Integrality gaps for Strong SDP Relaxations of Unique Games

Prasad Raghavendra David Steurer

May 7, 2009

Abstract

Khot's Unique Games Conjecture (UGC) [Kho02] has led to hardness of approximation results (often optimal) for several fundamental optimization problems such as Max Cut and Vertex Cover. The hardness results implied by the Unique Games Conjecture assert that simple semidefinite programs yield the best possible approximation ratios for a variety of problems. Yet, there is little evidence supporting the claim that stronger semidefinite programming (SDP) relaxations do not yield better approximations to these problems.

With the work of Khot and Vishnoi [KV05] 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 $2^{O(\log\log n)^{1/4}}$ 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 $R = O(\log\log n)^{1/4}$.

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 $2^{O(\log\log n)^{1/4}}$ 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 [DKSV06]. We also exhibit, for explicit constants $\gamma, \delta > 0$, an n-point negative-type metric which requires distortion $O(\log\log^{\gamma} n)$ to embed into ℓ_1, although all its subsets of size $2^{O(\log\log n)^{\delta}}$ embed isometrically into ℓ_1.

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 [Kho02] 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 fuelled many of the major developments in hardness of approximation. Starting with the work of Khot [Kho02] on Min-2SAT-Deletion, hardness of approximation results for several fundamental problems like Max Cut [KKMO07], Vertex Cover [KR08], non-uniform Sparsest Cut [CKK+06, KV05] and Max-2-SAT [Aus07] have been obtained assuming the Unique Games Conjecture. In more recent work [Rag08, GMR08, MNRS08, RS09], 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 explicit simple semidefinite program $\mathcal{SDP}$ (See section A).

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 [KV05], 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 [GMPT07], and k-CSPs [Sch08, Tul09], in all other cases, there are no strong SDP gaps supporting a UG hardness result.

Obtaining strong SDP gap that support a UGC based hardness result has been a difficult excercise. In fact, the work of [KV05] stemmed out of an effort in this direction for the SparsestCut problem. Specifically, the Goemans-Linial conjecture regarding embeddability of L_2^2 metrics in to L_1 was refuted in [KV05] 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 MaxCut problem which is fairly well studied [KKMO07, OW08], including an extra inequality on every set of 5 variables in to the standard semidefinite program yields a better approximation, thus disproving UGC.

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 [SA90]. Let $\mathsf{LH}_R(\Gamma)$ and $\mathsf{SA}_R(\Gamma)$ denote the optimum value of the corresponding SDP relaxations on the instance Γ. Further, let $\text{opt}(\Gamma)$ denote the value of the optimum labeling for Γ. For the LH and SA hierarchies, we show:

Theorem 1. *For all constants $\eta, \delta > 0$, there exists a Unique games instance Γ on N vertices such that $\mathsf{LH}_R(\Gamma) \geqslant 1 - \eta$ and $\text{opt}(\Gamma) \leqslant \delta$ for $R = O(2^{(\log\log N)^{\frac{1}{4}}})$*

Theorem 2. *For all constants $\eta, \delta > 0$, there exists a Unique games instance Γ on N vertices such that $\mathsf{SA}_R(\Gamma) \geqslant 1 - \eta$ and $\text{opt}(\Gamma) \leqslant \delta$ for $R = O((\log\log N)^{\frac{1}{4}})$*

Demonstrated for the first time in [KV05], and used in numerous later works [CMM09, STT07b, Tul09, Rag08, GMR08, MNRS08], 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 Γ for unique games, the resulting instance $\Phi(\Gamma)$ 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 [KKMO07, Aus07, Rag08, GMR08, MNRS08, RS09], one can obtain integrality gaps for LH_R and SA_R relaxations for several important problems. For the sake of succinctness, we will state the following general theorem:

Theorem 3. *Let Λ denote a problem in one of the following classes:*

- *A Generalized Constraint Satisfaction Problem [Rag08, Definition 3.1]: a generalization of CSPs permitting bounded payoff functions (positive or negative), instead of predicates.*
- *An Ordering Constraint Satisfaction Problem: a class of problems containing* Maximum Acyclic Subgraph, Betweenness *[CS98], with predicates/bounded payoff functions on orderings of elements.*

(See §A for definitions) Let $\mathcal{SDP}$ 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 $\mathcal{SDP}(\Upsilon) \geqslant c$ and $\text{opt}(\Upsilon) \leqslant s$, for every constant $\eta > 0$, there exists an instance Γ_η over N variables such that:

- $\mathsf{LH}_R(\Gamma_\eta) \geqslant c - \eta$ *and* $\text{opt}(\Gamma_\eta) \leqslant s + \eta$ *with* $R = O(2^{(\log\log N)^{1/4}})$.
- $\mathsf{SA}_R(\Gamma_\eta) \geqslant c - \eta$ *and* $\text{opt}(\Gamma_\eta) \leqslant s + \eta$ *with* $R = O((\log\log N)^{1/4})$.

The O notation in the number of rounds hides a constant depending on η.

The classes of problems for which the above result holds include Max Cut [KKMO07], Max 2-Sat [Aus07], Grothendieck Problem [RS09] k-way Cut [MNRS08] and Maximum Acyclic Subgraph [GMR08]. Notable exceptions that do not directly fall under this framework are Vertex Cover and Sparsest Cut.

Reductions from unique games to Sparsest Cut have been exhibited in [KV05] and [CKK+05]. With the integrality gap for LH_R relaxation of Unique games (Theorem 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 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 embedd in to L_1, while every set of size at most $O(2^{(\log\log N)^{\gamma}})$ embedds isometrically in to L_1.*

The Uniform Sparsest Cut problem is among the many important problems for which no Unique games reduction is known. In [DKSV06], the techniques of [KV05] were extended to obtain an integrality gap for Uniform Sparsest Cut for the SDP with triangle inequalities. Roughly speaking, the SDP gap construction in [DKSV06] 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 5. *For some absolute constants $\gamma, \delta > 0$, there exists an instance G on N vertices of* Balanced Separator *such that the ratio* $\mathrm{opt}(G)/\mathsf{LH}_R(G) \geqslant \Omega(\log\log N)^{\delta}$ *for* $R = O(\log\log N)^{\gamma}$.

1.2 Related Work

In a breakthrough result, Arora et al. [ARV04] used a strong semidefinite program with triangle inequalities to obtain $O(\sqrt{\log n})$ approximation for the Sparsest Cut problem. Inspired by this work, stronger semidefinite programs to obtain better approximation algorithms for certain graph coloring problems [Chl07, ACC06, CS08]. In particular, hierarchies of stronger SDP relaxations such as Lovász–Schriver [LS91], Lasserre [Las01], and Sherali–Adams hierarchies [SA90] (See [Lau03] for a comparison) have emerged as possible avenues to obtain better approximation ratios.

Considerable progress has been made in understanding the limits of linear programming hierarchies. Building on a sequence of works [ABL02, ABLT06, Tou06], Schoenebeck et al. [STT07b] obtained a $2-\varepsilon$-factor integrality gap for $\Omega(n)$ rounds of Lovász–Schriver LS hierarchy. More recently, Charikar et al. [CMM09] constructed integrality gaps for $\Omega(n^{\delta})$ rounds of Sherali–Adams hierarchy for several problems like MaxCut, Vertex Cover Sparsest Cut and Maximum Acyclic Subgraph. Furthermore, the same work also exhibits $\Omega(n^{\delta})$-round Sherali–Adams integrality gap for unique games, in turn obtaining a corresponding gap for every problem to which unique games is reduced to.

Lower bound results of this nature are fewer in the case of semidefinite programs. A $\Omega(n)$ LS_+ round lower bound for proving unsatisfiability of random 3-SAT formulae was obtained in [BOGH+06, AAT05]. In turn, this 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 $\frac{7}{6}$-integrality gap for $\Omega(n)$ rounds of LS_+ in [STT07a] falls in a regime where a matching NP-hardness result has been shown to hold. A significant exception is the result of Georgiou et al. [GMPT07] that exhibited a $2-\varepsilon$-integrality gap for $\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ rounds of LS_+ hierarchy. More recently, building on the beautiful work of [Sch08] on Lasserre integrality gaps for Random 3-SAT, Tulsiani [Tul09] obtained a $\Omega(n)$-round Lasserre integrality gap matching the corresponding UG hardness for k-CSP [ST06].

1.3 Overview of the Technique

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

The overall strategy in this work to construct SDP integrality gaps is along the lines of Khot–Vishnoi [KV05]. Let us suppose we wish to construct a SDP integrality gap for a problem Λ (say MaxCut). Let Φ_Λ be a reduction from Unique games to the problem Λ. The idea is to construct a SDP integrality gap Γ for unique games, and then execute the reduction Φ_Λ on the instance Γ, to obtain the SDP gap construction

$\Phi_\Lambda(\Gamma)$. Surprisingly, as demonstrated in [KV05], the SDP vector solution for Γ can be transformed through the reduction to obtain the SDP solution for $\Phi_\Lambda(\Gamma)$.

Although this technique has been used extensively in numerous works [CMM09, STT07b, Tul09, Rag08, GMR08, MNRS08] since [KV05], there is a crucial distinction between [KV05] and later works. In all other works, starting with an SDP gap Γ 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 [KV05], 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 Γ for a simple SDP relaxation of unique games, [KV05] obtain an SDP gap for MaxCut 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 [KV05]. One of the main contributions of this work is to conceptualize and simplify this aspect of [KV05]. Armed with the understanding of [KV05], 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 the reduction from Unique games to $\mathsf{E2Lin}_q$ obtained in [KKMO07]. Note that $\mathsf{E2Lin}_q$ is a special case of Unique games. Formally, we show the following reduction from a *weak gap* instance for Unique games over a large alphabet to a integrality gap for a strong SDP relaxation of $\mathsf{E2Lin}_q$.

Theorem 6. *(Weak Gaps for Unique games $\implies$ Strong gaps for $\mathsf{E2Lin}_q$)*
For a positive integer q, let Φ_{E2Lin_q} denote the reduction from unique games to $\mathsf{E2Lin}_q$. Given a $(1-\eta,\delta)$-weak gap instance Γ for Unique games, the $\mathsf{E2Lin}_q$ instance $\Phi_{\mathsf{E2Lin}_q}(\Gamma)$ is a $(1-2\gamma, 1/q^{\gamma/2}+o_\delta(1))$ SDP gap for the relaxation LH_R for $R = 2^{O(1/\eta^{1/4})}$. Further, $\Phi(\Gamma)$ is a $(1-\gamma,\delta)$ SDP gap for the relaxation SA_R for $R = O(1/\eta^{1/4})$.

Using the weak gap for Unique games constructed in [KV05], along with the above theorem, implies Theorems 1 and 2. As already pointed out, by now it is fairly straightforward to compose an R-round integrality gap for unique games, with reductions to obtain a R round integrality gaps for other problems. Hence, Theorem 3 is a fairly straightfoward consequence of Theorems 1 and 2.

1.3.1 Example of MaxCut

For the sake of exposition, we will describe the construction of an SDP integrality gap for Maxcut. To further simplify matters, we will exhibit an integrality gap for the basic Goemans-Williamson relaxation, augmented with the triangle inequalities on every three vectors. While an integrality gap of this nature is already part of the work of Khot–Vishnoi [KV05], our proof will be conceptual and amenable to generalization.

Let Γ be a SDP integrality gap for unique games on an alphabet $[n]$. For each vertex B in Γ, the SDP solution associates n orthogonal unit vectors $B = \{b_1, \cdots, b_n\}$. For the sake of clarity, we will refer to a vertex B in Γ and the set of vectors $B = \{b_1, \ldots, b_n\}$ associated with it as a "cloud". The clouds satisfy the following properties:

- (Matching Property) For every two clouds A, B, there is a unique matching $\pi_{B\leftarrow A}$ 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 $\langle a, \pi_{B\leftarrow A}(a)\rangle = \rho(A,B)$.
- (High objective value) For most edges $e = (A,B)$ in the unique games instance Γ, the maximal matching $\pi_{A\leftarrow B}$ is the same as the permutation π_e corresponding to the edge, and $\rho(A,B) \approx 1$.

Let $\Phi_{\text{MaxCut}}(\Gamma)$ be the MaxCut instance obtained by executing the reduction in [KKMO07] on Γ. The reduction Φ_{MaxCut} in [KKMO07] introduces a long code (2^n vertices indexed by $\{-1,1\}^n$) for every cloud in Γ. Hence the vertices of $\Phi_{\text{MaxCut}}(\Gamma)$ are given by pairs (B, x) where B is a cloud in Γ and $x \in \{-1,1\}^n$.

The SDP vectors we construct for the integrality gap instance resemble (somewhat simpler in this work) the vectors in [KV05]. Roughly speaking, for a vertex (B, x), we associate an SDP vector $\boldsymbol{V}^{B,x}$ defined as follows:

$$\boldsymbol{V}^{B,x} = \frac{1}{\sqrt{n}} \sum_{i \in [n]} x_i b_i^{\otimes t}$$

The point of departure from [KV05] is the proof that the vectors form a feasible solution for the stronger SDP. Instead of directly showing that the inequalities hold for the vectors, we exhibit a distribution over integral assignments whose second moments match the inner products. Specifically, to show that triangle inequality holds for three vertices $S = \{(A,x),(B,y),(C,z)\}$, we will exhibit a μ_S distribution over $\{\pm 1\}$ assignments to the three vertices, such that

$$\mathop{\mathbb{E}}_{\{Y^{A,x}, Y^{B,y}, Y^{C,z}\} \sim \mu_S} [Y^{A,x} Y^{B,y}] = \langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$$

The existence of an integral distribution matching the inner products shows that the vectors satisfy all valid inequalities on the three variables, including the triangle inequality. We shall construct the distribution μ_S over local assignments in three steps,

Local Distributions over Labelings for Unique Games For a subset of clouds $\mathcal{S}$ within the unique games instance Γ, we will construct a distribution $\mu_{\mathcal{S}}$ over labelings to the set $\mathcal{S}$. The distribution $\mu_{\mathcal{S}}$ over $[n]^{\mathcal{S}}$ will be "consistent" with the SDP solution to Γ. More precisely, if two clouds A and B are *highly correlated* ($\rho(A,B) \approx 1$), then when the distribution $\mu_{\mathcal{S}}$ assigns label ℓ to A, with high probability it assigns the corresponding label $\pi_{B\leftarrow A}(\ell)$ to B. Recall that $\rho(A,B)$ was defined as $\max_{a\in A, b\in B}\langle a, b\rangle$.

Consider a set $\mathcal{S}$ where every pair of clouds A, B are *highly correlated* ($\rho(A,B) \geqslant 0.9$). We will refer to such a set of clouds as *Consistent*. For a *Consistent* set $\mathcal{S}$,assigning a label ℓ for a cloud A in $\mathcal{S}$, forces the label of every other cloud B to $\pi_{B\leftarrow A}(\ell)$. Furthermore, it is easy to check that the resulting labeling satisfies consistency for every pair of clouds in $\mathcal{S}$. (see Lemma 3 for details) Hence, in this case, the distribution $\mu_{\mathcal{S}}$ could be simply obtained by picking the label ℓ for an arbitrary cloud in $\mathcal{S}$ uniformly at random, and assigning every other cloud the induced label.

Now consider a set $\mathcal{S}$ which is not consistent. Here the idea is to decompose the set of clouds $\mathcal{S}$ in to clusters, such that each cluster is consistent. Given a decomposition, for each cluster the labeling can be independently generated as described earlier.

In this work, we utilize two different clustering strategies.

– *Greedy Construction*: Fix a threshold θ, and construct a graph connecting every pair of clouds A, B whose correlation $\rho(A,B) > \theta$. The decomposition is given by connected components in this graph. (see Lemma 4)

– *Geometric Decomposition*: Somewhat surprisingly, the correlations $\rho(A,B)$ for clouds $A, B \in \mathcal{S}$, can be approximated well by a certain L_2^2 metric. More precisely, for each cloud A, we can associate a unit vector $\boldsymbol{v}_A = \sum_{a\in A} a^{\otimes s}$ such that the L_2^2 distance between $\boldsymbol{v}_A, \boldsymbol{v}_B$ is a good approximation of the quantity $1 - \rho(A,B)$.

  By using t random halfspace cuts on this geometric representation, we obtain a partition in to 2^t clusters. A pair of clouds A, B that are not highly correlated ($\rho(A,B) < 1 - 1/16$), are separated by the halfspaces with probability at least $1 - (3/4)^t$. Hence for a large enough t, all resulting clusters are consistent with high probability. (see Lemma 5).

A useful feature of the geometric clustering is that for two subsets $\mathcal{T} \subset \mathcal{S}$, the distribution over labelings $\mu_{\mathcal{T}}$ is equal to the marginal of the distribution $\mu_{\mathcal{S}}$ on $\mathcal{T}$. To see this, observe that the distribution over clusterings depends solely on the geometry of the associated vectors. On the downside, the geometric clustering produces inconsistent clusters with a very small but non-zero probability. (see Corollary 3).

The details of the construction of local distributions to Unique games are presented in 2.6.

Constructing Approximate Distributions Fix a set $S \subseteq \mathcal{S} \times \{\pm 1\}^n$ of vertices in the MaxCut instance $\Phi_{\text{MaxCut}}(\Gamma)$. We will now describe the construction of the local integral distribution μ_S.

