SDP Integrality Gaps with Local ℓ_1-Embeddability

Subhash Khot
NYU
khot@cs.nyu.edu

Rishi Saket
Georgia Tech
saket@cc.gatech.edu

June 9, 2009

Abstract

We construct integrality gap instances for SDP relaxation of the MAXIMUM CUT and the SPARSEST CUT problems. If the triangle inequality constraints are added to the SDP, then the SDP vectors naturally define an n-point negative type metric where n is the number of vertices in the problem instance. Our gap-instances satisfy a stronger constraint that every sub-metric on $t = O((\log\log\log n)^{\frac{1}{6}})$ points is isometrically embeddable into ℓ_1. The local ℓ_1-embeddability constraints are implied when the basic SDP relaxation is augmented with t rounds of the Sherali-Adams LP-relaxation.

For the MAXIMUM CUT problem, we obtain an optimal gap of $\alpha_{GW}^{-1} - \varepsilon$, where α_{GW} is the Goemans-Williamson constant [GW95] and $\varepsilon > 0$ is an arbitrarily small constant. For the SPARSEST CUT problem, we obtain a gap of $\Omega((\log\log\log n)^{\frac{1}{13}})$. The latter result can be rephrased as a construction of an n-point negative type metric such that every t-point sub-metric is isometrically ℓ_1-embeddable, but embedding the whole metric into ℓ_1 incurs distortion $\Omega((\log\log\log n)^{\frac{1}{13}})$.

1 Introduction

For several well-studied problems such as MAXIMUM CUT and SPARSEST CUT, the best known approximation algorithms are based on a Semi-definite Programming relaxation. For MAXIMUM CUT, the basic SDP relaxation suffices to achieve the best-known approximation guarantee whereas for the SPARSEST CUT problem, adding additional constraints called the *triangle-inequality constraints* provably improves the approximation guarantee. Once these constraints are added, the SDP vectors naturally define a so-called negative type (or squared-ℓ_2) metric, and such metrics can be embedded *well* into the class of ℓ_1 metrics. After the ℓ_1-embedding is carried out, it is straightforward to output a good cut since n-point ℓ_1 metrics are precisely the convex combinations of *cut-metrics*. In general, it is a worthwhile (and of great current interest) goal to investigate whether stronger LP/SDP relaxations help, by adding (say polynomially many) natural constraints that an integral solution must satisfy. One natural family of constraints is to require that the negative type metric defined by the SDP vectors has an additional property that every sub-metric on t points embeds isometrically into ℓ_1. This certainly makes sense for the SPARSEST CUT problem since we would like the metric to be as close to ℓ_1 as possible. The local ℓ_1-embeddability condition can be enforced by adding $n^{O(t)}$ LP-constraints and requiring that the LP solution is *consistent* with the SDP vectors. Concretely, it suffices to add all LP constraints generated by t rounds of the Sherali-Adams LP hierarchy. In this paper, we give evidence that this approach is unlikely to yield *good* approximations. Specifically, we construct integrality gap instances for the SDP relaxation augmented with t rounds of the Sherali-Adams LP hierarchy. Let us first describe such a relaxation.

SDP Augmented with Sherali-Adams LP

For a cut-problem such as MAXIMUM CUT, t-rounds of the Sherali-Adams LP hierarchy (or $O(t)$-rounds if a somewhat different formulation is used, see [dlVKM07]) amount to the following: on a graph $G(V, E)$, for every subset S of up to t vertices, there is a distribution $D(S)$ on $\{-1, 1\}^S$, thought of as a distribution over cuts on S. The distributions $\{D(S)\}_{S\subseteq V,|S|\leq t}$ are mutually consistent in the sense that if $T \subseteq S \subseteq V,\ |S| \leq t$, then $D(S)|_T = D(T)$, i.e. the marginal of $D(S)$ on the subset T is exactly equal to $D(T)$. The value of such a solution is average over all edges $(u, v) \in E$, of the probability $p_{u,v}$ that u and v are separated by a random cut on the set $S = \{u, v\}$ sampled according to the distribution $D(S)$. On the other hand, a basic SDP relaxation (one used by Goemans and Williamson [GW95]) amounts to assigning a unit vector $\mathbf{w}_u$ for every vertex $u \in V$ and the value of the solution is average over edges $(u, v) \in E$, of the quantity $\frac{1-\langle \mathbf{w}_u,\mathbf{w}_v\rangle}{2}$. We say that the SDP solution is consistent with the Sherali-Adams solution if $\forall u, v \in V,\ p_{u,v} = \frac{1-\langle \mathbf{w}_u,\mathbf{w}_v\rangle}{2}$. Finally, a (c, s)-integrality gap (or c/s-gap if concerned only with the ratio) for a LP/SDP relaxation is a graph along with a LP/SDP solution such that the relaxation has value at least c whereas the true (integral) optimum, i.e. the relative size of the maximum cut, is at most s. As is standard, existence of an integrality gap instance is taken as evidence that an algorithm based on such relaxation cannot yield an approximation guarantee better than c/s.

MAXIMUM CUT

For the MAXIMUM CUT problem, a break-through result of Goemans and Williamson [GW95] showed that the integrality gap of the basic SDP relaxation is at most α_{GW}^{-1} where $\alpha_{GW} \approx 0.878$ is the optimum of a certain trigonometric function. Feige and Schechtman [FS02] gave a matching integrality gap instance with gap $\alpha_{GW}^{-1} - \varepsilon$. Khot and Vishnoi [KV05] showed that even after adding the triangle inequality constraints, the integrality gap is still lower bounded by $\alpha_{GW}^{-1} - \varepsilon$. This result is quite involved and especially, the proof that the triangle inequality constraints hold, is by brute-force with little intuitive explanation. In an incomparable result, Charikar, Makarychev, and Makarychev [CMM09] gave $(1 - \varepsilon, \frac{1}{2} + \varepsilon)$-integrality gap for the Sherali-Adams hierarchy even with $n^{c(\varepsilon)}$ rounds. However, Goemans and Williamson showed that for the basic SDP relaxation, the gap cannot be stronger than $(1-\varepsilon, 1-\Omega(\sqrt{\varepsilon}))$ and thus the Sherali-Adams relaxation is qualitatively different from the SDP relaxation. This suggests a natural question: what is the integrality gap if we combine the SDP with t rounds of the Sherali-Adams LP hierarchy, as in the previous section, for some large constant or super-constant t? We resolve this question in this paper:

Theorem (Informal). *Let $\varepsilon > 0$ be an arbitrarily small constant. For the* MAXIMUM CUT *problem on a graph of n vertices, the SDP relaxation augmented with $O((\log\log\log n)^{\frac{1}{6}})$ rounds of Sherali-Adams LP hierarchy has an integrality gap at least $\alpha_{GW}^{-1} - \varepsilon$.*

Consider the distance $d(u,v) := \|\mathbf{w}_u - \mathbf{w}_v\|^2$ defined on the set of vertices by the SDP vector solution $\{\mathbf{w}_u\}_{u\in V}$. An easy and well-known observation is that (see the last paragraph in Section 3.1) if the vector solution is consistent with t rounds of Sherali-Adams solution, then for any set $S \subseteq V, |S| \le t$, the space $(S, d(\cdot,\cdot))$ embeds isometrically into ℓ_1. In particular, the distance $d(\cdot,\cdot)$ satisfies triangle inequality. As we mentioned, in [KV05], the proof that the triangle inequalities hold is very technical. On the other hand, our construction, though not necessarily simpler, is quite intuitive and there is a reasonable explanation why it works. Our construction does use techniques from [KV05].

SPARSEST CUT

For the (uniform) SPARSEST CUT problem on n-vertex graphs, the basic SDP relaxation is very poor and has an integrality gap of $\Omega(n)$. In a recent break-through, Arora, Rao, and Vazirani [ARV04] showed that the gap improves to $O(\sqrt{\log n})$ after adding the triangle inequality constraints (i.e. the distance $d(\cdot,\cdot)$ is required to be a metric). Arora, Lee, and Naor [ALN08] proved essentially the same upper bound even for the more general non-uniform SPARSEST CUT problem. In fact, it had been conjectured earlier by Goemans and Linial that the integrality gap for non-uniform SPARSEST CUT problem is at most a universal constant. This is equivalent to a conjecture that n-point negative type metrics embed into ℓ_1 with constant distortion. Khot and Vishnoi [KV05] disproved the conjecture by constructing an n-point negative type metric with ℓ_1-distortion at least $(\log\log n)^{\Omega(1)}$. The lower bound was subsequently improved to $\Omega(\log\log n)$ by Krauthgamer and Rabani [KR06], and $\Omega(\log\log n)$ for the uniform version by Devanur *et al.* [DKSV06]. Lee and Naor [LN06] proposed a different counter-example to the Goemans-Linial conjecture, and the works of Cheeger, Kleiner, and Naor [CK06a, CK06b, CKN09] showed that this counter-example gives a further improved lower bound of $(\log n)^{\Omega(1)}$ (the upper bound is $\tilde{O}(\sqrt{\log n})$ as mentioned before).

In light of the extensive research on the SPARSEST CUT integrality gap, it is natural to investigate whether the integrality gap becomes a constant if we require the negative type metric $d(\cdot,\cdot)$ to have the property that every sub-metric on t points embeds isometrically into ℓ_1. We provide a negative answer:

Theorem (Informal). *For the* SPARSEST CUT *problem on a graph of n vertices, the SDP relaxation augmented with $O((\log\log\log n)^{\frac{1}{6}})$ rounds of Sherali-Adams has an integrality gap at least $\Omega((\log\log\log n)^{\frac{1}{13}})$. Also, there is an n-point negative type metric such that every sub-metric on $O((\log\log\log n)^{\frac{1}{6}})$ points is isometrically ℓ_1-embeddable, but embedding the whole metric into ℓ_1 incurs distortion $\Omega((\log\log\log n)^{\frac{1}{13}})$.*

We note that in an incomparable result, Charikar, Makarychev, and Makarychev [CMM09] gave integrality gap of $\Omega\left(\sqrt{\frac{\log n}{\log t + \log\log n}}\right)$ for t rounds of the Sherali-Adams heirarchy (without the SDP). This amounts to an ℓ_1 lower bound for (general, not negative-type) metrics such that any sub-metric on t points is isometrically ℓ_1-embeddable.

Inapproximability Results via the Unique Games Conjecture

All the results mentioned so far are intimately connected with the Unique Games Conjecture of [Kho02]. The conjecture states that approximating the UNIQUE GAMES problem (see Definition 3) is NP-hard and is proposed as an avenue towards proving strong inapproximability results for many NP-hard problems. Indeed, assuming the conjecture, Khot *et al.* [KKMO07] proved that it is NP-hard to approximate the MAXIMUM CUT problem within any factor strictly less than α_{GW}^{-1}. This means that, assuming the UGC and that P $\neq$ NP, any LP/SDP relaxation for MAXIMUM CUT with polynomially many constraints, must have integrality gap arbitrarily close to α_{GW}. Thus integrality gap instances for potentially more and more powerful LP/SDP relaxations give more and more evidence towards the truth of the UGC.

[KV05] used this connection in the reverse direction to actually construct integrality gap instances for cut-problems. They (and independently Chawla *et al.* [CKK+06]) gave a *reduction* from the UNIQUE GAMES problem to SPARSEST CUT and showed that if the UGC is true, then SPARSEST CUT has no constant approximation and hence Goemans-Linial conjecture must be false. This observation led [KV05] to a construction of an integrality gap instance for the UNIQUE GAMES problem (see the SDP in Figure 1) and then they *translated* this instance into an integrality gap instance for SPARSEST CUT via the reduction alluded to before. A nice feature of the reduction is that it allows a translation of the SDP solution as well, i.e. starting with a vector solution for the UNIQUE GAMES SDP, one can construct a vector solution for SPARSEST CUT SDP in a natural way. However an unsatisfying feature of [KV05] is that there is no intuitive reason why the SPARSEST CUT vector solution obeys triangle inequalities. As we said, we are able to (at least partially) fix this, though following the same high-level methodology.

