The Unique Games Conjecture, Integrality Gap for Cut Problems and Embeddability of Negative Type Metrics into ℓ_1

Subhash A. Khot
College of Computing,
Georgia Institute of Technology,
801 Atlantic Drive, Atlanta GA 30332, USA.
Email: `khot@cc.gatech.edu`

Nisheeth K. Vishnoi
IBM India Research Lab,
Block-1 IIT Delhi, Hauz Khas,
New Delhi 110016, India.
Email: `nvishnoi@in.ibm.com`

Abstract

In this paper we disprove the following conjecture due to Goemans [16] and Linial [24] (also see [5, 26]): "Every negative type metric embeds into ℓ_1 with constant distortion." We show that for every $\delta > 0$, and for large enough n, there is an n-point negative type metric which requires distortion at-least $(\log\log n)^{1/6-\delta}$ to embed into ℓ_1.

Surprisingly, our construction is inspired by the Unique Games Conjecture (UGC) of Khot [19], establishing a previously unsuspected connection between PCPs and the theory of metric embeddings. We first prove that the UGC implies super-constant hardness results for (non-uniform) Sparsest Cut and Minimum Uncut problems. It is already known that the UGC also implies an optimal hardness result for Maximum Cut [20]. Though these hardness results rely on the UGC, we demonstrate, nevertheless, that the corresponding PCP reductions can be used to construct "integrality gap instances" for the respective problems. Towards this, we first construct an integrality gap instance for a natural SDP relaxation of Unique Games. Then, we "simulate" the PCP reduction, and "translate" the integrality gap instance of Unique Games to integrality gap instances for the respective cut problems! This enables us to prove a $(\log\log n)^{1/6-\delta}$ integrality gap for (non-uniform) Sparsest Cut and Minimum Uncut, and an optimal integrality gap for Maximum Cut. All our SDP solutions satisfy the so-called "triangle inequality" constraints. This also shows, for the first time, that the triangle inequality constraints do not add any power to the Goemans-Williamson's SDP relaxation of Maximum Cut.

The integrality gap for Sparsest Cut immediately implies a lower bound for embedding negative type metrics into ℓ_1. It also disproves the non-uniform version of Arora, Rao and Vazirani's Conjecture [5], asserting that the integrality gap of the Sparsest Cut SDP, with triangle inequality constraints, is bounded from above by a constant.

Contents

1 Introduction

In recent years, the theory of metric embeddings has played an increasing role in algorithm design. Best approximation algorithms for several NP-hard problems rely on techniques (and theorems) used to embed one metric space into another with *low distortion*.

Bourgain [7] showed that every n-point metric embeds into ℓ_1 (in fact into ℓ_2) with distortion $O(\log n)$. Independently, Aumann and Rabani [6] and Linial, London and Rabinovich [25] gave a striking application of Bourgain's Theorem: An $O(\log n)$ approximation algorithm for Sparsest Cut. The approximation ratio is exactly the distortion incurred in Bourgain's Theorem. This gave an alternate approach to the seminal work of Leighton and Rao [23], who obtained an $O(\log n)$ approximation algorithm for Sparsest Cut via a LP-relaxation based on muticommodity flows. It is well-known that an $f(n)$ factor algorithm for Sparsest Cut can be used iteratively to design an $O(f(n))$ factor algorithm for Balanced Separator: Given a graph that has a $(\frac{1}{2}, \frac{1}{2})$-partition cutting an α fraction of the edges, the algorithm produces a $(\frac{1}{3}, \frac{2}{3})$-partition that cuts at-most $O(f(n)\alpha)$ fraction of the edges. Such partitioning algorithms are very useful as sub-routines in designing graph theoretic algorithms via the divide-and-conquer paradigm.

The results of [6, 25] are based on the *metric* LP *relaxation* of Sparsest Cut. Given an instance $G(V, E)$ of Sparsest Cut, let d_G be the n-point metric obtained as a solution to this LP. The metric d_G is then embedded into ℓ_1 via Bourgain's Theorem. Since ℓ_1 metrics are non-negative linear combinations of cut metrics, an embedding into ℓ_1 essentially gives the desired sparse cut (up to an $O(\log n)$ approximation factor). Subsequent to this result, it was realized that one could write an SDP relaxation of Sparsest Cut, and enforce an additional condition, that the metric d_G belong to a special class of metrics, called the *negative type metrics* (denoted by ℓ_2^2). Clearly, if ℓ_2^2 embeds into ℓ_1 with distortion $g(n)$, then one would get a $g(n)$ approximation to Sparsest Cut.[1]

The results of [6, 25] led to the conjecture that ℓ_2^2 embeds into ℓ_1 with distortion C_{neg}, for some absolute constant C_{neg}. This conjecture has been attributed to Goemans [16] and Linial [24], see [5, 26]. This conjecture, which we will henceforth refer to as the $(\ell_2^2, \ell_1, O(1))$-Conjecture, if true, would have had tremendous algorithmic applications (apart from being an important mathematical result). Several problems, specifically cut problems (see [11]), can be formulated as optimization problems over the class of ℓ_1 metrics, and optimization over ℓ_1 is an NP-hard problem in general. However, one can optimize over ℓ_2^2 metrics in polynomial time via SDPs (and $\ell_1 \subseteq \ell_2^2$). Hence, if ℓ_2^2 was embeddable into ℓ_1 with constant distortion, one would get a computationally efficient approximation to ℓ_1 metrics.

However, no better embedding of ℓ_2^2 into ℓ_1, other than Bourgain's $O(\log n)$ embedding (that works for all metrics), was known until recently. A breakthrough result of Arora, Rao and Vazirani (ARV) [5] gave an $O(\sqrt{\log n})$ approximation to (uniform) Sparsest Cut by showing that the integrality gap of the SDP relaxation is $O(\sqrt{\log n})$ (see also [28] for an alternate perspective on ARV). Subsequently, ARV techniques were used by Chawla, Gupta and Räcke [9] to give an $O(\log^{3/4} n)$ distortion embedding of ℓ_2^2 metrics into ℓ_2, and hence, into ℓ_1. This result was further improved to $O(\sqrt{\log n}\log\log n)$ by Arora, Lee, and Naor [3]. The latter paper implies, in particular, that every n-point ℓ_1 metric embeds into ℓ_2 with distortion $O(\sqrt{\log n}\log\log n)$, almost matching decades old $\Omega(\sqrt{\log n})$ lower bound due to Enflo [12]. Techniques from ARV have also been applied, to obtain

[1] Algorithms based on metric embeddings (typically) work for the *non-uniform* version of Sparsest Cut, which is more general. The Leighton-Rao algorithm worked only for the uniform version.

$O(\sqrt{\log n})$ approximation to MINIMUM UNCUT and related problems [1], to VERTEX SEPARATOR [13], and to obtain a $2 - O(\frac{1}{\sqrt{\log n}})$ approximation to VERTEX COVER [18]. It was conjectured in the ARV paper, that the integrality gap of the SDP relaxation of SPARSEST CUT is bounded from above by an absolute constant (they make this conjecture only for the uniform version, and the $(\ell_2^2, \ell_1, O(1))$-Conjecture implies it also for the non-uniform version). Thus, if the $(\ell_2^2, \ell_1, O(1))$-Conjecture and/or the ARV-Conjecture were true, one would potentially get a constant factor approximation to a host of problems, and perhaps, an algorithm for VERTEX COVER with an approximation factor better than 2! Clearly, it is an important open problem to prove or disprove the $(\ell_2^2, \ell_1, O(1))$-Conjecture and/or the ARV-Conjecture. The main result in this paper is a disproval of the $(\ell_2^2, \ell_1, O(1))$-Conjecture and a disproval of the non-uniform version of the ARV-Conjecture, see Conjecture 6.16.[2] The disprovals follow from the construction of a super-constant integrality gap for the non-uniform version of BALANCED SEPARATOR (which implies the same gap for the non-uniform version of SPARSEST CUT). We also obtain integrality gap instances for MAXIMUM CUT and MINIMUM UNCUT. In the following sections, we describe our results in detail and present an overview of our ℓ_2^2 versus ℓ_1 lower bound.

2 Our Results

2.1 The Disproval of $(\ell_2^2, \ell_1, O(1))$-Conjecture

We prove the following theorem which follows from the integrality gap construction for non-uniform BALANCED SEPARATOR. See Section 6 for definitions and basic facts.

Theorem 2.1 *For every $\delta > 0$ and for all sufficiently large n, there is an n-point ℓ_2^2 metric which cannot be embedded into ℓ_1 with distortion less than $(\log \log n)^{1/6-\delta}$.*

Remark 2.2 *One of the crucial ingredients for obtaining the lower bound of $(\log \log n)^{1/6-\delta}$ in Theorems 2.1 and 2.3 is Bourgain's Junta Theorem [8]. A recent improvement of this theorem due to Mossel* et al. *[27] improves both of our lower bounds to $(\log \log n)^{1/4-\delta}$.*

2.2 Integrality Gap Instances for Cut Problems

SPARSEST CUT and BALANCED SEPARATOR (non-uniform versions), as well as MAXIMUM CUT and MINIMUM UNCUT are defined in Section 6.4. Natural SDP relaxations for these problems are also described there. All the SDPs include the so-called *triangle inequality* constraints: For every triple of vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in the SDP solution, $\|\mathbf{u} - \mathbf{v}\|^2 + \|\mathbf{v} - \mathbf{w}\|^2 \geq \|\mathbf{u} - \mathbf{w}\|^2$. Note that these constraints are always satisfied by the *integral* solutions, i.e., $+1, -1$ valued solutions. We prove the following two theorems:

Theorem 2.3 SPARSEST CUT, BALANCED SEPARATOR *(non-uniform versions of both) and* MINIMUM UNCUT *have an integrality gap of at-least $(\log \log n)^{1/6-\delta}$, where $\delta > 0$ is arbitrary. The integrality gap holds for standard* SDP*s with triangle inequality constraints.*

Theorem 2.4 *Let α_{GW} (≈ 0.878) be the approximation ratio obtained by Goemans-Williamson's algorithm for* MAXIMUM CUT *[17]. For every $\delta > 0$, the Goemans-Williamson's* SDP *has an integrality gap of at-least $\alpha_{\mathsf{GW}} + \delta$, even after including the triangle inequality constraints.*

[2] We believe that even the uniform version of the ARV-Conjecture is false.

This theorem relies on a Fourier analytic result called *Majority is Stablest Theorem* due to Mossel *et al.* [27].

We note that without the triangle inequality constraints, Feige and Schechtman [15] already showed an $\alpha_{\mathsf{GW}} + \delta$ integrality gap. One more advantage of our result is that it is an explicit construction, where as Feige and Schechtman's construction is randomized (they need to pick random points on the unit sphere). Our result shows that adding the triangle inequality constraints does not add any power to the Goemans-Williamson's SDP. This nicely complements the result of Khot *et al.* [20], where it is shown that, assuming the Unique Games Conjecture (UGC), it is NP-hard to approximate Maximum Cut within a factor better than $\alpha_{\mathsf{GW}} + \delta$.