In the reduction Φ_{MaxCut}, the labeling ℓ to a cloud B in the unique games instance is encoded as choosing the ℓth dictator cut in the long code corresponding to cloud B. Specifically, assigning the label ℓ to a cloud B should translate in to assigning x_ℓ for every vertex (B, x) in the long code of B. Hence, a straightforward approach to define the distribution μ_S would be the following:

- Sample a labeling $\ell : \mathcal{S} \to [n]$ from the distribution $\mu_{\mathcal{S}}$,
- For every vertex $(B, x) \in \mathcal{S}$, assign $x_{\ell(B)}$.

Although inspired by this, our actual construction of μ_S is slightly more involved. First, we make the following additional assumption regarding the unique games instance Γ :

> *Assumption*: All the SDP vectors for the integrality gap instance Γ are $\{\pm 1\}$-vectors (have all their coordinates from $\{\pm 1\}$).

The SDP gap instance for unique games constructed in [KV05] satisfies this additional requirement. Furthermore, we outline a generic transformation to convert an arbitrary unique games SDP gap in to one that satisfies the above property (see Observation 1). A $\{\pm 1\}$-vector is to be thought of as a distribution over $\{\pm 1\}$ assignments. It is easy to see that tensored powers of $\{\pm 1\}$-vectors yield $\{\pm 1\}$-vectors. Let T denote the number of coordinates in the vectors $\boldsymbol{V}^{B,x}$. The distribution μ_S is defined as follows,

- Sample a labeling $\ell : \mathcal{S} \to [n]$ from the distribution $\mu_{\mathcal{S}}$, and a coordinate $i \in [T]$ uniformly at random.
- For every vertex $(B, x) \in S$, assign $Y^{B,x}$ to be the ith coordinate of the vector $x_{\ell(B)} b_{\ell(B)}^{\otimes t}$.

We will now argue that the first two moments of the local distributions μ_S defined above, approximately match the corresponding inner products between SDP vectors.

Consider the inner product $\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$ of the SDP vectors corresponding to some pair of vertices (A, x) and (B, y) in S. The inner product consists of n^2 terms of the form $\langle x_i a_i^{\otimes t}, y_j b_j^{\otimes t} \rangle$. The crucial observation we will utilize is that the inner product $\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$ is approximately determined by the n terms corresponding to the matching $\pi_{B \leftarrow A}$. In other words, we have

$$\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle \approx \frac{1}{n} \sum_{\ell \in [n]} x_\ell y_{\pi_{B \leftarrow A}(\ell)} \langle a_\ell^{\otimes t}, b_{\pi_{B \leftarrow A}(\ell)}^{\otimes t} \rangle \qquad \leqslant \quad \rho(A, B)^t$$

(see Section 2.5.1 for details)

If $\rho(A, B) < 0.9$, then with high probability the clustering would place the clouds A, B in different clusters. Hence the labels assigned to A, B would be completely independent of each other, and so would the assignments to (A, x) and (B, y). Hence, we would have $\mathbb{E}[Y^{A,x} Y^{B,y}] = 0$. On the other hand, by the above inequality the inner product $\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle \leqslant 0.9^t \approx 0$. Therefore, for clouds A, B that are not *highly correlated*, the inner product of vectors $\boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y}$ agree approximately with the distribution over local assignments.

At the other extreme, if $\rho(A, B) \approx 1$, then with high probability the clustering would not separate A from B. If A, B are not separated, then the distribution $\mu_{\mathcal{S}}$ over labelings will respect the matching between A and B. Specifically, whenever A is assigned label ℓ by $\mu_{\mathcal{S}}$, with high probability B is assigned the label $\pi_{B\leftarrow A}(\ell)$. Consequently, in this case we have

$$\mathop{\mathbb{E}}_{\mu_S}[Y^{A,x}Y^{B,y}] = \frac{1}{n}\sum_{\ell\in[n]}\langle x_\ell a_\ell^{\otimes t}, y_{\pi_{B\leftarrow A}(\ell)} b_{\pi_{B\leftarrow A}(\ell)}^{\otimes t}\rangle \approx \langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y}\rangle$$

Smoothening 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 LH_R and SA_R 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 $u_i = (1-\varepsilon)v_i + \sqrt{2\varepsilon - \varepsilon^2}w_i$. 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.

More abstractly, the set of vectors $\{w_1, \ldots, w_n\}$ correspond to the uniform distribution over all $\{\pm 1\}$ assignments to n variables. In other words, the SDP solution $\{w_i\}_{i=1}^n$ correspond to the center of the integral hull. The SDP solution $\{v_i\}_{i=1}^n$ is within ε distance from the polytope corresponding to LH_R or SA_R as the case may be. Hence, averaging with $\{w_i\}_{i=1}^n$ moves the SDP solution $\{v_i\}_{i=1}^n$ towards the center of the polytope and in to the feasible region. In general, the smoothening operation takes a convex combination of the approximately feasible SDP solution, with an SDP solution corresponding to uniform distribution over all integral assignments.

The claims about the robustness of solutions to LH_R and SA_R (Theorems 8, 7) are presented in Section 2.3 while the proofs are presented in 7

Extending to $\mathsf{E2Lin}_q$ The above argument for MaxCut can be made precise. However, to obtain an SDP gap for larger number of rounds, we use a slightly more involved construction of SDP vectors.

$\{\pm 1\}$-vectors were natural in the above discussion, since MaxCut is a CSP over $\{0, 1\}$. For $\mathsf{E2Lin}_q$, it is necessary to work with vectors whose coordinates are from $\mathbb{F}_q$, as opposed to $\{\pm 1\}$. The tensoring operation for $\mathbb{F}_q$-integral vectors is to be appropriately defined to ensure that while the behaviour of the inner products resemble traditional tensoring, the tensored vectors are $\mathbb{F}_q$-integral themselves (see Section 2.4 for details).

For the case of MaxCut, we used a gap instance Γ for unique games all of whose SDP vectors where $\{\pm 1\}$-vectors. In case of $\mathsf{E2Lin}_q$, the SDP vectors corresponding to the unique games instance Γ would have to be $\mathbb{F}_q$-integral vectors. We outline a generic transformation to convert an arbitrary unique games SDP gap in to one that satisfies this property (see Observation 4).

1.4 Organization

We begin by presenting the formal definition of the SDP hierarchies LH_R and SA_R in Section 2.2. In the next two sections, we introduce some of the machinery concerning robustness of SDP solutions (§2.3) and properties of integral vectors (§2.4). In Section 2.5, we abstract out the properties of a SDP solutions to unique games. Stating certain lemmas in this abstraction has the advantage of being applicable directly to the Balanced Separator gap where there is no explicit unique games instance within. We construct local distributions over labelings for unique games in Section 2.6.

With the requisite machinery already developed in previous sections, we describe the construction of approximately feasible solutions for LH_R and SA_R relaxations of $\mathsf{E2Lin}_q$ in Section 3. Finally, we outline the proof of Theorem 6 in Section 4.

The construction of SDP vectors and local distributions for Balanced Separator are presented in Section 5. This is followed by the overall description of the integrality gap instance in Section 6.

The sections about gaps for $\mathsf{E2Lin}_q$ and the sections about gaps for Balanced Separator only assume the preliminaries and can be read independently of each other. The construction for Balanced Separator separated is less complicated but maybe more ad-hoc in the sense that the gap instance is not obtained through a UG-hardness reduction (at least not explicitly).

Subsequently, we describe the detailed proof of the effectiveness of smoothening operation in converting approximately feasible solutions in to perfectly feasible ones in Section 7. Finally, some of the missing problem definitions and SDP relaxations are included in Appendix A.

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

$$\mathop{\mathbb{E}}_{x\sim\mu} f(x) \stackrel{\text{def}}{=} \sum_{x\in\Sigma^S} \mu(x) f(x)\,.$$

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

$$\mathop{\mathbb{P}\mathrm{r}}_{x\sim\mu} \mathcal{E} \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{x\sim\mu} \mathbf{1}_{\mathcal{E}}(x) = \sum_{x\in\mathcal{E}} \mu(x)\,.$$

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* $\mathsf{margin}_T\,\mu\colon \Sigma^T \to \mathbb{R}_+$ as

$$\mathsf{margin}_T\,\mu(x) \stackrel{\text{def}}{=} \sum_{y\in\Sigma^{S\setminus T}} \mu(x,y)\,.$$

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 will define certain families of strong SDP relaxations that are closely related to the Sherali–Adams hierarchy.

$\{\pm 1\}$-relaxations We begin by describing the relaxations for combinatorial optimization problems over the domain $\{\pm 1\}$ such as Max Cut and Sparsest Cut. Let $\mathfrak{I}$ be a problem instance over a set of variables $\mathcal{V}$. An *SDP solution* for the instance $\mathfrak{I}$ consists of the following:

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

The intention is that the local distributions μ_S arise as marginal distributions of a global distribution $\mu\colon \{\pm 1\}^{\mathcal{V}} \to \mathbb{R}_+$ over $\{\pm 1\}$-assignments to $\mathcal{V}$. In the intended vector solution, all vectors have only $\{\pm 1\}$-coordinates (so that $\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \mathbb{E}_{x\sim\mu} x_i x_j$ for some distribution $\mu \in \Delta(\{\pm 1\}^{\mathcal{V}})$). We assume that the objective function of the instance $\mathfrak{I}$ can be expressed as a linear function in the local distributions, say $\sum_{S\subseteq\mathcal{V},|S|\leqslant R}\sum_{x\in\{\pm 1\}^S} c_{S,x}\,\mu_S(x)$. In the relaxation LH_R, we require that the second-moments of the local distributions μ_S match the inner products of the global vector solution.

LH_R-Relaxation:

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

$$\text{subject to} \quad \langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \mathop{\mathbb{E}}_{x\sim\mu_S} x_i x_j \qquad (S\subseteq\mathcal{V},\ |S|\leqslant R,\ i,j\in S)\,, \tag{2}$$

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_0\rangle = \mathop{\mathbb{E}}_{x\sim\mu_S} x_i \qquad (S\subseteq\mathcal{V},\ |S|\leqslant R,\ i\in S)\,. \tag{3}$$

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

The above relaxation succinctly encodes all possible inequalties on up to R vectors. The next remark makes this observation precise.

Remark 1. A linear inequality on the inner products of a subset of vectors $\{\boldsymbol{v}_i\}_{i\in S}$ for $S\subseteq\mathcal{V}$ is *valid* if it inequality if it holds for all $\{\pm 1\}$-assignments to the variables S. A feasible solution to the LH_R-relaxation satisfies all valid inequalities on sets of up to R vectors.

Notice that the local distributions in the LH_R-relaxation have redundancies. Specifically, consider two sets $A, B\subseteq\mathcal{V}$ such that $A\subset B$ and $|A|, |B|\leqslant R$. The local distribution μ_B induces a distribution $\mathsf{margin}_A\,\mu_B$ over assignments to the set A, since $A\subset B$. (Here, $\mathsf{margin}_A\,\mu_B$ denotes the marginal of μ_B on the set A as defined in §2.1.) It is but natural to enforce that $\mathsf{margin}_A\,\mu_B$ and μ_A be the same distribution.

The following stronger relaxation denoted SA_R also imposes consistency between the local distributions.

SA_R-Relaxation.

$$\text{maximize} \quad \sum_{S\subseteq\mathcal{V},|S|\leqslant R} \sum_{x\in\{\pm1\}^S} c_{S,x}\,\mu_S(x) \tag{4}$$

$$\text{subject to} \quad \langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \mathop{\mathbb{E}}_{x\sim\mu_S} x_i x_j \qquad (S\subseteq\mathcal{V},\ |S|\leqslant R,\ i,j\in S)\,, \tag{5}$$

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_0\rangle = \mathop{\mathbb{E}}_{x\sim\mu_S} x_i \qquad (S\subseteq\mathcal{V},\ |S|\leqslant R,\ i\in S)\,, \tag{6}$$

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

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

Remark 2. The SA_R relaxation is closely related to the R^{th} level of the Sherali–Adams hierarchy. In fact, SA_R is obtained from the basic SDP relaxation by R-rounds Sherali–Adams lift-and-project. In other words, we are optimizing over the intersection of the basic SDP relaxation and the Sherali–Adams relaxation.

$\mathbb{F}_q$-relaxations In the following, we consider optimization problems over the alphabet $\mathbb{F}_q$. Let $\mathfrak{I}$ be a problem instance over a set of variables $\mathcal{V}$.

An *SDP solution* for the instance $\mathfrak{I}$ 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 again 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. Again, we assume that the objective of the instance $\mathfrak{I}$ 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.

$$\text{maximize} \quad \sum_{S\subseteq\mathcal{V},|S|\leqslant R} \sum_{x\in\mathbb{F}_q^S} c_{S,x}\,\mu_S(x) \tag{8}$$

$$\text{subject to} \quad \langle \boldsymbol{v}_{i,a}, \boldsymbol{v}_{j,b}\rangle = \mathop{\mathbb{P}\mathrm{r}}_{x\sim\mu_S}\left\{x_i = a, x_j = b\right\} \qquad (S\subseteq\mathcal{V},\ |S|\leqslant R,\ i,j\in S,\ a,b\in[q])\,, \tag{9}$$

$$\langle \boldsymbol{v}_{i,a}, \boldsymbol{v}_0\rangle = \mathop{\mathbb{P}\mathrm{r}}_{x\sim\mu_S}\left\{x_i = a\right\} \qquad (S\subseteq\mathcal{V},\ |S|\leqslant R,\ i\in S,\ a\in[q])\,. \tag{10}$$

Here, $\boldsymbol{v}_0\in\mathbb{R}^d$ is again an arbitrary fixed unit vector. We say that an SDP solution $\{\mu_S\}, \{\boldsymbol{v}_{i,a}\}$ for $\mathfrak{I}$ is *feasible for* LH_R if it satisfies the constraints (9)–(10). We denote by $\mathsf{LH}_R(\mathfrak{I})$ the value of an optimal solution to this relaxation.

Remark 3. The feasible solution for LH_R satisfies all valid linear inequalities on the inner products of the vectors corresponding to up to R variables in $\mathcal{V}$.

As in the case of $\{\pm1\}$-relaxation, we define the following stronger SDP relaxation with additional consistency requirement.

SA_R-Relaxation:

$$\text{maximize} \quad \sum_{S\subseteq\mathcal{V},|S|\leqslant R}\ \sum_{x\in\mathbb{F}_q^S} c_{S,x}\,\mu_S(x) \tag{11}$$

$$\text{subject to} \quad \langle \boldsymbol{v}_{i,a},\boldsymbol{v}_{j,b}\rangle = \Pr_{x\sim\mu_S}\left\{x_i=a,x_j=b\right\} \qquad (S\subseteq\mathcal{V},\ |S|\leqslant R,\ i,j\in S,\ a,b\in[q])\,, \tag{12}$$

$$\langle \boldsymbol{v}_{i,a},\boldsymbol{v}_0\rangle = \Pr_{x\sim\mu_S}\left\{x_i=a\right\} \qquad (S\subseteq\mathcal{V},\ |S|\leqslant R,\ i\in S,\ a\in[q])\,, \tag{13}$$

$$\left\|\mathsf{margin}_{A\cap B}\,\mu_A - \mathsf{margin}_{A\cap B}\,\mu_B\right\|_1 = 0 \qquad (A,B\subseteq\mathcal{V},\ |A|,|B|\leqslant R)\,. \tag{14}$$

We say an SDP solution $\{\mu_S\},\{\boldsymbol{v}_{i,a}\}$ is *feasible for* SA_R if it satisfies the constraints (12)–(14). We denote by $\mathsf{SA}_R(\mathfrak{I})$ the value of an optimal solution to this relaxation.

2.3 Smoothening

In this section, we give formal statements of the robustness properties of the SDP relaxation defined in the previous section (§2.2). We first introduce a precise notion of approximately feasible SDP solutions. We will 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). All proofs are deferred to Section 7. We state these results only for $\mathbb{F}_q$-relaxations. Analogous results hold for $\{\pm1\}$-relaxations.

Definition 1. 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

- for every set $S\subseteq\mathcal{V}$ with $|S|\leqslant 2$, the distribution μ_S is consistent with the vector solution, that is constraints (9)–(10) of LH_R (or constraints (12)–(14) of SA_R) are satisifed,
- for every set $S\subseteq\mathcal{V}$ with $2<|S|\leqslant R$, the constraints (9)–(10) of LH_R (or the constraints (12)–(14) of SA_R) are satisfied up to an additive error of ε.

