Lemma 3.7)}.\end{aligned}$$

Combining the above expression with (14), we get

$$\begin{aligned}&\mathop{\mathbb{E}}_{P,H,L,r}[\langle\psi(F^{A,x}+h_A-i),\psi(F^{B,y}+h_B-j)\rangle \mid \mathrm{Cons}(\mathcal{S},P)]\\ &=\mathop{\mathbb{E}}_{P}[\mathbf{1}_{(\alpha_A=\alpha_B)} \mid \mathrm{Cons}(\mathcal{S},P)]\langle W_i^{A,x},W_j^{B,y}\rangle\pm(|S|^2(3/4)^t+2\cdot(3/4)^{t/2})\\ &=\Pr_P\left[P(A)=P(B)\right]\langle W_i^{A,x},W_j^{B,y}\rangle\pm 10|\mathcal{S}|^2(3/4)^{t/2}\,.\end{aligned}$$

In the last step, we used $\Pr\{\mathrm{Cons}(\mathcal{S},P)\}\geqslant 1-|S|^2(3/4)^t$ and $|\langle W_i^{A,x},W_j^{B,y}\rangle|\leqslant 1$. Finally, substituting back in (12) yields the desired result. ■

Lemma 5.4: Let $\mathcal{S}'\subset\mathcal{S}$ be two subsets of $\mathcal{B}$ and let $S'=\mathcal{S}'\times\mathbb{F}_q^n$ and $S=\mathcal{S}\times\mathbb{F}_q^n$. Then, $\|\mu_{S'}-\mathrm{margin}_{S'}\,S\|_1\leqslant 2|\mathcal{S}|^2(3/4)^t$.

Proof: For a partition $P\in\mathcal{P}$, let $\mu_{S|P}$ denote the distribution μ_S conditioned on the choice of partition P. Firstly, we will show the following claim:

Claim 5.5: If $\mathrm{Cons}(\mathcal{S}',P)$ and $\mathrm{Cons}(\mathcal{S},P)$, then $\mu_{S'|P}=\mathrm{margin}_{S'}\,\mu_{S|P}$.

Proof: Let $\{\mathcal{S}_\alpha\}$ and $\{\mathcal{S}'_\alpha\}$ denote the partitions induced by P on the sets $\mathcal{S}$ and $\mathcal{S}'$ respectively. Since $\mathcal{S}'\subseteq\mathcal{S}$, we have $\mathcal{S}'_\alpha\subseteq\mathcal{S}_\alpha$ for all $\alpha\in[T]$. By our assumption, each of the sets $\mathcal{S}'_\alpha$ are *consistent* in that $\rho(A,B)\geqslant 1-1/16$ for all $A,B\in\mathcal{S}'_\alpha$. Similarly, the sets $\mathcal{S}_\alpha$ are also *consistent*.

Let us consider the pair of sets $\mathcal{S}'_\alpha\subset\mathcal{S}_\alpha$ for some $\alpha\in[T]$. Intuitively, the vectors within these sets fall in to n distinct clusters. Thus the distribution over the choice of consistent representatives are the same in $\mu_{S'|P}$ and $\mathrm{margin}_{S'}\,\mu_{S|P}$. Formally, we have two sets of bijections $\Pi_{\mathcal{S}'_\alpha}=\{\pi'_A\mid A\in\mathcal{S}'_\alpha\}$ and $\Pi_{\mathcal{S}_\alpha}=\{\pi_A\mid A\in\mathcal{S}_\alpha\}$ satisfying the following property for all $A,B\in\mathcal{S}'_\alpha$ $\pi_{A\to B}\circ\pi'_A=\pi'_B$ and $\pi_{A\to B}\circ\pi_A=\pi_B$.

Fix a collection $A\in\mathcal{S}'_\alpha$. Let $\sim$ denote that two sets of random variables are identically distributed.

$$\begin{aligned}&\{\pi'_B(\ell_\alpha)\mid B\in\mathcal{S}'_\alpha\}\sim\{\pi_{A\to B}\circ\pi'_A(\ell_\alpha)\mid B\in\mathcal{S}'_\alpha\}\\ &\sim\{\pi_{A\to B}(a)\mid B\in\mathcal{S}'_\alpha,\ a\text{ is uniformly random in }A\}\\ &\sim\{\pi_{A\to B}\circ\pi_A(\ell_\alpha)\mid B\in\mathcal{S}'_\alpha\}\sim\{\pi_B(\ell_\alpha)\mid B\in\mathcal{S}'_\alpha\}\,.\end{aligned}$$