2.3 Hardness Results for Sparsest Cut and Balanced Separator Assuming the UGC

Our starting point is the hardness of approximation results for cut problems assuming the UGC (see Section 8 for the statement of the conjecture). We prove the following result:

Theorem 2.5 *Assuming the* UGC*,* Sparsest Cut *and* Balanced Separator *(non-uniform versions) are* NP*-hard to approximate within any constant factor.*

This particular result was also proved[3] by Chawla *et al.* [10]. Similar result for Minimum Uncut is implicit in [19], where the author formulated the UGC and proved the hardness of approximating Min-2SAT-Deletion. As mentioned before, Khot *et al.* [20] proved that the UGC implies $\alpha_{\mathsf{GW}}+\delta$ hardness result for Maximum Cut. As an aside, we note that the UGC also implies optimal $2-\delta$ hardness result for Vertex Cover, as shown in [22].

Therefore, assuming the UGC, all of the above problems are NP-hard to approximate within respective factors, and hence, the corresponding integrality gap examples must exist (unless P=NP). In particular, if the UGC is true, then the $(\ell_2^2, \ell_1, O(1))$-Conjecture is false. This is a rather peculiar situation, because the UGC is still unproven, and may very well be false. Nevertheless, we are able to disprove the $(\ell_2^2, \ell_1, O(1))$-Conjecture *unconditionally* (which may be taken as an argument supporting the UGC). Indeed, the UGC plays a crucial role in our disproval. Let us outline the basic approach we take. First, we build an integrality gap instance for a natural SDP relaxation of Unique Games (see Figure 6). Surprisingly, we are then able to *translate* this integrality gap instance into an integrality gap instance of Sparsest Cut, Balanced Separator, Maximum Cut and Minimum Uncut. This translation *mimics* the PCP reduction from the UGC to these problems (note that the same reduction also proves hardness results assuming the UGC)! We believe that this novel approach will have several applications in the future. Already, inspired by our work, Khot and Naor [21] have proved several non-embeddability results (e.g. Edit Distance into ℓ_1), and Arora *et al.* [2] have constructed integrality gap instances for the MaxQP problem.

2.4 Integrality Gap Instance for the Unique Games SDP Relaxation

As mentioned above, we construct an integrality gap instance for a natural SDP relaxation of Unique Games (see Figure 6). Here, we choose to provide an informal description of this construction (the reader should be able to understand this construction without even looking at the SDP formulation).

[3] We would like to stress that our work was completely independent, and no part of our work was influenced by their paper.

Theorem 2.6 *(Informal statement) Let N be an integer and $\eta > 0$ be a parameter (think of N as large and η as very tiny). There is a graph $G(V, E)$ of size $2^N/N$ with the following properties: Every vertex $u \in V$ is assigned a set of unit vectors $B(u) := \{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ that form an orthonormal basis for the space $\mathbb{R}^N$. Further,*

1. *For every edge $e = (u, v) \in E$, the set of vectors $B(u)$ and $B(v)$ are almost the same upto some small perturbation. To be precise, there is a permutation $\pi_e : [N] \mapsto [N]$, such that $\forall\ 1 \leq i \leq N,\ \langle \mathbf{u}_{\pi_e(i)}, \mathbf{v}_i \rangle \geq 1 - \eta$. In other words, for every edge $(u, v) \in E$, the basis $B(u)$ moves "smoothly/continuously" to the basis $B(v)$.*

2. *For any* labeling *$\lambda : V \mapsto [N]$, i.e., assignment of an integer $\lambda(u) \in [N]$ to every $u \in V$, for at-least $1 - \frac{1}{N^\eta}$ fraction of the edges $e = (u, v) \in E$, we have $\pi_e(\lambda(u)) \neq \lambda(v)$. In other words, no matter how we choose to assign a vector $\mathbf{u}_{\lambda(u)} \in B(u)$ for every vertex $u \in V$, the movement from $\mathbf{u}_{\lambda(u)}$ to $\mathbf{v}_{\lambda(v)}$ is "discontinuous" for almost all edges $(u, v) \in E$.*

3. *All vectors in $\cup_{u \in V} B(u)$ have co-ordinates in the set $\{\frac{1}{\sqrt{N}}, \frac{-1}{\sqrt{N}}\}$, and hence, any three of them satisfy the triangle inequality constraint.*

The construction is rather non-intuitive: One can walk on the graph G by changing the basis $B(u)$ continuously, but as soon as one picks a *representative vector* for each basis, the motion becomes discontinuous almost everywhere! Of course, one can pick these representatives in a continuous fashion for any small enough local sub-graph of G, but there is no way to pick representatives in a global fashion. This construction eventually leads us to a ℓ_2^2 metric which, roughly speaking, is locally ℓ_1-embeddable, but globally, it requires super-constant distortion to embed into ℓ_1 (such local versus global phenomenon has also been observed by Arora *et al.* [4]).

3 Difficulty in Proving ℓ_2^2 vs. ℓ_1 Lower Bound

In this section, we describe the difficulties in constructing ℓ_2^2 metrics that do not embed well into ℓ_1. This might partly explain why one needs an unusual construction as the one in this paper. Our discussion here is informal, without precise statements or claims.

Difficulty in constructing ℓ_2^2 metrics: To the best of our knowledge, no natural families of ℓ_2^2 metrics are known other than the Hamming metric on $\{-1, 1\}^k$. The Hamming metric is an ℓ_1 metric, and hence, not useful for the purposes of obtaining ℓ_1 lower bounds. Certain ℓ_2^2 metrics can be constructed via Fourier analysis, and one can also construct some by solving SDPs explicitly. The former approach has a drawback that metrics obtained via Fourier methods typically embed into ℓ_1 isometrically. The latter approach has limited scope, since one can only hope to solve SDPs of moderate size. Feige and Schechtman [15] show that selecting an appropriate number of points from the unit sphere gives a ℓ_2^2 metric. However, in this case, most pairs of points have distance $\Omega(1)$ and hence, the metric is likely to be ℓ_1-embeddable with low distortion.

Difficulty in proving ℓ_1 lower bounds: To the best of our knowledge, there is no standard technique to prove a lower bound for embedding a metric into ℓ_1. The only interesting (super-constant) lower bound that we know is due to [6, 25], where it is shown that the shortest path metric on a constant degree expander requires $\Omega(\log n)$ distortion to embed into ℓ_1.

General theorems regarding group norms: A *group norm* is a distance function $d(\cdot, \cdot)$ on a group $(G, \circ)$, such that $d(x, y)$ depends only on the group difference $x \circ y^{-1}$. Using Fourier methods, it is possible to construct group norms that are ℓ_2^2 metrics. However, it is known that any group norm on $\mathbb{R}^k$, or on any group of characteristic 2, is isometrically ℓ_1-embeddable (see [11]). It is also known (among the experts in this area) that such a result holds for every abelian group. Therefore, any approach, just via group norms would be unlikely to succeed, as long as the underlying group is abelian. (But, only in the abelian case, the Fourier methods work well.)

The best known lower bounds for the ℓ_2^2 versus ℓ_1 question were due to Vempala ($\frac{10}{9}$ for a metric obtained by a computer search), and Goemans (1.024 for a metric based on the Leech Lattice), see [30]. Thus, it appeared that an entirely new approach was needed to resolve the $(\ell_2^2, \ell_1, O(1))$-Conjecture. In this paper, we present an approach based on tools from complexity theory, namely, the UGC, PCPs, and Fourier analysis of boolean functions. Interestingly, Fourier analysis is used both to construct the ℓ_2^2 metric, as well as, to prove the ℓ_1 lower bound.

4 Overview of Our ℓ_2^2 vs. ℓ_1 Lower Bound

In this section, we present a high level idea of our ℓ_2^2 versus ℓ_1 lower bound (see Theorem 2.1). Given the construction of Theorem 2.6, it is fairly straight-forward to describe the candidate ℓ_2^2 metric: Let $G(V, E)$ be the graph, and $B(u)$ be the orthonormal basis for $\mathbb{R}^N$ for every $u \in V$, as in Theorem 2.6. Fix $s = 4$. For $u \in V$ and $\mathbf{x} = (x_1, \ldots, x_N) \in \{-1, 1\}^N$, define the vector $\mathbf{V}_{u,s,\mathbf{x}}$ as follows:

$$\mathbf{V}_{u,s,\mathbf{x}} := \frac{1}{\sqrt{N}} \sum_{i=1}^{N} x_i \mathbf{u}_i^{\otimes 2s} \tag{1}$$

Note that since $B(u) = \{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ is an orthonormal basis for $\mathbb{R}^N$, every $\mathbf{V}_{u,s,\mathbf{x}}$ is a unit vector. Fix t to be a large odd integer, for instance $2^{240} + 1$, and consider the set of unit vectors $\mathcal{S} = \{\mathbf{V}_{u,s,\mathbf{x}}^{\otimes t} \mid u \in V,\ \mathbf{x} \in \{-1, 1\}^N\}$. Using, essentially, the fact that the vectors in $\cup_{u \in V} B(u)$ are a *good* solution to the SDP relaxation of Unique Games, we are able to show that every triple of vectors in $\mathcal{S}$ satisfy the triangle inequality constraint and, hence, $\mathcal{S}$ defines a ℓ_2^2 metric. One can also directly show that this ℓ_2^2 metric does not embed into ℓ_1 with distortion less than $(\log N)^{1/6-\delta}$.

However, we choose to present our construction in a different and a quite indirect way. The (lengthy) presentation goes through the Unique Games Conjecture, and the PCP reduction from Unique Games integrality gap instance to Balanced Separator. Hopefully, our presentation will bring out the intuition as to why and how we came up with the above set of vectors, which happened to define a ℓ_2^2 metric. At the end, the reader will recognize that the idea of taking all $+/-$ linear combinations of vectors in $B(u)$ (as in Equation (1)) is directly inspired by the PCP reduction. Also, the proof of the ℓ_1 lower bound will be hidden inside the *soundness analysis* of the PCP!

The overall construction can be divided into three steps:

1. A PCP reduction from Unique Games to Balanced Separator.
2. Constructing an integrality gap instance for a natural SDP relaxation of Unique Games.
3. Combining these two to construct an integrality gap instance of Balanced Separator. This also gives a ℓ_2^2 metric that needs $(\log \log n)^{1/6-\delta}$ distortion to embed into ℓ_1.