We now explain how our result fits with Raghavendra's recent result [Rag08]. Raghavendra shows that for every *constraint satisfaction problem*, there is a certain generic relaxation such that, any integrality gap instance for this relaxation with gap α, can be translated into a UGC based hardness result with the hardness factor same as α. The relaxation he uses is exactly the combination of a basic SDP and a constant number of rounds of the Sherali-Adams LP (the number of rounds is at most $O(k+q)$ for a k-ary CSP over q-ary alphabet)! An implication of his result is that (assuming UGC and that P $\neq$ NP) adding more constraints to the generic relaxation does not help. We believe that the results in our paper can be generalized to arbitrary CSPs though we haven't attempted this so as to keep the presentation readable. If so, it would partially confirm Raghavendra's implication, namely that adding more Sherali-Adams rounds to the generic relaxation does not help.

Other LP and SDP Hierarchies

Finally, a few words about other LP and SDP hierarchies are in order. Recent works have obtained integrality gap results for many different problems (cut problems, vertex cover, independent set, 3SAT etc.) for LP and/or SDP relaxations in different hierarchies, i.e. Lovász-Schrijver, Sherali-Adams, and Lasserre. A full overview of these results is beyond the scope of this paper and we do not attempt it here. We would like to mention however that the Lasserre hierarchy is the most powerful one and it remains a challenging open problem to prove Lasserre integrality gaps. A t-round Lasserre includes, for example, the basic SDP as well as t rounds of Sherali-Adams LP. It is conceivable that the techniques in our paper could be applied towards obtaining strong Lasserre integrality gaps.

2 Overview of Our Construction

The constructions for the MAXIMUM CUT and the SPARSEST CUT integrality gaps are very similar (one only needs to change a certain *perturbation parameter*) and therefore, for the sake of exposition we shall focus only the MAXIMUM CUT integrality gap.

High level strategy

Our construction relies in large part on the work of Khot and Vishnoi [KV05] who gave SDP integrality gap examples for UNIQUE GAMES and cut-problems including MAXIMUM CUT. Their overall approach was to follow the reduction from UNIQUE GAMES to the target problem (say) MAXIMUM CUT. They first construct an integrality gap example for the UNIQUE GAMES SDP, i.e. an instance with low optimum (i.e. no good labeling) and a vector solution with high objective value. Using the reduction from [KKMO07], they convert the instance of UNIQUE GAMES with low optimum to an instance of MAXIMUM CUT, also with low optimum. The same reduction also transforms the vector solution for the UNIQUE GAMES SDP into a vector solution for the MAXIMUM CUT SDP. The transformation ensures that the MAXIMUM CUT SDP solution has a high objective value, thereby providing an integrality gap. In this work, we observe that there is also a natural way to construct a *good* solution to the Sherali-Adams LP relaxation for the UNIQUE GAMES instance constructed in [KV05]. This solution can then be transformed into one for the Sherali-Adams LP relaxation for the MAXIMUM CUT instance, via the same reduction as before. Again, the transformation ensures that the objective value of the Sherali-Adams solution remains high. Moreover, for any set of two vertices, the Sherali-Adams solution is *almost* consistent with the SDP vector solution. We then *massage* these solutions so that they are exactly consistent, yielding the integrality gap for MAXIMUM CUT SDP augmented with super-constant rounds of Sherali-Adams LP. The next few paragraphs give an informal description of the construction.

Sherali-Adams solution (labeling) to Unique Games instance

We start with the UNIQUE GAMES instance $\mathcal{U}$ constructed by Khot and Vishnoi [KV05]. Let $G(V, E)$ be its constraint graph and $[N]$ be the label set. The first step is to construct Sherali-Adams solution for $\mathcal{U}$. Specifically, we construct for every set $U \subseteq V,\ |U| \leq t$, a distribution $D(U)$ over labelings $\sigma : U \mapsto [N]$ such that:

- The distributions are mutually consistent, i.e. for any $W \subseteq U \subseteq V,\ |U| \leq t$, $D(U)|_W = D(W)$.
- The objective value of the solution is high, i.e. if $(u, v) \in E$ is a UNIQUE GAMES constraint, then a random labeling $\sigma : \{u, v\} \mapsto [N]$ from $D(\{u, v\})$ satisfies the constraint with probability close to 1.

Towards this end, we look at the [KV05] example closely, and observe that one can define a metric $\rho(\cdot, \cdot)$ on the vertex set V such that any two vertices with an edge/constraint between them are very close w.r.t. ρ. Moreover for any set $U \subseteq V,\ |U| \leq t$ that has low diameter w.r.t. ρ, it is possible to assign a randomized labeling $\sigma : U \mapsto [N]$ that satisfies all the constraints inside U. The labeling has a very strong consistency property that we do not describe here. This property ensures that for any subset $W \subseteq U$ (it also has a low diameter), the randomized labeling $\tau : W \mapsto [N]$ is same as $(\sigma : U \mapsto [N])|_W$ in distribution. In other words, we construct mutually consistent Sherali-Adams distributions $D(U)$ for all sets U having low ρ-diameter.

However, the Sherali-Adams relaxation requires us to define a randomized labeling $D(U)$ for *every* set of size at most t. Here is a natural idea: for an arbitrary set U, partition it (possibly in a randomized way) into sets of *low* ρ-diameter (call these clusters), and then label each cluster as earlier. Such partitioning schemes are well-known in the literature on metric embeddings. For us, the issue however is the consistency between sets. For $W \subseteq U$, we desire that the partition of W on its own is same as partition of W induced by a partition of U (in distribution if the partitioning scheme is randomized). At this point, we observe that the metric ρ can be chosen to be an ℓ_2 metric on points of a unit sphere. The sphere has unrestricted dimension, but if look only at a set $U \subseteq V,\ |U| \leq t$, then U can be thought of as embedded onto $(t-1)$-dimensional unit sphere $\mathbb{S}^{t-1}$ via a random orthogonal transformation. Now we partition $\mathbb{S}^{t-1}$ into clusters with low diameter using a well-known partitioning scheme and that automatically gives a partition of U into low diameter clusters. Since the partition of U depends only on its ℓ_2 geometry, it follows that if $W \subseteq U$, then partition of W is consistent in distribution with that induced from a partition of U!

A somewhat magical part is coming up with the ℓ_2 metric ρ. It turns out that the metric can be constructed from the SDP solution to the Unique Games instance. The solution consists (up to a normalization) of an orthonormal tuple $\{\mathbf{T}_{u,j}\}_{j\in[N]}$ for every vertex $u \in V$. Roughly speaking, desired metric ρ should capture the closeness between these tuples. Defining a single unit vector $\mathbf{T}_u$ from the tuple by $\mathbf{T}_u := \frac{1}{\sqrt{N}} \sum_{j\in[N]} \mathbf{T}_{u,j}^{\otimes 4}$, the ℓ_2 metric $\|\mathbf{T}_u - \mathbf{T}_v\|$ captures the closeness between tuples. This is the metric ρ that we desire.

Sherali-Adams solution to Maximum Cut

It is quite straightforward to translate the t-round Sherali-Adams solution for the Unique Games instance $\mathcal{U}$ into a t-round Sherali-Adams solution for the Maximum Cut instance $\mathcal{I}$. In the reduction of [KV05, KKMO07], a Unique Games vertex is replaced by a N-dimensional boolean hypercube where the N labels correspond to the N dimensions of the hypercube. Roughly speaking, the Sherali-Adams solution to Unique Games instance defines a labeling to its vertices. Each label corresponds to a dimension of a hypercube and the hypercube can be cut along that dimension. This yields Sherali-Adams solution for the Maximum Cut instance.

Approximately consistent SDP solution to Maximum Cut

In a similar way to [KV05], the vector solution for $\mathcal{U}$ can be transformed into one for $\mathcal{I}$ via a certain tensoring operation. We need to ensure that for the instance $\mathcal{I}$, the Sherali-Adams solution at the second level and the SDP vector solution are consistent, at least approximately. Unfortunately, we do not know whether this is true. We get around this problem in the following manner (which is possibly another place where some magic happens):

The Sherali-Adams solution for the Unique Games instance (and therefore the Maximum Cut instance) is parameterized by r, that specifies how *low* the diameter of the clusters is. On the other hand, the SDP solution for Maximum Cut instance is parameterized by an integer s, that specifies how *large* a tensor power is used. We appropriately choose a large number of pairs $\{(r_i, s_i)\}_{i=1}^{\Delta}$. For every choice of index i, we have a Sherali-Adams solution and the SDP solution parameterized by the diameter parameter r_i and the tensor-power parameter s_i. Finally, we define overall Sherali-Adams and SDP solutions to be the *combinations* of i^{th} solutions for $i \in \{1, \ldots, \Delta\}$. The crux of our argument is to show that for all but two values of $i \in [\Delta]$, the i^{th} Sherali-Adams and SDP solutions are almost consistent. Choosing Δ large, we see that the overall Sherali-Adams and SDP solutions are almost (i.e. approximately) consistent.

Correction step

Finally we massage the Sherali-Adams and the SDP solutions for Maximum Cut and ensure that the two are perfectly consistent with each other. The change in the LP/SDP objective value is negligible.

Organization of the paper

In Section 3, we formally define the problems Unique Games, Maximum Cut and Sparsest Cut, describe the relaxations we consider, and state our results. In Section A, we describe the construction of *local* labelings to sets of Unique Games vertices with low diameter under the appropriate metric ρ. In Section B, the Maximum Cut instance is derived from Unique Games instance via the same reduction as in [KV05]. Section C contains the construction of Sherali-Adams and SDP solutions to the MAXIMUM CUT instance that are approximately consistent. In Section D the approximate solution is transformed to a feasible one and the value of the integrality gap is computed. The construction for SPARSEST CUT is very similar to the one for MAXIMUM CUT and we only sketch it in Section E.

3 Preliminaries

We first formally define the MAXIMUM CUT, SPARSEST CUT, and UNIQUE GAMES problems.

Definition 1. *(*MAXIMUM CUT*) For a weighted graph $G = (V, E)$ with non-negative weights $\mathbf{wt}(e)$ for each edge $e \in E$, the goal is to find a cut that maximizes the weight of crossing edges, i.e. to maximize the following objective function,*

$$\max_{\emptyset \neq S \subseteq V} \sum_{e \in E(S, \overline{S})} \mathbf{wt}(e).$$

Definition 2. *(*non-uniform SPARSEST CUT*) Given a graph $G = (V, E)$ with non-negative weights $\mathbf{wt}(e)$ and demands $\mathbf{dem}(e)$ for each edge, the goal is to find a cut to minimize the following,*

$$\min_{\emptyset \neq S \subseteq V} \frac{\sum_{e \in E(S, \overline{S})} \mathbf{wt}(e)}{\sum_{e \in E(S, \overline{S})} \mathbf{dem}(e)}.$$

Definition 3. *An instance of* UNIQUE GAMES *$\mathcal{U}(G(V, E), [N], \{\pi_e\}_{e \in E})$ is a constraint satisfaction problem. For every edge $e = (u, v)$ in the graph, there is a bijection $\pi_e : [N] \mapsto [N]$ on the label set $[N]$. A labeling $\sigma : V \mapsto [N]$ satisfies an edge $e = (u, v) \in E$ iff $\pi_e(\sigma(u)) = \sigma(v)$. The goal is to find a labeling that satisfies maximum fraction of edges.*

The Unique Games Conjecture of Khot [Kho02] states the following:

Conjecture 1. *For arbitrarily small constants $\varepsilon, \delta > 0$, there is a positive integer $N = N(\varepsilon, \delta)$ such that, given an instance $\mathcal{U}$ of* UNIQUE GAMES *with label set $[N]$, it is* NP*-hard to distinguish between the following two cases:*

YES Case: There is a labeling to the vertices of $\mathcal{U}$ that satisfies at least $1 - \varepsilon$ fraction of the edges.