The variables $L=\{\ell_\alpha\}$ are independent of each other. Therefore, $\{\pi'_B(\ell_B)\mid B\in\mathcal{S}'\}\sim\{\pi_B(\ell_B)\mid B\in\mathcal{S}'\}$. Notice that the choice of $r\in\mathcal{R}$, H and κ are independent of the set $\mathcal{S}$. Hence, the final assignments $\{Z^{B,x}\mid B\in\mathcal{S}',x\in\mathbb{F}_q^n\}$ are identically distributed in both cases. ■

Returning to the proof of Lemma 5.4, we can write

$$\begin{aligned}&\|\mu_{S'}-\mathrm{margin}_{S'}\,\mu_S\|_1=\|\mathbb{E}_P\mu_{S'|P}-\mathbb{E}_P\mathrm{margin}_{S'}\,\mu_{S|P}\|_1\\ &\leqslant\mathbb{E}_P[\|\mu_{S'|P}-\mathrm{margin}_{S'}\,\mu_{S|P}\|_1]\quad\text{(by Jensen's inequality)}\\ &=\Pr\{\mathrm{Incons}(\mathcal{S},P)\}\mathop{\mathbb{E}}_{P}[\|\mu_{S'|P}-\mathrm{margin}_{S'}\,\mu_{S|P}\|_1\mid\mathrm{Incons}(\mathcal{S},P)]\,.\end{aligned}$$

The first step uses that the operator $\mathsf{margin}_{S'}$ is linear. The final step in the above calculation makes use of Claim 5.5. The lemma follows by observing that $\Pr[\mathrm{Incons}(\mathcal{S},P)]\leqslant|\mathcal{S}|^2(3/4)^t$ and $\|\mu_{S'|P}-\mathrm{margin}_{S'}\,\mu_{S|P}\|_1\leqslant 2$. ■

The next corollary follows from the previous lemma (Lemma 5.4) and the triangle inequality.

Corollary 5.6: Let $\mathcal{S},\mathcal{S}'$ be two subsets of $\mathcal{B}$ and let $S'=\mathcal{S}'\times\mathbb{F}_q^n$ and $S=\mathcal{S}\times\mathbb{F}_q^n$. Then,

$$\|\mathrm{margin}_{S\cap S'}\,\mu_S-\mathrm{margin}_{S\cap S'}\,\mu_{S'}\|_1\leqslant 4\max(|\mathcal{S}|^2,|\mathcal{S}'|^2)(3/4)^t\,.$$

6. Putting It Together

Proof of Theorem 1.6: Using Lemma 3.3, we obtain a weak $\mathbb{F}_q$-integral SDP solution $\mathcal{B}=\{B_u\}_{u\in V}$ of value $1-O(\sqrt{\eta\log q})$ for Υ. We construct a vector solution $\{V_i^{B,x}\mid i\in\mathbb{F}_q,B\in\mathcal{B},x\in\mathbb{F}_q^n\}$ and local distributions $\{\mu_S\mid S\subseteq\mathcal{B}\times\mathbb{F}_q^n\}$ as defined in the previous section (§5).

Lemma 5.3 and Corollary 5.6 show that this SDP solution is ε-infeasible for SA_R and LH_R, where $\varepsilon=O(R^2\cdot(3/4)^{t/2})$.

For $S = \mathcal{S} \times \mathbb{F}_q^n$, the local distribution μ_S over assignments $\mathbb{F}_q^S$ is defined by the following sampling procedure:

Partitioning:

1) Sample a partition $P = \{P_\alpha\}_{\alpha=1}^T$ of $\mathcal{B}$ from the distribution $\mathcal{P}$ obtained by Lemma 3.15. Let α_A, α_B denote the indices of sets in the partition P that contain $A, B \in \mathcal{S}$ respectively.
2) If $\mathrm{Incons}(\mathcal{S}, P)$ then output a uniform random $\mathbb{F}_q$-assignment to $S = \mathcal{S} \times \mathbb{F}_q^n$. Specifically, set

$$Z^{(B,x)} = \text{uniformly random in } \mathbb{F}_q \quad \forall B \in \mathcal{S}, x \in \mathbb{F}_q^n .$$

Choosing consistent representatives:

4) If $\mathrm{Cons}(\mathcal{S}, P)$ then by Corollary 3.14, for every part $\mathcal{S}_\alpha = P_\alpha \cap \mathcal{S}$, there exists bijections $\Pi_{\mathcal{S}_\alpha} = \{\pi_B : [n] \to B \mid B \in \mathcal{S}_\alpha\}$ such that for every $A, B \in \mathcal{S}_\alpha$, $\pi_A = \pi_{A \leftarrow B} \circ \pi_B$.
5) Sample $L = \{\ell_\alpha\}_{\alpha=1}^T$ by choosing each ℓ_α uniformly at random from $[n]$. For every cloud $B \in \mathcal{S}$, define $\ell_B = \ell_{\alpha_B}$. The choice of L determines a set of representatives for each $B \in \mathcal{S}$. Specifically, the representative of B is fixed to be $\pi_B(\ell_B)$.

Sampling assignments:

5) Sample $r \in \mathcal{R}$ from the corresponding probability measure and assign

$$F^{B,x}(P, L, r) = x_{\pi_B(\ell_B)} + \pi_B(\ell_B)^{\otimes t}(r) .$$

6) Sample $H = \{h_\alpha\}_{\alpha=1}^T$ by choosing each h_α uniformly at random from $[q]$. For every cloud $B \in \mathcal{B}$, define $h_B = h_{\alpha_B}$.
7) Sample κ uniformly at random from $[q]$.
8) For each $B \in \mathcal{S}_\alpha$ and $x \in \mathbb{F}_q^n$, set

$$Z^{B,x}(P, L, r, H, \kappa) = F^{B,x}(P, L, r) + h_B + \kappa .$$

9) Output the $\mathbb{F}_q$-assignment $\{Z^{B,x}\}_{(B,x)\in S}$.

Figure 2. Local distribution over $\mathbb{F}_q$-assignments

The value of the SDP solution for $\Phi_{\gamma,q}(\Upsilon)$ (see Fig. 1) is given by

$$\mathop{\mathrm{E}}_{v\in V} \mathop{\mathrm{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}} \mathop{\mathrm{E}}_{\{x,y\}} \mathop{\mathrm{E}}_{c\in\mathbb{F}_q} \textstyle\sum_{i=1}^q \langle \boldsymbol{V}_i^{w,(x\circ\pi+c\cdot\mathbb{1})}, \boldsymbol{V}_{i-c}^{w',y\circ\pi'} \rangle .$$

Using Eq. (10)–(11),

$$\begin{aligned}&\langle \boldsymbol{V}_i^{w,(x\circ\pi+c\cdot\mathbb{1})}, \boldsymbol{V}_{i-c}^{w',y\circ\pi'} \rangle = \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \Pr_{P\sim\mathcal{P}} \{P(B_w) = P(B_{w'})\}\\ &\cdot \tfrac{1}{n} \sum_{\ell,\ell'\in[n]} \langle \psi(x_{\pi(\ell)} + c - i), \psi(y_{\pi'(\ell')} - (i-c)) \rangle \langle b_{w,\ell}, b_{w',\ell'} \rangle_\psi^t .\end{aligned}$$

Note that $\langle \psi(x_{\pi(\ell)} + c - i), \psi(y_{\pi'(\ell')} - (i - c)) \rangle = \langle \psi(x_{\pi(\ell)}), \psi(y_{\pi'(\ell')}) \rangle$. Using Observation 3.6, we have $\pi_{(w,v)}(\ell) = \pi_{(w',v)}(\ell')$ if and only if $\ell = \pi_{B_w \leftarrow B_{w'}}(\ell')$. Hence, by Lemma 3.7,

$$\begin{aligned}&\tfrac{1}{n} \sum_{\ell,\ell'\in[n]} \langle \psi(x_{\pi(\ell)}), \psi(y_{\pi'(\ell')}) \rangle \langle b_{w,\ell}, b_{w',\ell'} \rangle_\psi^t\\ &= \tfrac{1}{n} \sum_{\ell} \langle \psi(x_{\pi(\ell)}), \psi(y_{\pi(\ell)}) \rangle \langle b_{w,\pi(\ell)}, b_{w',\pi(\ell)} \rangle_\psi^t \pm 2 \cdot R^2 (3/4)^{t/2}\\ &= \tfrac{1}{n} \sum_{\ell} \langle \psi(x_\ell), \psi(y_\ell) \rangle \rho(B_w, B_{w'})^t \quad \pm O(\varepsilon) .\end{aligned}$$