We present an overview of each of these steps in three separate sections. Before we do that, let us summarize the precise notion of an integrality gap instance of BALANCED SEPARATOR. To keep things simple in this exposition, we will pretend as if our construction works for the uniform version of BALANCED SEPARATOR as well. (Actually, it doesn't. We have to work with the non-uniform version and it complicates things a little.)

4.1 SDP Relaxation of BALANCED SEPARATOR

Given a graph $G'(V', E')$, BALANCED SEPARATOR asks for a $(\frac{1}{2}, \frac{1}{2})$-partition of V' that cuts as few edges as possible. (However, the algorithm is allowed to output a roughly balanced partition, say $(\frac{1}{4}, \frac{3}{4})$-partition.) Following is an SDP relaxation of BALANCED SEPARATOR:

$$\text{Minimize} \quad \frac{1}{|E'|} \sum_{e'=\{i,j\} \in E'} \frac{1}{4} \|\mathbf{v}_i - \mathbf{v}_j\|^2 \tag{2}$$

Subject to

$$\forall\, i \in V' \qquad \|\mathbf{v}_i\|^2 = 1 \tag{3}$$

$$\forall\, i, j, l \in V' \quad \|\mathbf{v}_i - \mathbf{v}_j\|^2 + \|\mathbf{v}_j - \mathbf{v}_l\|^2 \geq \|\mathbf{v}_i - \mathbf{v}_l\|^2 \tag{4}$$

$$\textstyle\sum_{i<j} \|\mathbf{v}_i - \mathbf{v}_j\|^2 \geq |V'|^2 \tag{5}$$

Figure 1: SDP relaxation of the uniform version of BALANCED SEPARATOR

Note that a $\{+1, -1\}$-valued solution represents a true partition, and hence, this is an SDP relaxation. Constraint (4) is the triangle inequality constraint and Constraint (5) stipulates that the partition be balanced. The notion of integrality gap is summarized in the following definition:

Definition 4.1 *An integrality gap instance of* BALANCED SEPARATOR *is a graph* $G'(V', E')$ *and an assignment of unit vectors* $i \mapsto \mathbf{v}_i$ *to its vertices such that:*

- *Every almost balanced partition (say* $(\frac{1}{4}, \frac{3}{4})$*-partition; the choice is arbitrary) of* V' *cuts at-least* α *fraction of edges.*
- *The set of vectors* $\{\mathbf{v}_i |\ i \in V'\}$ *satisfy* (3)-(5)*, and the* SDP *objective value in Equation* (2) *is at-most* γ*.*

The integrality gap is defined to be α/γ*. (Thus, we desire that* $\gamma \ll \alpha$*.)*

The next three sections describe the three steps involved in constructing an integrality gap instance of BALANCED SEPARATOR. Once that is done, it follows from a folk-lore result that the resulting ℓ_2^2 metric (defined by vectors $\{\mathbf{v}_i |\ i \in V'\}$) requires distortion at-least $\Omega(\alpha/\gamma)$ to embed into ℓ_1. This would prove Theorem 2.1 with an appropriate choice of parameters.

4.2 The PCP Reduction from UNIQUE GAMES to BALANCED SEPARATOR

An instance $\mathcal{U}(G(V, E), [N], \{\pi_e\}_{e \in E})$ of UNIQUE GAMES consists of a graph $G(V, E)$ and permutations $\pi_e : [N] \mapsto [N]$ for every edge $e = (u, v) \in E$. The goal is to find a *labeling* $\lambda : V \mapsto [N]$ that *satisfies* as many edges as possible. An edge $e = (u, v)$ is satisfied if $\pi_e(\lambda(u)) = \lambda(v)$. Let $\mathsf{OPT}(\mathcal{U})$ denote the maximum fraction of edges satisfied by any labeling.

UGC *(Informal Statement): It is* NP*-hard to decide whether an instance* $\mathcal{U}$ *of* Unique Games *has* $\mathsf{OPT}(\mathcal{U}) \geq 1 - \eta$ *(*YES *instance) or* $\mathsf{OPT}(\mathcal{U}) \leq \zeta$ *(*NO *instance), where* $\eta, \zeta > 0$ *can be made arbitrarily small by choosing* N *to be a sufficiently large constant.*

It is possible to construct an instance of Balanced Separator $G'_\varepsilon(V', E')$ from an instance $\mathcal{U}(G(V, E), [N], \{\pi_e\}_{e \in E})$ of Unique Games. We describe only the high level idea here. The construction is parameterized by $\varepsilon > 0$. The graph G'_ε has a block of 2^N vertices for every $u \in V$. This block contains one vertex for every point in the boolean hypercube $\{-1, 1\}^N$. Denote the set of these vertices by $V'[u]$. More precisely,

$$V'[u] := \{(u, \mathbf{x}) \mid \mathbf{x} \in \{-1, 1\}^N\}.$$

We let $V' := \cup_{u \in V} V'[u]$. For every edge $e = (u, v) \in E$, the graph G'_ε has edges between the blocks $V'[u]$ and $V'[v]$. These edges are supposed to capture the constraint that the labels of u and v are consistent (i.e. $\pi_e(\lambda(u)) = \lambda(v)$). Roughly speaking, a vertex $(u, \mathbf{x}) \in V'[u]$ is connected to a vertex $(v, \mathbf{y}) \in V'[v]$ if and only if, after identifying the co-ordinates in $[N]$ via the permutation π_e, the Hamming distance between the bit-strings $\mathbf{x}$ and $\mathbf{y}$ is at-most εN.

This reduction has the following two properties:

Theorem 4.2 *(*PCP *reduction: Informal statement)*

1. *(Completeness/*YES *case): If* $\mathsf{OPT}(\mathcal{U}) \geq 1 - \eta$*, then the graph* G'_ε *has a* $(\frac{1}{2}, \frac{1}{2})$*-partition that cuts at-most* $\eta + \varepsilon$ *fraction of its edges.*

2. *(Soundness/*NO *Case): If* $\mathsf{OPT}(\mathcal{U}) \leq 2^{-O(1/\varepsilon^2)}$*, then every* $(\frac{1}{4}, \frac{3}{4})$*-partition of* G'_ε *cuts at-least* $\sqrt{\varepsilon}$ *fraction of its edges.*

Remark 4.3 *We were imprecise on two counts: (1) The soundness property holds only for those partitions that partition a constant fraction of the blocks* $V'[u]$ *in a roughly balanced way. We call such partitions* piecewise balanced*. This is where the issue of uniform versus non-uniform version of* Balanced Separator *arises. (2) For the soundness property, we can only claim that every piecewise balanced partition cuts at least* ε^t *fraction of edges, where any* $t > \frac{1}{2}$ *can be chosen in advance. Instead, we write* $\sqrt{\varepsilon}$ *for the simplicity of notation.*

4.3 Integrality Gap Instance for the Unique Games SDP Relaxation

This has already been described in Theorem 2.6. The graph $G(V, E)$ therein along with the orthonormal basis $B(u)$, for every $u \in V$, can be used to construct an instance $\mathcal{U}(G(V, E), [N], \{\pi_e\}_{e \in E})$ of Unique Games. For every edge $e = (u, v) \in E$, we have an (unambiguously defined) permutation $\pi_e : [N] \mapsto [N]$, where $\langle \mathbf{u}_{\pi_e(i)}, \mathbf{v}_i \rangle \geq 1 - \eta$, for all $1 \leq i \leq N$.

Theorem 2.6 implies that $\mathsf{OPT}(\mathcal{U}) \leq \frac{1}{N^\eta}$. On the other hand, the fact that for every edge $e = (u, v)$, the bases $B(u)$ and $B(v)$ are very close to each other means that the SDP objective value for $\mathcal{U}$ is at-least $1 - \eta$. (Formally, the SDP objective value is defined to be $E_{e=(u,v) \in E} \left[\frac{1}{N} \sum_{i=1}^{N} \langle \mathbf{u}_{\pi_e(i)}, \mathbf{v}_i \rangle \right]$.)

Thus, we have a concrete instance of Unique Games with optimum at most $\frac{1}{N^\eta} = o(1)$, and which has an SDP solution with objective value at-least $1 - \eta$. This is what an integrality gap example means: The SDP solution *cheats* in an unfair way!

4.4 Integrality Gap Instance for the Balanced Separator SDP Relaxation

Now we combine the two modules described above. We take the instance $\mathcal{U}(G(V,E),[N],\{\pi_e\}_{e\in E})$ as above, and run the PCP reduction on it. This gives us an instance $G'(V',E')$ of Balanced Separator. We show that this is an integrality gap instance in the sense of Defintion 4.1.

Since $\mathcal{U}$ is a NO instance of Unique Games (i.e. $\mathsf{OPT}(\mathcal{U}) = o(1)$), Theorem 4.2 implies that every (piecewise) balanced partition of G' must cut at-least $\sqrt{\varepsilon}$ fraction of the edges. We need to have $1/N^{\eta} \le 2^{-O(1/\varepsilon^2)}$ for this to hold.

On the other hand, we can construct an SDP solution for the Balanced Separator instance which has an objective value of at-most $O(\eta+\varepsilon)$. Note that a typical vertex of G' is $(u,\mathbf{x})$, where $u\in V$ and $\mathbf{x}\in\{-1,1\}^N$. To this vertex, we attach the unit vector $\mathbf{V}_{u,s,\mathbf{x}}^{\otimes t}$ (for $s=4, t=2^{240}+1$), where

$$\mathbf{V}_{u,s,\mathbf{x}} := \frac{1}{\sqrt{N}}\sum_{i=1}^{N} x_i \mathbf{u}_i^{\otimes 2s}.$$

It can be shown that the set of vectors $\{\mathbf{V}_{u,s,\mathbf{x}}^{\otimes t} \mid u\in V,\ \mathbf{x}\in\{-1,1\}^N\}$ satisfy the triangle inequality constraint, and hence, defines a ℓ_2^2 metric. Vectors $\mathbf{V}_{u,s,\mathbf{x}}^{\otimes t}$ and $\mathbf{V}_{u,s,-\mathbf{x}}^{\otimes t}$ are antipodes of each other, and hence, the SDP Constraint (5) is also satisfied. Finally, we show that the SDP objective value (Expression (2)) is $O(\eta+\varepsilon)$. It suffices to show that for every edge $((u,\mathbf{x}),(v,\mathbf{y}))$ in $G'(V',E')$, we have

$$\left\langle \mathbf{V}_{u,s,\mathbf{x}}^{\otimes t}, \mathbf{V}_{v,s,\mathbf{y}}^{\otimes t}\right\rangle \ \ge\ 1 - O(st(\eta+\varepsilon)).$$

This holds because, whenever $((u,\mathbf{x}),(v,\mathbf{y}))$ is an edge of G', we have (after identifying the indices via the permutation $\pi_e : [N]\mapsto[N]$): (a) $\langle \mathbf{u}_{\pi_e(i)}, \mathbf{v}_i\rangle \ge 1-\eta$, for all $1\le i\le N$. (b) The Hamming distance between $\mathbf{x}$ and $\mathbf{y}$ is at-most εN.