NO Case: There is no labeling that satisfies even δ fraction of the edges of $\mathcal{U}$.

Let $\mathcal{U}$ be the instance as described in Definition 3. Figure 1 gives a natural SDP relaxation SDP-UG. The relaxation is over the vector variables $\mathbf{x}_{u,i}$ for every vertex u of the graph G and label $i \in [N]$. Regarding the integrality gap of this relaxation, Khot and Vishnoi [KV05] proved the following Theorem. We will make use of their gap example.

Theorem 2. *There is a* UNIQUE GAMES *instance $\mathcal{U}_\eta(G(V, E), [N], \{\pi_e\}_{e \in E})$ where $\eta > 0$ is a parameter, such that any labeling to $\mathcal{U}_\eta$ satisfies at most $\frac{1}{N^\eta}$ fraction of the edges, whereas there exists a solution to the relaxation* SDP-UG *with an objective value of at least $1 - 4\eta$.*

3.1 Relaxations for MAXIMUM CUT and SPARSEST CUT

The relaxation we consider for the MAXIMUM CUT and the SPARSEST CUT problems is a combination of a basic SDP and t rounds of the Sherali-Adams LP hierarchy. Let $G = (V, E, \mathbf{wt})$ be a weighted graph.

The relaxation for the MAXIMUM CUT problem, which we denote by SDP-MC(t), is given in Figure 2. The SDP component consists of a unit vector $\mathbf{w}_u$ for every vertex $u \in V$. The LP component consists

$$\max \sum_{e=(u,v)\in E} \sum_{i\in[N]} \left\langle \mathbf{x}_{u,i}, \mathbf{x}_{v,\pi_e(i)} \right\rangle$$

Subject to,

$$\forall u \in V \qquad \textstyle\sum_{i\in[N]} \|\mathbf{x}_{u,i}\|^2 = 1 \quad (1)$$

$$\forall u \in V, i,j \in [N], i \neq j \qquad \langle \mathbf{x}_{u,i}, \mathbf{x}_{u,j} \rangle = 0 \quad (2)$$

$$\forall u,v \in V, i,j \in [N] \qquad \langle \mathbf{x}_{u,i}, \mathbf{x}_{v,j} \rangle \geq 0 \quad (3)$$

Figure 1: Relaxation `SDP-UG` for UNIQUE GAMES.

of, for every set $S \subseteq V,\ |S| \leq t$, a distribution $D(S)$ over $\{-1,1\}$-assignments to S. The distribution is specified by the probabilities $\{x_{S,\sigma} \mid \sigma \in \{-1,1\}^S\}$ and it can be thought of as a distribution on cuts of S. We ensure the consistency between any sets $T \subseteq S,\ |S| \leq t$, i.e. the distribution of cuts on T is same as the one induced by a distribution of cuts on S. Finally, we ensure that the SDP solution is consistent with the LP solution for every set $S = \{u,v\}$ of size two. Specifically, let y_u and y_v be the marginals of the distribution $D(S)$ on $\{-1,1\}^S$ onto the co-ordinates u and v respectively. Constraint (4) states the consistency requirement:

$$\langle \mathbf{w}_u, \mathbf{w}_v \rangle = \mathrm{E}_{D(S)}[y_u y_v].$$

For the SPARSEST CUT problem we have an additional parameter $\mathbf{dem}(e)$ for each edge e in the graph. In this case the objective function is the following.

$$\min \frac{\sum_{e=\{u,v\}\in E} \mathbf{wt}(e) \left(\frac{1-\langle \mathbf{w}_u, \mathbf{w}_v\rangle}{2}\right)}{\sum_{e=\{u,v\}\in E} \mathbf{dem}(e) \left(\frac{1-\langle \mathbf{w}_u, \mathbf{w}_v\rangle}{2}\right)}$$

We normalize the denominator to 1 and add this as a constraint. Figure 3 gives the relaxation `SDP-SC`(t) for the SPARSEST CUT problem.

Local ℓ_1-Embeddability: We observe that Constraints (1)-(4) imply that the distance function $d(u,v) := \|\mathbf{w}_u - \mathbf{w}_v\|^2$ defines a metric such that any sub-metric on at most t points is isometrically embeddable into ℓ_1. Indeed, fix any set $S \subseteq V, |S| \leq t$. Constraint (4) implies for any pair $u,v \in S$, $\langle \mathbf{w}_u, \mathbf{w}_v \rangle = \mathrm{E}_{D(S)}[y_u y_v]$, where y_u is the marginal of the distribution $D(S)$ onto u. Thus the mapping $u \mapsto y_u$ gives the isometric ℓ_1-embedding of the sub-metric $(S, d(\cdot,\cdot))$.

3.2 Our Results

We prove the following two theorems about the integrality gaps of the relaxations `SDP-MC`(t) and `SDP-SC`(t). The first theorem is proved in Sections A through D, whereas for the proof of the second theorem, we give a brief sketch in Section E.

$$\max \sum_{e=\{u,v\}\in E} \mathbf{wt}(e) \frac{1-\langle \mathbf{w}_u, \mathbf{w}_v\rangle}{2}$$

Subject to,

$$\forall u \in V \qquad \|u\|^2 = 1 \qquad (1)$$

$$\forall S \subseteq V, \sigma \in \{-1,1\}^S \text{ s.t. } |S| \le t \qquad 0 \le x_{S,\sigma} \le 1,\ \textstyle\sum_{\sigma\in\{-1,1\}^S} x_{S,\sigma} = 1 \qquad (2)$$

$$\forall T \subseteq S \subseteq V, |S| \le t, \tau \in \{-1,1\}^T \qquad \sum_{\substack{\sigma \in \{-1,1\}^S \\ \sigma|_T = \tau}} x_{S,\sigma} = x_{T,\tau} \qquad (3)$$

$$\forall u, v \in V \qquad \sum_{\sigma\in\{-1,1\}^{\{u,v\}}} \sigma(u)\sigma(v) \cdot x_{\{u,v\},\sigma} = \langle \mathbf{w}_u, \mathbf{w}_v\rangle \qquad (4)$$

Figure 2: Relaxation $\mathsf{SDP\text{-}MC}(t)$ for MAXIMUM CUT

Theorem 3. *For all $\varepsilon > 0$, there is an instance $\mathcal{I}$ of* MAXIMUM CUT *on (sufficiently large) n vertices, such that for $t = O((\log\log\log n)^{\frac{1}{6}})$,*

$$\frac{\mathsf{FRAC}(\mathcal{I})}{\mathsf{OPT}(\mathcal{I})} \ge \alpha_{GW}^{-1} - \varepsilon,$$

where $\mathsf{OPT}(\mathcal{I})$ is the optimum value of MAXIMUM CUT *on $\mathcal{I}$, $\mathsf{FRAC}(\mathcal{I})$ is the optimal objective value of $\mathsf{SDP\text{-}MC}(t)$ on $\mathcal{I}$, and α_{GW} is the Goemans-Williamson constant, i.e. $\alpha_{GW} = \min_{\rho\in[-1,1]} \frac{\arccos(\rho)/\pi}{(1-\rho)/2}$.*

Theorem 4. *There is an instance $\mathcal{I}$ of* SPARSEST CUT *on (sufficiently large) n vertices, such that for $t = O((\log\log\log n)^{\frac{1}{6}})$,*

$$\frac{\mathsf{OPT}(\mathcal{I})}{\mathsf{FRAC}(\mathcal{I})} \ge \Omega((\log\log\log n)^{\frac{1}{13}}),$$

where $\mathsf{OPT}(\mathcal{I})$ is the optimum value of SPARSEST CUT *on the instance $\mathcal{I}$, $\mathsf{FRAC}(\mathcal{I})$ is the optimal objective value of $\mathsf{SDP\text{-}SC}(t)$ on $\mathcal{I}$.*

$$\min \sum_{e=\{u,v\}\in E} \mathbf{wt}(e) \frac{1-\langle \mathbf{w}_u, \mathbf{w}_v\rangle}{2}$$

Subject to,

$$\forall u \in V \qquad \|u\|^2 = 1 \qquad (1)$$

$$\forall S \subseteq V, \sigma \in \{-1,1\}^S \text{ s.t. } |S| \le t \qquad 0 \le x_{S,\sigma} \le 1,\ \textstyle\sum_{\sigma\in\{-1,1\}^S} x_{S,\sigma} = 1 \qquad (2)$$

$$\forall T \subseteq S \subseteq V, |S| \le t, \tau \in \{-1,1\}^T \qquad \sum_{\substack{\sigma \in \{-1,1\}^S \\ \sigma|_T = \tau}} x_{S,\sigma} = x_{T,\tau} \qquad (3)$$

$$\forall u, v \in V \qquad \sum_{\sigma\in\{-1,1\}^{\{u,v\}}} \sigma(u)\sigma(v) \cdot x_{\{u,v\},\sigma} = \langle \mathbf{w}_u, \mathbf{w}_v\rangle \qquad (4)$$

$$\sum_{e=\{u,v\}\in E} \mathbf{dem}(e) \left(\frac{1-\langle \mathbf{w}_u, \mathbf{w}_v\rangle}{2}\right) = 1 \qquad (5)$$

Figure 3: Relaxation $\mathsf{SDP\text{-}SC}(t)$ for SPARSEST CUT

References

[ALN08] S. Arora, James R. Lee, and A. Naor. Euclidean distortion and the sparsest cut. *J. AMS*, 21(1):1–21, 2008.

[ARV04] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings and graph partitioning. In *Proc.* 36^{th} *ACM STOC*, pages 222–231, 2004.

[CK06a] J. Cheeger and B. Kleiner. Differentiating maps into L^1 and the geometry of BV functions. Preprint, 2006.

[CK06b] J. Cheeger and B. Kleiner. On the differentiation of Lipschitz maps from metric measure spaces to Banach spaces. Preprint, 2006.

[CKK+06] S. Chawla, R. Krauthgamer, R. Kumar, Y. Rabani, and D. Sivakumar. On the hardness of approximating multicut and sparsest-cut. *Computational Complexity*, 15(2):94–114, 2006.

[CKN09] J. Cheeger, B. Kleiner, and A. Naor. Quantitative bounds on the rate of central collapse of Lipschitz maps on the Heisenberg group. Preprint, 2009.

[CMM09] M. Charikar, K. Makarychev, and Y. Makarychev. Integrality gaps for Sherali-Adams relaxations. To Appear, 2009.

[DKSV06] N. Devanur, S. Khot, R. Saket, and N. Vishnoi. Integrality gaps for sparsest cut and minimum linear arrangement problems. In *Proc.* 38^{th} *ACM STOC*, pages 537–546, 2006.

[dlVKM07] W. F. de la Vega and C. Kenyon-Mathieu. Linear programming relaxations of maxcut. In *ACM SODA*, pages 53–61, 2007.

[FS02] U. Feige and G. Schechtman. On the optimality of the random hyperplane rounding technique for max cut. *Random Struct. Algorithms*, 20(3):403–440, 2002.

[GKL03] A. Gupta, R. Krauthgamer, and J. R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proc.* 44^{th} *IEEE FOCS*, 2003.

[GW95] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995.

[Kho02] S. Khot. On the power of unique 2-prover 1-round games. In *Proc.* 34^{th} *ACM STOC*, pages 767–775, 2002.

[KKMO07] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007.

[KR06] R. Krauthgamer and Y. Rabani. Improved lower bounds for embeddings into l_1. In *ACM SODA*, pages 1010–1017, 2006.

[KV05] S. Khot and N. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into l_1. In *Proc.* 46^{th} *IEEE FOCS*, pages 53–62, 2005.

[LN06] J. R. Lee and A. Naor. l_p metrics on the Heisenberg group and the Goemans-Linial conjecture. In *Proc.* 47^{th} *IEEE FOCS*, pages 99–108, 2006.

[Rag08] P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proc.* 40^{th} *ACM STOC*, pages 245–254, 2008.