Theorem 7. *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 $\{\boldsymbol{v}'_{i,a}\},\{\mu'_S\}_{S\subset\mathcal{V},|S|\leqslant R}$ for LH_R such that for all subsets $S\subseteq\mathcal{V},|S|\leqslant R$,*

$$\left\|\mu_S-\mu'_S\right\|_1\leqslant \mathrm{poly}(q)\cdot R^2\varepsilon\,.$$

Theorem 8. *Given a 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 $S\subseteq\mathcal{V},|S|\leqslant R$,*

$$\left\|\mu_S-\mu'_S\right\|_1\leqslant \mathrm{poly}(q)\cdot\varepsilon\cdot q^R\,.$$

Fact 1. *Suppose the vector solution $\{\boldsymbol{v}_{i,a}\}_{i\in\mathcal{V},a\in\mathbb{F}_q}$ satisfies the properties*

- $\langle\boldsymbol{v}_{i,a},\boldsymbol{v}_{i,b}\rangle=0$ *for all $i\in\mathcal{V}$ and $a\neq b\in[q]$,*
- $\sum_{a\in[q]}\boldsymbol{v}_{i,a}=\boldsymbol{v}_0$ *for all $i\in\mathcal{V}$,*
- $\langle\boldsymbol{v}_{i,a},\boldsymbol{v}_{j,b}\rangle\geqslant 0$ *for all $i,j\in\mathcal{V}$ and $a,b\in[q]$.*

Then for every set $S \subseteq \mathcal{V}$ with $|S| \leqslant 2$, there exists a unique distribution μ_S on $[q]^S$ that is consistent with the vector solution.

This fact implies that given an SDP solution $\{\boldsymbol{v}_{i,a}\}_{i\in\mathcal{V},a\in\mathbb{F}_q}$, $\{\mu_S\}_{S\subset\mathcal{V},|S|\leqslant R}$ that satisfies all constraints of LH_R or SA_R up to an additive error of ε and the vector solution satisfies the properties above, we can change the second-level local distributions such that they are consistent with the vector solution. Hence, the resulting solution is ε-infeasible and we can apply one of the Theorems 7– 8.

2.4 Integral Vectors

In this section, we will develop tools to create and manipulate vectors all of whose coordinates are "integral".

$\{\pm 1\}$-integral vectors We begin by defining our notion of a $\{\pm 1\}$-integral vector.

Definition 2. Let $\mathcal{R} = (\Omega, \mu)$ be a probability space. A function $u \colon \mathcal{R} \to \{\pm 1\}$ is called an $\{\pm 1\}$-*integral vector*. In other words, u is a $\{\pm 1\}$-valued random variable defined on the probability space $\mathcal{R}$. We define an inner product of functions $u, v \colon \mathcal{R} \to \{\pm 1\}$ by

$$\langle u, v\rangle = \mathop{\mathbb{E}}_{r\sim\mathcal{R}} u(r)v(r)\,.$$

In our construction, we often start with $\{\pm 1\}$-integral vectors given by the hypercube $\{\pm 1\}^n$. In the terminology of $\{\pm 1\}$-integral vectors, we can think of the hypercube $\{\pm 1\}^n$ as the set of $\{\pm 1\}$-integral vectors where $\mathcal{R}$ is the uniform distribution over $\{1, \ldots, n\}$.

The following lemma shows how the Goemans–Williamson [GW95] rounding scheme can be thought of as a procedure to "round" arbitrary real vectors to $\{\pm 1\}$-integral vectors.

Observation 1. *Given a family of unit vectors $\{v_1, \ldots, v_n\} \in \mathbb{R}^d$, define the set of $\{\pm 1\}$-valued functions $v_1^*, \ldots, v_n^* : \mathcal{R} \to \{\pm 1\}$ with $\mathcal{R} = \mathcal{G}^d$ - the Gaussian space of appropriate dimension as follows:*

$$v_i^*(g) = \mathsf{sign}(\langle v_i, g\rangle)$$

for $g \in \mathcal{G}^d$. The $\{\pm 1\}$-valued functions $\{v_i^*\}$ satisfy $\langle v_1^*, v_2^*\rangle = 2\arccos(\langle v_1, v_2\rangle)/\pi$. Specifically, this operation obeys the following properties:

$$\langle u, v\rangle = 0 \Longleftrightarrow \langle u^*, v^*\rangle = 0 \qquad \langle u, v\rangle = 1 - \varepsilon \Longrightarrow \langle u^*, v^*\rangle \geqslant 1 - O(\sqrt{\varepsilon})$$

The tensor product operation on $\{\pm 1\}$-integral vectors, yields a $\{\pm 1\}$-integral vector.

Definition 3. Given two $\{\pm 1\}$-valued functions $u : \mathcal{R}_1 \to \{\pm 1\}$ and $v : \mathcal{R}_2 \to \{\pm 1\}$, the tensor product $u \otimes v : \mathcal{R}_1 \times \mathcal{R}_2 \to \{\pm 1\}$ is defined as $u \otimes v(r_1, r_2) = u(r_1)v(r_2)$.

Observation 2. *For $u, u' : \mathcal{R}_1 \to \{\pm 1\}$ and $v, v' : \mathcal{R}_2 \to \{\pm 1\}$, we have*

$$\begin{aligned}\langle u \otimes v, u' \otimes v'\rangle &= \mathop{\mathbb{E}}_{r_1,r_2}[u \otimes v(r_1, r_2)u' \otimes v'(r_1, r_2)] \\ &= \mathop{\mathbb{E}}_{r_1}[u(r_1)u'(r_1)]\mathop{\mathbb{E}}_{r_2}[v(r_2)v'(r_2)] = \langle u, u'\rangle\langle v, v'\rangle\end{aligned}$$

$\mathbb{F}_q$-integral vectors Let q be a prime. Now, we will define $\mathbb{F}_q$-integral vectors and their tensor products.

Definition 4. A $\mathbb{F}_q$-integral vector $v\colon \mathcal{R} \to \mathbb{F}_q$ is a function from a measure 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. Here are few examples of functions that will be relevant to us:

1. The function $\chi\colon \mathbb{F}_q \to \mathbb{C}^{q-1}$ given by

$$\chi(i) \stackrel{\text{def}}{=} \frac{1}{\sqrt{q-1}}(\omega^{1\cdot i}, \ldots, \omega^{j\cdot i}, \ldots, \omega^{(q-1)i}),$$

   where ω is a primitive q^{th} root of unity. The vector $\chi(i) \in \mathbb{C}^{q-1}$ is the restriction of the i^{th} character function of the group $\mathbb{Z}_q$ to the set $\mathbb{F}_q^*$. It is easy to see that

$$\langle \chi(a), \chi(b)\rangle = \mathop{\mathbb{E}}_{t\in\mathbb{F}_q^*}\left[\omega^{ta} \cdot \omega^{-tb}\right] = \begin{cases} 1 & \text{if } a = b\,, \\ -\frac{1}{q-1} & \text{if } a \neq b\,. \end{cases}$$

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

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

Remark 4. A $\mathbb{F}_q$-integral vector $v \in \mathbb{F}_q^N$ can be thought of as a $\mathbb{F}_q$-valued function over the measure space $([N], \mu)$ where μ is the uniform distribution over $[N]$.

Remark 5. The following notions are equivalent : Collection of $\mathbb{F}_q$-valued functions on some measure space $\mathcal{R} \Longleftrightarrow$ Collection of jointly-distributed, $\mathbb{F}_q$-valued random variables $\Longleftrightarrow$ 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 5. 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

$$(u \otimes_q u')\,(r, t, r', t') \stackrel{\text{def}}{=} t \cdot u(r) + t' \cdot u'(r')\,.$$

Lemma 1. *For any $\mathbb{F}_q$-valued functions $u, v\colon \mathcal{R} \to \mathbb{F}_q$ and $u', v'\colon \mathcal{R}' \to \mathbb{F}_q$,*

$$\langle \psi(u \otimes_q u'), \psi(v \otimes_q v')\rangle = \langle \psi(u), \psi(v)\rangle\langle \psi(u'), \psi(v')\rangle\,.$$

Proof.

$$\begin{aligned}
&\langle \psi(u \otimes_q u'), \psi(v' \otimes_q v')\rangle \\
&= \langle \chi(u \otimes_q u'), \chi(v' \otimes_q v')\rangle \qquad (\text{using } \langle \psi_a, \psi_b\rangle = \langle \chi(a), \chi(b)\rangle) \\
&= \mathop{\mathbb{E}}_{(r,t)} \mathop{\mathbb{E}}_{(r',t')} \mathop{\mathbb{E}}_{\ell \in \mathbb{F}_q^*} \omega^{\ell t u(r) + \ell t' u'(r')} \cdot \omega^{-\ell t v(r) - \ell t' v'(r')} \qquad (\text{by definitons of } \otimes_q \text{ and } \chi) \\
&= \mathop{\mathbb{E}}_{\ell \in \mathbb{F}_q^*} \left(\mathop{\mathbb{E}}_{(r,t)} \omega^{\ell t u(r) - \ell t v(r)} \right) \cdot \left(\mathop{\mathbb{E}}_{(r',t')} \omega^{\ell t' u'(r') - \ell t' v'(r')} \right) \\
&= \mathop{\mathbb{E}}_{\ell \in \mathbb{F}_q^*} \left(\mathop{\mathbb{E}}_{r} \langle \chi(\ell u(r)), \chi(\ell v(r))\rangle \right) \cdot \left(\mathop{\mathbb{E}}_{r'} \langle \chi(\ell u'(r')), \chi(\ell v'(r'))\rangle \right) \\
&= \mathop{\mathbb{E}}_{\ell \in \mathbb{F}_q^*} \langle \chi(\ell u), \chi(\ell v)\rangle \langle \chi(\ell u'), \chi(\ell v')\rangle \\
&= \langle \chi(u), \chi(v)\rangle \langle \chi(u'), \chi(v')\rangle \qquad (\text{using } \langle \chi(\ell a), \chi(\ell b)\rangle = \langle \chi(a), \chi(b)\rangle \text{ for } \ell \in \mathbb{F}_q^*) \\
&= \langle \psi(u), \psi(v)\rangle \langle \psi(u'), \psi(v')\rangle \qquad (\text{using } \langle \psi_a, \psi_b\rangle = \langle \chi(a), \chi(b)\rangle)
\end{aligned}$$

■

Remark 6. Unlike the ordinary tensor operation, the q-ary tensor operation we defined is not associative. Formally, we define the tensoring operation to be right-associative

$$u_1 \otimes_q u_2 \otimes_q \ldots \otimes_q u_{k-1} \otimes_q u_k \stackrel{\text{def}}{=} u_1 \otimes_q \left(u_2 \otimes_q (\ldots (u_{k-1} \otimes_q u_k) \cdots) \right).$$

The lack of associativity will never be an issue in our constructions.

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

Observation 3. *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 $\mathbb{E}_r \langle \psi(a + u(r)), \psi(b + v(r))\rangle = \langle a \otimes u, b \otimes v\rangle$.*

Proof. Using the symmetry assumption, we see that

$$\Pr_{r \sim \mathcal{R}, t, t' \in \mathbb{F}_q^*} \left\{ ta + t'u(r) = tb + t'v(r) \right\} = \Pr_{r \sim \mathcal{R}, t \in \mathbb{F}_q^*} \left\{ a - b = t \cdot \left(v(r) - u(r) \right) \right\} = \Pr_{r \sim \mathcal{R}} \left\{ a - b = v(r) - u(r) \right\} \tag{15}$$

If we let ρ denote this probability, then we have $\langle a \otimes u, b \otimes v\rangle = \rho - (1 - \rho)/(q - 1)$ (using the left-hand side of Eq. (15) as well as $\mathbb{E}_r \langle \psi(a + u(r)), \psi(b + v(r))\rangle = \rho - (1 - \rho)/(q - 1)$ (using the right-hand side of Eq. (15)). ■

The following procedure[2] yields a way to generate $\mathbb{F}_q$-integral vectors from arbitrary vectors. The transformation is inspired by the rounding scheme for Unique Games in Charikar et al. [CMM06].

Observation 4. *Define the function $\zeta \colon \mathcal{G}^q \to \mathbb{F}_q$ on the Gaussian domain as follows:*

$$\zeta(x_1, \ldots, x_q) = \operatorname{argmax}_{i \in [q]} x_i \tag{16}$$

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,

[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.

1. $\langle u, v\rangle = 0 \implies \langle \psi(u^*), \psi(v^*)\rangle = 0$,

2. $\langle u, v\rangle = 1 - \varepsilon \implies \langle \psi(u^*), \psi(v^*)\rangle = 1 - f(\varepsilon, q) = 1 - O(\sqrt{\varepsilon \log q})$.

Proof. To see (1), 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,

$$\langle \psi(u^*), \psi(v^*)\rangle = \mathop{\mathbb{E}}_{r \in \mathcal{G}^{dq}} \left[\psi(u^*(r))\right] \cdot \mathop{\mathbb{E}}_{r \in \mathcal{G}^{dq}} \left[\psi(u^*(r))\right] = 0\,.$$

Assertion 2 follows from Lemma $C.8$ in [CMM06]. ∎

2.5 Properties of Unique games SDP vectors

In this section, we will abstract the definitions and properties of SDP solutions to Unique games. More precisely, the collections of vectors we consider are solutions to the basic SDP relaxation [FL92] for Γ-Max-2Lin(n). Such an SDP solution consists of a collection of sets of n orthonormal vectors. The following definition abstracts the properties of these vectors that are of interest here.

Definition 6. A *nice system of clouds* is a collection $\mathcal{B}$ of subsets of $\mathbb{R}^d$ with the following properties:

- Every set $B \in \mathcal{B}$ consists of n unit vectors. We refer to the vector sets $B \in \mathcal{B}$ as *clouds*.
- *Near Orthogonality:* For every $B \in \mathcal{B}$, and for every unit vector $u \in \mathbb{R}^d$, we have $\sum_{v \in B} \langle u, v\rangle^2 \leqslant 3/2$.
- *Matching Property:* For every pair of clouds $A, B \in \mathcal{B}$, there exists a perfect matching M between A and B such that for every matched pair $(u, v) \in M$, we have $|\langle u, v\rangle| = \max_{a \in A, b \in B} |\langle a, b\rangle|$.
- *(Symmetrized) ℓ_2^2-triangle inequality:* Every triple $u, v, w \in \bigcup \mathcal{B}$ satisfies the ℓ_2^2-triangle inequality, that is, $\|u - v\|^2 \leqslant \|u - w\|^2 + \|w - v\|^2$. If two vectors in $\bigcup \mathcal{B}$ have negative inner product, we require in addition that every triple in $\pm \bigcup \mathcal{B}$ satisfies the ℓ_2^2-triangle inequality. Here, $\pm \bigcup \mathcal{B}$ denotes the union of $\mathcal{B}$ and its reflection $- \bigcup \mathcal{B} = \{-u \mid u \in \bigcup \mathcal{B}\}$.

Remark 7.
- *Near Orthogonality:* For most of our applications, the vectors in every cloud will be completely orthogonal. For the results of this section, the weaker condition of near-orthogonality will not complicate the calculations.
- *Matching Property:* Any SDP solution for a Γ-Max-2Lin(n) instance can be made to satisfy this property for symmetry reasons.
- *(Symmetrized) ℓ_2^2-triangle inequality:* In our applications, this property will always be satisifed for simple reasons (all vectors will be integral in the sense of § 2.4). For this section, the ℓ_2^2-triangle inequality is not essential but it simplifies some computations.

Stating the lemmas in the above abstraction makes them useful even for the Balanced Separator integrality gap where there is no explicit unique games instance. The abstraction also highlights the properties of a Unique Games SDP solution that are relevant for our construction.

Definition 7. 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 near-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 2. *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 $a \in A$.*

Remark 8. The collection $\mathcal{B}$ succinctly encodes a Unique Games instance. For a graph $G = (\mathcal{B}, E)$ on $\mathcal{B}$, the goal is to find a labeling $\{\ell_A \in A\}_{A\in\mathcal{B}}$ (a labeling can be seen as a system of representatives for the clouds in $\mathcal{B}$) so as to maximize the probability

$$\Pr_{(A,B)\in E}\left\{\ell_A = \pi_{A\leftarrow B}(\ell_B)\right\}.$$

2.5.1 Tensoring

Lemma 2. *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 near-orthogonality, $\sum_{a\in B}\langle a,b\rangle^2 \leqslant 3/2$ for every $b \in B$. Hence, $\langle a,b\rangle^2 \leqslant 3/4$ for all $a \neq \pi_{A\leftarrow B}(b)$. Thus,

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

■

The notation $X = Y \pm Z$ means that $|X - Y| \leqslant Z$.

Corollary 1. *For $t \in \mathbb{N}$ and every pair of clouds $A, B \in \mathcal{B}$,*

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

Remark 9. The left-hand side in the corollary is the inner product of the vectors $1/\sqrt{n}\sum_{u\in A} u^{\otimes t}$ and $1/\sqrt{n}\sum_{v\in B} v^{\otimes t}$. If t is even, then we can replace the right-hand side by $\rho(A,B)^t$. This fact that the functional $\rho(A,B)^t$ is closely approximated by inner products averaged-tensored vectors has implicitly been used in [KV05] and was explicitly noted in [AKK+08, Lemma 2.2].

2.6 Local Distributions for Unique Games

In this section, we will construct local distribution over labelings to a unique games instance. We will state these construction in the abstract setting of a collection of vectors as in Definition 6. Stating the results at this level of abstraction makes them useful even for the Balanced Separator integrality gap where there is no explicit unique games instance.

The following facts are direct consequences of the (symmetrized) ℓ_2^2-triangle inequality.

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

Fact 3. *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 4. *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 [Tre08, AKK+08].