4.5 Quantitative Parameters

It follows from above discussion (see also Definition 4.1) that the integrality gap for Balanced Separator is $\Omega(1/\sqrt{\varepsilon})$ provided that $\eta\approx\varepsilon$, and $N^{\eta} > 2^{O(1/\varepsilon^2)}$. We can choose $\eta\approx\varepsilon\approx(\log N)^{-1/3}$. Since the size of the graph G' is at-most $n = 2^{2N}$, we see that the integrality gap is $\approx(\log\log n)^{1/6}$ as desired.

4.6 Proving the Triangle Inequality

As mentioned above, one can show that the set of vectors $\{\mathbf{V}_{u,s,\mathbf{x}}^{\otimes t} \mid u\in V,\ \mathbf{x}\in\{-1,1\}^N\}$ satisfy the triangle inequality constraints. This is the most technical part of the paper, but we would like to stress that this is where the "magic" happens. In our construction, all vectors in $\cup_{u\in V} B(u)$ happen to be points of the hypercube $\{-1,1\}^N$ (upto a normalizing factor of $1/\sqrt{N}$), and therefore, they define an ℓ_1 metric. The apparently *outlandish* operation of taking their $+/-$ combinations combined with tensoring, miraculously leads to a metric that is (ℓ_2^2 and) non-ℓ_1-embeddable.

5 Organization of the Main Body of the Paper

In Sections 6.1 and 6.2 we recall important definitions and results about metric spaces. Section 6.4 defines the cut optimization problems we will be concerned about: Sparsest Cut, Balanced

Separator, Maximum Cut and Minimum Uncut. We also give their SDP relaxations for which we will construct integrality gap instances. Section 6.5 presents useful tools from Fourier analysis.

In Section 7, we present our overall strategy for disproving the $(\ell_2^2, \ell_1, O(1))$-Conjecture. We give a disproval of the $(\ell_2^2, \ell_1, O(1))$-Conjecture assuming an appropriate integrality gap construction for Balanced Separator.

Section 8 presents the UGC of Khot [19]. We also present a natural SDP relaxation of Unique Games in this section. In Section 9 we present the integrality gap instance for the SDP relaxation of Unique Games. We then abstract out the key properties of the instance in Sections 9.3 and 9.3.1.

We build on the Unique Games integrality gap instance in Section 9 to obtain the integrality gap instances for Balanced Separator, Maximum Cut and Minimum Uncut. These are presented in Section 10. This section has two parts: In the first part (Section 10.1), we present the graphs, and in the second part (Section 10.2), we present the corresponding SDP solutions. We establish the soundness of the instances in Section 11 by presenting the corresponding PCP reductions.

Section 10.4 is the most technical part of the paper and this is where we establish that the SDP solutions we construct satisfy the triangle inequality constraint.

6 Preliminaries

6.1 Metric Spaces

Definition 6.1 (X, d) *is a metric space, or* d *is a metric on* X *if:*

1. *For all* $x \in X$, $d(x, x) = 0$.
2. *For all* $x, y \in X, x \neq y$ $d(x, y) > 0$.
3. *For all* $x, y \in X$, $d(x, y) = d(y, x)$.
4. *For all* $x, y, z \in X$, $d(x, y) + d(y, z) \geq d(x, z)$.

(X, d) *is said to be a finite metric space if* X *is finite.* (X, d) *is called a semi-metric space if one allows* $d(x, y) = 0$ *even when* $x \neq y$.

Definition 6.2 (X_1, d_1) *embeds with distortion at-most* Γ *into* (X_2, d_2) *if there exists a map* $\phi : X_1 \mapsto X_2$ *such that for all* $x, y \in X$

$$d_1(x, y) \leq d_2(\phi(x), \phi(y)) \leq \Gamma \cdot d_1(x, y).$$

If $\Gamma = 1$, *then* (X_1, d_1) *is said to* **isometrically** *embed in* (X_2, d_2).

The metrics we would be concerned with in this paper are:

1. ℓ_p *metrics*: For $X \subseteq \mathbb{R}^m$, for some $m \geq 1$, and $\mathbf{x}, \mathbf{y} \in X$, $\ell_p(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^m |x_i - y_i|^p\right)^{1/p}$. Here, $p \geq 1$, and the metric $\ell_\infty(\mathbf{x}, \mathbf{y}) = \max_{i=1}^m |x_i - y_i|$.

2. *Cut (semi-)metrics*: A cut metric δ_S on a set X, defined by the set $S \subseteq X$ is:

$$\delta_S(x,y) = \begin{cases} 1 & \text{if } |\{\text{x}, \text{y}\} \cap \text{S}| = 1 \\ 0 & \text{otherwise} \end{cases}$$

The cut-cone (denoted CUT_n) is the cone generated by cut metrics on an n-point set X. Formally,

$$\mathsf{CUT}_n := \left\{ \sum_S \lambda_S \delta_S : \lambda_S \geq 0 \text{ for all S} \subseteq \text{X} \right\}.$$

To avoid referring to the dimension, denote $\mathsf{CUT} := \cup_n \mathsf{CUT}_n$.

3. *Negative type metrics*: A metric space (X, d) is said to be of negative type if $(X, \sqrt{d})$ embeds isometrically into ℓ_2. Formally, there is an integer m and a vector $\mathbf{v}_x \in \mathbb{R}^m$ for every $x \in X$, such that $d(x,y) = \|\mathbf{v}_x - \mathbf{v}_y\|^2$. Equivalently, for a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^m$, $d(i,j) := \|\mathbf{v}_i - \mathbf{v}_j\|^2$ defines a negative type metric provided that for every triple $\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_k$, the angle between the vectors $\mathbf{v}_i - \mathbf{v}_j$ and $\mathbf{v}_k - \mathbf{v}_j$ is at-most $\pi/2$. The class of all negative type metrics is denoted by ℓ_2^2. A metric d on $\{1, \ldots, n\}$ is of negative type if and only if the matrix $\mathbf{Q}$, defined as $\mathbf{Q}[i,j] := \frac{1}{2}\left(d(i,n) + d(j,n) - d(i,j)\right)$, is positive semi-definite. One can optimize (any linear function) over the class of negative type metrics efficiently via semi-definite programming. In particular, one can efficiently calculate the best distortion needed to embed a given metric into ℓ_2^2.

6.2 Facts about Metric Spaces

Fact 6.3 *[11] Any finite metric space isometrically embeds into ℓ_∞.*

Fact 6.4 *[11] (X, d) is ℓ_1 embeddable if and only if $d \in \mathsf{CUT}$.*

Fact 6.5 *[11] Every ℓ_1 metric is of negative type (i.e. $\ell_1 \subseteq \ell_2^2$).*

Theorem 6.6 (Bourgain's Embedding Theorem [7]) *Any n-point metric space embeds into ℓ_1 with distortion at-most $C_{\mathrm{b}} \log n$, for some absolute constant C_{b}.*

Fact 6.7 *[6, 25] There is an n-point metric, any embedding of which into ℓ_1, requires $\Omega(\log n)$ distortion.*

6.3 The $(\ell_2^2, \ell_1, O(1))$-Conjecture

Conjecture 6.8 ($(\ell_2^2, \ell_1, O(1))$-Conjecture, [16, 24]) *Every negative type metric can be embedded into ℓ_1 with distortion at-most C_{neg}, for some absolute constant $C_{\mathrm{neg}} \geq 1$.*

6.4 Cut Problems and their SDP Relaxations

In this section, we define the cut problems that we study in the paper and present their SDP relaxations. All graphs are complete undirected graphs with non-negative *weights* or *demands* associated to its edges. For a graph $G = (V, E)$, and $S \subseteq V$, let $E(S, \overline{S})$ denote the set of edges with one endpoint in S and other in $\overline{S}$. A cut $(S, \overline{S})$ is called non-trivial if $S \neq \emptyset$ and $\overline{S} \neq \emptyset$.

Remark 6.9 *The versions of* Sparsest Cut *and* Balanced Separator *that we define below are non-uniform versions with demands. The uniform version has all demands equal to* 1 *(i.e. unit demand for every pair of vertices).*

6.4.1 The Sparsest Cut Problem

Definition 6.10 (Sparsest Cut**)** *For a graph $G = (V, E)$ with a weight $\mathbf{wt}(e)$, and a demand $\mathbf{dem}(e)$ associated to each edge $e \in E$, the goal is to optimize*

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

It follows from Fact 6.4 that the objective function above is the same as

$$\min_{d \text{ is } \ell_1 \text{ embeddable}} \frac{\sum_{e=\{x,y\} \in E} \mathbf{wt}(e) d(x,y)}{\sum_{e=\{x,y\} \in E} \mathbf{dem}(e) d(x,y)}.$$

Denote this minimum for $\{G, \mathbf{wt}, \mathbf{dem}\}$ by $\psi_1(G)$. Consider the following two quantities associated to $\{G, \mathbf{wt}, \mathbf{dem}\}$:

$$\psi_\infty(G) := \min_{d \text{ is } \ell_\infty \text{ embeddable}} \frac{\sum_{e=\{x,y\} \in E} \mathbf{wt}(e) d(x,y)}{\sum_{e=\{x,y\} \in E} \mathbf{dem}(e) d(x,y)}, \quad \text{and}$$

$$\psi_{\text{neg}}(G) := \min_{d \text{ is negative type}} \frac{\sum_{e=\{x,y\} \in E} \mathbf{wt}(e) d(x,y)}{\sum_{e=\{x,y\} \in E} \mathbf{dem}(e) d(x,y)}.$$

Facts 6.5 and 6.3 imply that $\psi_1(G) \geq \psi_{\text{neg}}(G) \geq \psi_\infty(G)$. In addition, Bourgain's Embedding Theorem (Theorem 6.6) can be used to show that $\psi_1(G) \leq O(\log n) \cdot \psi_\infty(G)$, where $n := |V|$. Fact 6.7 implies that this factor of $O(\log n)$ is tight upto a constant.