A UNIQUE GAMES Instance

In this section we state the relevant properties of the Unique Games instance and the corresponding SDP solution constructed by Khot and Vishnoi [KV05]. For parameters $\eta > 0$ and $N = 2^m$ for some $m \in \mathbb{Z}^+$, Khot and Vishnoi [KV05] construct the UNIQUE GAMES instance $\mathcal{U}_\eta(G(V,E),[N],\{\pi_e\}_{e\in E})$ where the number of vertices $|V| = 2^N/N$. [1] The instance has no good labeling, i.e. has low optimum.

Lemma 5. *Any labeling to the vertices of the* UNIQUE GAMES *instance* $\mathcal{U}_\eta(G(V,E),[N],\{\pi_e\}_{e\in E})$ *satisfies at most* $\frac{1}{N^\eta}$ *fraction of the edges.*

In construction of [KV05] the elements of $[N]$ are identified with the additive group $(\mathbb{F}[2]^m, \oplus)$. The authors construct a vector solution that consists of unit vectors $\mathbf{T}_{u,i}$ for every vertex $u \in V$ and label $i \in [N]$. These vectors (up to a normalization) form the solution to the UNIQUE GAMES SDP relxation SDP-UG. We highlight the important properties of the SDP solution below:

Properties of the Unique Games SDP Solution

- **(Orthonormality)** $\forall\, u \in V,\ \ \forall\, i \neq j \in [N]$,

$$\|\mathbf{T}_{u,i}\| = 1, \quad \langle \mathbf{T}_{u,i}, \mathbf{T}_{u,j}\rangle = 0. \tag{1}$$

- **(Non-negativity)** $\forall\, u, v \in V,\ \ \forall\, i, j \in [N]$,

$$\langle \mathbf{T}_{u,i}, \mathbf{T}_{v,j}\rangle \geq 0. \tag{2}$$

- **(Symmetry)** $\forall\, u, v \in V,\ \ \forall\, i, j, k \in [N]$,

$$\langle \mathbf{T}_{u,i}, \mathbf{T}_{v,j}\rangle = \langle \mathbf{T}_{u,k\oplus i}, \mathbf{T}_{v,k\oplus j}\rangle \tag{3}$$

  where '$\oplus$' is the group operation on $[N]$ as described above.

- **(High SDP Value)** For every edge $e = (v,w) \in E$,

$$\forall\, i \in [N], \quad \left\langle \mathbf{T}_{v,i}, \mathbf{T}_{w,\pi_e(i)}\right\rangle \geq 1 - 4\eta. \tag{4}$$

  In fact, there is $k_e \in [N]$ such that $\forall\, i \in [N],\ \ \pi_e(i) = k_e \oplus i$.

We now define for every vertex $u \in V$ a unit vector $\mathbf{T}_u$ as follows (it is a unit vector due to orthonormality condition (1)),

$$\forall u \in V \quad \mathbf{T}_u := \frac{1}{\sqrt{N}} \sum_{i\in[N]} \mathbf{T}_{u,i}^{\otimes 4}. \tag{5}$$

Our main idea is that the Euclidean distances between the vectors $\{\mathbf{T}_u\}_{u\in V}$ are a measure of the 'closeness' between the orthonormal tuples $\{\mathbf{T}_{u,i} \mid i \in [N]\}$. Specifically:

Lemma 6. *For every* $u, v \in V$,

$$\min_{i,j\in[N]} \|\mathbf{T}_{u,i} - \mathbf{T}_{v,j}\| \ \leq\ \|\mathbf{T}_u - \mathbf{T}_v\| \ \leq\ 2 \cdot \min_{i,j\in[N]} \|\mathbf{T}_{u,i} - \mathbf{T}_{v,j}\|. \tag{6}$$

[1] For the sake of simplicity, we have slightly altered the presentation from [KV05].

Proof: Note that

$$1-\frac{1}{2}\|\mathbf{T}_u-\mathbf{T}_v\|^2=\langle\mathbf{T}_u,\mathbf{T}_v\rangle=\left\langle\frac{1}{\sqrt{N}}\sum_{i\in[N]}\mathbf{T}_{u,i}^{\otimes 4},\frac{1}{\sqrt{N}}\sum_{j\in[N]}\mathbf{T}_{v,j}^{\otimes 4}\right\rangle=\frac{1}{N}\sum_{i\in[N]}\left(\sum_{j\in[N]}\langle\mathbf{T}_{u,i},\mathbf{T}_{v,j}\rangle^4\right).$$

Due to symmetry (i.e. condition (3)), the inner sum above is the same for every index $i\in[N]$. Therefore fixing some $i_0\in[N]$,

$$1-\frac{1}{2}\|\mathbf{T}_u-\mathbf{T}_v\|^2=\sum_{j\in[N]}\langle\mathbf{T}_{u,i_0},\mathbf{T}_{v,j}\rangle^4. \tag{7}$$

Since $\{\mathbf{T}_{v,j}\}_{j\in[N]}$ is an orthonormal set (and using non-negativity condition (2)), we have

$$\sum_{j\in[N]}\langle\mathbf{T}_{u,i_0},\mathbf{T}_{v,j}\rangle^4\le\max_{j\in[N]}\langle\mathbf{T}_{u,i_0},\mathbf{T}_{v,j}\rangle=1-\frac{1}{2}\min_{j\in[N]}\|\mathbf{T}_{u,i_0}-\mathbf{T}_{v,j}\|^2. \tag{8}$$

Combining (7) and (8), we get the left inequality in (6). On the other hand,

$$\begin{aligned}\sum_{j\in[N]}\langle\mathbf{T}_{u,i_0},\mathbf{T}_{v,j}\rangle^4 &\ge \max_{j\in[N]}\langle\mathbf{T}_{u,i_0},\mathbf{T}_{v,j}\rangle^4=\left(1-\frac{1}{2}\min_{j\in[N]}\|\mathbf{T}_{u,i_0}-\mathbf{T}_{v,j}\|^2\right)^4\\ &\ge 1-2\min_{j\in[N]}\|\mathbf{T}_{u,i_0}-\mathbf{T}_{v,j}\|^2. \end{aligned}\tag{9}$$

Combining (7) and (9), and using symmetry, we get the right inequality in (6). ∎

A.1 Local Consistency

Lemma 7. *Suppose $u,v\in V$ are such that $\|\mathbf{T}_u-\mathbf{T}_v\|\le\alpha\le 0.1$. Then there is a unique $k_{u,v}\in[N]$ such that*

$$\forall i\in[N],\quad \|\mathbf{T}_{u,i}-\mathbf{T}_{v,k_{u,v}\oplus i}\|\le\alpha. \tag{10}$$

Proof: Since $\|\mathbf{T}_u-\mathbf{T}_v\|\le\alpha$, by Lemma 6, there exist $i_0,j_0\in[N]$ such that $\|\mathbf{T}_{u,i_0}-\mathbf{T}_{v,j_0}\|\le\alpha$. Defining $k_{u,v}=i_0\oplus j_0$ and using symmetry, we satisfy the hypothesis of the lemma. For the uniqueness, suppose that $k_{u,v},k'_{u,v}$ both satisfy the hypothesis of the lemma. Then for any i,

$$\|\mathbf{T}_{v,k_{u,v}\oplus i}-\mathbf{T}_{v,k'_{u,v}\oplus i}\|\le\|\mathbf{T}_{v,k_{u,v}\oplus i}-\mathbf{T}_{u,i}\|+\|\mathbf{T}_{u,i}-\mathbf{T}_{v,k'_{u,v}\oplus i}\|\le\alpha+\alpha=2\alpha$$

Since $\{\mathbf{T}_{v,j}\mid j\in[N]\}$ is an orthonormal set, the distance between any two distinct vectors in this set is exactly $\sqrt{2}$. So one must have $k_{u,v}\oplus i=k'_{u,v}\oplus i$ and hence $k_{u,v}=k'_{u,v}$. ∎

Definition 4. *A set of vertices $V'\subseteq V$ is called 0.1-local if $\forall\, u,v\in V'$, $\|\mathbf{T}_u-\mathbf{T}_u\|\le 0.1$.*

Lemma 7 states that whenever two vertices u and v are close (in terms of the distance $\|\mathbf{T}_u-\mathbf{T}_v\|$), there is a unique matching $i\mapsto k_{u,v}\oplus i$ such that the orthonormal tuples $\{\mathbf{T}_{u,i}\mid i\in[N]\}$ and $\{\mathbf{T}_{v,j}\mid j\in[N]\}$ are close via this matching. The next lemma shows that for a set V' that is 0.1 local, the matchings induced between every pair of vertices in V' are consistent with each other.

Lemma 8 (Local Consistency). *Suppose a set V' is 0.1-local and $u,v,w\in V'$. Let $k_{u,v},k_{u,w},k_{v,w}\in[N]$ be the elements given by Lemma 7, i.e. $\forall i\in[N]$,*

$$\|\mathbf{T}_{u,i}-\mathbf{T}_{v,k_{u,v}\oplus i}\|\le 0.1,\quad \|\mathbf{T}_{u,i}-\mathbf{T}_{w,k_{u,w}\oplus i}\|\le 0.1,\quad \|\mathbf{T}_{v,i}-\mathbf{T}_{w,k_{v,w}\oplus i}\|\le 0.1.$$

Then $k_{v,w}=k_{u,v}\oplus k_{u,w}$.

Proof: By triangle inequality,

$$\|\mathbf{T}_{v,i} - \mathbf{T}_{w,k_{u,v}\oplus k_{u,w}\oplus i}\| \leq \|\mathbf{T}_{v,i} - \mathbf{T}_{u,k_{u,v}\oplus i}\| + \|\mathbf{T}_{u,k_{u,v}\oplus i} - \mathbf{T}_{w,k_{u,v}\oplus k_{u,w}\oplus i}\| \leq 0.1 + 0.1 = 0.2.$$

Since $\|\mathbf{T}_{v,i} - \mathbf{T}_{w,k_{v,w}\oplus i}\| \leq 0.1$, it follows that

$$\|\mathbf{T}_{w,k_{v,w}\oplus i} - \mathbf{T}_{w,k_{u,v}\oplus k_{u,w}\oplus i}\| \leq 0.3.$$

Now note that the set $\{\mathbf{T}_{w,j}\}_{j\in[N]}$ is orthonormal, so the distance between any two distinct vectors in this set is exactly $\sqrt{2}$. Therefore one must have $k_{v,w} \oplus i = k_{u,v} \oplus k_{u,w} \oplus i$, and hence $k_{v,w} = k_{u,v} \oplus k_{u,w}$. ∎

A.2 Construction of local labelings

Now we construct a (randomized) labeling $L_{V'}$ for any 0.1-local set $V' = \{u_1, \ldots, u_\ell\} \subseteq V$. Choose u_1 as the *pivot* vertex. We pick the label of u_1 to be a random $i \in [N]$ and let the label of every other vertex to be the *mate* of i via the induced matching between u_1 and that vertex. Thanks to Lemma 8, the labeling $L_{V'}$ does not depend on the choice of the pivot vertex. Formally, the labeling $L_{V'}$ is obtained as:

- Pick one vertex from V', say u_1.
- Choose the label of u_1 to be a random element $i \in [N]$.
- For $2 \leq p \leq \ell$, set the label of u_p to be $i \oplus k_{u_1,u_p}$.

A.3 Construction of labelings to arbitrary size-t sets

Let t be the universal parameter denoting the number of levels of Sherali-Adams relaxation our solution satisfies. We will now describe a procedure UG-LABEL which, given a parameter $r \leq 0.1$ and a subset $U \subseteq V$, $|U| \leq t$, outputs a (randomized) labeling to the vertices of U. Note that U need not be 0.1-local and is completely arbitrary. The idea is to first partition U into clusters such that each cluster is 0.1-local, and then each cluster is labeled according to (local) labeling procedure described in Section A.2. The algorithm UG-LABEL outputs the partition of U as well, along with a labeling to U.

The following Theorem can be inferred from [GKL03, Theorem 3.2] applied to the Euclidean unit sphere.