Definition 8. A set $\mathcal{S} \subseteq \mathcal{B}$ is *consistent* if

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

Lemma 3. *If* $\mathcal{S} \subseteq \mathcal{B}$ *is consistent, there exists bijections* $\{\pi_A : [n] \to A\}_{A \in \mathcal{S}}$ *such that*

$$\forall A, B \in \mathcal{S}. \qquad \pi_B = \pi_{B \leftarrow A} \circ \pi_A\,.$$

Proof. We can construct the bijections in a greedy fashion: Start with an arbitrary cloud $C \in \mathcal{S}$ and choose an arbitrary bijection $\pi_C : [n] \to C$. 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). Since $(1 - 2\eta)^2 > 1 - 4\eta = 3/4$, Fact 2 (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.

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 definiton:

Definition 9. A set $\mathcal{S} \subseteq \mathcal{B}$ is *consistent* with respect to a partition P of $\mathcal{B}$ (denoted $\mathsf{Consistent}(\mathcal{S}, P)$) if

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

We use $\mathsf{Inconsistent}(\mathcal{S}, P)$ to denote the event that $\mathcal{S}$ is not consistent with P. The following is a corollary of Lemma 3.

Corollary 2. *Let* P *be a partition of* $\mathcal{B}$ *and let* $\mathcal{S} \subseteq \mathcal{B}$*. If* $\mathsf{Consistent}(\mathcal{S}, P)$*, then there exists bijections* $\{\pi_A : [n] \to A \mid A \in \mathcal{S}\}$ *such that*

$$\forall C \in P. \quad \forall A, B \in C \cap \mathcal{S}. \qquad \pi_B = \pi_{B \leftarrow A} \circ \pi_A\,.$$

2.6.1 Greedy Construction

Let $\mathcal{S} \subseteq \mathcal{B}$ with $|S| \leqslant R$. Consider the graph G on $\mathcal{S}$ in which two clouds $A, B \in \mathcal{S}$ are connected by an edge if $\rho(A, B) \geqslant 1 - 1/16R$. Let P denote the partition of $\mathcal{S}$ induced by the connected components of G. Every pair of clouds A, B in the same connected component are connected by a path of length at most R, hence by triangle inequality (Fact 4) we have $\rho(A, B) \geqslant 1 - 1/16$ and therefore $\mathcal{S}$ is consistent with respect to the partition P. Along with Corollary 2, this observation implies the following lemma.

Lemma 4. *For* $R \in \mathbb{N}$*, let* $\mathcal{S} \subseteq \mathcal{B}$ *be a set of at most* R *clouds. Then, there exists bijections* $\{\pi_A : [n] \to A\}_{A \in \mathcal{S}}$ *such that for all* $A, B \in \mathcal{S}$*,*

$$\rho(A, B) \geqslant 1 - 1/16R \quad \Longrightarrow \quad \pi_B = \pi_{B \leftarrow A} \circ \pi_A\,.$$

2.6.2 Local Distributions via Geometric Decomposition

The lemma relies on the fact that the correlations $\rho(A,B)$ behave up to a small errors like inner products of real vectors. This insight has also been used in UG algorithms[AKK+08].

Lemma 5. *For every $t \in \mathbb{N}$, there exists a distribution over partitions P of $\mathcal{B}$ such that*

– *if $\rho(A,B) \geqslant 1-\varepsilon$, then*
$$\mathbb{P}\mathrm{r}\{P(A)=P(B)\} \geqslant 1-O(t\sqrt{\varepsilon})\,.$$

– *if $\rho(A,B) \leqslant 1-1/16$, then*
$$\mathbb{P}\mathrm{r}\{P(A)=P(B)\} \leqslant (3/4)^t\,.$$

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
$$\boldsymbol{v}_B := \tfrac{1}{\sqrt{n}} \sum_{v\in B} v^{\otimes s}\,.$$
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 $H_1,\ldots,H_t$).

Since s is even, Corollary 1 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,
$$\mathbb{P}\mathrm{r}\{P(A)=P(B)\} \leqslant (3/4)^t\,.$$

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,
$$\mathbb{P}\mathrm{r}\{P(A)=P(B)\} \geqslant \left(1-O(\sqrt{\varepsilon})\right)^t \geqslant 1-O(t\sqrt{\varepsilon})\,.$$

■

Remark 10. Using a more sophisticated construction, we can improve the bound $1-O(t\sqrt{\varepsilon})$ to $1-O(\sqrt{t\varepsilon})$.

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

Corollary 3. *The distribution over partitions from Lemma 5 satisfies the following property: For every set $\mathcal{S}\subseteq\mathcal{B}$,*
$$\mathbb{P}\mathrm{r}\left\{\mathsf{Inconsistent}(\mathcal{S},P)\right\} \leqslant |S|^2\cdot(3/4)^t$$

Remark 11. Using a slightly more refined argument (triangle inequality), we could improve the bound $R^2\cdot(3/4)^t$ to $R\cdot(3/4)^t$.

3 Construction of SDP Solutions for E2LIN(q)

Let $\mathcal{B}$ be a collections as in Definition 6 with the following additional properties:

– All vectors in $\bigcup \mathcal{B}$ are $\mathbb{F}_q$-integral. We endow the vectors in $\mathcal{B}$ with the inner product

$$\langle u, v\rangle_\psi \stackrel{\text{def}}{=} \langle \psi(u), \psi(v)\rangle .$$

– all clouds of $\mathcal{B}$ are exactly orthogonal (as opposed to nearly orthogonal as in Section).

In this section, we construct SDP vectors and local distributions for $\mathcal{B} \times \mathbb{F}_q^n$. The set $\mathcal{B} \times \mathbb{F}_q^n$ correspond to the set of vertices in the instance obtained by applying a q-ary long code based reduction on the unique games instance encoded by $\mathcal{B}$. 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 5 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.

3.1 Vector Solution

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

$$\boldsymbol{W}_j^{B,x}(r) = \tfrac{1}{\sqrt{n}} \sum_{b\in B} \psi(x_b - j + b^{\otimes t}(r)) \tag{17}$$

$$\boldsymbol{U}_j^{B,x}(P, \alpha, r) = P_\alpha(B) \cdot \boldsymbol{W}_j^{B,x}(r) \tag{18}$$

$$\boldsymbol{V}_j^{B,x} = \tfrac{1}{q}\boldsymbol{V}_0 + \tfrac{\sqrt{q-1}}{q}\boldsymbol{U}_j^{B,x} \tag{19}$$

Here $\mathcal{R}$ 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 $B \in P_\alpha$. Further, $\boldsymbol{V}_0$ is a unit vector orthogonal to all the vectors $\boldsymbol{U}_j^{B,x}$.

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)

$$\begin{aligned}
\langle \boldsymbol{V}_i^{A,x}, \boldsymbol{V}_j^{B,y}\rangle &= \tfrac{1}{q^2} + \tfrac{q-1}{q^2}\langle \boldsymbol{U}_i^{A,x}, \boldsymbol{U}_j^{B,y}\rangle \\
&= \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \mathop{\mathbb{E}}_{P\sim\mathcal{P}} \textstyle\sum_{\alpha=1}^T P_\alpha(A)P_\alpha(B)\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle \\
&= \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \mathop{\Pr}_{P\sim\mathcal{P}} \{P(A) = P(B)\} \langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle
\end{aligned} \tag{20}$$

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

$$\begin{aligned}
\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle &= \tfrac{1}{n} \sum_{a\in A, b\in B} \mathop{\mathbb{E}}_{r\sim\mathcal{R}} \langle x_a - i + a^{\otimes t}(r), y_b - j + b^{\otimes t}(r)\rangle_\psi \\
&= \tfrac{1}{n} \sum_{a\in A, b\in B} \langle (x_a - i) \otimes a^{\otimes t}, (y_b - j) \otimes b^{\otimes t}\rangle_\psi \quad \text{(by Observation 3)} \\
&= \tfrac{1}{n} \sum_{a\in A, b\in B} \langle \psi(x_a - i), \psi(y_b - j)\rangle \langle a, b\rangle_\psi^t \quad \text{(by Lemma 1)}
\end{aligned} \tag{21}$$

In the next lemma, we verify that the vector solution $\{\boldsymbol{V}_j^{B,x}\}$ satisfies the conditions in Fact 1 (which need in order to apply the smoothening theorems from Section 2.3)

Lemma 6. *The vector solution* $\{\boldsymbol{V}_j^{A,x}\}_{A\in\mathcal{B},x\in\mathbb{F}_q^A,j\in\mathbb{F}_q}$ *satisfies the following properties:*

1. *For every vertex* (B,x)*, we have* $\sum_i \boldsymbol{V}_i^{B,x} = \boldsymbol{V}_0$*.*
2. *For every vertex* (B,x)*, we have* $\langle \boldsymbol{V}_i^{B,x}, \boldsymbol{V}_j^{B,x}\rangle = \frac{1}{q}\cdot \mathbf{1}[i=j]$ *for all* $i,j\in\mathbb{F}_q$*.*
3. *For any two vertices* (A,x) *and* (B,y)*, we have* $\langle \boldsymbol{V}_i^{A,x}, \boldsymbol{V}_j^{B,y}\rangle \geqslant 0$ *for all* $i,j\in\mathbb{F}_q$*.*

Proof. Assertion 1: Observe that for all $r\in\mathcal{R}$,

$$\left(\sum_{j\in\mathbb{F}_q} \boldsymbol{W}_i^{B,x}\right)(r) = \frac{1}{\sqrt{n}}\sum_{b\in B}\left(\sum_{i\in[q]}\psi\Big(x_b - j + b^{\otimes t}(r)\Big)\right) = 0\,.$$

Here we use the fact that $\sum_{j\in\mathbb{F}_q}\psi(a-j)=0$ for all $a\in\mathbb{F}_q$. From the above equation and definition of $\boldsymbol{V}_i^{B,x}$, it follows that $\sum_i \boldsymbol{V}_i^{B,x} = \boldsymbol{V}_0$.
Assertion 2: From (21) and the orthonormality of B, we see that

$$\langle \boldsymbol{W}_i^{B,x}, \boldsymbol{W}_j^{B,x}\rangle = \langle \psi(x_b-i),\psi(x_b-j)\rangle = \langle\psi(i),\psi(j)\rangle\,.$$

Hence, using (20), we get

$$\langle \boldsymbol{V}_i^{B,x}, \boldsymbol{V}_j^{B,x}\rangle = \tfrac{1}{q^2} + \tfrac{q-1}{q^2}\langle\psi(i),\psi(j)\rangle = \tfrac{1}{q}\mathbf{1}[i=j]\,.$$

Assertion 3: We can lower bound the inner products $\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle$ by

$$\begin{aligned}\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle &\geqslant -\tfrac{1}{q-1}\cdot\tfrac{1}{n}\sum_{a\in A,b\in B}\langle a,b\rangle_\psi^t \qquad \text{(using (21) and } \langle\psi(\cdot),\psi(\cdot)\rangle \geqslant -\tfrac{1}{q-1}\text{)}\\ &\geqslant -\tfrac{1}{q-1} \qquad \text{(using orthonormality of } A \text{ and } B \text{ and } t\geqslant 2)\,.\end{aligned}$$

Therefore, $\langle \boldsymbol{V}_i^{A,x}, \boldsymbol{V}_j^{B,y}\rangle \overset{(20)}{=} \frac{1}{q^2} + \frac{q-1}{q^2}\Pr_P\{P(A)=P(B)\}\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle \geqslant \frac{1}{q^2} - \frac{1}{q^2}\Pr_P\{P(A)=P(B)\}\geqslant 0$. ■

3.2 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 1). 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.

Remark 12. In the construction in Figure 1, the steps 6–7 are not strictly necessary, but they simplify some of the following calculations. Specifically, we could use the $\mathbb{F}_q$-assignment $\{F^{B,x}\}_{(B,x)\in S}$ to define the local distribution for the vertex set S. The resulting collection of local distributions could be extended to an approximately feasible SDP solution (albeit using a slightly different vector solution).

We need the following two simple observations.

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 5. Let α_A, α_B denote the indices of sets in the partition P that contain $A, B \in \mathcal{S}$ respectively.

2. If $\mathsf{Inconsistent}(\mathcal{S}, P)$ then output a uniform random $\mathbb{F}_q$-assignment to $S = \mathcal{S} \times \mathbb{F}_q^n$. Specifically, set

$$Z^{(B,x)} = \text{ uniform random element from } \mathbb{F}_q \qquad \forall B \in \mathcal{S}, x \in \mathbb{F}_q^n .$$

Choosing Consistent Representatives:

4. If $\mathsf{Consistent}(\mathcal{S}, P)$ then by Corollary 2, for every part $\mathcal{S}_\alpha = P_\alpha \cap \mathcal{S}$, there exists bijections $\Pi_{\mathcal{S}_\alpha} = \{\pi_B \colon [n] \to B \mid B \in \mathcal{S}_\alpha\}$ 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 1: Local distribution over $\mathbb{F}_q$-assignments

Observation 5. *For all $a, b \in \mathbb{F}_q$, we have*

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

Proof. If $a - i = b - j$ then both LHS and RHS are equal to $1/q$, otherwise both are equal to 0. ■

Observation 6. *Fix $a, b \in \mathbb{F}_q$, over a random choice of $h_1, h_2 \in \mathbb{F}_q$,*

$$\mathop{\mathbb{E}}_{h_1, h_2 \in \mathbb{F}_q}[\langle \psi(a + h_1), \psi(b + h_2)\rangle] = 0 \,.$$

Proof. Follows easily from the fact that $\langle \psi(i), \psi(j)\rangle = 1$ if $i = j$ and $-1/q-1$ otherwise. ■

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

Lemma 7. *For any two vertices $(A, x), (B, y) \in S$,*

$$\Pr_{Z \sim \mu_S}\left[Z^{A,x} = i \wedge Z^{B,y} = j\right] = \langle \boldsymbol{V}_i^{A,x}, \boldsymbol{V}_j^{B,y}\rangle \quad \pm 10|\mathcal{S}|^2 (3/4)^{t/2} \,.$$

Proof. Firstly, since $\Pr[\mathsf{Consistent}(\mathcal{S}, P)] \geqslant 1 - |S|^2 (3/4)^t$ (by Corollary 3),

$$\Pr_{\mu_S}\left[Z^{A,x} = i \wedge Z^{B,y} = j\right] = \Pr_{\mu_S}\left[Z^{A,x} = i \wedge Z^{B,y} = j \,\middle|\, \mathsf{Consistent}(\mathcal{S}, P)\right] \quad \pm |S|^2 (3/4)^t \,. \tag{22}$$

Using Observation 5, and the definition of $Z^{A,x}$ and $Z^{B,y}$ we can write

$$\begin{aligned}
&\Pr_{\mu_S}\left[Z^{A,x} = i \wedge Z^{B,y} = j \,\middle|\, \mathsf{Consistent}(\mathcal{S}, P)\right] \\
&\quad = \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \mathop{\mathbb{E}}_{P,H,L,r}\left[\langle \psi(F^{A,x} + h_A - i), \psi(F^{B,y} + h_B - j)\rangle \,\middle|\, \mathsf{Consistent}(\mathcal{S}, P)\right] \,.
\end{aligned} \tag{23}$$

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 6, we can write

$$\begin{aligned}
&\mathop{\mathbb{E}}_{P,H,L,r}\left[\langle \psi(F^{A,x} + h_A - i), \psi(F^{B,y} + h_B - j)\rangle \,\middle|\, \mathsf{Consistent}(\mathcal{S}, P)\right] \\
&\quad = \mathop{\mathbb{E}}_{P,L,r}\left[\mathbf{1}(\alpha_A = \alpha_B)\langle \psi(F^{A,x} - i), \psi(F^{B,y} - j)\rangle \,\middle|\, \mathsf{Consistent}(\mathcal{S}, P)\right] \,.
\end{aligned} \tag{24}$$

Let P be a partition such that $\mathsf{Consistent}(\mathcal{S}, P)$ and $\alpha_A = \alpha_B = \alpha$. The bijections π_A, π_B (see step 4 Figure 1) 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,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{L}\mathop{\mathbb{E}}_{r}\left[\langle \psi(F^{A,x}(P, L, r) - i), \psi(F^{B,y}(P, L, r) - j)\rangle\right] \\
&= \mathop{\mathbb{E}}_{\ell_\alpha}\mathop{\mathbb{E}}_{r}\left[\langle \psi(x_{\pi_A(\ell_\alpha)} - i + \pi_A(\ell_\alpha)^{\otimes t}(r)), \psi(y_{\pi_B(\ell_\alpha)} - j + \pi_B(\ell_\alpha)^{\otimes t}(r))\rangle\right] \\
&= \frac{1}{n} \sum_{\substack{a \in A, b \in B \\ a = \pi_{A \leftarrow B}(b)}} \mathop{\mathbb{E}}_{r}\langle \psi(x_a - i + a^{\otimes t}(r)), \psi(y_b - j + b^{\otimes t}(r))\rangle \qquad (\text{using } \pi_A = \pi_{A \leftarrow B} \circ \pi_B) \\
&= \frac{1}{n} \sum_{\substack{a \in A, b \in B \\ a = \pi_{A \leftarrow B}(b)}} \langle \psi(x_a - i), \psi(y_b - j)\rangle \cdot \langle a, b\rangle_\psi^t \qquad (\text{using Observation 3 and Lemma 1}) \\
&= \langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle \pm 2 \cdot (3/4)^{t/2} \qquad (\text{using Eq. (21) and Lemma 2}) \,.
\end{aligned}$$