It is also the case that $\psi_{\text{neg}}(G)$ is efficiently computable using the semi-definite program (SDP) of Figure 2. Let the vertex set be $V = \{1, 2, \ldots, n\}$. For a metric d on V, let $\mathbf{Q} := \mathbf{Q}(d)$ be the matrix whose (i,j)-th entry is $\mathbf{Q}[i,j] := \frac{1}{2}\left(d(i,n) + d(j,n) - d(i,j)\right)$.

$$\text{Minimize} \quad \sum_{e=\{i,j\}} \mathbf{wt}(e) d(i,j) \tag{6}$$

Subject to

$$\forall i,j \in V \qquad d(i,j) \geq 0 \tag{7}$$

$$\forall i,j \in V \qquad d(i,j) = d(j,i) \tag{8}$$

$$\forall\ i,j,k \in V \qquad d(i,j) + d(j,k) \geq d(i,k) \tag{9}$$

$$\textstyle\sum_{e=\{i,j\}} \mathbf{dem}(e) d(i,j) = 1 \tag{10}$$

$$\mathbf{Q}(d) \quad \text{is positive semidefinite} \tag{11}$$

Figure 2: SDP relaxation of Sparsest Cut

Fact 6.11 *Suppose that every n-point negative type metric embeds into ℓ_1 with distortion $f(n)$. Then, $\psi_1(G) \le f(n) \cdot \psi_{neg}(G)$, and* Sparsest Cut *can be approximated within factor $f(n)$. In particular, if the $(\ell_2^2, \ell_1, O(1))$-Conjecture (Conjecture 6.8) is true, then there is a constant factor approximation algorithm for* Sparsest Cut.

6.4.2 The Balanced Separator Problem

Definition 6.12 (Balanced Separator) *For a graph $G = (V, E)$ with a weight $\mathbf{wt}(e)$, and a demand $\mathbf{dem}(e)$ associated to each edge $e \in E$, let $D := \sum_{e \in E} \mathbf{dem}(e)$ be the total demand. Let a* **balance** *parameter B be given where $D/4 \le B \le D/2$. The goal is to find a non-trivial cut $(S, \overline{S})$ that minimizes*

$$\sum_{e \in E(S,\overline{S})} \mathbf{wt}(e),$$

subject to

$$\sum_{e \in E(S,\overline{S})} \mathbf{dem}(e) \ge B.$$

*The cuts that satisfy $\sum_{e \in E(S,\overline{S})} \mathbf{dem}(e) \ge B$ are called B-***balanced cuts**.

Figure 3 is an SDP relaxation of Balanced Separator with parameter B.

$$\text{Minimize} \quad \frac{1}{4} \sum_{e=\{x,y\}} \mathbf{wt}(e) \|\mathbf{v}_x - \mathbf{v}_y\|^2 \tag{12}$$

Subject to

$$\forall x \in V \qquad \|\mathbf{v}_x\|^2 = 1 \tag{13}$$

$$\forall\, x, y, z \in V \quad \|\mathbf{v}_x - \mathbf{v}_y\|^2 + \|\mathbf{v}_y - \mathbf{v}_z\|^2 \ge \|\mathbf{v}_x - \mathbf{v}_z\|^2 \tag{14}$$

$$\tfrac{1}{4} \textstyle\sum_{e=\{x,y\}} \mathbf{dem}(e) \|\mathbf{v}_x - \mathbf{v}_y\|^2 \ge B \tag{15}$$

Figure 3: SDP relaxation of Balanced Separator

We need the following two (folk-lore) results stating that one can find a balanced partition in a graph by iteratively finding (approximate) sparsest cut in the graph.

Theorem 6.13 *Suppose $x \mapsto \mathbf{v}_x$ is a solution for* SDP *of Figure 3 with objective value*

$$\frac{1}{4} \sum_{e=\{x,y\}} \mathbf{wt}(e) \|\mathbf{v}_x - \mathbf{v}_y\|^2 \le \varepsilon.$$

Assume that the negative type metric defined by the vectors $\{\mathbf{v}_x \,|\; x \in V\}$ embeds into ℓ_1 with distortion $f(n)$ ($n = |V|$). Then, there exists a B'-balanced cut $(S, \overline{S})$, $B' \ge B/3$ such that

$$\sum_{e \in E(S,\overline{S})} \mathbf{wt}(e) \le O(f(n) \cdot \varepsilon).$$

Lemma 6.14 *If there is factor $f(n)$ approximation for* Sparsest Cut*, then there is a $O(f(n))$ pseudo-approximation to* Balanced Separator*. To be precise, given a* Balanced Separator *instance which has a $D/2$-balanced partition (D is the total demand) that cuts edges with weight α, the algorithm finds a $D/6$-balanced partition that cuts edges with weight at-most $O(f(n) \cdot \alpha)$.*

6.4.3 The ARV-Conjecture

Conjecture 6.15 (Uniform Version) *The integrality gap of the* SDP *of Figure 1 is $O(1)$.*

Conjecture 6.16 (Non-Uniform Version) *The integrality gap of the* SDP *of Figure 3 is $O(1)$.*

Fact 6.17 *The $(\ell_2^2, \ell_1, O(1))$-Conjecture implies the non-uniform* ARV*-Conjecture (Conjecture 6.16).*

In this paper, we disprove the $(\ell_2^2, \ell_1, O(1))$-Conjecture by disproving the non-uniform ARV-Conjecture.

6.4.4 The Maximum Cut Problem

Definition 6.18 (Maximum Cut**)** *For a graph $G = (V, E)$ with a weight $\mathbf{wt}(e)$ associated to each edge $e \in E$, the goal is to optimize*

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

Without loss of generality we may assume that $\sum_{e \in E} \mathbf{wt}(e) = 1$. Figure 4 is a semi-definite relaxation of Maximum Cut.

$$\text{Maximize} \quad \frac{1}{4} \sum_{e=\{x,y\}} \mathbf{wt}(e) \|\mathbf{v}_x - \mathbf{v}_y\|^2 \tag{16}$$

Subject to

$$\forall x \in V \qquad \|\mathbf{v}_x\|^2 = 1 \tag{17}$$

$$\forall\, x, y, z \in V \quad \|\mathbf{v}_x - \mathbf{v}_y\|^2 + \|\mathbf{v}_y - \mathbf{v}_z\|^2 \geq \|\mathbf{v}_x - \mathbf{v}_z\|^2 \tag{18}$$

Figure 4: SDP for Maximum Cut

Goemans and Williamson [17] gave α_{GW} (≈ 0.878) approximation algorithm for Maximum Cut. They showed that every SDP solution with objective value γ_{SDP} can be rounded to a cut in the graph that cuts edges with weight $\geq \alpha_{\mathsf{GW}} \cdot \gamma_{\mathsf{SDP}}$. We note here that their rounding procedure does not make use of the triangle inequality constraints.

6.4.5 The Minimum Uncut Problem

Definition 6.19 (Minimum Uncut**)** *Given a graph $G = (V, E)$ with a weight $\mathbf{wt}(e)$ associated to each edge $e \in E$, the goal is to optimize*

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

$$\text{Minimize} \quad \left(1 - \quad \frac{1}{4} \sum_{e=\{x,y\}} \mathbf{wt}(e)\|\mathbf{v}_x - \mathbf{v}_y\|^2\right) \tag{19}$$

Subject to

$$\forall x \in V \qquad \|\mathbf{v}_x\|^2 = 1 \tag{20}$$

$$\forall\, x, y, z \in V \quad \|\mathbf{v}_x - \mathbf{v}_y\|^2 + \|\mathbf{v}_y - \mathbf{v}_z\|^2 \geq \|\mathbf{v}_x - \mathbf{v}_z\|^2 \tag{21}$$

Figure 5: SDP relaxation of Minimum Uncut

Without loss of generality we may assume that $\sum_{e \in E} \mathbf{wt}(e) = 1$. Figure 5 is a semi-definite relaxation of Minimum Uncut. This is similar to that for Maximum Cut.

Goemans and Williamson [17] showed that every SDP solution for Minimum Uncut with objective value β_{SDP} can be rounded to a cut in the graph, such that the weight of edges left uncut is at-most $O(\sqrt{\beta_{\mathsf{SDP}}})$. We note again that their rounding procedure does not make use of the triangle inequality constraints.

6.5 Fourier Analysis

Consider the real vector space of all functions $f : \{-1,1\}^n \mapsto \mathbb{R}$, where addition of two functions is defined as pointwise addition. For $f, g : \{-1,1\}^n \mapsto \mathbb{R}$, define the following inner product: $\langle f, g \rangle_2 := 2^{-n} \sum_{\mathbf{x} \in \{-1,1\}^n} f(\mathbf{x}) g(\mathbf{x})$.[4] For a set $S \subseteq [n]$, define the *Fourier character* $\chi_S(\mathbf{x}) := \prod_{i \in S} x_i$. It is well-known (and easy to prove) that the set of all Fourier characters form an orthonormal basis with respect to the above inner product. Hence, every function $f : \{-1,1\}^n \mapsto \mathbb{R}$ has a (unique) representation as $f = \sum_{S \subseteq [n]} \hat{f}_S \chi_S$, where the Fourier coefficient $\hat{f}_S := \langle f, \chi_S \rangle_2$.

Fact 6.20 (Parseval's Identity) *For any* $f : \{-1,1\}^n \mapsto \{-1,1\}$, $\sum_{S \subseteq [n]} \hat{f}_S^2 = 1$.

Definition 6.21 (Long Code) *The Long Code over a domain* $[N]$ *is indexed by all inputs* $\mathbf{x} \in \{-1,1\}^N$. *The Long Code* f *of an element* $j \in [N]$ *is defined as* $f(\mathbf{x}) := \chi_{\{j\}}(\mathbf{x}) = x_j$, *for all* $\mathbf{x} = (x_1, \ldots, x_N) \in \{-1,1\}^N$.

Thus, a Long Code is simply a boolean function that is a dictatorship, i.e., it depends only on one co-ordinate. In particular, if f is the Long Code of $j \in [N]$, then $\hat{f}_{\{j\}} = 1$ and all other Fourier coefficients are zero.

Definition 6.22 *For* $\rho \in (-1,1)$,

$$S_\rho(f) := \sum_{S \subseteq [n]} \hat{f}_S^2 \rho^{|S|}.$$

[4] Notice that this inner product is the same as the standard inner product of truth-tables of f and g upto the normalizing factor of $1/2^n$.

Theorem 6.23 (Majority Is Stablest [27]) *For every $\rho \in (-1, 0]$ and $\varepsilon > 0$, there is a small enough $\delta := \delta(\varepsilon, \rho) > 0$, and a large enough $k := k(\varepsilon, \rho)$, such that, for any $f : \{-1,1\}^n \mapsto [-1,1]$, if $\mathrm{Inf}_i^{\le k}(f) \le \delta$ for all $1 \le i \le n$, then*

$$S_\rho(f) \ge 1 - \frac{2}{\pi} \arccos \rho - \varepsilon.$$

Here, $\mathrm{Inf}_i^{\le k}(f) := \sum_{S : i \in S, |S| \le k} \widehat{f}_S^2$.

For $f : \{-1,1\}^n \mapsto \mathbb{R}$, and $p \ge 1$, let $\|f\|_p := \left(\frac{1}{2^n} \sum_{\mathbf{x} \in \{-1,1\}^n} |f(\mathbf{x})|^p\right)^{1/p}$.