Theorem 9 ([GKL03]). *Let $\mathbb{S}^{t-1}$ denote the $(t-1)$ dimensional unit sphere. For every $r > 0$ there is a randomized partition $\tilde{P}(r)$ of $\mathbb{S}^{t-1}$ into disjoint clusters such that,*

1. *For every cluster $\tilde{C} \in \tilde{P}(r)$, $\tilde{C} \subseteq \mathbb{S}^{t-1}$, $\mathtt{diam}(\tilde{C}) \leq r$.*
2. *For any pair of points $x, y \in \mathbb{S}^{t-1}$ such that $\|x - y\| = \beta \leq \frac{r}{4}$,*

$$\Pr_{\tilde{P}(r)}\Big[x \text{ and } y \text{ fall into different clusters}\Big] \leq \frac{100\beta t}{r}.$$

Here is our randomized algorithm that outputs a labeling to an arbitrary set $U \subseteq V$ of size at most t, along with its partition into 0.1-local clusters.

Algorithm UG-LABEL (U, r), $r \leq 0.1$.

1. Embed the set of at most t unit vectors $\{\mathbf{T}_v \mid v \in U\}$ isometrically into the $(t-1)$-dimensional unit sphere $\mathbb{S}^{t-1}$ via a random orthogonal transformation.
2. Let $\tilde{P}(r)$ be the partition of $\mathbb{S}^{t-1}$ given by Theorem 9. This naturally induces a partition $P(r)$ of the set U via the above embedding.
3. Since every cluster $\tilde{C} \in \tilde{P}(r)$ has diameter at most 0.1, the corresponding cluster $C \in P(r)$ in the induced partition of U is 0.1-local (C is possibly empty).
4. To every non-empty cluster $C \in P, C \subseteq U$, assign the labeling L_C as in Section A.2.

Consistency between sets: For a parameter $r \leq 0.1$, the algorithm UG-LABEL defines a distribution $D_{UG,r}(U)$ over labelings to the vertices of U, for every subset $U \subseteq V$ such that $|U| \leq t$. From the algorithm it is clear that the labeling to U depends only on the (geometric configuration of the) corresponding vectors $\{\mathbf{T}_u\}_{u \in U}$. It follows that for any two sets $W \subseteq U \subseteq V$ such that $|U| \leq t$, $D_{UG,r}(U)|_W = D_{UG,r}(W)$. Therefore these distributions define a solution to t rounds of Sherali-Adams relaxation for UNIQUE GAMES.

B Construction of MAXIMUM CUT Instance

The MAXIMUM CUT instance is essentially the same as constructed in [KV05]. We describe it in brief. Let $\rho \in (-1, 0)$ be a parameter[2] and denote the instance constructed as $\mathcal{I}_\rho(V^*, E^*)$. We start with the UNIQUE GAMES instance $\mathcal{U}_\eta(G(V,E), [N], \{\pi_e\}_{e \in E})$ and replace each vertex $v \in V$ by a block of vertices $(v, \mathbf{x})$ where $\mathbf{x} \in \{-1,1\}^N$. Thus each block is an N-dimensional boolean hypercube. Let $\mathbf{x} \in_p \{-1,1\}^N$ denote a random string chosen from the p-biased distribution, i.e. every co-ordinate of $\mathbf{x}$ is chosen independently to be -1 with probability p and 1 with probability $1-p$.

For every pair of edges $e = (v,w), e' = (v,w') \in E$, there are (all possible) weighted edges between the blocks $(w, \cdot)$ and $(w', \cdot)$ in the instance $\mathcal{I}_\rho(V^*, E^*)$. The weight of an edge e^* between $(w, \mathbf{x})$ and $(w', \mathbf{y})$ is defined as:

$$\mathbf{wt}(e^*) := \Pr_{\substack{\mathbf{z} \in_{1/2} \{-1,1\}^N \\ \boldsymbol{\mu} \in_{\frac{1-\rho}{2}} \{-1,1\}^N}} \left[(\mathbf{x} = \mathbf{z} \circ \pi_e^{-1}) \wedge (\mathbf{y} = \mathbf{z}\boldsymbol{\mu} \circ \pi_{e'}^{-1}) \right],$$

where $\mathbf{z} \circ \pi := (z_{\pi(1)}, \ldots, z_{\pi(N)})$. The following theorem is proved in [KKMO07, KV05].

Theorem 10. *For any constants $\rho \in (-1,0)$ and $\lambda > 0$, there is a constant $c(\rho, \lambda)$ such that the following holds: Let $\mathcal{U}_\eta(G(V,E), [N], \{\pi_e\}_{e \in E})$ be an instance of* UNIQUE GAMES *with $OPT(\mathcal{U}_\eta) \leq c(\rho, \lambda)$, then the corresponding instance $\mathcal{I}_\rho$ as defined above satisfies the property that*

$$\mathsf{OPT}(\mathcal{I}_\rho) \leq \frac{1}{\pi} \arccos \rho + \lambda$$

where $\mathsf{OPT}(\mathcal{I}_\rho)$ is the normalized value of the maximum cut.

C Construction of Approximate Solution $\mathcal{A}$ to SDP-MC(t)

In this section we will describe the construction of an *approximate* solution for the relaxation SDP-MC(t) for the MAXIMUM CUT instance $\mathcal{I}_\rho(V^*, E^*)$. The parameter t is a superconstant which we shall explicitly define later. Our solution will satisfy all constraints of SDP-MC(t) except for the Constraint (4) which will

[2] For the MAXIMUM CUT problem, $\rho < 0$ will be chosen so that $\alpha_{GW} := \min_{\rho \in [-1,1]} \frac{2 \cdot \arccos(\rho)}{\pi(1-\rho)}$ is attained. For the SPARSEST CUT problem, $\rho = 1 - \delta$ will be close to 1.

be satisfied only approximately. More precisely, the solution $\mathcal{A}$ has two components $\mathcal{A} = (D_{\mathcal{A}}(\cdot), G_{\mathcal{A}})$ where for every set $S \subseteq V^*$ of size at most t, $D_{\mathcal{A}}(S)$ is a distribution over $\{-1, 1\}$-assignments over S and $G_{\mathcal{A}}$ is an assignment of unit vectors to V^*. The distributions $D_{\mathcal{A}}(S)$ satisfy the consistency property of the Sherali-Adams relaxation, i.e. for $T \subseteq S \subseteq V^*$, $|S| \leq t$, we have $D_{\mathcal{A}}(S) \mid_T = D_{\mathcal{A}}(T)$. Moreover, the vector solution $G_{\mathcal{A}}$ is approximately consistent with the Sherali-Adams solution at the second level, i.e. for any two vertices $a, b \in V^*$, if y_a, y_b are the marginals of the distribution $D_{\mathcal{A}}(\{a, b\})$ on either co-ordinate, then $\mathrm{E}[y_a y_b] \approx \langle G_{\mathcal{A}}(a), G_{\mathcal{A}}(b) \rangle$.

We first describe the distributions $D_{\mathcal{A}}(S)$. This is done in two stages. In the first stage, for a parameter $r \leq 0.1$, we construct distributions $D_{\mathcal{A},r}(S)$ and then in the second stage, we let $D_{\mathcal{A}}(S)$ to be the average of $D_{\mathcal{A},r_i}(S)$ for appropriately chosen sequence of parameters $\{r_i \mid i \in [\Delta]\}$. We will ensure that the distributions $D_{\mathcal{A},r_i}(S)$ (and therefore their average $D_{\mathcal{A},r}(S)$) satisfy the consistency property of the Sherali-Adams solution.

C.1 Construction of the Sherali-Adams solution $D_{\mathcal{A},r}(\cdot)$

Fix a parameter $r \leq 0.1$. For every set $S \subseteq V^*, |S| \leq t$, the distribution $D_{\mathcal{A},r}(S)$ is given by the following algorithm (i.e. the algorithm outputs a $\{-1, 1\}$-assignment to S in a randomized manner):

1. Let $U \subseteq V$ be defined as $U := \{v \mid (v, \mathbf{x}) \in S\}$ (recall that V is the set of vertices of the Unique Games instance from which the MAXIMUM CUT instance is derived). Clearly $|U| \leq t$.
2. Run UG-LABEL(U, r) to obtain a random labeling $\sigma : U \mapsto [N]$ and a partition $P = P(r)$ of U.
3. For every cluster $C \in P$ choose a value $\omega_C \in \{-1, 1\}$ at random uniformly and independently.
4. For every vertex $(v, \mathbf{x}) \in S$ such that $v \in C$, assign it the value $\mathbf{x}(\sigma(v)) \cdot \omega_C$.

Observe that the distributions $D_{\mathcal{A},r}(\cdot)$ satisfy the consistency property of the Sherali-Adams relaxation. This is inherited from the consistency property of the UG-LABEL algorithm.

C.2 Construction of the Sherali-Adams solution $D_{\mathcal{A}}(\cdot)$.

Let $\Delta := t^4$ and for $i \in [\Delta]$, define a decreasing sequence of radii:

$$r_i = 2^{-it}. \tag{11}$$

For any set $S \subseteq V^*, |S| \leq t$, the following algorithm defines the distribution $D_{\mathcal{A}}(S)$ over $\{-1, 1\}$-assignments to S.

1. Choose a random index $i \in [\Delta]$.
2. Output a random $\{-1, 1\}$-assignment to S according to the distribution $D_{\mathcal{A},r_i}(S)$.

C.3 Construction of vector solution $G_{\mathcal{A}}$

Finally we construct the vector solution $G_{\mathcal{A}}$ and show that it is approximately consistent with the Sherali-Adams solution $D_{\mathcal{A}}$ at the second level. For $i \in [\Delta]$, define an increasing sequence of integers s_i as,

$$s_i = 8 \cdot 2^{2it}. \tag{12}$$

Roughly speaking, for every $i \in [\Delta]$, there will be a vector solution $G_{\mathcal{A},s_i}$ parameterized by integer s_i, that approximately agrees with the Sherali-Adams solution $D_{\mathcal{A},r_i}$. However, as it turns out, this is not necessarily true for *every* i, but for *most* $i \in [\Delta]$ (in fact for all but two values). The values of i for which the approximation fails may depend on the pair of vertices under consideration. We will then define the overall vector solution $G_{\mathcal{A}}$ to be the combination (direct sum) of the solutions $G_{\mathcal{A},s_i}$ for $i \in \Delta$. Since $D_{\mathcal{A}}$ is an average of $D_{\mathcal{A},r_i}$, and $D_{\mathcal{A},r_i}$ approximately agrees with $G_{\mathcal{A},s_i}$ for most $i \in [\Delta]$, it would follow that $D_{\mathcal{A}}$ approximately agrees with $G_{\mathcal{A}}$.

Now we formally describe the construction. Let $(u, \mathbf{x}) \in V^*$ where $u \in V$ is a vertex of the UNIQUE GAMES instance and $\mathbf{x} \in \{-1, 1\}^N$.

For every $i \in [\Delta]$ we define the following unit vector:

Solution $G_{\mathcal{A},s_i}$:

$$\mathbf{G}^i_{(u,\mathbf{x})} := \frac{1}{\sqrt{N}} \sum_{k \in [N]} \mathbf{x}(k) \cdot \mathbf{T}^{\otimes s_i}_{u,k}. \tag{13}$$

Finally, we take direct sum of these vectors to construct the following unit vector:

Solution $G_{\mathcal{A}}$:

$$\mathbf{G}_{(u,\mathbf{x})} := \frac{1}{\sqrt{\Delta}} \left(\bigoplus_{i=1}^{\Delta} \mathbf{G}^i_{(u,\mathbf{x})} \right). \tag{14}$$

The following is the main theorem showing that the vector solution $G_{\mathcal{A}}$ approximately agrees with the Sherali-Adams solution $D_{\mathcal{A}}(\cdot)$ at the second level.