Combining the last equation with the previous equations (22)–(24), we can finish the proof

$$\begin{aligned}
&\Pr_{\mu_S}\left[Z^{A,x} = i \wedge Z^{B,y} = j\right] \\
&\quad = \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \mathop{\mathbb{E}}_{P}\left[\mathbf{1}(\alpha_A = \alpha_B) \,\middle|\, \mathsf{Consistent}(\mathcal{S}, P)\right] \langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle \quad \pm (|\mathcal{S}|^2(3/4)^t + 2\cdot(3/4)^{t/2}) \\
&\quad = \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \Pr_{P}\left[P(A) = P(B)\right] \cdot \langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle \quad \pm 10|\mathcal{S}|^2(3/4)^{t/2} \\
&\qquad\qquad\qquad (\text{using } \Pr\{\mathsf{Consistent}(\mathcal{S}, P)\} \geqslant 1 - |\mathcal{S}|^2(3/4)^t \text{ and } |\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle| \leqslant 1) \\
&= \langle \boldsymbol{V}_i^{A,x}, \boldsymbol{V}_j^{B,y}\rangle \quad \pm 10|\mathcal{S}|^2(3/4)^{t/2} \qquad (\text{using Eq. (20)}).
\end{aligned}$$

■

Lemma 8. *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,*

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

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 1. *If $\mathsf{Consistent}(\mathcal{S}', P)$ and $\mathsf{Consistent}(\mathcal{S}, P)$, then $\mu_{S'|P} = \mathsf{margin}_{S'}\, \mu_{S|P}$.*

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 $\mathsf{margin}_{S'}\, \mu_{S|P}$. 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:

$$\pi_{A\to B} \circ \pi'_A(\ell) = \pi'_B(\ell) \qquad\qquad \pi_{A\to B} \circ \pi_A(\ell) = \pi_B(\ell) \qquad\qquad \forall A, B \in \mathcal{S}'_\alpha, \ell \in [n] .$$

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 8, we can write

$$\begin{aligned}
\|\mu_{S'} - \mathsf{margin}_{S'}\, \mu_S\|_1 &= \left\|\mathop{\mathbb{E}}_{P} \mu_{S'|P} - \mathop{\mathbb{E}}_{P} \mathsf{margin}_{S'}\, \mu_{S|P}\right\|_1 \\
&\leqslant \mathop{\mathbb{E}}_{P}\left[\|\mu_{S'|P} - \mathsf{margin}_{S'}\, \mu_{S|P}\|_1\right] \qquad (\text{using Jensen's inequality}) \\
&= \Pr[\mathsf{Inconsistent}(\mathcal{S}, P)] \cdot \mathop{\mathbb{E}}_{P}\left[\|\mu_{S'|P} - \mathsf{margin}_{S'}\, \mu_{S|P}\|_1 \,\middle|\, \mathsf{Inconsistent}(\mathcal{S}, P)\right].
\end{aligned}$$

The first step uses that the operator $\text{margin}_{S'}$ is linear. The final step in the above calculation makes use of Claim 1. The lemma follows by observing that $\Pr[\text{Inconsistent}(\mathcal{S}, P)] \leqslant |\mathcal{S}|^2(3/4)^t$ and $\|\mu_{S'|P} - \text{margin}_{S'}\mu_{S|P}\|_1 \leqslant 2$. ■

The next corollary follows from the previous lemma (Lemma 8) and the triangle inequality.

Corollary 4. *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,*

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

4 Integrality Gap Instance for Unique Games

In this section, we will exhibit a strong SDP integrality gap for the E2Lin_q problem. Recall that the E2Lin_q problem is a special case of Unique games. To this end, we follow the approach of Khot–Vishnoi [KV05] to construct the gap instance.

Khot et al. [KKMO07] show a UGC-based hardness result for the E2Lin_q problem. Specifically, they exhibit a reduction $\Phi_{\gamma,q}$ that maps a Unique Games instance Γ to an E2Lin_q instance $\Phi_{\gamma,q}(\Gamma)$ such that the following holds: For every $\gamma > 0$ and all $q \geqslant q_0(\gamma)$,

- Completeness: If Γ is $1-\eta$-satisfiable then $\Phi_{\gamma,q}(\Gamma)$ is $1-\gamma-o_{\eta,\delta}(1)$ satisfiable.
- Soundness: If Γ 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 $\Phi_{\gamma,q}(\Gamma)$.

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

Our starting point is an integrality gap instance Γ for the basic semidefinite program for Unique Games. Surprisingly, on executing the [KKMO07]-reduction on the instance Γ yields the strong SDP gap instance $\Phi_{\gamma,q}(\Gamma)$ for E2Lin_q. (For the sake of completeness, we describe in Figure 2 this reduction from Unique Games to E2Lin_q.)

Definition 10 (Weak SDP solutions and weak gap instances)**.** Let $\Gamma = (V, E, \{\pi_e : [n] \to [n]\}_{e\in E})$. We say a collection $\mathcal{B} = \{B_u\}_{u\in V}$ is a *weak SDP solution of value* $1-\eta$ for Γ 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. 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 $\pi^{(1)}, \ldots, \pi^{(n)} : B_u \to B_v$ such that for all $i \in [n], b, b' \in B_u$, we have $\langle b, \pi^{(i)}(b)\rangle = \langle b', \pi^{(i)}(b')\rangle$.
3. For every edge $e = (u, v) \in E$, the vector sets B_u and B_v have signifcant correlation under the permutation $\pi = \pi_e$. Specifically,
$$\forall \ell \in [n]. \quad \langle b_{u,\ell}, b_{v,\pi(\ell)}\rangle^2 \geqslant 0.99.$$
4. The collection $\mathcal{B}$ of orthornormal sets is a good SDP solution for Γ, in the sense that
$$\mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}} \frac{1}{n}\sum_{\ell\in[n]}\langle b_{w,\pi(\ell)}, b_{w',\pi'(\ell)}\rangle \geqslant 1-\eta.$$

E2Lin$_q$ Hardness Reduction [KKMO07]

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

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

- Pick a random vertex $v \in V$. Choose two random neighbours $w, w' \in N(v) \subseteq V$. 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:

$$y_i = \begin{cases} x_i \text{ with probability } 1-\gamma \\ \text{uniform random element from } \mathbb{F}_q \text{ with probability } \gamma \end{cases}$$

- 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 2: Reduction from Unique Games to E2Lin$_q$

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

Remark 13. The weak gap instances defined here are fairly natural objects. In fact, if Υ is an instance of Γ-Max-2Lin(n) with $\mathrm{sdp}(\Upsilon) \geqslant 1-\eta$ and $\mathrm{opt}(\Upsilon) \leqslant \delta$, it is easy to construct a corresponding weak gap instance Υ'. The idea is to start with an optimal SDP solution for Υ, symmetrize it (with respect to the group Γ), and delete all edges of Υ that contribute less than $\sqrt{3/4}$ to the SDP objective.

We observe the following consequence of Fact 2 and item 3 of Definition 10.

Observation 7. *If $\mathcal{B} = \{B_u\}_{u \in V}$ is a weak SDP solution for $\Gamma = (V, E, \{\pi_e\}_{e \in E})$, 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. 7) give rise to the same matching between the vector sets B_w and $B_{w'}$,*

$$\pi(i) = j \quad \iff \quad \pi_{B_{w'} \leftarrow B_w}(b_{w,i}) = b_{w',j}\,.$$

The previous observation implies that in a weak gap instance Γ the collection of permutations $\{\pi_e\}_{e \in E}$ is already determined by the geometry of the vector sets in a weak SDP solution $\mathcal{B}$.

There are a few explicit constructions of weak gap instances of Unique Games, most prominently the Khot–Vishnoi instance [KV05]. In particular, the following observation is a restatement of Theorem 9.2 and Theorem 9.3 in [KV05].

Observation 8. *For all $\eta, \delta > 0$, there exists a weak $(1-\eta, \delta)$-gap instance with $2^{2^{O(\log(1/\delta)/\eta)}}$ vertices.*

Observation 4 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. Here we use again $\langle\cdot,\cdot\rangle_\psi := \langle\psi(\cdot), \psi(\cdot)\rangle$ as inner product for $\mathbb{F}_q$-integral vectors.

Lemma 9. *Let $\Gamma = (V, E, \{\pi_e\}_{e\in E})$ 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 Γ' which is obtained from Γ by deleting $O(\sqrt{\eta \log q})$ edges.*

Proof. Let $\mathcal{B}$ be a weak SDP solution for Γ of value $1-\eta$. By applying the transformation from Observation 4 to the vectors in $\mathcal{B}$, we obtain a collection $\mathcal{B}' = \{B'_u\}_{u\in V}$ of sets of $\mathbb{F}_q$-integral vectors. For every $u \in V$, the vectors in B'_u are orthonormal. Furthermore, any two sets B'_u, B'_v in $\mathcal{B}'$ satisfy the strong matching property (using the facts that the original sets B_u, B_v satisfy this property and that $\langle b'_{u,i}, b'_{v,j}\rangle_\psi$ is a function of $\langle b_{u,i}, b_{v,j}\rangle$).

Let $\eta_{v,w,w',\ell} = 1 - \langle b_{w,\pi(\ell)}, b_{w',\pi'(\ell)}\rangle$. Using Jensen's inequality, we can verify that the value of the SDP solution $\mathcal{B}'$ is high,

$$\mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}} \frac{1}{n}\sum_{\ell\in[n]} \langle b'_{w,\pi(\ell)}, b'_{w',\pi'(\ell)}\rangle_\psi \geqslant \mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}} \frac{1}{n}\sum_{\ell\in[n]} 1 - O\Big(\sqrt{\eta_{v,w,w',\ell}\log q}\Big) \quad \text{(by Obs. 4)}$$

$$\geqslant 1 - O(\sqrt{\eta \log q}) \qquad \text{(using Jensen's inequality)}\,.$$

So far, we verified that $\mathcal{B}'$ satisfies all requirements of a weak SDP solution besides item 3 of Definition 10. We can ensure that this condition is also satisfied by deleting all edges from E where the condition is violated. Using standard averaging arguments, it is easy to see that the matching property and the high SDP value imply that this condition is satisfied for all but at most an $O(\sqrt{\eta \log q})$ fraction of edges. ∎

Theorem 9. *Let Γ 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 $\Phi_{\gamma,q}(\Gamma)$ such that*

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

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

Proof. Suppose Γ is given by the vertex set V, the edge set $E \subseteq V \times V$, and the collection of permutations $\{\pi_e\}_{e\in E}$. Using Lemma 9, 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 Γ.

We construct a vector solution $\{\boldsymbol{V}_i^{B,x} \mid i \in \mathbb{F}_q, B \in \mathcal{B}, x \in \mathbb{F}_q^n\}$ and local distributions $\{\mu_S \mid S \subseteq \mathcal{B} \times \mathbb{F}_q^n\}$ as defined in the previous section (§3).

Note that since each set $B \in \mathcal{B}$ correspond to a vertices in $u \in V$, we can view these vectors and local distributions as an SDP solution for the $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Gamma)$. Specifically, we make the identifications $\boldsymbol{V}_i^{u,x} := \boldsymbol{V}_i^{B_u,x}$ and $\mu_S := \mu_{\{(B_u,x)\mid(u,x)\in S\}}$ for all $u \in V$, $x \in \mathbb{F}_q^n$, and sets $S \subseteq V \times \mathbb{F}_q^n$.

Lemma 7 and Corollary 4 show that this SDP solution is ε-infeasible for SA_R and LH_R, where $\varepsilon = O(R^2 \cdot (3/4)^{t/2})$. The value of the SDP solution for $\Phi_{\gamma,q}(\Gamma)$ (see Fig. 2) is given by

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

Using Eq. (20)–(21),

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

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 7, 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 2,

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

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

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

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

$$\begin{aligned}&\mathop{\mathbb{E}}_{v\in V}\ \mathop{\mathbb{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}}\ \mathop{\mathbb{E}}_{\{x,y\}}\ \mathop{\mathbb{E}}_{c\in\mathbb{F}_q}\sum_{i=1}^q\langle \boldsymbol{V}_i^{w,(x\circ\pi+c\cdot\mathbb{1})}, \boldsymbol{V}_{i-c}^{w',y\circ\pi'}\rangle\\ &= \mathop{\mathbb{E}}_{v\in V}\ \mathop{\mathbb{E}}_{w,w'\in N(v)}\left[\tfrac{1}{q^2}+\tfrac{q-1}{q^2}\Pr_{P\sim\mathcal{P}}[P(B_w)=P(B_{w'})]\cdot q\cdot\rho(B_w,B_{w'})^t(1-\gamma)\right] \qquad \pm O(\varepsilon)\\ &\geqslant (1-\gamma)\mathop{\mathbb{E}}_{v\in V}\ \mathop{\mathbb{E}}_{w,w'\in N(v)}\Pr_{P\sim\mathcal{P}}[P(B_w)=P(B_{w'})]\,\rho(B_w,B_{w'})^t \qquad \pm O(\varepsilon)\\ &\geqslant (1-\gamma)\mathop{\mathbb{E}}_{v\in V}\ \mathop{\mathbb{E}}_{w,w'\in N(v)}(1-O(t\sqrt{\eta_{w,w'}})) \qquad \pm O(\varepsilon) \qquad \text{(using Lemma 5)}\end{aligned}$$

Using Jensen's inequality and the fact that $\mathbb{E}_{v,w,w'}\,\eta_{v,w,w'} = O(\sqrt{\eta\log q})$ (Lemma 9), we see that the 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 smoothening the SDP solution using Theorem 7, 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 smoothening the SDP solution using Theorem 8, 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 1–2. Using Theorem 9 with the Khot–Vishnoi integrality gap instance (Lemma 8), 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 $R = O((\log\log N)^{1/4})$.

5 Construction of SDP Solutions for Balanced Separator

Let $\mathcal{B}$ be a collection of sets of vectors as in Definition 6 with the following additional property:

Integrality: all vectors in $\bigcup\mathcal{B}$ are $\{\pm1\}$-valued functions on a common measure space $\mathcal{R}$.

In this section, for $R \in \mathbb{N}$, we will construct a global vector solution $\{\boldsymbol{v}_B\}_{B \in \mathcal{B}}$ for the clouds in $\mathcal{B}$, and a collection of local distributions $\{\mu_{\mathcal{S}} \colon \{\pm 1\}^{\mathcal{S}} \to \mathbb{R}_+ \mid \mathcal{S} \subseteq \mathcal{B}, |\mathcal{S}| \leqslant R\}$ such that the inner products of the vectors approximately match the second-moments of the local distributions. Using the smoothening trick, we will be able to get closely-related vectors and local distributions such that the inner products exactly match the second-moments (Lemma 10).

In the next section (§6), we will show how to use these vectors and local distributions to construct integrality gaps for a strong SDP relaxation of Balanced Separator.