Note that the distribution of $\{x, y\}$ is independent of the vertices v, w, w', and

$$\mathop{\mathrm{E}}_{\{x,y\}} \tfrac{1}{n} \sum_{\ell\in[n]} \langle \psi(x_\ell), \psi(y_\ell) \rangle = 1 - \gamma .$$

Therefore, if we let $\eta_{w,w'} = \rho(B_w, B_{w'})$, we can lower bound the value of the SDP solution as follows

$$\begin{aligned}&\mathop{\mathrm{E}}_{v\in V} \mathop{\mathrm{E}}_{w,w'\in N(v)} \mathop{\mathrm{E}}_{\{x,y\}} \mathop{\mathrm{E}}_{c\in\mathbb{F}_q} \sum_{i=1}^q \langle \boldsymbol{V}_i^{w,(x\circ\pi_{w,v}+c\cdot\mathbb{1})}, \boldsymbol{V}_{i-c}^{w',y\circ\pi_{w',v}} \rangle\\ &= \mathop{\mathrm{E}}_{v,w,w'} \left[\tfrac{1}{q^2} + \tfrac{q-1}{q^2} \Pr_{P\sim\mathcal{P}} [P(B_w)=P(B_{w'})] \cdot q \cdot \rho(B_w,B_{w'})^t (1-\gamma) \right] \pm O(\varepsilon)\\ &\geqslant (1-\gamma) \mathop{\mathrm{E}}_{v,w,w'} \Pr_{P\sim\mathcal{P}} [P(B_w) = P(B_{w'})] \rho(B_w, B_{w'})^t \quad \pm O(\varepsilon)\\ &\geqslant (1-\gamma) \mathop{\mathrm{E}}_{v,w,w'} (1 - O(t\sqrt{\eta_{w,w'}})) \pm O(\varepsilon) \quad \text{(using Lemma 3.15)}\end{aligned}$$

Using Jensen's inequality and the fact that $\mathrm{E}_{v,w,w'} \eta_{w,w'} = O(\sqrt{\eta \log q})$ (Lemma 3.3), we see that the value of our SDP solution is at least $1 - \gamma - O(\varepsilon + t\eta^{1/4})$ (recall that we assume q to be constant).

On smoothing the SDP solution using Theorem 2.3, we lose $O(R^2\varepsilon) = O(R^4 (3/4)^t)$ in the SDP value. Thus we can set $t = o(\eta^{-1/4})$ and $R = (3/4)^{t/10}$ in order to get a feasible SDP solution for LH_R with value $1 - \gamma - o_{\eta,\delta}(1)$.

On smoothing the SDP solution using Theorem 2.4, we lose $O(q^R\varepsilon) = O(q^R (3/4)^t)$ in the SDP value. Thus we can set, $t = o(\eta^{-1/4})$ and $R = t/\log^2 q$, we would get a feasible SDP solution for SA_R with value $1 - \gamma - o_{\eta,\delta}(1)$. ■

Proof of Theorems 1.1–1.2.: Using Theorem 1.6 with the Khot–Vishnoi integrality gap instance (Lemma 3.2), we have $N = 2^{2^{\log(1/\delta)/\eta}}$ and thus $R = 2^{O((\log\log N)^{1/4})}$. Similarly for SA_R, we get $R = O((\log\log N)^{1/4})$.

References

[1] M. Alekhnovich, S. Arora, and I. Tourlakis, "Towards strong nonapproximability results in the Lovász–Schrijver hierarchy," in *STOC*, 2005, pp. 294–303.

[2] S. Arora, B. Bollobás, and L. Lovász, "Proving integrality gaps without knowing the linear program," in *FOCS*, 2002, pp. 313–322.

[3] S. Arora, B. Bollobás, L. Lovász, and I. Tourlakis, "Proving integrality gaps without knowing the linear program," *Theory of Computing*, vol. 2, no. 1, pp. 19–51, 2006.

[4] S. Arora, S. Khot, A. Kolla, D. Steurer, M. Tulsiani, and N. K. Vishnoi, "Unique games on expanding constraint graphs are easy: extended abstract," in *STOC*, 2008, pp. 21–28.

[5] P. Austrin, "Towards sharp inapproximability for any 2-CSP," in *FOCS*, 2007, pp. 307–317.