Theorem 11. *Let $(u, \mathbf{x})$ and $(v, \mathbf{y})$ be any two vertices of V^* where $u, v \in V$ and $\mathbf{x}, \mathbf{y} \in \{-1, 1\}^N$. Let $y_{(u,\mathbf{x})}$ and $y_{(v,\mathbf{y})}$ be the marginals of the $\{-1, 1\}$-assignment to the pair $S = \{(u, \mathbf{x}), (v, \mathbf{y})\}$, either under the distribution $D_{\mathcal{A}}(S)$ or under the distribution $D_{\mathcal{A},r_i}(S)$ (it will be clear from the context). Then,*

$$\left| \mathrm{E}_{D_{\mathcal{A}}} \left[y_{(u,\mathbf{x})} y_{(v,\mathbf{y})} \right] - \left\langle \mathbf{G}_{(u,\mathbf{x})}, \mathbf{G}_{(v,\mathbf{y})} \right\rangle \right| \le 2 \cdot 2^{-t/2} + \frac{2}{\Delta}. \tag{15}$$

Proof: Since $D_{\mathcal{A}}(\cdot)$ is an average of $D_{\mathcal{A},r_i}(\cdot)$, we have

$$\mathrm{E}_{D_{\mathcal{A}}} \Big[y_{(u,\mathbf{x})} y_{(v,\mathbf{y})} \Big] = \mathrm{E}_{i \in [\Delta]} \left[\mathrm{E}_{D_{\mathcal{A},r_i}} [y_{(u,\mathbf{x})} y_{(v,\mathbf{y})}] \right]. \tag{16}$$

Similarly, since the vector (14) is (up to normalization) direct sum of vectors in (13),

$$\left\langle \mathbf{G}_{(u,\mathbf{x})}, \mathbf{G}_{(v,\mathbf{y})} \right\rangle = \mathrm{E}_{i \in [\Delta]} \left[\left\langle \mathbf{G}^i_{(u,\mathbf{x})}, \mathbf{G}^i_{(v,\mathbf{y})} \right\rangle \right]. \tag{17}$$

We want to show that the left hand sides of (16) and (17) are close. We will achieve this by showing that for all but two values of $i \in [\Delta]$, after fixing i, the right hand sides of (16) and (17) are close, i.e. within $2 \cdot 2^{-t/2}$ of each other. Towards this end, let $r_0 = \sqrt{2}$ and $r_{\Delta+1} = 0$, so that we have a decreasing sequence of radii

$$\sqrt{2} = r_0 > r_1 > \dots r_\Delta > r_{\Delta+1} = 0.$$

Let $0 \le p \le \Delta$ be the unique index such that $r_p \ge \|\mathbf{T}_u - \mathbf{T}_v\| \ge r_{p+1}$. We will show that the right hand sides of (16) and (17) are close except possibly for $i = p, p + 1$.

Case 1: $p+2 \le i \le \Delta$.

In this case, we show that the right hand sides of (16) and (17) are essentially zero. First consider the right hand side of (16). The procedure UG-LABEL with parameter r_i produces clusters with diameter at most r_i and therefore always places u and v into different clusters since $\|\mathbf{T}_u - \mathbf{T}_v\| \ge r_{p+1} > r_i$. Therefore, it outputs labelings to u and v uniformly at random and independent of each other. Moreoever, for any cluster C, the variable ω_C is uniformly distributed in $\{-1,1\}$. Hence, in this case $y_{(u,\mathbf{x})}$ and $y_{(v,\mathbf{y})}$ are independent uniform $\{-1,1\}$ random variables and therefore,

$$\mathrm{E}_{D_{\mathcal{A},r_i}}[y_{(u,\mathbf{x})} y_{(v,\mathbf{y})}] = 0. \tag{18}$$

Now consider the right hand side of (17). We bound it by e^{-2^t}.

$$\begin{aligned} \left|\left\langle \mathbf{G}^i_{(u,\mathbf{x})}, \mathbf{G}^i_{(v,\mathbf{y})} \right\rangle\right| &= \left|\left\langle \frac{1}{\sqrt{N}} \sum_{j\in[N]} \mathbf{x}(j)\cdot \mathbf{T}^{\otimes s_i}_{u,j}, \frac{1}{\sqrt{N}} \sum_{\ell\in[N]} \mathbf{y}(\ell)\cdot \mathbf{T}^{\otimes s_i}_{v,\ell} \right\rangle\right| \\ &\le \frac{1}{N}\sum_{j\in[N]} \left(\sum_{\ell\in[N]} \langle \mathbf{T}_{u,j}, \mathbf{T}_{v,\ell}\rangle^{s_i} \right) \end{aligned}$$

By symmetry, the inner sum is the same for every $j \in [N]$, so we may fix some $j_0 \in [N]$. Since $\{\mathbf{T}_{v,\ell} \mid \ell \in [N]\}$ is an orthomormal set

$$\sum_{\ell\in[N]} \langle \mathbf{T}_{u,j_0}, \mathbf{T}_{v,\ell}\rangle^{s_i} \le \max_{\ell\in[N]} \langle \mathbf{T}_{u,j_0}, \mathbf{T}_{v,\ell}\rangle^{s_i-2} = \left(1 - \frac{1}{2}\min_{\ell\in[N]} \|\mathbf{T}_{u,j_0} - \mathbf{T}_{v,\ell}\|^2\right)^{s_i-2}.$$

The last term can be bounded (using Lemma 6),

$$\left(1 - \frac{1}{8}\|\mathbf{T}_u - \mathbf{T}_v\|^2\right)^{s_i-2} \le \left(1 - \frac{1}{8} r_{p+1}^2\right)^{s_{p+2}-2} = \left(1 - \frac{1}{8} 2^{-2(p+1)t}\right)^{8\cdot 2^{2(p+2)t}-2} \le e^{-2^t}.$$

Case 2: $1 \le i \le p-1$.

This case is more subtle. In this case $\|\mathbf{T}_u - \mathbf{T}_v\| \le r_p \le r_2 < 0.1$. By Lemma 7, there is a unique $k^* = k_{u,v}$ such that the orthonormal tuples $\{\mathbf{T}_{u,j} \mid j \in [N]\}$ and $\{\mathbf{T}_{v,\ell} \mid \ell \in [N]\}$ are close via the matching $j \mapsto k^* \oplus j$. In other words,

$$\forall j \in [N], \quad \|\mathbf{T}_{u,j} - \mathbf{T}_{v,k^*\oplus j}\| \le r_p. \tag{19}$$

We will show that the right hand sides of (16) and (17) are both close to $\frac{1}{N}\sum_{j\in[N]} \mathbf{x}(j)\cdot \mathbf{y}(k^*\oplus j)$.

Towards this end, first consider the right hand side of (16). Let Φ be the event that u and v are not separated into two clusters in the procedure UG-LABEL$(\{u,v\}, r_i)$. From Theorem 9 we have,

$$\Pr[\neg\Phi] \le \frac{100\cdot \|\mathbf{T}_u - \mathbf{T}_v\|\cdot t}{r_i} \le \frac{100\cdot r_p \cdot t}{r_{p-1}} = 100\cdot 2^{-t}\cdot t \le 2^{-t/2}. \tag{20}$$

In the event Φ, both u and v lie in the same cluster C. The procedure UG-LABEL picks $j \in [N]$ uniformly at random, assigns label $\sigma(u) = j$ and label $\sigma(v) = k^* \oplus j$. In the construction of $D_{\mathcal{A},r_i}$ (see section C.1), the vertex $(u,\mathbf{x})$ gets assigned $\mathbf{x}(\sigma(u))\cdot\omega_C = \mathbf{x}(j)\cdot\omega_C$ and the vertex $(v,\mathbf{y})$ gets assigned $\mathbf{y}(\sigma(v))\cdot\omega_C = \mathbf{y}(k^*\oplus j)\cdot\omega_C$. Therefore,

$$\mathrm{E}_{D_{\mathcal{A},r_i}}[y_{(u,\mathbf{x})}y_{(v,\mathbf{y})} \mid \Phi] = \frac{1}{N}\sum_{j\in[N]} \mathbf{x}(j)\mathbf{y}(k^*\oplus j) \tag{21}$$

And using (20) we obtain,

$$\left|\mathrm{E}_{D_{\mathcal{A},r_i}}[y_{(u,\mathbf{x})}y_{(v,\mathbf{y})}] - \frac{1}{N}\sum_{j\in[N]} \mathbf{x}(j)\mathbf{y}(k^*\oplus j)\right| \;\le\; 2^{-t/2}. \tag{22}$$

Now consider the right hand side of (17).

$$\begin{aligned}\left\langle \mathbf{G}^i_{(u,\mathbf{x})}, \mathbf{G}^i_{(v,\mathbf{y})}\right\rangle &= \left\langle \frac{1}{\sqrt{N}}\sum_{j\in[N]} \mathbf{x}(j)\cdot \mathbf{T}^{\otimes s_i}_{u,j}, \frac{1}{\sqrt{N}}\sum_{\ell\in[N]} \mathbf{y}(\ell)\cdot \mathbf{T}^{\otimes s_i}_{v,\ell}\right\rangle \\ &= \left(\frac{1}{N}\sum_{j\in[N]} \mathbf{x}(j)\mathbf{y}(k^*\oplus j)\left\langle \mathbf{T}_{u,j}, \mathbf{T}_{v,k^*\oplus j}\right\rangle^{s_i}\right) + \frac{1}{N}\sum_{\ell\neq k^*\oplus j} \mathbf{x}(j)\mathbf{y}(\ell)\left\langle \mathbf{T}_{u,j}, \mathbf{T}_{v,\ell}\right\rangle^{s_i}.\end{aligned}$$

We show that the second term is negligible and in the first term, $\langle \mathbf{T}_{u,j}, \mathbf{T}_{v,k^*\oplus j}\rangle^{s_i}$ is essentially equal to 1. This would imply that $\left\langle \mathbf{G}^i_{(u,\mathbf{x})}, \mathbf{G}^i_{(v,\mathbf{y})}\right\rangle$ is very close to $\frac{1}{N}\sum_{j\in[N]}\mathbf{x}(j)\mathbf{y}(k^*\oplus j)$ as desired. Indeed by equation (19),

$$\begin{aligned}\langle \mathbf{T}_{u,j}, \mathbf{T}_{v,k^*\oplus j}\rangle^{s_i} &= \left(1 - \frac{1}{2}\,\|\mathbf{T}_{u,j} - \mathbf{T}_{v,k^*\oplus j}\|^2\right)^{s_i} \ge \left(1 - \frac{1}{2}\,r_p^2\right)^{s_{p-1}} \\ &= \left(1 - \frac{1}{2}\,2^{-2pt}\right)^{8\cdot 2^{2(p-1)t}} \quad \ge 1 - 4\cdot 2^{-2t}.\end{aligned}$$

On the other hand, if $\ell \neq k^*\oplus j$, then

$$\|\mathbf{T}_{u,j} - \mathbf{T}_{v,\ell}\| \ge |\mathbf{T}_{v,k^*\oplus j} - \mathbf{T}_{v,\ell}\| - \|\mathbf{T}_{v,k^*\oplus j} - \mathbf{T}_{u,j}\| \ge \sqrt{2} - r_p \ge 1,$$

and hence by non-negativity (condtition (2)) $0 \le \langle \mathbf{T}_{u,j}, \mathbf{T}_{v,\ell}\rangle \le \frac{1}{2}$. This implies that

$$\left|\frac{1}{N}\sum_{\ell\neq k^*\oplus j} \mathbf{x}(j)\mathbf{y}(\ell)\left\langle \mathbf{T}_{u,j}, \mathbf{T}_{v,\ell}\right\rangle^{s_i}\right| \;\le\; \frac{1}{N}\sum_{j\in[N]}\left(\sum_{\ell\neq k^*\oplus j}\left\langle \mathbf{T}_{u,j}, \mathbf{T}_{v,\ell}\right\rangle^{s_i}\right),$$

and for every $j\in[N]$, since $\{\mathbf{T}_{v,\ell} \mid \ell\neq k^*\oplus j\}$ is an orthonormal set,

$$\sum_{\ell\neq k^*\oplus j}\left\langle \mathbf{T}_{u,j}, \mathbf{T}_{v,\ell}\right\rangle^{s_i} \le \max_{\ell\neq k^*\oplus j}\left\langle \mathbf{T}_{u,j}, \mathbf{T}_{v,\ell}\right\rangle^{s_i-2} \le \left(\frac{1}{2}\right)^{s_1-2} \le 2^{-2^t}.$$

Combining everything, we see that the right hand sides of (16) and (17) are within $2\cdot 2^{-t/2}$ of each other.