Definition 6.24 (Hyper-contractive Operator) *For each $\rho \in [-1, 1]$, the Bonami-Beckner operator T_ρ is a linear operator that maps the space of functions $\{-1,1\}^n \mapsto \mathbb{R}$ into itself via*

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

Fact 6.25 *For a boolean function $f : \{-1,1\}^n \mapsto \{-1,1\}$, and $\rho \in [0,1]$,*

$$\|T_{\sqrt{\rho}}[f]\|_2^2 = S_\rho(f).$$

Theorem 6.26 (Bonami-Beckner Inequality [29]) *Let $f : \{-1,1\}^n \mapsto \mathbb{R}$ and $1 < p < q$. Then*

$$\|T_\rho[f]\|_q \le \|f\|_p$$

for all $0 \le \rho \le \left(\frac{p-1}{q-1}\right)^{1/2}$.

Theorem 6.27 (Bourgain's Junta Theorem [8]) *Fix any $\frac{1}{2} < t < 1$. Then, there exists a constant $c_t > 0$, such that, for all integers k, for all $\gamma > 0$ and for all boolean functions $f : \{-1,1\}^n \mapsto \{-1,1\}$,*

$$\text{If} \sum_{S \,:\, |S| > k} \widehat{f}_S^2 < c_t k^{-t} \quad \text{then} \qquad \sum_{S \,:\, |\widehat{f}_S| \le \gamma 4^{-k^2}} \widehat{f}_S^2 < \gamma^2.$$

6.6 Standard Inequalities

The following are some standard inequalities which will be used in the soundness analysis of the PCPs without any explicit referencing. We state them here for completeness.

Fact 6.28 *Let X be a random variable such that $\mathbf{Pr}[0 \le X \le 1] = 1$, and $0 \le \delta \le 1$. If $\mathbf{E}[X] \ge \delta$, then $\mathbf{Pr}[X \ge \delta/2] \ge \delta/2$.*

Fact 6.29 (Markov's Inequality) *Let X be a non-negative random variable. For any $t > 0$, $\mathbf{Pr}[X \ge t] \le \frac{\mathbf{E}[X]}{t}$.*

Fact 6.30 (Jensen's Inequality) *For a random variable X, $\mathbf{E}[X^2] \ge (\mathbf{E}[X])^2$.*

7 Overall Strategy for Disproving the $(\ell_2^2, \ell_1, O(1))$-Conjecture

We describe the high-level approach to our disproval of the $(\ell_2^2, \ell_1, O(1))$-Conjecture in this section. We construct an integrality gap instance of non-uniform BALANCED SEPARATOR to disprove the non-uniform ARV-Conjecture, and that suffices to disprove the $(\ell_2^2, \ell_1, O(1))$-Conjecture using the (folk-lore) Fact 6.17. The instance has two parts: (1) The graph and (2) The SDP solution. The graph construction is described in Section 10.1, while the SDP solution appears in Section 10.2.

We construct a complete weighted graph $G(V, \mathbf{wt})$, with vertex set V and weight $\mathbf{wt}(e)$ on edge e, and with $\sum_e \mathbf{wt}(e) = 1$. The vertex set is partitioned into sets $V_1, V_2, \ldots, V_r$, each of size $|V|/r$ (think of $r \approx \sqrt{|V|}$).

A cut A in the graph is viewed as a function $A : V \mapsto \{-1, 1\}$. We are interested in cuts that cut *many* sets V_i in a *somewhat balanced* way.

Definition 7.1 *For $0 \le \theta \le 1$, a cut A is called θ-piecewise balanced if*

$$\mathbf{E}_{i \in_R [r]} \Big[\Big| \mathbf{E}_{x \in_R V_i} [A(x)] \Big| \Big] \le \theta.$$

We also assign a unit vector to every vertex in the graph. Let $\mathbf{v}_x$ denote the vector assigned to vertex x. Our construction of the graph $G(V, \mathbf{wt})$ and the vector assignment $x \mapsto \mathbf{v}_x$ can be summarized as follows:

Theorem 7.2 *Fix any $\frac{1}{2} < t < 1$. For every sufficiently small $\varepsilon > 0$, there exists a graph $G(V, \mathbf{wt})$, with a partition $V = \cup_{i=1}^r V_i$, and a vector assignment $x \mapsto \mathbf{v}_x$ for every $x \in V$, such that*

1. $|V| \le 2^{2^{O(1/\varepsilon^3)}}$.
2. *Every $\frac{5}{6}$-piecewise balanced cut A must cut ε^t fraction of edges, i.e., for any such cut*
$$\sum_{e \in E(A, \overline{A})} \mathbf{wt}(e) \ge \varepsilon^t.$$
3. *The unit vectors $\{\mathbf{v}_x \mid x \in V\}$ define a negative type metric, i.e., the following triangle inequality is satisfied:*
$$\forall\ x, y, z \in V,\ \|\mathbf{v}_x - \mathbf{v}_y\|^2 + \|\mathbf{v}_y - \mathbf{v}_z\|^2 \ge \|\mathbf{v}_x - \mathbf{v}_z\|^2 .$$
4. *For each part V_i, the vectors $\{\mathbf{v}_x \mid x \in V_i\}$ are* **well-separated**, *i.e.,*
$$\frac{1}{2}\mathbf{E}_{x,y \in_R V_i} \left[\|\mathbf{v}_x - \mathbf{v}_y\|^2 \right] = 1.$$
5. *The vector assignment gives a* **low** SDP *objective value, i.e.,*
$$\frac{1}{4} \sum_{e=\{x,y\}} \mathbf{wt}(e) \|\mathbf{v}_x - \mathbf{v}_y\|^2 \le \varepsilon.$$

Theorem 7.3 *The $(\ell_2^2, \ell_1, O(1))$-Conjecture is false. In fact, for every $\delta > 0$, for all sufficiently large n, there are n-point negative type metrics that require distortion at-least $(\log \log n)^{1/6-\delta}$ to embed into ℓ_1.*

Proof: Suppose that the negative type metric defined by vectors $\{\mathbf{v}_x | \ x \in V\}$ in Theorem 7.2 embeds into ℓ_1 with distortion Γ. We will show that $\Gamma = \Omega\left(\frac{1}{\varepsilon^{1-t}}\right)$ using Theorem 6.13.

Construct an instance of Balanced Separator as follows. The graph $G(V, \mathbf{wt})$ is as in Theorem 7.2. The demands $\mathbf{dem}(e)$ depend on the partition $V = \cup_{i=1}^r V_i$. We let $\mathbf{dem}(e) = 1$ if e has both endpoints in the same part V_i for some $1 \le i \le r$, and $\mathbf{dem}(e) = 0$ otherwise. Clearly, $D := \sum_e \mathbf{dem}(e) = r \cdot \binom{|V|/r}{2}$.

Now, $x \mapsto \mathbf{v}_x$ is an assignment of unit vectors that satisfy the triangle inequality constraints. This will be a solution to the SDP of Figure 3. Property (4) of Theorem 7.2 guarantees that

$$\frac{1}{4} \sum_{e=\{x,y\}} \mathbf{dem}(e) \|\mathbf{v}_x - \mathbf{v}_y\|^2 = \frac{1}{4} \cdot r \cdot \binom{|V|/r}{2} \cdot 2 = D/2 =: B.$$

Thus, the SDP solution is $D/2$-balanced and its objective value is at-most ε. Using Theorem 6.13, we get a B'-balanced cut $(A, \overline{A})$, $B' \ge D/6$ such that $\sum_{e \in E(A, \overline{A})} \mathbf{wt}(e) \le O(\Gamma \cdot \varepsilon)$.

Claim: The cut $(A, \overline{A})$ must be a $\frac{5}{6}$-piecewise balanced cut.

Proof of Claim: Let $p_i := \mathbf{Pr}_{x \in V_i}[A(x) = 1]$. The total demand cut by $(A, \overline{A})$ is equal to $\sum_{i=1}^r p_i(1-p_i)|V_i|^2$. This is at-least $B' \ge D/6$ since $(A, \overline{A})$ is B'-balanced. Hence

$$\sum_{i=1}^r p_i(1-p_i) \cdot |V|^2/r^2 \ge \frac{1}{6}\ r \cdot \binom{|V|/r}{2}.$$

Thus, $\sum_{i=1}^r p_i(1-p_i) \ge \frac{r}{12}$. By Cauchy-Schwartz,

$$\mathbf{E}_{i \in_R [r]}\Big[\Big|\mathbf{E}_{x \in_R V_i}[A(x)]\Big|\Big] = \frac{1}{r}\sum_{i=1}^r |1-2p_i| \le \sqrt{\frac{1}{r}\sum_{i=1}^r (1-2p_i)^2} = \sqrt{1 - \frac{4}{r}\sum_{i=1}^r p_i(1-p_i)} \le \sqrt{\frac{2}{3}} < \frac{5}{6}.$$

Hence, $(A, \overline{A})$ must be a $\frac{5}{6}$-piecewise balanced cut.

However, Property (2) of Theorem 7.2 says that such a cut must cut at-least ε^t fraction of edges. This implies that $\Gamma = \Omega(\frac{1}{\varepsilon^{1-t}})$. The theorem follows by noting that $t > \frac{1}{2}$ is arbitrary and $n := |V| \le 2^{2^{O(1/\varepsilon^3)}}$. ∎

8 The Unique Games Conjecture (UGC)

In this section, we present the UGC due to Khot [19]. We also present the SDP relaxation of Unique Games considered by Khot in [19] in Section 8.3. This SDP was inspired by a paper of Feige and Lovasz [14]. We start by defining Unique Games and presenting some notations.

8.1 Unique Games

Definition 8.1 (Unique Games) *An instance $\mathcal{U}(G(V,E), [N], \{\pi_e\}_{e \in E}, \mathbf{wt})$ of* Unique Games *is defined as follows: $G = (V, E)$ is a a graph with a set of vertices V and a set of edges E, with possibly parallel edges. An edge e whose endpoints are v and w is written as $e\{v, w\}$. For every $e \in E$, there is a bijection $\pi_e : [N] \mapsto [N]$, and a weight $\mathbf{wt}(e) \in \mathbb{R}^+$. For an edge $e\{v, w\}$, we think* *of* π_e *as a pair of permutations* $\{\pi_e^v, \pi_e^w\}$*, where* $\pi_e^w = (\pi_e^v)^{-1}$*.* π_e^v *is a mapping that takes a label of vertex* w *to a label of vertex* v*. The goal is to assign one* label *to every vertex of the graph from the set* $[N]$*. The labeling is supposed to satisfy the constraints given by bijective maps* π_e*. A labeling* $\lambda : V \mapsto [N]$ satisfies *an edge* $e\{v, w\}$*, if* $\lambda(v) = \pi_e^v(\lambda(w))$*. Define the indicator function* $I^\lambda(e)$*, which is* 1 *if* e *is satisfied by* λ *and* 0 *otherwise. The optimum* $\mathsf{OPT}(\mathcal{U})$ *of the* Unique Games *instance is defined to be*

$$\max_\lambda \sum_{e \in E} \mathbf{wt}(e) \cdot I^\lambda(e).$$

Without loss of generality, we assume that $\sum_{e \in E} \mathbf{wt}(e) = 1$*.*

Conjecture 8.2 (UGC **[19]**) *For every pair of constants* $\eta, \zeta > 0$*, there exists a sufficiently large constant* $N := N(\eta, \zeta)$*, such that it is* NP*-hard to decide whether a* Unique Games *instance* $\mathcal{U}\left(G(V, E), [N], \{\pi_e\}_{e \in E}, \mathbf{wt}\right)$*, has* $\mathsf{OPT}(\mathcal{U}) \geq 1 - \eta$*, or* $\mathsf{OPT}(\mathcal{U}) \leq \zeta$*.*

8.2 Notations

The following notations will be used throughout the paper.