[6] J. Buresh-Oppenheim, N. Galesi, S. Hoory, A. Magen, and T. Pitassi, "Rank bounds and integrality gaps for cutting planes procedures," *Theory of Computing*, vol. 2, no. 1, pp. 65–90, 2006.

[7] M. Charikar, K. Makarychev, and Y. Makarychev, "Near-optimal algorithms for unique games," in *STOC*, 2006, pp. 205–214.

[8] ——, "Integrality gaps for Sherali–Adams relaxations," in *STOC*, 2009, to appear.

[9] S. Chawla, R. Krauthgamer, R. Kumar, Y. Rabani, and D. Sivakumar, "On the hardness of approximating multicut and sparsest-cut," *Computational Complexity*, vol. 15, no. 2, pp. 94–114, 2006.

[10] B. Chor and M. Sudan, "A geometric approach to betweenness," *SIAM J. Discrete Math.*, vol. 11, no. 4, pp. 511–523 (electronic), 1998.

[11] N. R. Devanur, S. Khot, R. Saket, and N. K. Vishnoi, "Integrality gaps for sparsest cut and minimum linear arrangement problems," in *STOC*, 2006, pp. 537–546.

[12] K. Georgiou, A. Magen, T. Pitassi, and I. Tourlakis, "Integrality gaps of 2 - o(1) for vertex cover SDPs in the Lovász–Schrijver hierarchy," in *FOCS*, 2007, pp. 702–712.

[13] V. Guruswami, R. Manokaran, and P. Raghavendra, "Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph," in *FOCS*, 2008, pp. 573–582.

[14] S. Khot, "On the power of unique 2-prover 1-round games," in *STOC*, 2002, pp. 767–775.

[15] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell, "Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?" *SIAM J. Comput.*, vol. 37, no. 1, pp. 319–357, 2007.

[16] S. Khot and O. Regev, "Vertex cover might be hard to approximate to within $2-\varepsilon$," *J. Comput. Syst. Sci.*, vol. 74, no. 3, pp. 335–349, 2008.

[17] S. Khot and R. Saket, "Sdp integrality gaps with local ℓ_1-embeddability," in *FOCS*, 2009.

[18] S. Khot and N. K. Vishnoi, "The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1," in *FOCS*, 2005, pp. 53–62.

[19] R. Manokaran, J. Naor, P. Raghavendra, and R. Schwartz, "SDP gaps and UGC hardness for multiway cut, 0-extension, and metric labeling," in *STOC*, 2008, pp. 11–20.

[20] R. O'Donnell and Y. Wu, "An optimal SDP algorithm for max-cut, and equally optimal long code tests," in *STOC*, 2008, pp. 335–344.

[21] P. Raghavendra, "Optimal algorithms and inapproximability results for every CSP?" in *STOC*, 2008, pp. 245–254.

[22] P. Raghavendra and D. Steurer, "How to round any csp?" in *FOCS*, 2009.

[23] ——, "Towards computing the Grothendieck constant," in *SODA*, 2009, pp. 525–534.

[24] A. Samorodnitsky and L. Trevisan, "Gowers uniformity, influence of variables, and PCPs," in *STOC*, 2006, pp. 11–20.

[25] G. Schoenebeck, "Linear level Lasserre lower bounds for certain k-CSPs," in *FOCS*, 2008, pp. 593–602.

[26] G. Schoenebeck, L. Trevisan, and M. Tulsiani, "A linear round lower bound for Lovász-Schrijver SDP relaxations of vertex cover," in *CCC*, 2007, pp. 205–216.

[27] ——, "Tight integrality gaps for Lovász–Schrijver LP relaxations of vertex cover and max cut," in *STOC*, 2007, pp. 302–310.

[28] H. D. Sherali and W. P. Adams, "A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems," *SIAM J. Discrete Math.*, vol. 3, no. 3, pp. 411–430, 1990.

[29] I. Tourlakis, "New lower bounds for vertex cover in the Lovász–Schrijver hierarchy," in *CCC*, 2006, pp. 170–182.

[30] L. Trevisan, "Approximation algorithms for unique games," *Theory of Computing*, vol. 4, no. 1, pp. 111–128, 2008.

[31] M. Tulsiani, "CSP gaps and reductions in the Lasserre hierarchy," in *STOC*, 2009, to apperar.