Completing the proof of Theorem 11

Now we can complete the proof of Theorem 11. We have shown that the right hand sides of (16) and (17) are within $2 \cdot 2^{-t/2}$ of each other for all $i \in \{p+2, \ldots, \Delta\} \cup \{1, \ldots, p-1\}$, i.e. for every $i \in [\Delta]$ except possibly $i = p, p+1$. Clearly, the expressions on (16) and (17) are within $2 \cdot 2^{-t/2} + \frac{2}{\Delta}$ of each other. ∎

We have shown in this section an approximate solution $\mathcal{A}$ to $\texttt{SDP-MC}(t)$ on the instance $\mathcal{I}_\rho$. The solution satisfies all constraints except Constraint (4) of the relaxation, which is only approximately satisfied, up to an error of $2 \cdot 2^{-t/2} + \frac{2}{\Delta}$ that can be made sufficiently small with the choice of t, Δ. In the next section we show how to eliminate this error and obtain the final solution to the relaxation.

D Final solution $\mathcal{F}$ to $\texttt{SDP-MC}(t)$

This section describes the construction of our final feasible solution $\mathcal{F}$ to $\texttt{SDP-MC}(t)$. In the next subsection we first prove a crucial theorem which shows that given an approximate solution of a certain kind to the relaxation $\texttt{SDP-MC}(t)$, it is possible to derive from it a feasible solution to the relaxation with only a negligible loss in the objective value.

D.1 Deriving a feasible solution from an approximate solution

The following is the generic theorem we shall require for our construction.

Theorem 12. *Let $t \in \mathbb{Z}^+$ be any (large enough) parameter and let $\mathcal{I}(V^{\mathcal{I}}, E^{\mathcal{I}})$ be an instance of* MAXIMUM CUT. *Suppose there is a (possibly infeasible) solution $\mathcal{A}$ to the relaxation $\texttt{SDP-MC}(t)$ where $\mathcal{A}$ consists of a collection of distributions $\{D_{\mathcal{A}}(S)\}_{S \subseteq V^{\mathcal{I}}, |S| \le t}$ on $\{-1,1\}$-assignments to sets of vertices S with size at most t, and a vector solution $G_{\mathcal{A}}$ consisting of unit vectors $\{\mathbf{G}_a\}_{a \in V^{\mathcal{I}}}$. Suppose that for $T \subseteq S \subseteq V^{\mathcal{I}}$, $|S| \le t$, the distributions $D_{\mathcal{A}}(S)$ and $D_{\mathcal{A}}(T)$ are consistent. Further, for any two vertices $a, b \in V^{\mathcal{I}}$, if y_a and y_b are the marginals of the $\{-1,1\}$-assignments to $\{a,b\}$ given by the distribution $D_{\mathcal{A}}(\{a,b\})$, then*

$$\left| \mathrm{E}_{D_{\mathcal{A}}}[y_a y_b] - \langle \mathbf{G}_a, \mathbf{G}_b \rangle \right| \le \frac{1}{t^2}. \tag{23}$$

Then there exists a feasible solution $\mathcal{F}$ to the relaxation $\texttt{SDP-MC}(t)$, consisting of a collection of distributions $\{D_{\mathcal{F}}(S)\}_{S \subseteq V^{\mathcal{I}}, |S| \le t}$ and a vector solution $H_{\mathcal{F}}$ of unit vectors $\{\mathbf{H}_a\}_{a \in V^{\mathcal{I}}}$ such that for any two vertices $a, b \in V^{\mathcal{I}}$,

$$\langle \mathbf{H}_a, \mathbf{H}_b \rangle = \left(1 - O\left(\frac{1}{t}\right)\right) \langle \mathbf{G}_a, \mathbf{G}_b \rangle \tag{24}$$

Proof: We start by defining a collection of distributions $D_{\mathcal{F}}(S)$.

Construction of $D_{\mathcal{F}}(S)$: For every pair of distinct vertices $a, b \in V^{\mathcal{I}}$, we construct a "correcting" distribution $\Gamma_{\{a,b\}}$ over $\{-1,1\}$-assignments to the set $\{a,b\}$. We will explicitly define these distributions later. We note for now that the marginals of $\Gamma_{a,b}$ on either co-ordinate is uniform.

Let $S \subseteq V^{\mathcal{I}}$ be such that $|S| \le t$. The dsitribution $D_{\mathcal{F}}(S)$ on $\{-1,1\}$-assignments to the vertices of S is given by the following randomized procedure:

1. Let $W := S \cup \{z_1, \ldots, z_{t-|S|}\}$ where $z_1, \ldots, z_{t-|S|}$ are dummy vertices. Thus $|W| = t$.
2. From the set W, select uniformly at random a pair of distinct vertices $I = \{w_1, w_2\}$.
3. Using the distribution $D_{\mathcal{A}}(S \setminus I)$, sample a $\{-1,1\}$-assignment γ to vertices of $S \setminus I$.

4. If $I \cap S = \emptyset$, we are done.

5. If $I \cap S = \{a\}$, then the vertex a is assigned a value from $\{-1, 1\}$ uniformly at random.

6. If $I \cap S = \{a, b\}$, then the assignment to set $\{a, b\}$ is sampled from the distribution $\Gamma_{\{a,b\}}$.

A case analysis shows that for $T \subseteq S, |S| \leq t$, the distributions $D_{\mathcal{F}}(S)$ and $D_{\mathcal{F}}(T)$ are consistent. One uses the fact that in Step (3), the assignment γ is sampled from $D_{\mathcal{A}}(S \setminus I)$, and these distributions are mutually consistent. Moreover, the marginals of $\Gamma_{a,b}$ are uniform. We skip a formal proof.

Before we define the corresponding vector solution we analyse the distributions $D_{\mathcal{F}}(S)$ corresponding to sets of size two. This analysis will be useful later in the proof.

Analyzing $D_{\mathcal{F}}(S)$ for $|S| = 2$: Let $S = \{a, b\} \subseteq V^{\mathcal{I}}$. Let y_a and y_b denote the marginals of the distribution $D_{\mathcal{F}}(S)$ (or $D_{\mathcal{A}}(S)$ or Γ_S depending on the context). For $J \subseteq S$, let E_J denote the event that $S \cap I = J$, where I is as chosen in Step (2) of the construction of $D_{\mathcal{F}}(S)$. The following are easy to see:

$$\Pr\left[E_{\emptyset}\right] = 1 - \frac{2t-3}{\binom{t}{2}}. \tag{25}$$

$$\Pr\left[E_{\{a\}}\right] = \Pr\left[E_{\{b\}}\right] = \frac{t-2}{\binom{t}{2}}. \tag{26}$$

$$\Pr\left[E_S\right] = \frac{1}{\binom{t}{2}}. \tag{27}$$

We also have,

$$\begin{aligned} \mathrm{E}_{D_{\mathcal{F}}(S)}\left[y_a y_b\right] \;&=\; \mathrm{E}_{D_{\mathcal{F}}(S)}\left[y_a y_b \mid E_{\emptyset}\right] \cdot \Pr\left[E_{\emptyset}\right] + \mathrm{E}_{D_{\mathcal{F}}(S)}\left[y_a y_b \mid E_{\{a\}}\right] \cdot \Pr\left[E_{\{a\}}\right] \\ &\quad + \mathrm{E}_{D_{\mathcal{F}}(S)}\left[y_a y_b \mid E_{\{b\}}\right] \cdot \Pr\left[E_{\{b\}}\right] + \mathrm{E}_{D_{\mathcal{F}}(S)}\left[y_a y_b \mid E_S\right] \cdot \Pr\left[E_S\right]. \end{aligned}$$

If the event $E_{\{a\}}$ occurs then y_a is chosen uniformly at random from $\{-1, 1\}$ independent of y_b and therefore $\mathrm{E}_{D_{\mathcal{F}}(S)}[y_a y_b \mid E_{\{a\}}] = 0$. Similarly, $\mathrm{E}_{D_{\mathcal{F}}(S)}[y_a y_b \mid E_{\{b\}}] = 0$. Moreover, given event $E_{\emptyset}$, $D_{\mathcal{F}}(S)$ is identical to $D_{\mathcal{A}}(S)$. Similarly, given the event E_S, $D_{\mathcal{F}}(S)$ is identical to Γ_S. Therefore,

$$\mathrm{E}_{D_{\mathcal{F}}(S)}[y_a y_b] = \left(1 - \frac{2t-3}{\binom{t}{2}}\right) \mathrm{E}_{D_{\mathcal{A}}(S)}[y_a y_b] + \frac{1}{\binom{t}{2}} \mathrm{E}_{\Gamma_S}[y_a y_b]. \tag{28}$$

Construction of vector solution $H_{\mathcal{F}}$: We now construct the final vector solution $H_{\mathcal{F}}$ as follows.

1. Let $\zeta := 1 - \left(1 - \frac{2(2t-3)}{t(t-1)}\right)$.

2. Construct a pairwise orthonormal set of vectors $\{\mathbf{h}_a\}_{a \in V^{\mathcal{I}}}$ such that for every vertex $a \in V^{\mathcal{I}}$, the vector $\mathbf{h}_a$ is orthogonal to the set of vectors $\{\mathbf{G}_b\}_{b \in V^{\mathcal{I}}}$ comprising the solution $G_{\mathcal{A}}$.

3. In the vector solution $H_{\mathcal{F}}$, for any vertex $a \in V^{\mathcal{I}}$, define the unit vector

$$\mathbf{H}_a := \left(\sqrt{1-\zeta}\right) \mathbf{G}_a + \left(\sqrt{\zeta}\right) \mathbf{h}_a. \tag{29}$$

For any two vertices $a, b \in V^{\mathcal{I}}$, we have that $\mathbf{h}_a \perp \mathbf{h}_b$ and $\mathbf{h}_a$ and $\mathbf{h}_b$ are each orthogonal to both $\mathbf{G}_a$ and $\mathbf{G}_b$. Therefore,

$$\begin{aligned}
\langle \mathbf{H}_a, \mathbf{H}_b \rangle &= \left\langle \left(\sqrt{1-\zeta}\right) \mathbf{G}_a + \left(\sqrt{\zeta}\right) \mathbf{h}_a, \left(\sqrt{1-\zeta}\right) \mathbf{G}_b + \left(\sqrt{\zeta}\right) \mathbf{h}_b \right\rangle \\
&= (1-\zeta) \langle \mathbf{G}_a, \mathbf{G}_b \rangle \\
&= \left(1 - O\left(\frac{1}{t}\right)\right) \langle \mathbf{G}_a, \mathbf{G}_b \rangle
\end{aligned}$$

which satisfies the desired condition of equation (24).

Finally, we show that there is a way to define the distributions $\Gamma_{\{a,b\}}$ so that the solution $\mathcal{F}$ satisfies the Constraint (4) of the relaxation $\mathtt{SDP\text{-}MC}(t)$.