Global vector solution. For every cloud $B \in \mathcal{B}$, define a vector $\boldsymbol{v}_B$ as

$$\boldsymbol{v}_B := \tfrac{1}{\sqrt{n}} \sum_{v \in B} v^{\otimes t} .$$

Local distributions. Let $\mathcal{S} \subseteq B$ with $|\mathcal{S}| \leqslant R$. By Lemma 4, there exists bijections $\{\pi_B \colon [n] \to B \mid B \in \mathcal{S}\}$ such that for all $A, B \in \mathcal{S}$

$$\rho(A,B) \geqslant 1 - 1/16R \quad \Longrightarrow \quad \pi_B = \pi_{B \leftarrow A} \circ \pi_A . \tag{25}$$

For $t \in \mathbb{N}$, we define a distribution $\mu_{\mathcal{S}}$ over assignments $\{\pm 1\}^{\mathcal{S}}$ for $\mathcal{S}$:

1. Pick r uniformly at random from $[n]$.
2. Independenly, draw t samples $r_1, \ldots, r_t$ from the distribution $\mathcal{R}$.
3. For every cloud $B \in \mathcal{S}$, set

$$v_B := \pi_B(r) \quad \text{and} \quad x_B := v_B(r_1) \cdot \ldots \cdot v_B(r_t) .$$

4. Output the assignment $x = (x_B)_{B \in \mathcal{S}} \in \{\pm 1\}^{\mathcal{S}}$.

Let us compute the second-moment matrix of this distribution: For $A, B \in \mathcal{S}$, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{x \sim \mu_{\mathcal{S}}} x_A x_B &= \mathop{\mathbb{E}}_{r \in [n]} \left[\mathop{\mathbb{E}}_{r_1,\ldots,r_t \sim \mathcal{R}} \left[\textstyle\prod_{\tau=1}^t v_A(r_\tau) \cdot v_B(r_\tau) \,\middle|\, v_A = \pi_A(r), v_B = \pi_B(r) \right] \right] \quad \text{(by construction)} \\
&= \mathop{\mathbb{E}}_{r \in [n]} \left[\textstyle\prod_{\tau=1}^t \mathop{\mathbb{E}}_{r_\tau \sim \mathcal{R}} \left[v_A(r_\tau) \cdot v_B(r_\tau) \,\middle|\, v_A = \pi_A(r), v_B = \pi_B(r) \right] \right] \quad \text{(using independence of } r_1, \ldots, r_t\text{)} \\
&= \mathop{\mathbb{E}}_{r \in [n]} \left[\left(\mathop{\mathbb{E}}_{r' \sim \mathcal{R}} \left[v_A(r') \cdot v_B(r') \,\middle|\, v_A = \pi_A(r), v_B = \pi_B(r) \right] \right)^t \right] \quad \text{(same distribution for all } \tau \in [t]\text{)} \\
&= \mathop{\mathbb{E}}_{r \in [n]} \langle \pi_A(r), \pi_B(r) \rangle^t \quad \text{(by definition } \langle v_A, v_B \rangle = \mathop{\mathbb{E}}_{r' \sim \mathcal{R}} v_A(r') v_B(r') \text{)} \\
&= \tfrac{1}{n} \sum_{r=1}^n \langle \pi_A(r), \pi_B(r) \rangle^t .
\end{aligned} \tag{26}$$

Recall that the bijections $\{\pi_B\}_{B \in \mathcal{S}}$ depend on the set $\mathcal{S}$. However, using (25), we see that up to a small error $\mathbb{E}_{x \sim \mu_{\mathcal{S}}} x_A x_B$ is independent of $\mathcal{S}$, namely

$$\tfrac{1}{n} \sum_{r=1}^n \langle \pi_A(r), \pi_B(r) \rangle^t = \tfrac{1}{n} \sum_{a \in A} \langle a, \pi_{B \leftarrow A}(a) \rangle^t \quad \pm 2(1 - 1/16R)^t . \tag{27}$$

Let us argue how (25) implies (27): If $\rho(A,B) < 1 - 1/16R$, both quantities $\frac{1}{n}\sum_{a\in A}\langle a, \pi_{B\leftarrow A}(a)\rangle^t$ and $\frac{1}{n}\sum_{r=1}^n \langle \pi_A(r), \pi_B(r)\rangle^t$ are bounded by $(1-1/16R)^t$ in absolute value. Otherwise, if $\rho(A,B) \geqslant 1 - 1/16R$, then (25) asserts $\pi_B(r) = \pi_{B\leftarrow A}(\pi_A(r))$. Thus, in this case, both quantities $\frac{1}{n}\sum_{a\in A}\langle a, \pi_{B\leftarrow A}(a)\rangle^t$ and $\frac{1}{n}\sum_{r=1}^n \langle \pi_A(r), \pi_B(r)\rangle^t$ are equal. From Corollary 1 (also see Remark 9), it follows that the second-moments of the distribution μ_S approximately match the inner products of the vectors $\{\boldsymbol{v}_B\}_{B\in\mathcal{B}}$,

$$\begin{aligned}
\mathop{\mathbb{E}}_{x\sim\mu_S} x_A x_B &= \frac{1}{n}\sum_{r=1}^n \langle \pi_A(r), \pi_B(r)\rangle^t && \text{(by (26))}\\
&= \frac{1}{n}\sum_{a\in A}\langle a, \pi_{B\leftarrow A}(a)\rangle^t \quad \pm 2(1-1/16R)^t && \text{(by (27))}\\
&= \frac{1}{n}\sum_{a\in A, b\in B}\langle a, b\rangle^t \quad \pm 2(3/4)^{t/2} + 2(1-1/16R)^t && \text{(using Corollary 1)}\\
&= \langle \boldsymbol{v}_A, \boldsymbol{v}_B\rangle \quad \pm 4e^{-t/16R}\,.
\end{aligned}$$

Smoothening. Combining the construction above and lemma analogous to 18 (omitted from this preliminary version of the paper), we get the next lemma.

Lemma 10. *For $R, t \in \mathbb{N}$, let $\varepsilon = R^2\cdot 10e^{-t/16R}$. Then, there exists a collection of distributions $\{\mu'_{\mathcal{S}} \colon \{0,1\}^{\mathcal{S}} \to \mathbb{R}_+ \mid \mathcal{S}\subseteq\mathcal{B}, |\mathcal{S}| \leqslant R\}$ such that*

$$\mathop{\mathbb{E}}_{x\sim\mu'_{\mathcal{S}}} x_A x_B = \langle \boldsymbol{u}_A, \boldsymbol{u}_B\rangle \qquad (\mathcal{S}\subseteq\mathcal{B}, |\mathcal{S}|\leqslant R,\ A, B\in\mathcal{S})$$

where the collection of vectors $\{\boldsymbol{u}_A\}_{A\in\mathcal{B}}$ is defined as

$$\boldsymbol{u}_B := \sqrt{1-\varepsilon}\cdot \mathsf{normal}\left(\frac{1}{\sqrt{n}}\sum_{v\in B} v^{\otimes t}\right) \quad + \sqrt{\varepsilon}\cdot \boldsymbol{v}_B^{\perp}\,.$$

Here, $\{\boldsymbol{v}_B^{\perp}\}_{B\in\mathcal{B}}$ is an orthonormal set of vectors orthogonal to all vectors $v^{\otimes t}$ with $v \in \bigcup\mathcal{B}$. We use $\mathsf{normal}(v)$ to denote the unit vector in the direction of v.

6 Integrality Gap Instance for Uniform Sparsest Cut

In this section, we will show how to use Lemma 10 to obtain an integrality gap for a strong SDP relaxation of Balanced Separator. The gap instance is the same as in [DKSV06].

Let $n \in \mathbb{N}$ and let T_ρ be the boolean noise graph on $\{\pm 1\}^n$. Let $\sigma\colon [n] \to [n]$ be the cyclic shift permutation, $\sigma(i) = i + 1 \bmod n$. We let the group generated by σ (with composition as group operation) act on $\{\pm 1\}^n$ by permuting coordinates. For $u \in \{\pm 1\}^n$ and $c \in [n]$, we denote by $\sigma^c.u$ the vector with r^{th}-coordinate $(\sigma^c.u)_r = u_{r+c \bmod n}$. Let $\{\pm 1\}^n/(\sigma)$ denote the partition of $\{\pm 1\}^n$ into the orbits of this group action. We define $T_\rho/(\sigma)$ to be the graph on $\{\pm 1\}^n/(\sigma)$ obtained by identifying vertices of T_ρ that are cyclic shifts of each other.

Theorem 10 ([Bou02, MOO05])**.** *There exists a constant c such that for all $\gamma \geqslant (c\log\log n)/\log n$, every $\Omega(1)$-balanced bipartition of $\{\pm 1\}/(\sigma)$ cuts a $\Omega(\sqrt{\gamma})$ fraction of the edges of $T_{1-\gamma}/(\sigma)$.*

We will show that a natural SDP relaxation for Balanced Separator has objective value $O(\gamma \log(1/\gamma))$ on the instance $T_{1-\gamma}/(\sigma)$.

Let $d = n^3$. For $x \in \{\pm 1\}^n$, let $B(x) := \{(1/\sqrt{n} \cdot \sigma^i.x)^{\otimes 3} \mid i \in [n]\} \subseteq \mathbb{R}^d$. Note that the elements of $B(x)$ are unit vectors. We say x is *good* if $B(x)$ is near-orthogonal in the sense of Definition 6, that is, for every unit vector $u \in \mathbb{R}^d$, we have $\sum_{v \in B(x)} \langle u, v\rangle^2 \leqslant 3/2$. We let $\mathcal{B}$ be the collection of clouds $B(x)$ for all good $x \in \{\pm 1\}^n$. We identify the vectors in the clouds $B(x)$ by $\{\pm 1\}$-valued funcions on a probability space $\mathcal{R}$. Here, $\mathcal{R}$ is the uniform distribution over $[n]^3$. Also note that since $\bigcup \mathcal{B}$ consists only of $\{\pm 1\}$-integral vectors, it satisfies the (symmetrized) ℓ_2^2-triangle inequality.

In the following lemma, we gather further useful properties of the clouds $B(x)$.

Lemma 11. *1. All but an $O(1/n^3)$ fraction of the vectors $x \in \{\pm 1\}^n$ are good.*

2. For $x, y \in \{\pm 1\}^n$, the clouds $B(x)$ and $B(y)$ satisfy the matching property.

Proof. Assertion (1): A straight-forward Chernoff bound argument [DKSV06] shows that a random vector $x \in \{\pm 1\}^n$ satisfies with probability at least $1 - 4/n^3$,

$$\max_{1 \leqslant i \leqslant n} |\langle x, \sigma^i.x\rangle| \leqslant 4\sqrt{n \log n}\,.$$

(The reason being that for a fixed $i \in [n]$, the standard deviation of $|\langle x, \sigma^i.x\rangle|$ is $O(\sqrt{n})$.) We claim that every x that satisfies this bound is good. Let $v_i = (1/\sqrt{n} \cdot \sigma^i.x)^{\otimes 3}$. The maximum value of $\sum_{i=1}^n \langle u, v_i\rangle^2$ over all unit vectors u is equal to the maximum eigenvalue of the Gram matrix $(\langle v_i, v_j\rangle)_{i,j \in [n]}$. We can upper bound the largest eigenvalue of this matrix by the maximum ℓ_1-norm of a row,

$$\sum_{j=1}^n \langle u, v_j\rangle^2 \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^n |\langle v_i, v_j\rangle| \leqslant 1 + (n-1)\left(4\sqrt{\frac{\log n}{n}}\right)^3 = 1 + O\left(\frac{\log^{3/2} n}{\sqrt{n}}\right).$$

For large enough n, the right-hand side is less than $3/2$.

Assertion (2): For $x, y \in \{\pm 1\}^n$, let $u_i = (1/\sqrt{n} \cdot \sigma^i.x)^{\otimes 3}$ and $v_i = (1/\sqrt{n} \cdot \sigma^i.y)^{\otimes 3}$. Suppose $\rho(B(x), B(y)) = |\langle u_i, v_j\rangle|$. Then the matching $M = \{(u_{i+c}, v_{j+c}) \mid c \in [n]\}$ shows that $B(x)$ and $B(y)$ satisfy the matching property. (Here, indices are modulo n). ■

We conclude that the collection $\mathcal{B}$ satisfies all necessary conditions for the construction in §5. Lemma 10 suggest the following vector solution for the instance $T_{1-\gamma}/(\sigma)$: For $R, t \in \mathbb{N}$ with t odd, let $\varepsilon = R^2 \cdot 4e^{-t/16R}$. Fix an arbitrary good vector $x_0 \in \{\pm 1\}^n$. For every vector $x \in \{\pm 1\}^n$, define

$$\boldsymbol{u}_x := \begin{cases} \boldsymbol{u}_{B(x)} = \sqrt{1-\varepsilon} \cdot \mathsf{normal}\left(\frac{1}{\sqrt{n}} \sum_{v \in B(x)} v^{\otimes t}\right) + \sqrt{\varepsilon} \cdot \boldsymbol{v}_B^{\perp} & \text{if } x \text{ is good,} \\ \boldsymbol{u}_{B(x_0)} & \text{otherwise.} \end{cases}$$

Note that this vector solution is invariant under cyclic shifts in the sense that $\boldsymbol{u}_x = \boldsymbol{u}_{\sigma.x}$ for all $x \in \{\pm 1\}^n$. Hence, we can think of $\{\boldsymbol{u}_x\}_{x \in \{\pm 1\}^n}$ as an assignment of vectors to $\{\pm 1\}^n/(\sigma)$. In the following lemma we show that this vector assignment achieves an objective value of roughly $O(t\gamma + \varepsilon)$ for a strong SDP relaxation of the balanced separator problem.

Lemma 12. *1. There exsits a collection of distributions $\{\mu_S : \{\pm 1\}^S \to \mathbb{R}_+ \mid S \subseteq \{\pm 1\}^n, |S| \leqslant R\}$ such that*

$$\mathop{\mathbb{E}}_{z \sim \mu_S} z_x z_y = \langle \boldsymbol{u}_x, \boldsymbol{u}_y\rangle \qquad (S \subseteq \{\pm 1\}^n, |S| \leqslant R,\ x, y \in \mathcal{S})$$

2.

$$\mathop{\mathbb{E}}_{x\sim_{1-\gamma}y} \tfrac{1}{4}\|\boldsymbol{u}_x - \boldsymbol{u}_y\|^2 \leqslant O(t\gamma) \; + \varepsilon + O(1/n^3 + e^{-t/16R})$$

Here, the notation $x \sim_\rho y$ means that we sample from the edge distribution of T_ρ.

3.

$$\mathop{\mathbb{E}}_{x,y\in\{\pm1\}^n} \tfrac{1}{4}\|\boldsymbol{u}_x - \boldsymbol{u}_y\|^2 \geqslant \tfrac{1}{2} \; - O(1)(3/4)^{t/2}$$

Proof. Assertion (1): We can use the distributions provided by Lemma 10. (These distributions are only over assignments to the clouds $B \in \mathcal{B}$. We extend those to distributions over assignments to vectors $x \in \{\pm1\}^n$ by assigning to a vector the value of its cloud with probability 1.)
Assertion (2): Let x and y be good. Then

$$\begin{aligned}
\tfrac{1}{2}\|\boldsymbol{u}_x - \boldsymbol{u}_y\|^2 &= 1 - \langle \boldsymbol{u}_x, \boldsymbol{u}_y\rangle \\
&= 1 - (1-\varepsilon)\cdot \tfrac{1}{n} \sum_{u\in B(x), v\in B(y)} \langle u, v\rangle^t \; \pm O(e^{-t/16R}) \\
&= 1 - (1-\varepsilon)\cdot \sum_i (\tfrac{1}{n}\langle x, \sigma^i.y\rangle)^{3t} \; \pm O(e^{-t/16R})\,.
\end{aligned}$$

For all $x, y \in \{\pm1\}^n$, denote $f(x,y) = 1 - (1-\varepsilon)\cdot\sum_i(\frac{1}{n}\langle x, \sigma^i.y\rangle)^{3t}$. We have

$$\mathop{\mathbb{E}}_{x\sim_\rho y} \tfrac{1}{2}\|\boldsymbol{u}_x - \boldsymbol{u}_y\|^2 \leqslant \mathop{\mathbb{E}}_{x\sim_\rho y} f(x,y) + 2\cdot \mathop{\mathbb{P}\mathrm{r}}_{x\in\{\pm1\}^n} \{\; x \text{ not good }\} \pm O(e^{-t/16R})$$

It remains to bound

$$\begin{aligned}
\mathop{\mathbb{E}}_{x\sim_\rho y} f(x,y) &= 1 - (1-\varepsilon)\cdot \mathop{\mathbb{E}}_{x\sim_\rho y}\sum_i (\tfrac{1}{n}\langle x, \sigma^i.y\rangle)^{3t} = 1 - (1-\varepsilon)\cdot \mathop{\mathbb{E}}_{x\sim_\rho y} (\tfrac{1}{n}\langle x, y\rangle)^{3t} \leqslant 1 + \varepsilon - \mathop{\mathbb{E}}_{x\sim_\rho y} (\tfrac{1}{n}\langle x, y\rangle)^{3t} \\
&\leqslant 1 + \varepsilon + 2 \mathop{\mathbb{P}\mathrm{r}}_{x\sim_\rho y}\{\langle x, y\rangle < 0\} - \mathop{\mathbb{E}}_{x\sim_\rho y}(\tfrac{1}{n}|\langle x, y\rangle|)^{3t} \leqslant 1 + \varepsilon + 2 \mathop{\mathbb{P}\mathrm{r}}_{x\in\{\pm1\}^n}\{\langle x, y\rangle < 0\} - \left(\mathop{\mathbb{E}}_{x\sim_\rho y} \tfrac{1}{n}|\langle x, y\rangle|\right)^{3t}
\end{aligned}$$

If $x \sim_\rho y$, then $\mathbb{E}\,\frac{1}{n}\langle x, y\rangle = \rho = 1 - \gamma$. Since ρ is positive (actually close to 1), the probability that $\langle x, y\rangle < 0$ for ρ-correlated x and y is exponentially small in n. We conclude

$$\mathop{\mathbb{E}}_{x\sim_\rho y} \tfrac{1}{2}\|\boldsymbol{u}_x - \boldsymbol{u}_y\|^2 \leqslant 1 + \varepsilon - (1-\gamma)^{3t} + O(1/n^3 + e^{-t/16R}) \leqslant 3t\gamma + \varepsilon + O(1/n^3 + e^{-t/16R})$$

Assertion (3): Recycling the notation for the proof of assertion (2), we have

$$\mathop{\mathbb{E}}_{x,y\in\{\pm1\}^n} \tfrac{1}{2}\|\boldsymbol{u}_x - \boldsymbol{u}_y\|^2 \geqslant \mathop{\mathbb{E}}_{x,y\in\{\pm1\}^n} f(x,y) - 2\cdot \mathop{\mathbb{P}\mathrm{r}}_{x\in\{\pm1\}^n}\{\; x \text{ not good }\} \pm O(e^{-t/16R})\,.$$

It is easy to see that $\mathbb{E}_{x,y\in\{\pm1\}^n} f(x,y) = 1$ because the exponent $3t$ is odd. It follows that

$$\mathop{\mathbb{E}}_{x,y\in\{\pm1\}^n} \tfrac{1}{2}\|\boldsymbol{u}_x - \boldsymbol{u}_y\|^2 \geqslant 1 - O(1/n^3) - O(e^{-t/16R})\,.$$

■

Remark 14. By modifying the definition of the vector solution $\{\boldsymbol{u}_x\}$ slighly, assertion (3) in the previous lemma can be strengthened to

$$\mathop{\mathbb{E}}_{x,y\in\{\pm 1\}^n} \tfrac{1}{4}\|\boldsymbol{u}_x - \boldsymbol{u}_y\|^2 = \tfrac{1}{2}\,.$$

We change the definition of $\boldsymbol{u}_x$ only for those x that are not good. If we assume that n is odd, then an even number of orbits $B(x)$ are not good (a vector x is good if and only if $-x$ is good, and for n odd, we have always $B(x) \neq B(-x)$). Hence, we can partition the non-good orbits into two sets of equal size. We assign $\boldsymbol{u}_{B(x_0)}$ to the orbits in one set and $\boldsymbol{u}_{B(-x_0)}$ to the orbits in the other set.