1. Since the graph can have parallel edges, we denote an edge e with end-points $\{v, w\}$ by $e\{v, w\}$.

2. For an edge $e\{v, w\}$, π_e^v takes a label assigned to w to a label for v, and $\pi_e^w := (\pi_e^v)^{-1}$. We will often use π_e when we do not want to refer to either π_e^v or π_e^w.

3. We always assume (without loss of generality) that $\sum_{e \in E} \mathbf{wt}(e) = 1$. This is to be thought of as a probability distribution over the edges of the graph.

4. $\Gamma(v)$ will denote the set of edges adjacent to a vertex v.

5. $p_v := \frac{1}{2} \sum_{e \in \Gamma(v)} \mathbf{wt}(e)$. Since $\sum_{e \in E} \mathbf{wt}(e) = 1$, this defines a probability distribution on the vertices of the graph.

6. For a vertex v and $e \in \Gamma(v)$, let $\Psi_v(e) := \frac{\mathbf{wt}(e)}{2p_v}$. This defines a probability distribution on the edges adjacent to v.

7. With the above notation, the following two sampling procedures are equivalent:

   - Pick an edge $e\{v, w\}$ with probability $\mathbf{wt}(e)$.
   - Pick a vertex v with probability p_v and then pick an edge $e \in \Gamma(v)$ with probability $\Psi_v(e)$.

8.3 SDP Relaxation for Unique Games

Consider a Unique Games instance $\mathcal{U} = \left(G(V, E), [N], \{\pi_e\}_{e \in E}, \mathbf{wt}\right)$. Khot [19] proposed the SDP in Figure 6. Here, for every $u \in V$, we associate a set of N orthogonal vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$. The intention is that if $i_0 \in [N]$ is a label for vertex $u \in V$, then $\mathbf{u}_{i_0} = \sqrt{N}\mathbf{1}$, and $\mathbf{u}_i = \mathbf{0}$ for all $i \neq i_0$. Here, $\mathbf{1}$ is some fixed unit vector and $\mathbf{0}$ is the zero-vector. However, once we take the SDP relaxation, this may no longer be true and $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N\}$ could be any set of orthogonal vectors. In fact, in our construction, they will form an orthonormal basis of $\mathbb{R}^N$.

$$\text{Maximize} \quad \sum_{e\{u,v\}\in E} \mathbf{wt}(e)\cdot \frac{1}{N}\left(\sum_{i=1}^{N}\left\langle \mathbf{u}_{\pi_e^u(i)}, \mathbf{v}_i\right\rangle\right) \tag{22}$$

Subject to

$$\forall\ u \in V \quad \langle \mathbf{u}_1, \mathbf{u}_1\rangle + \cdots + \langle \mathbf{u}_N, \mathbf{u}_N\rangle = N \tag{23}$$

$$\forall\ u \in V\ \ \forall\ \ i \neq j \qquad \langle \mathbf{u}_i, \mathbf{u}_j\rangle = 0 \tag{24}$$

$$\forall\ u, v \in V\ \ \forall\ \ i, j \qquad \langle \mathbf{u}_i, \mathbf{v}_j\rangle \geq 0 \tag{25}$$

$$\forall\ u, v \in V \qquad \textstyle\sum_{1\leq i,j\leq N}\ \langle \mathbf{u}_i, \mathbf{v}_j\rangle = N \tag{26}$$

Figure 6: SDP for Unique Games

Theorem 8.3 (Khot [19]) *If the* SDP *in Figure 6 has a solution with objective value* $1-\delta$*, then that solution can be rounded to get a labeling that satisfies edges with weight* $1-O\left(N^2\delta^{1/5}\sqrt{\log(\frac{1}{\delta})}\right)$.

9 Integrality Gap Instance for the SDP of Unique Games

In this section, we construct an integrality gap instance for the SDP in Figure 6. To be precise, for parameters N and η, we will construct an instance $\mathcal{U}\left(G(V,E),[N],\{\pi_e\}_{e\in E},\mathbf{wt}\right)$ of Unique Games such that (a) $\mathsf{OPT}(\mathcal{U}) \leq \frac{1}{N^\eta}$ and (b) There is an SDP solution with objective value at-least $1-\eta$. This construction will be later used to construct integrality gap instances for cut problems.

Let $\mathcal{F}$ denote the family of all boolean functions on $\{-1,1\}^k$. For $f, g \in \mathcal{F}$, define the product fg as $(fg)(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$. Consider the equivalence relation $\equiv$ on $\mathcal{F}$ defined as $f \equiv g$ if and only if there is a $S \subseteq [k]$, such that $f = g\chi_S$. This relation partitions $\mathcal{F}$ into equivalence classes $\mathcal{P}_1, \ldots, \mathcal{P}_m$. We denote by $[\mathcal{P}_i]$, the function $f_i \in \mathcal{P}_i$ which is the smallest among all functions in $\mathcal{P}_i$ with respect to $\preccurlyeq$. Thus, by definition,

$$\mathcal{P}_i = \{[\mathcal{P}_i]\chi_S \mid S \subseteq [k]\}.$$

It follows from the orthogonality of the characters $\{\chi_S\}_{S\subseteq[k]}$, that all the functions in any partition are also orthogonal. Further, for a function $f \in \mathcal{F}$, let $\mathcal{P}(f)$ denote the $\mathcal{P}_i$ in which f belongs.

For $\rho > 0$, let $f \in_\rho \mathcal{F}$ denote a random boolean function on $\{-1,1\}^k$ where for every $\mathbf{x} \in \{-1,1\}^k$, independently, $f(\mathbf{x}) = 1$ with probability $1-\rho$, and -1 with probability ρ. For a parameter $\eta > 0$ and boolean functions $f, g \in \mathcal{F}$, let $\mathbf{wt}_\eta(\{f,g\})$ denote the following:

$$\mathbf{wt}_\eta(\{f,g\}) := \mathbf{Pr}_{f'\in_{1/2}\mathcal{F},\, \boldsymbol{\mu}\in_\eta\mathcal{F}}\left[(f = f' \wedge g = f'\boldsymbol{\mu}) \vee (g = f' \wedge f = f'\boldsymbol{\mu})\right].$$

First, notice that this probability does not change if we interchange the roles of f and g. Further, for any $S \subseteq [k]$, $\mathbf{wt}_\eta(\{f,g\}) = \mathbf{wt}_\eta(\{f\chi_S, g\chi_S\})$.

We are now ready to define the Unique Games instance for which we will establish the integrality gap. Fix $\eta > 0$. First, we define the graph $G = (V,E)$. The set of vertices is $V := \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$. For every $f, g \in \mathcal{F}$, there is an edge in E between the vertices $\mathcal{P}(f)$ and $\mathcal{P}(g)$ with weight $\mathbf{wt}_\eta(\{f,g\})$. The set of labels for the Unique Games instance will be $2^{[k]} := \{S : S \subseteq [k]\}$, i.e., the set of labels $[N]$ is identified with the set $2^{[k]}$ (and thus, $N = 2^k$). This identification will be used for the rest of the paper. The bijection π_e, for the edge $e\{f,g\}$ corresponding to the pair of functions f and g can now be defined: If $f = [\mathcal{P}(f)]\chi_S$ and $g = [\mathcal{P}(g)]\chi_T$, for some $S, T \subseteq [k]$, then

$$\pi_e^{\mathcal{P}(f)}(T \triangle U) := S \triangle U, \quad \forall\, U \subseteq [k].$$

Here, $\triangle$ is the symmetric difference operator on sets. Note that $\pi_e^{\mathcal{P}(f)} : 2^{[k]} \mapsto 2^{[k]}$ is a permutation on the set of allowed labels.

Remark 9.1 *It is readily verified (for the graph defined above) that for a vertex $v \in V$, the quantity*

$$p_v := \frac{1}{2} \sum_{e \in \Gamma(v)} \mathbf{wt}_\eta(e),$$

is independent of v. The weight function $\mathbf{wt}_\eta$ should be thought of as a probability distribution on the edges, while the function p_v should be thought of as the probability distribution on the vertices, induced by $\mathbf{wt}_\eta$.

9.1 Soundness: No Good Labeling

We show that any labeling to the Unique Games instance described above achieves an objective of at-most $\frac{1}{N^\eta}$. Consider a labeling $\mathsf{R} : \{\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_m\} \mapsto 2^{[k]}$ which assigns to each $\mathcal{P}_i$, a subset of $[k]$. We *extend* this labeling (denote this extended labeling by $\mathcal{R}$) to all $g \in \mathcal{F}$ as follows: For $g = [\mathcal{P}(g)]\chi_T$, define $\mathcal{R}(g) := \mathsf{R}(\mathcal{P}(g)) \triangle T$. We claim that the objective of the Unique Games instance for this labeling is exactly equal to the following probability:

$$\mathbf{Pr}_{f \in_R \mathcal{F},\ \boldsymbol{\mu} \in_\eta \mathcal{F}} \left[\mathcal{R}(f) = \mathcal{R}(f\boldsymbol{\mu})\right]. \tag{27}$$

The reason is this: Let $e\{f, g\}$ be an edge of the Unique Games instance with $f = [\mathcal{P}(f)]\chi_S$, $g = [\mathcal{P}(g)]\chi_T$, and $g = f\boldsymbol{\mu}$. Then

$$\begin{aligned} \pi_e^{\mathcal{P}(f)}(\mathsf{R}(\mathcal{P}(g))) = \mathsf{R}(\mathcal{P}(f)) \quad &\iff \quad \mathsf{R}(\mathcal{P}(g))\Delta T = \mathsf{R}(\mathcal{P}(f))\Delta S \\ &\iff \quad \mathcal{R}(g) = \mathcal{R}(f) \iff \mathcal{R}(f\boldsymbol{\mu}) = \mathcal{R}(f) \end{aligned}$$

Thus, the labeling R satisfies the edge $e\{f, g\}$ if and only if $\mathcal{R}(f) = \mathcal{R}(f\mu)$. Hence, the objective of the Unique Games instance (i.e. total weight of its edges satisfied) equals (27).