Lemma 13. *For every two distinct vertices $a, b \in V^{\mathcal{I}}$, there is a distribution Γ_S (where $S = \{a, b\}$) such that,*

$$\langle \mathbf{H}_a, \mathbf{H}_b \rangle = \mathrm{E}_{D_{\mathcal{F}}(S)}[y_a y_b]. \tag{30}$$

Proof: From the construction of $H_{\mathcal{F}}$ we have,

$$\begin{aligned}
\langle \mathbf{H}_a, \mathbf{H}_b \rangle &= \left\langle \left(\sqrt{1-\zeta}\right) \mathbf{G}_a + \left(\sqrt{\zeta}\right) \mathbf{h}_a, \left(\sqrt{1-\zeta}\right) \mathbf{G}_b + \left(\sqrt{\zeta}\right) \mathbf{h}_b \right\rangle \\
&= (1-\zeta) \langle \mathbf{G}_a, \mathbf{G}_b \rangle . \qquad (31)
\end{aligned}$$

Equation (28) and substituting the value of ζ in it gives us,

$$\mathrm{E}_{D_{\mathcal{F}}(S)}[y_a y_b] = (1-\zeta)\mathrm{E}_{D_{\mathcal{A}}(S)}[y_a y_b] + \frac{1}{\binom{t}{2}} \mathrm{E}_{\Gamma_S}[y_a y_b].$$

Since we desire that equation (30) holds, it suffices to set

$$(1-\zeta) \langle \mathbf{G}_a, \mathbf{G}_b \rangle = (1-\zeta)\mathrm{E}_{D_{\mathcal{A}}(S)}[y_a y_b] + \frac{1}{\binom{t}{2}} \mathrm{E}_{\Gamma_S}[y_a y_b], \quad \text{i.e.}$$

$$\mathrm{E}_{\Gamma_S}[y_a y_b] = (1-\zeta) \binom{t}{2} \left(\langle \mathbf{G}_a, \mathbf{G}_b \rangle - \mathrm{E}_{D_{\mathcal{A}}(S)}[y_a y_b] \right).$$

Due to the bound (23), the right hand side above is in $[-1, 1]$. Therefore we can define Γ_S appropriately, with the additional property that its marginal on either co-ordinates is uniform. This completes the proof of Lemma 13 as well as Theorem 12. ∎

∎

Applying the above Theorem to the (possibly infeasible) solution $\mathcal{A}$ constructed in section C, we obtain a feasible solution $\mathcal{F}$ to the relaxation $\mathtt{SDP\text{-}MC}(t)$ for the instance $\mathcal{I}_\rho$ of MAXIMUM CUT. The solution $\mathcal{F}$ consists of a collection of distributions $\{D_{\mathcal{F}}(S)\}_{S \subseteq V^*, |S| \leq t}$ and a vector solution $H_{\mathcal{F}}$ with unit vectors $\{\mathbf{H}_{(u,\mathbf{x})}\}_{(u,\mathbf{x}) \in V^*}$. The theorem guarantees that for any two vertices $(u, \mathbf{x}), (v, \mathbf{y}) \in V^*$,

$$\left\langle \mathbf{H}_{(u,\mathbf{x})}, \mathbf{H}_{(v,\mathbf{y})} \right\rangle = \left(1 - O\left(\frac{1}{t}\right)\right) \left\langle \mathbf{G}_{(u,\mathbf{x})}, \mathbf{G}_{(v,\mathbf{y})} \right\rangle . \tag{32}$$

D.2 Computation of the Integrality Gap

We start by setting the parameters $\eta = (\log N)^{-0.99}$ and $t = (\log\log N)^{\frac{1}{6}}$. The optimum of the unique games instance is at most $\frac{1}{N^\eta} = 2^{-(\log N)^{0.01}}$. The size of the unique games instance is $|V| = 2^N/N$ whereas the size of the max-cut instance is $n := |V| \cdot 2^N = 2^{2N}/N$. The value of $\rho \in (-1, 0)$ is chosen so that $\alpha_{GW}^{-1} := \max_{\rho\in[-1,1]} \frac{\pi(1-\rho)}{2\cdot\arccos(\rho)}$ is attained.

We shall first show the following. Fix a vertex $v \in V$. Let $e(v, w)$ and $e'(v, w') \in E(v)$ any two edges incident on v. Let $\mathbf{x} \in_{1/2} \{-1, 1\}^N$, and $\boldsymbol{\mu} \in_{\frac{1-\rho}{2}} \{-1, 1\}^N$. Then, with probability at least $1 - \eta$,

$$\left\langle \mathbf{H}_{(w,\mathbf{x}\circ\pi_e^{-1})}, \mathbf{H}_{(w',\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1})} \right\rangle = \rho \pm O\left(\frac{1}{t}\right) \tag{33}$$

Using Chernoff bound we can make the following observation.

Observation 14. *The following event takes place with probability at least* $1 - \eta$*,*

$$\mathrm{E}_{i\in_R[N]}[\boldsymbol{\mu}(i)] \in [\rho - \eta, \rho + \eta] \,.$$

Using the **High SDP Value** property (condition (4)) of the UNIQUE GAMES SDP solution, we have that for any $\ell \in [N]$,

$$\langle \mathbf{T}_{v,\ell}, \mathbf{T}_{w,k_e\oplus\ell}\rangle \geq 1 - 4\eta, \qquad \langle \mathbf{T}_{v,\ell}, \mathbf{T}_{w',k_{e'}\oplus\ell}\rangle \geq 1 - 4\eta.$$

From the above and using the triangle inequality, we have,

$$\|\mathbf{T}_{w,k_e\oplus\ell} - \mathbf{T}_{w',k_{e'}\oplus\ell}\| \leq 4\sqrt{2\eta}$$

Using the **Symmetry** property (condition (3)) the above can be restated as follows. For all $\ell \in [N]$,

$$\|\mathbf{T}_{w,\ell} - \mathbf{T}_{w',(k_e\oplus k_{e'})\oplus\ell}\| \leq 4\sqrt{2\eta}. \tag{34}$$

Combining the above with Lemma 6 we obtain,

$$\|\mathbf{T}_w - \mathbf{T}_{w'}\| \leq 8\sqrt{2\eta}. \tag{35}$$

Combining equations (34) and (35) with Lemma 7 we obtain $k_{w,w'} = k_e \oplus k_{e'}$.

From our choice of the parameters $\sqrt{\eta} \ll r_\Delta := 2^{-t^5}$, and thus,

$$r_\Delta > \|\mathbf{T}_w - \mathbf{T}_{w'}\| \;\geq\; r_{\Delta+1} = 0, \tag{36}$$

where r_i ($i \in \{0, \ldots, \Delta + 1\}$) are as defined in Section C. Moreover, combining equation (36) with Case 2 of the proof of Theorem 11, we obtain that for all indices i such that $1 \leq i \leq \Delta - 1$,

$$\left\langle \mathbf{G}^i_{w,\mathbf{x}\circ\pi_e^{-1}}, \mathbf{G}^i_{w',\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1}} \right\rangle = \frac{1}{N}\sum_{j\in[N]} \left(\mathbf{x}\circ\pi_e^{-1}(j)\right)\left(\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1}(k_{w,w'}\oplus j)\right) \;\pm 2^{-t}$$

Since $k_{w,w'} = k_e \oplus k_{e'}$, we can rewrite the above as,

$$\left\langle \mathbf{G}^i_{w,\mathbf{x}\circ\pi_e^{-1}}, \mathbf{G}^i_{w',\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1}} \right\rangle = \frac{1}{N}\sum_{j\in[N]} \left(\mathbf{x}\circ\pi_e^{-1}(k_e\oplus j)\right)\left(\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1}(k_{e'}\oplus j)\right) \pm\; 2^{-t}$$

From the **High SDP Value** property we have that $k_e(j) = \pi_e(j)$ and $k_{e'}(j) = \pi_{e'}(j)$. Substituting in the above equation we get that for all indices $1 \le i \le \Delta - 1$,

$$\begin{aligned}\left\langle \mathbf{G}^i_{w,\mathbf{x}\circ\pi_e^{-1}}, \mathbf{G}^i_{w',\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1}} \right\rangle &= \frac{1}{N}\sum_{j\in[N]} \left(\mathbf{x}\circ\pi_e^{-1}(\pi_e(j))\right)\left(\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1}(\pi_{e'}(j))\right) \pm 2^{-t} \\ &= \frac{1}{N}\sum_{j\in[N]} \mathbf{x}(j)\mathbf{x}\boldsymbol{\mu}(j) \pm 2^{-t} \\ &= \frac{1}{N}\sum_{j\in[N]} \boldsymbol{\mu}(j) \pm 2^{-t}\end{aligned}$$

Since the above holds for all $i \in \{1, \ldots, \Delta - 1\}$, by equation (17) we have,

$$\left\langle \mathbf{G}_{w,\mathbf{x}\circ\pi_e^{-1}}, \mathbf{G}_{w',\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1}} \right\rangle = \frac{1}{N}\sum_{j\in[N]} \boldsymbol{\mu}(j) \pm O\left(\frac{1}{\Delta} + 2^{-t}\right) \tag{37}$$

Combining the above with Observation 14, we obtain that with probability at least $1 - \eta$,

$$\left\langle \mathbf{G}_{w,\mathbf{x}\circ\pi_e^{-1}}, \mathbf{G}_{w',\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1}} \right\rangle = \rho \pm O\left(\frac{1}{\Delta}\right), \tag{38}$$

and from equation (32),

$$\left\langle \mathbf{H}_{w,\mathbf{x}\circ\pi_e^{-1}}, \mathbf{H}_{w',\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1}} \right\rangle = \rho \pm O\left(\frac{1}{t}\right)$$

which proves the condition given by equation (33). Since equation (33) holds for all $v, e = (v, w)$ and $e' = (v, w')$ this implies that the normalized objective value of $\texttt{SDP-MC}(t)$ on $\mathcal{I}_\rho$ is,

$$\mathsf{FRAC}(\mathcal{I}_\rho) \;\ge\; \mathrm{E}_{\mathbf{x},\boldsymbol{\mu}}\left[\frac{1 - \left\langle \mathbf{H}_{(w,\mathbf{x}\circ\pi_e^{-1})}, \mathbf{H}_{(w',\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^{-1})} \right\rangle}{2}\right] \;\ge\; \frac{1-\rho}{2} - O\left(\frac{1}{t}\right) - O(\eta).$$

Applying Theorem 10, by choosing the optimum of the unique games instance $2^{-(\log N)^{0.01}}$ low enough, we see that the (normalized) value of the best cut in $\mathcal{I}_\rho$ is,

$$\mathsf{OPT}(\mathcal{I}_\rho) \le \frac{1}{\pi}\arccos\rho + \frac{\varepsilon}{4}, \tag{39}$$

where $\varepsilon > 0$ is the constant in Theorem 3. Therefore, the Integrality Gap of $\texttt{SDP-MC}(t)$ is,

$$\frac{\mathsf{FRAC}(\mathcal{I}_\rho)}{\mathsf{OPT}(\mathcal{I}_\rho)} \;\ge\; \frac{\pi(1-\rho)}{2\cdot\arccos\rho} - O\left(\frac{1}{t}\right) - O(\eta) - \frac{\varepsilon}{2} \;\ge\; \alpha_{GW}^{-1} - \varepsilon.$$

This proves Theorem 3 (note that $1/t$ and η are sub-constant).

E Integrality gap for SPARSEST CUT

We give a brief overview of the construction of the integrality gap example for the SPARSEST CUT relaxation $\texttt{SDP-SC}(t)$. As in the construction of Khot and Vishnoi [KV05], we actually construct an integrality gap example for a similar relaxation for the BALANCED SEPARATOR problem. For this the only change we need to make to the construction for MAXIMUM CUT is the setting of the parameter ρ. We choose set ρ to be $1-\delta$, where $\delta = 1/t$. It was shown in [KV05] that the instance of BALANCED SEPARATOR thus obtained has an optimum of $\Omega(\delta^c)$ (any exponent $c > \frac{1}{2}$ works, say $c = \frac{7}{13}$), provided that the soundness of the unique games instance is at most $2^{-O(1/\delta^2)}$. On the other hand, the SDP value is at most $O(\delta + 1/t) = O(\delta)$. This gives us an integrality gap of $\Omega((1/\delta)^{1-c})$ which, on substituting the value of the chosen parameters, is $\Omega((\log \log \log n)^{\frac{1}{13}})$.