The next theorem is a consequence of the previous lemma.

Theorem 11. *Let $R \in \mathbb{N}$ and $G = T_\rho(\sigma)$ with $\rho = 1-\gamma$. Then, there exists a solution to the SDP relaxation for the $(1/2 - O(\gamma))$-balanced separator problem on G that achieves objective value*

$$O\left(R^2\gamma\log(1/\gamma)\right)\,.$$

Furthermore, this SDP solution satisfies all valid inequalities for vertex subset of size up to R.

7 Smoothening

Let Σ be a finite alphabet of size q. Let $\{\chi_1,\ldots,\chi_q\}$ be an orthonormal basis for the vector space $\{f\colon \Sigma\to\mathbb{R}\}$ such that $\chi_1(a) = 1$ for all $a\in\Sigma$. (Here, orthonormal means $\mathbb{E}_{a\in\Sigma}\chi_i(a)\chi_j(a) = \delta_{ij}$ for all $i,j\in[q]$.) For $R\in\mathbb{N}$, let $\left\{\chi_\sigma \mid \sigma\in[q]^R\right\}$ be the orthonormal basis of the vector space $\left\{f\colon \Sigma^R\to\mathbb{R}\right\}$ defined by

$$\chi_\sigma(x) \stackrel{\text{def}}{=} \chi_{\sigma_1}(x_1)\cdot\cdots\cdot\chi_{\sigma_R}(x_R)\,, \tag{28}$$

where $\sigma = (\sigma_1,\ldots,\sigma_R)\in[q]^R$ and $x = (x_1,\ldots,x_R)\in\Sigma^R$.

For a function $f\colon\Sigma^R\to\mathbb{R}$, we denote

$$\hat{f}(\sigma) \stackrel{\text{def}}{=} \sum_{x\in\Sigma^R} f(x)\chi_\sigma(x)\,. \tag{29}$$

Using the fact $\mathbb{E}_{\sigma\in[q]^R}\chi_\sigma(x)\chi_\sigma(y) = \delta_{xy}$ for all $x,y\in\Sigma^R$, we see that

$$f = \mathop{\mathbb{E}}_{\sigma\in[q]^R}\hat{f}(\sigma)\chi_\sigma\,.$$

We define the following norm for functions $\hat{f}\colon [q]^R\to\mathbb{R}$,

$$\|\hat{f}\|_1 \stackrel{\text{def}}{=} \sum_{\sigma\in[q]^R}|\hat{f}(\sigma)|\,.$$

We say $f\colon\Sigma^R\to\mathbb{R}$ is a *distribution* if $f(x)\geqslant 0$ for all $x\in\Sigma^R$ and $\sum_{x\in\Sigma^R} f(x) = 1$. We define

$$K \stackrel{\text{def}}{=} \max_{\sigma\in[q]^R, x\in\Sigma^R}|\chi_\sigma(x)|\,.$$

In the next lemma, we give a proof of the following intuitive fact: If a function $g\colon\Sigma^R\to\mathbb{R}$ satisfies the normalization constraint $\sum_{x\in\Sigma^R} g(x) = 1$ and it is close to a distribution in the sense that there exists a distribution f such that $\|\hat{f}-\hat{g}\|$ is small, then g can be made to a distribution by "smoothening" it. Here, smoothening means to move slightly towards the uniform distribution (where every assignment has probability q^{-R}).

Lemma 13. *Let $f, g\colon \Sigma^R \to \mathbb{R}$ be two functions with $\hat{f}(\mathbb{1}) = \hat{g}(\mathbb{1}) = 1$. Suppose f is a distribution. Then, the following function is also a distribution*

$$(1-\varepsilon)g + \varepsilon q^{-R} \quad \textit{where } \varepsilon = \|\hat{f} - \hat{g}\|_1 \cdot K\,.$$

Proof. It is clear that the function $h = (1-\varepsilon)g + \varepsilon q^{-R}$ satisfies the constraint $\hat{h}(\mathbb{1}) = 1$. For every $x \in \Sigma^R$, we have

$$\begin{aligned}
h(x) &= (1-\varepsilon)g(x) + \varepsilon q^{-R} \\
&\geqslant (1-\varepsilon)\Big(g(x) - f(x)\Big) + \varepsilon q^{-R} \qquad \text{(using } f(x) \geqslant 0\text{)} \\
&= \varepsilon q^{-R} + (1-\varepsilon) \mathop{\mathbb{E}}_{\sigma \in [q]^R} \Big(\hat{g}(\sigma) - \hat{f}(\sigma)\Big)\chi_\sigma(x) \\
&\geqslant \varepsilon q^{-R} - (1-\varepsilon) \mathop{\mathbb{E}}_{\sigma \in [q]^R} \Big|\hat{g}(\sigma) - \hat{f}(\sigma)\Big| \cdot K \\
&= \varepsilon q^{-R} - (1-\varepsilon)K\|\hat{f} - \hat{g}\|_1 \cdot q^{-R} \\
&\geqslant 0\,. \qquad \text{(by our choice of } \varepsilon\text{)}
\end{aligned}$$

■

Let V be a set. For a function $f\colon \Sigma^V \to \mathbb{R}$ and a subset $S \subseteq V$, we define the function $\mathsf{margin}_S\, f\colon \Sigma^S \to \mathbb{R}$ as

$$\mathsf{margin}_S\, f(x) \stackrel{\text{def}}{=} \sum_{y \in \Sigma^{V \setminus S}} f(x, y)\,.$$

Note that if f is a distribution over Σ-assignments to V then $\mathsf{margin}_S\, f$ is its marginal distribution over Σ-assignments to T.

Lemma 14. *For every $f\colon \Sigma^V \to \mathbb{R}$ and $S \subseteq V$,*

$$\mathsf{margin}_S\, f = \mathop{\mathbb{E}}_{\sigma \in [q]^S} \hat{f}(\sigma, \mathbb{1})\chi_\sigma\,.$$

Here, $\sigma, \mathbb{1}$ denotes the Σ-assignment to V that agrees with σ on S and assigns 1 to all variables in $V \setminus S$.

Proof.

$$\begin{aligned}
\mathsf{margin}_S\, f(x) &= \sum_{y \in \Sigma^{V \setminus S}} f(x, y) \\
&= \sum_{y \in \Sigma^{V \setminus S}} \mathop{\mathbb{E}}_{\sigma \in [q]^V} \hat{f}(\sigma)\chi_\sigma(x, y) \\
&= \sum_{y \in \Sigma^{V \setminus S}} \mathop{\mathbb{E}}_{\sigma \in [q]^S} \mathop{\mathbb{E}}_{\sigma' \in [q]^{V \setminus S}} \hat{f}(\sigma)\chi_\sigma(x)\chi_{\sigma'}(y) \\
&= \mathop{\mathbb{E}}_{\sigma \in [q]^S} \mathop{\mathbb{E}}_{\sigma' \in [q]^{V \setminus S}} \hat{f}(\sigma, \sigma')\chi_\sigma(x) \cdot \sum_{y \in \Sigma^{V \setminus S}} \chi_{\sigma'}(y) \\
&= \mathop{\mathbb{E}}_{\sigma \in [q]^S} \hat{f}(\sigma, \mathbb{1})\chi_\sigma(x)\,. \quad \text{(using } \textstyle\sum_{y \in [q]^{V \setminus S}} \chi_{\sigma'}(y) = 0 \text{ for } \sigma' \neq \mathbb{1}.\text{)}
\end{aligned}$$

■

The margin operator has the following useful property (which is clear from its definiton).

Lemma 15. *For every function $f\colon \Sigma^V \to \mathbb{R}$ and any sets $T \subseteq S \subseteq V$,*

$$\mathsf{margin}_T\, \mathsf{margin}_S\, f = \mathsf{margin}_T\, f\,.$$

Lemma 16. *Let V be a set and let $\left\{\mu_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ be a collection of distributions. Suppose that for all sets $A, B \subseteq V$ with $|A|, |B| \leqslant R$,*

$$\|\mathsf{margin}_{A\cap B}\, \mu_A - \mathsf{margin}_{A\cap B}\, \mu_B\|_1 \leqslant \eta\,.$$

Then, there exists a collection of distributions $\left\{\mu'_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ such that

– *for all $A, B \subseteq V$ with $|A|, |B| \leqslant R$,*

$$\mathsf{margin}_{A\cap B}\, \mu'_A = \mathsf{margin}_{A\cap B}\, \mu'_B\,.$$

– *for all $S \subseteq V$ with $|S| \leqslant R$,*

$$\|\mu'_S - \mu_S\|_1 \leqslant O(\eta q^R K^2)\,,$$

The previous lemma is not enough to establish the robustness of our SDP relaxations. The issue is that we not only require that the distributions are consistent among themselves but they should also also be consistent with the SDP vectors.

The following lemma allows us to deal with this issue.

Lemma 17. *Let V be a set and let $\left\{\mu_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ be a collection of distributions. Suppose that*

– *for all sets $A, B \subseteq V$ with $|A|, |B| \leqslant R$,*

$$\|\mathsf{margin}_{A\cap B}\, \mu_A - \mathsf{margin}_{A\cap B}\, \mu_B\|_1 \leqslant \eta\,.$$

– *for all sets $A, B \subseteq V$ with $|A|, |B| \leqslant 2$,*

$$\mathsf{margin}_{A\cap B}\, \mu_A = \mathsf{margin}_{A\cap B}\, \mu_B\,.$$

Then, for $\varepsilon \geqslant q^R K^2 \eta$, there exists a collection of distributions $\left\{\mu'_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ such that

– *for all $A, B \subseteq V$ with $|A|, |B| \leqslant R$,*

$$\mathsf{margin}_{A\cap B}\, \mu'_A = \mathsf{margin}_{A\cap B}\, \mu'_B\,. \tag{30}$$

– *for all $S \subseteq V$ with $|S| \leqslant R$,*

$$\|\mu'_S - \mu_S\|_1 \leqslant O(K^2 \eta q^R)\,, \tag{31}$$

– *for all $S \subseteq V$ with $|S| \leqslant 2$,*

$$\mu'_S = (1-\varepsilon)\mu_S + \varepsilon \cdot q^{-|S|}\,. \tag{32}$$

Proof. For $\sigma \in [q]^V$, let $\operatorname{supp}(\sigma)$ denote the set of coordinates of σ not equal to 1, and let $|\sigma|$ denote the number of such coordinates,

$$\operatorname{supp}(\sigma) \stackrel{\text{def}}{=} \{i \in V \mid \sigma_i \neq 1\} \quad \text{and} \quad |\sigma| \stackrel{\text{def}}{=} \left|\operatorname{supp}(\sigma)\right| .$$

For every $\sigma \in [q]^V$ with $|\sigma| \leqslant R$, we define

$$\hat{f}(\sigma) := \mathop{\mathbb{E}}_{x \sim \mu_S} \chi_\sigma(x) \qquad \text{where } S = \operatorname{supp}(\sigma) .$$

For every σ with $|\sigma| > R$, we set $\hat{f}(\sigma) := 0$. We define μ'_S in terms of $f = \mathbb{E}_\sigma \hat{f}(\sigma)\chi_\sigma$,

$$\mu'_S := \mathsf{margin}_S\, (1-\varepsilon) f + \varepsilon q^{-|V|} .$$

By Lemma 15, this choice of μ'_S satisfies condition (30).

First, let us argue that the functions μ'_S are distributions. Let $S \subseteq V$ with $|S| \leqslant R$. For $\sigma \in [q]^S$ with $T := \operatorname{supp}(\sigma) \subseteq S$, we have

$$\begin{aligned} |\hat{f}(\sigma, \mathbb{1}) - \mathop{\mathbb{E}}_{x \sim \mu_S} \chi_\sigma(x)| &= |\mathop{\mathbb{E}}_{x \sim \mu_T} \chi_\sigma(x) - \mathop{\mathbb{E}}_{x \sim \mu_S} \chi_\sigma(x)| \\ &\leqslant \|\mu_T - \mathsf{margin}_T\, \mu_S\|_1 \cdot \max|\chi_\sigma| \\ &\leqslant \eta \cdot K . \end{aligned} \tag{33}$$

Let f_S denote the function $\mathsf{margin}_S\, f$. By Lemma 14, $\hat{f}_S(\sigma) = \hat{f}(\sigma, \mathbb{1})$ for all $\sigma \in [q]^S$. Hence, $\|\hat{f} - \hat{\mu}_S\|_1 \leqslant q^R \cdot K\eta$. It follows that for $\varepsilon \geqslant q^R K^2 \eta$, the function $\mu'_S = (1-\varepsilon) f_S + \varepsilon q^{-|S|}$ is a distribution (using Lemma 13).

Next, let us verify that (31) holds. We have

$$\begin{aligned} \|\mu'_S - \mu_S\|_1 &\leqslant O(\varepsilon) + \|\mathsf{margin}_S\, f - \mu_S\|_1 \\ &\stackrel{\text{La. 14}}{=} O(\varepsilon) + \left\| \mathop{\mathbb{E}}_{\sigma \in [k]^S} \left(\hat{f}(\sigma, \mathbb{1}) - \mathop{\mathbb{E}}_{x \sim \mu_S} \chi_\sigma(x) \right) \chi_\sigma \right\|_1 \\ &\stackrel{(33)}{\leqslant} O\left(\eta K^2 \cdot k^R\right) \quad (\text{using } |\hat{f}(\sigma, \mathbb{1}) - \hat{\mu}_S(\sigma)| \leqslant \eta K \text{ and } |\chi_\sigma(x)| \leqslant K) . \end{aligned}$$

Finally, we show that the new distributions satisfy (32). Let $S \subseteq V$ be a set of size at most 2. It follows from the consistency assumption that for all $\sigma \in [k]^S$, we have $\hat{f}(\sigma, \mathbb{1}) = \hat{\mu}_S(\sigma)$. Hence, $f_S = \mu_S$, which implies (32). ∎

Lemma 18. *Let V be a set and let $\left\{\mu_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ be a collection of distributions. Suppose that*

– *for all sets $A, B \subseteq V$ with $|A|, |B| \leqslant R$,*

$$\|\mathsf{margin}_{A \cap B}\, \mu_A - \mathsf{margin}_{A \cap B}\, \mu_B\|_1 \leqslant \eta .$$

– *for all sets $A, B \subseteq V$ with $|A|, |B| \leqslant 2$,*

$$\mathsf{margin}_{A \cap B}\, \mu_A = \mathsf{margin}_{A \cap B}\, \mu_B .$$

Then, for $\varepsilon \geqslant kR^2K^2\eta$, there exists a collection of distributions $\left\{\mu'_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ such that

– *for all $A, B \subseteq V$ with $|A|, |B| \leqslant R$ with $|A \cap B| \leqslant 2$,*

$$\mathsf{margin}_{A\cap B}\,\mu'_A = \mathsf{margin}_{A\cap B}\,\mu'_B\,. \tag{34}$$

– *for all $S \subseteq V$ with $|S| \leqslant R$,*

$$\|\mu'_S - \mu_S\|_1 \leqslant O(K^2\eta k R^2)\,, \tag{35}$$

– *for all $S \subseteq V$ with $|S| \leqslant 2$,*

$$\mu'_S = (1-\varepsilon)\mu_S + \varepsilon \cdot k^{-|S|}\,. \tag{36}$$

Proof. The proof is along the lines of the proof of the previous lemma.

Define $\hat{f} \colon [k]^R \to \mathbb{R}$ as before. We define new functions $\{\mu^*_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\}$ such that

$$\hat{\mu}^*_S(\sigma) = \begin{cases} \hat{\mu}_S(\sigma) & \text{if } \operatorname{supp}(\sigma) > 2, \\ \hat{f}(\sigma, \mathbb{1}) & \text{if } 1 \leqslant \operatorname{supp}(\sigma) \leqslant 2, \\ 1 & \text{otherwise.} \end{cases}$$

Since $|\hat{f}(\sigma, \mathbb{1}) - \hat{\mu}_S(\sigma)| \leqslant K\eta$ (see proof of previous lemma), we can upper bound $\|\hat{\mu}^*_S - \hat{\mu}_S\|_1 \leqslant kR^2 \cdot K\eta$ (there are not more than kR^2 different $\sigma \in [k]^S$ with $\hat{f}(\sigma, \mathbb{1}) \neq \hat{\mu}_S(\sigma)$). By Lemma 13, for $\varepsilon \geqslant kR^2K^2\eta$, the functions $\{\mu'_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\}$ defined by $\mu'_S := (1-\varepsilon)\mu^*_S + \varepsilon k^{-|S|}$ are the desired distributions. We can check that the assertions of the lemma are satisfied in the same way as for the proof of the previous lemma. ∎

Proofs of Theorem 7 and Theorem 8 (Sketch). We apply Lemma 18 or Lemma 17 to the local distributions $\{\mu_S\}$ of the ε-infeasible LH_R or SA_R solution, respectively. We get a new set of local distributions $\{\mu'_S\}$ that have the desired consistency properties. It remains to change the vectors so that their inner product match the corresponding probabilities in the local distributions. Suppose $\{\boldsymbol{v}_{i,a}\}$ is the original vector assignment. Let $\{\boldsymbol{u}_{i,a}\}$ be the vector assignment that corresponds to the uniform distribution over all possible assignments to the variables (this vector assignment is the geometric center of the set of all vector assignments). Then, we define the new vector assignment $\{\boldsymbol{v}'_{i,a}\}$ as

$$\boldsymbol{v}'_{i,a} = \sqrt{1-\delta} \cdot \boldsymbol{v}_{i,a} \oplus \sqrt{\delta}\boldsymbol{u}_{i,a}\,,$$

where δ is the smoothening parameter in Lemma 18 or Lemma 17. It is easy to verify that $\{\boldsymbol{v}'_{i,a}\}$ together with $\{\mu'_S\}$ form a feasible LH_R or SA_R solution. ∎

References

[AAT05] Michael Alekhnovich, Sanjeev Arora, and Iannis Tourlakis, *Towards strong nonapproximability results in the Lovász–Schrijver hierarchy*, STOC, 2005, pp. 294–303. 3

[ABL02] Sanjeev Arora, Béla Bollobás, and László Lovász, *Proving integrality gaps without knowing the linear program*, FOCS, 2002, pp. 313–322. 3

[ABLT06] Sanjeev Arora, Béla Bollobás, László Lovász, and Iannis Tourlakis, *Proving integrality gaps without knowing the linear program*, Theory of Computing **2** (2006), no. 1, 19–51. 3

[ACC06] Sanjeev Arora, Eden Chlamtac, and Moses Charikar, *New approximation guarantee for chromatic number*, STOC, 2006, pp. 215–224. 3

[AKK+08] Sanjeev Arora, Subhash Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K. Vishnoi, *Unique games on expanding constraint graphs are easy: extended abstract*, STOC, 2008, pp. 21–28. 16, 17, 18

[ARV04] Sanjeev Arora, Satish Rao, and Umesh V. Vazirani, *Expander flows, geometric embeddings and graph partitioning*, STOC, 2004, pp. 222–231. 3

[Aus07] Per Austrin, *Towards sharp inapproximability for any 2-CSP*, FOCS, 2007, pp. 307–317. 1, 2

[BOGH+06] Joshua Buresh-Oppenheim, Nicola Galesi, Shlomo Hoory, Avner Magen, and Toniann Pitassi, *Rank bounds and integrality gaps for cutting planes procedures*, Theory of Computing **2** (2006), no. 1, 65–90. 3

[Bou02] J. Bourgain, *On the distributions of the Fourier spectrum of Boolean functions*, Israel J. Math. **131** (2002), 269–276. 29

[Chl07] Eden Chlamtac, *Approximation algorithms using hierarchies of semidefinite programming relaxations*, FOCS, 2007, pp. 691–701. 3

[CKK+05] Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani, and D. Sivakumar, *On the hardness of approximating multicut and sparsest-cut*, IEEE Conference on Computational Complexity, 2005, pp. 144–153. 2

[CKK+06] ———, *On the hardness of approximating multicut and sparsest-cut*, Computational Complexity **15** (2006), no. 2, 94–114. 1

[CMM06] Moses Charikar, Konstantin Makarychev, and Yury Makarychev, *Near-optimal algorithms for unique games*, STOC, 2006, pp. 205–214. 14, 15

[CMM09] ———, *Integrality gaps for Sherali–Adams relaxations*, STOC, 2009, To appear. 2, 3, 4

[CS98] Benny Chor and Madhu Sudan, *A geometric approach to betweenness*, SIAM J. Discrete Math. **11** (1998), no. 4, 511–523 (electronic). 2

[CS08] Eden Chlamtac and Gyanit Singh, *Improved approximation guarantees through higher levels of SDP hierarchies*, APPROX-RANDOM, 2008, pp. 49–62. 3

[DKSV06] Nikhil R. Devanur, Subhash Khot, Rishi Saket, and Nisheeth K. Vishnoi, *Integrality gaps for sparsest cut and minimum linear arrangement problems*, STOC, 2006, pp. 537–546. 1, 3, 29, 30

[FL92] Uriel Feige and László Lovász, *Two-prover one-round proof systems: Their power and their problems (extended abstract)*, STOC, 1992, pp. 733–744. 15

[GMPT07] Konstantinos Georgiou, Avner Magen, Toniann Pitassi, and Iannis Tourlakis, *Integrality gaps of 2 - o(1) for vertex cover SDPs in the Lovész–Schrijver hierarchy*, FOCS, 2007, pp. 702–712. 1, 3

[GMR08] Venkatesan Guruswami, Rajsekar Manokaran, and Prasad Raghavendra, *Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph*, FOCS, 2008, pp. 573–582. 1, 2, 4

[GW95] Michel X. Goemans and David P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. ACM **42** (1995), no. 6, 1115–1145. 12

[Kho02] Subhash Khot, *On the power of unique 2-prover 1-round games*, STOC, 2002, pp. 767–775. 1

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell, *Optimal inapproximability results for max-cut and other 2-variable csps?*, SIAM J. Comput. **37** (2007), no. 1, 319–357. 1, 2, 4, 5, 24, 25

[KR08] Subhash Khot and Oded Regev, *Vertex cover might be hard to approximate to within* $2-\varepsilon$, J. Comput. Syst. Sci. **74** (2008), no. 3, 335–349. 1

[KV05] Subhash Khot and Nisheeth K. Vishnoi, *The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1*, FOCS, 2005, pp. 53–62. 1, 2, 3, 4, 5, 6, 16, 24, 25

[Las01] Jean B. Lasserre, *An explicit exact SDP relaxation for nonlinear 0-1 programs*, IPCO, 2001, pp. 293–303. 3

[Lau03] M. Laurent, *A comparison of the Sherali–Adams, Lovász–Schrijver and Lasserre relaxations for* $0-1$ *programming*, Mathematics of Operations Research **28** (2003), no. 3, 470–496. 3

[LS91] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and* 0-1 *optimization*, SIAM J. Optim. **1** (1991), no. 2, 166–190. 3

[MNRS08] Rajsekar Manokaran, Joseph Naor, Prasad Raghavendra, and Roy Schwartz, *SDP gaps and UGC hardness for multiway cut, 0-extension, and metric labeling*, STOC, 2008, pp. 11–20. 1, 2, 4

[MOO05] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz, *Noise stability of functions with low in.uences invariance and optimality*, FOCS, 2005, pp. 21–30. 29

[OW08] Ryan O'Donnell and Yi Wu, *An optimal SDP algorithm for max-cut, and equally optimal long code tests*, STOC, 2008, pp. 335–344. 1

[Rag08] Prasad Raghavendra, *Optimal algorithms and inapproximability results for every csp?*, STOC, 2008, pp. 245–254. 1, 2, 4

[RS09] Prasad Raghavendra and David Steurer, *Towards computing the Grothendieck constant*, SODA, 2009, pp. 525–534. 1, 2

[SA90] Hanif D. Sherali and Warren P. Adams, *A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems*, SIAM J. Discrete Math. **3** (1990), no. 3, 411–430. 2, 3

[Sch08] Grant Schoenebeck, *Linear level Lasserre lower bounds for certain k-CSPs*, FOCS, 2008, pp. 593–602. 1, 3

[ST06] Alex Samorodnitsky and Luca Trevisan, *Gowers uniformity, influence of variables, and PCPs*, STOC, 2006, pp. 11–20. 3

[STT07a] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani, *A linear round lower bound for Lovász-Schrijver SDP relaxations of vertex cover*, IEEE Conference on Computational Complexity, 2007, pp. 205–216. 3

[STT07b] ———, *Tight integrality gaps for Lovász–Schrijver LP relaxations of vertex cover and max cut*, STOC, 2007, pp. 302–310. 2, 3, 4

[Tou06] Iannis Tourlakis, *New lower bounds for vertex cover in the Lovász-Schrijver hierarchy*, IEEE Conference on Computational Complexity, 2006, pp. 170–182. 3

[Tre08] Luca Trevisan, *Approximation algorithms for unique games*, Theory of Computing **4** (2008), no. 1, 111–128. 17

[Tul09] Madhur Tulsiani, *CSP gaps and reductions in the Lasserre hierarchy*, STOC, 2009, To apperar. 1, 2, 3, 4

A Problem Definitions

A.1 Constraint Satisfaction Problems

let Λ be a family of *payoff functions* $P\colon [q]^k \to [-1,1]$. We say Λ has *arity* k and *alphabet* $[q] \stackrel{\text{def}}{=} \{1,\ldots,q\}$. A function $P'\colon [q]^V \to [-1,1]$ has *type* Λ if for some $P \in \Lambda$ and some $i_1,\ldots,i_k \in V$, we have $P'(x) = P(x_{i_1},\ldots,x_{i_k})$ for all $x \in [q]^V$. We define $V(P') \subseteq V$ to be the set of coordinates that P' depends on. In other words, if $P'(x) = P(x_{i_1},\ldots,x_{i_k})$, then $V(P') = \{i_1,\ldots,i_k\}$. In particular, $|V(P')| \leqslant k$ for any function P' of type Λ. A Λ-*CSP instance $\mathcal{P}$ with variable set V* is a distribution over payoff functions $P\colon [q]^V \to [-1,1]$ of type Λ.

Problem 1 (Λ-ConstraintSatisfactionProblem (CSP))**.** Given a variable set V and a distribution $\mathcal{P}$ over payoff functions $P\colon [q]^V \to [-1,1]$ of type Λ, the goal is to find an assignment $x \in [q]^V$ so as to maximize $\mathbb{E}_{P\sim\mathcal{P}} P(x)$. We define the value opt($\mathcal{P}$) as

$$\mathrm{opt}(\mathcal{P}) \stackrel{\mathrm{def}}{=} \max_{x\in[q]^V} \mathop{\mathbb{E}}_{P\sim\mathcal{P}} P(x)\,.$$

Observe that if the payoff functions P are predicates, then maximizing the payoff amounts to maximizing the number of constraints satisfied.

Given an instance $\mathcal{P}$ with variable set V, the goal is to find a collection of vectors $\{\boldsymbol{v}_{i,a}\}_{i\in V, a\in[q]} \subseteq \mathbb{R}^d$ and a collection $\{\mu_P\}_{P\in\mathrm{supp}(\mathcal{P})}$ of distributions over local assignments. For each payoff $P \in \mathcal{P}$, the distribution μ_P is a distribution over $[q]^{V(P)}$ corresponding to assignments for the variables $V(P)$. We will write $\mathbb{P}\mathrm{r}_{x\in\mu_P}\{E\}$ to denote the probability of an event E with under the distribution μ_P.

$\mathcal{SDP}$ Relaxation

$$\text{maximize} \quad \mathop{\mathbb{E}}_{P\sim\mathcal{P}} \mathop{\mathbb{E}}_{x\sim\mu_P} P(x) \tag{37}$$

$$\text{subject to} \quad \langle \boldsymbol{v}_{i,a}, \boldsymbol{v}_{j,b}\rangle = \mathop{\mathbb{P}\mathrm{r}}_{x\sim\mu_P}\left\{x_i = a, x_j = b\right\} \qquad (P \in \mathrm{supp}(\mathcal{P}),\ i,j \in V(P),\ a,b \in [q])\,, \tag{38}$$

$$\langle \boldsymbol{v}_{i,a}, \boldsymbol{v}_0\rangle = \mathop{\mathbb{P}\mathrm{r}}_{x\sim\mu_P}\left\{x_i = a\right\} \qquad (P \in \mathrm{supp}(\mathcal{P}),\ i \in V(P),\ a \in [q])\,. \tag{39}$$

Here $\boldsymbol{v}_0$ can be any fixed unit vector in $\mathbb{R}^d$, and d can be any sufficiently large number, say $d = q|V|$.

A.2 Ordering CSP

Definition 11. An Ordering Constraint Satisfaction Problem (OCSP) Λ is specified by a family of *payoff functions* $P : \Pi_k \to [-1,1]$ on the set Π_k of permutations on k elements. The integer k is referred to as the arity of the OCSP Λ.

Definition 12 (Λ-OrderingConstraintSatisfactionProblem (OCSP))**.** An instance Φ of Ordering Constraint Satisfaction Problem Λ is given by $\Phi = (\mathcal{V}, \mathcal{P})$ where

- $\mathcal{V} = \{y_1, \ldots, y_m\}$ is the set of variables that need to be ordered. Thus an ordering $\mathcal{O}$ is a one to one map from $\mathcal{V}$ to natural numbers $\mathbb{N}$.
- $\mathcal{P}$ is a probability distribution over constraints/payoffs applied to subsets of at most k variables from $\mathcal{V}$. More precisely, a sample $P \sim \mathcal{P}$ would be a payoff function from Λ, applied on a sequence of variables $y_S = (y_{s_1}, \ldots, y_{s_k})$. If $\mathcal{O}_{|S}$ denotes the injective map from $y_S \to \mathbb{N}$ obtained by restricting $\mathcal{O}$ to y_S, then the payoff returned is $P(\mathcal{O}_{|S})$.

For a payoff $P \in \mathcal{P}$, we define $V(P) \in \mathcal{V}$ to denote the set of variables on which P is applied. The objective is to find an ordering $\mathcal{O}$ of the variables that maximizes the total weighted payoff/expected payoff, i.e.,

$$\mathop{\mathbb{E}}_{P\sim\mathcal{P}}\left[P(\mathcal{O}_{|P})\right]$$

Here $O_{|P}$ denotes the ordering O restricted to the variables in $\mathcal{V}(P)$. We define the value $\text{opt}(\mathcal{P})$ as

$$\text{opt}(\Phi) \stackrel{\text{def}}{=} \max_{O:\Pi_{\mathcal{V}\to\mathbb{N}}} \mathop{\mathbb{E}}_{P\sim\mathcal{P}} P(O_{|P}) .$$

A.3 Balanced Separator Problem

Problem 2 (b-Balanced Separator)**.** Given a graph G on vertex set V, the goal is to find a set $S \subseteq V$ with $b \leqslant |S|/|V| \leqslant 1/2$ so as to minimize the fraction of edges cut by S.

We consider the following SDP relaxation for this problem:

LH_R Relaxation for b-Balanced Separator:

$$\text{maximize} \quad \mathop{\mathbb{E}}_{(i,j)\in E} \tfrac{1}{4}\|v_i - v_j\|^2 \tag{40}$$

$$\text{subject to} \quad \mathop{\mathbb{E}}_{i,j\in V} \tfrac{1}{4}\|v_i - v_j\|^2 \geqslant 2b(1-b) \tag{41}$$

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \mathop{\mathbb{E}}_{x\sim\mu_S} x_i x_j \qquad (S \subseteq \mathcal{V}, |S| \leqslant R,\ i,j \in S)\,, \tag{42}$$

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_0 \rangle = \mathop{\mathbb{E}}_{x\sim\mu_S} x_i \qquad (S \subseteq \mathcal{V}, |S| \leqslant R,\ i \in S)\,. \tag{43}$$

A.4 Unique Games, Γ-Max-2Lin(n), $\mathsf{E2Lin}_q$

Problem 3 ($\mathsf{E2Lin}_q$)**.** Given a variable set $\mathcal{V}$ and a system of linear equations over the finite field $\mathbb{F}_q$, with equation of the form $x_i - x_j = c_{ij}$ for variables $i, j \in \mathcal{V}$, the goal is to find an $\mathbb{F}_q$-assignment x to $\mathcal{V}$ that satisfies the maximum number of equations.

Note that Max Cut is a slight generalization of $\mathsf{E2Lin}_2$.

Problem 4 (Γ-Max-2Lin)**.** Given a variable set $\mathcal{V}$ and a list of constraints of the form $x_i x_j^{-1} = c_{ij}$ over the group Γ for variables $i, j \in \mathcal{V}$, i, j, the goal is to find a Γ-assignment to $\mathcal{V}$ so as to maximize the number of satisfied constraints.

Note that $\mathsf{E2Lin}_q$ is the same problem as Γ-Max-2Lin for $\Gamma = \mathbb{F}_q$. We sometimes use the notation Γ-Max-2Lin(k) to refer to the more general problem where a group Γ of order k is given as part of the input.

Problem 5 (Unique Games(k))**.** Given a variable set $\mathcal{V}$ and a list of constraints of the form $x_u = \pi_{uv}(x_v)$ where $u, v \in \mathcal{V}$ are two variables and π_{uv} is a permutation of $[k]$, the goal is to find a $[k]$-assignment to $\mathcal{V}$ so as to maximize the number of satisfied constraints.

For a problem instance Υ, let $\text{opt}(\Upsilon)$ denote the fraction of constraints satisfied by an optimum solution