We will upper-bound this probability using the Bonami-Beckner Inequality. For a set $S \subseteq [k]$, let $\mathcal{R}^S(f)$ be the indicator function, which is 1 if $\mathcal{R}(f) = S$, and 0 otherwise. Clearly, for any extended labeling $\mathcal{R}$, and any $S \subseteq [k]$, $\mathbf{E}_{f \in_R \mathcal{F}}[\mathcal{R}^S(f)] = \frac{1}{N}$.

$$\mathbf{Pr}_{f \in_R \mathcal{F},\ \boldsymbol{\mu} \in_\eta \mathcal{F}} \left[\mathcal{R}(f) = \mathcal{R}(f\boldsymbol{\mu}) = S\right] = \mathbf{E}_{f \in_R \mathcal{F},\ \boldsymbol{\mu} \in_\eta \mathcal{F}} \left[\mathcal{R}^S(f)\mathcal{R}^S(f\boldsymbol{\mu})\right].$$

Via the Fourier expansion of the function $\mathcal{R}^S : \{-1, 1\}^N \mapsto \{0, 1\}$, we get that the above expectation is

$$\sum_{\alpha \subseteq [N]} \left(\widehat{\mathcal{R}^S_\alpha}\right)^2 (1 - 2\eta)^{|\alpha|} = \left\| T_{\sqrt{1-2\eta}}[\mathcal{R}^S] \right\|_2^2 \le \left\|\mathcal{R}^S\right\|_{2-2\eta}^2 = \frac{1}{N^{\frac{1}{1-\eta}}} \le \frac{1}{N^{1+\eta}}.$$

Here, the second last inequality uses the Bonami-Beckner Inequality (see Theorem 6.26). Hence, by a union bound over the sets $S \subseteq [k]$, it follows that

$$\mathbf{Pr}_{f \in_R \mathcal{F}, \boldsymbol{\mu} \in_\eta \mathcal{F}} \left[\mathcal{R}(f) = \mathcal{R}(f\boldsymbol{\mu})\right] \le \frac{1}{N^\eta}.$$

This proof, using the Bonami-Beckner Inequality, was suggested by Ryan O'Donnell.

9.2 Completeness: A Good SDP Solution

Let $N := 2^k$ and $\mathcal{F}$ be the set of functions $f : \{-1,1\}^k \mapsto \{-1,1\}$ as before. For $f \in \mathcal{F}$, let $\mathbf{u}_f$ denote the unit vector (with respect to the ℓ_2 norm) corresponding to the truth-table of f. Formally, indexing the vector $\mathbf{u}_f$ with coordinates $\mathbf{x} \in \{-1,1\}^k$, $(\mathbf{u}_f)_{\mathbf{x}} := \frac{f(\mathbf{x})}{\sqrt{N}}$.

Recall that in the SDP relaxation of Unique Games, for every vertex in V, we need to assign a set of orthogonal vectors. Let $f = [\mathcal{P}_i]$ be the representative boolean function for the vertex $\mathcal{P}_i$. With $\mathcal{P}_i$, we associate the set of vectors $\left\{\mathbf{u}_{f\chi_S}^{\otimes 2}\right\}_{S\subseteq[k]}$. The following facts are easily verified.

1. $\sum_{S\subseteq[k]} \left\langle \mathbf{u}_{f\chi_S}^{\otimes 2}, \mathbf{u}_{f\chi_S}^{\otimes 2}\right\rangle = \sum_{S\subseteq[k]} \langle \mathbf{u}_{f\chi_S}, \mathbf{u}_{f\chi_S}\rangle^2 = N$.
2. For $S \neq T \subseteq [k]$, $\left\langle \mathbf{u}_{f\chi_S}^{\otimes 2}, \mathbf{u}_{f\chi_T}^{\otimes 2}\right\rangle = \langle \mathbf{u}_{f\chi_S}, \mathbf{u}_{f\chi_T}\rangle^2 = \langle \mathbf{u}_{\chi_S}, \mathbf{u}_{\chi_T}\rangle^2 = 0$.
3. For $f, g \in \mathcal{F}$ and $S, T \subseteq [k]$, $\left\langle \mathbf{u}_{f\chi_S}^{\otimes 2}, \mathbf{u}_{g\chi_T}^{\otimes 2}\right\rangle = \langle \mathbf{u}_{f\chi_S}, \mathbf{u}_{g\chi_T}\rangle^2 \geq 0$.
4. For $f \in \mathcal{P}_i$, $g \in \mathcal{P}_j$ for $i \neq j$,

$$\sum_{S,T\subseteq[k]} \left\langle \mathbf{u}_{f\chi_S}^{\otimes 2}, \mathbf{u}_{g\chi_T}^{\otimes 2}\right\rangle = \sum_{S,T\subseteq[k]} \langle \mathbf{u}_{f\chi_S}, \mathbf{u}_{g\chi_T}\rangle^2 = \sum_{T\subseteq[k]} \|\mathbf{u}_{g\chi_T}\|^2 = N.$$

Here, the second last equality follows from the fact that, for any $f \in \mathcal{F}$, $\{\mathbf{u}_{f\chi_S}\}_{S\subseteq[k]}$ forms an orthonormal basis for $\mathbb{R}^N$.

Hence, all the conditions of the SDP for Unique Games are satisfied. Next, we show that this vector assignment has an objective at-least $1-9\eta$. Notice that most of the weight is concentrated on edges corresponding to pairs of boolean functions f, g which differ at at-most 2η fraction of points. More formally, it follows from a Chernoff type bound that for any $f \in \mathcal{F}$, $\mathbf{Pr}_{\boldsymbol{\mu}\in_\eta\mathcal{F}}[\mathbf{dist}(f, f\boldsymbol{\mu}) \geq 2\eta N] \leq (0.9)^{\eta N}$. Here, $\mathbf{dist}(f,g)$ is defined as the number of points where f and g differ. Hence, for $N \geq \Omega\left(\frac{1}{\eta}\log\frac{1}{\eta}\right)$, this probability is at-most η. Further, if $\mathbf{dist}(f,g) \leq 2\eta N$, $\langle \mathbf{u}_f, \mathbf{u}_g\rangle \geq 1 - 4\eta$. Hence, $\left\langle \mathbf{u}_f^{\otimes 2}, \mathbf{u}_g^{\otimes 2}\right\rangle \geq (1-4\eta)^2 \geq (1-8\eta)$. This implies that the objective of this SDP solution is at-least $(1-\eta)(1-8\eta) \geq 1-9\eta$.

9.3 Abstracting the Unique Games Integrality Gap Instance

Letting $n = 2^N$, the following theorem summarizes our integrality gap example for the SDP of Unique Games.

Theorem 9.2 *There is a constant $c > 0$, such that for any $0 < \eta < 1/2$, for all $n > 2^{\frac{c}{\eta}\log\frac{1}{\eta}}$, there is a* Unique Games *instance $\mathcal{U}_\eta\left(G(V,E), [N], \{\pi_e\}_{e\in E}, \mathbf{wt}\right)$ satisfying properties in Fig. 7. Moreover, this instance has an SDP solution as described in Theorem 9.3.*

Property	
Vertex Set V	$\lvert V\rvert = n/\log n$
Label Set $[N]$	$[N]$ is identified with $2^{[k]}$. Here $N = 2^k,\ n = 2^N$.
Optimum	$\mathsf{OPT}(\mathcal{U}_\eta) \leq \log^{-\eta} n$
Prob. distr.	
Edge Weights	$\mathbf{wt} : E \mapsto \mathbb{R}^+$, $\sum_{e\in E} \mathbf{wt}(e) = 1$
$\Gamma(v)$	Set of edges adjacent to v
$p_v,\ v \in V$	$p_v = \frac{1}{2}\sum_{e\in\Gamma(v)} \mathbf{wt}(e) = \frac{1}{\lvert V\rvert}$
$\Psi_v(e),\ v \in V$	$\Psi_v(e) = \frac{\mathbf{wt}(e)}{2p_v}$

Figure 7: Abstracting the Unique Games Instance

9.3.1 Abstracting the SDP solution

In this section, we summarize the key properties of the SDP solution presented in Section 9.2. For every vertex $v \in V$ of the Unique Games instance of Theorem 9.2, there is an associated set of vectors: $\{\mathbf{v}_S\}_{S\subseteq[k]}$. For instance, if the vertex v corresponds to the partition $\mathcal{P}_i$, and $f = [\mathcal{P}_i]$ is the representative, then we define $\mathbf{v}_S := \mathbf{u}_{f\chi_S}$. The set of vectors for vertices $v, w, \ldots$ will be denoted by their boldface counterparts $\{\mathbf{v}_S\}_{S\subseteq[k]}, \{\mathbf{w}_T\}_{T\subseteq[k]}, \ldots$ respectively.

Theorem 9.3 *For every $v \in V$ of the* Unique Games *instance of Theorem 9.2, there is a set of vectors $\{\mathbf{v}_S\}_{S\subseteq[k]}$ which satisfy the following properties:*

1. **Orthonormal Basis**
The set of vectors $\{\mathbf{v}_S\}_{S\subseteq[k]}$ forms an orthonormal basis for the space $\mathbb{R}^{2^k}$. Hence, for any vector $\mathbf{w} \in \mathbb{R}^{2^k}$, $\|\mathbf{w}\|^2 = \sum_{S\subseteq[k]}\langle \mathbf{w}, \mathbf{v}_S\rangle^2$.

2. **Triangle Inequality**
For any $u, v, w \in V$, and any $S, T, U \subseteq [k]$, $1 + \langle \mathbf{u}_S, \mathbf{v}_T\rangle \geq \langle \mathbf{u}_S, \mathbf{w}_U\rangle + \langle \mathbf{v}_T, \mathbf{w}_U\rangle$.

3. **Matching Property**
For any $v, w \in V$, and $T_1, T_2, S \subseteq [k]$, $\langle \mathbf{v}_{T_1}, \mathbf{w}_{T_2}\rangle = \langle \mathbf{v}_{T_1\triangle S}, \mathbf{w}_{T_2\triangle S}\rangle$.

4. **Closeness Property**
For $e\{v, w\}$ picked with probability $\mathbf{wt}(e)$, with probability at-least $1-\eta$, there are sets $S, S' \subseteq [k]$ such that $\langle \mathbf{v}_S, \mathbf{w}_{S'}\rangle \geq 1 - 4\eta$. Moreover, if π_e is the bijection corresponding to this edge, then $\pi_e^v(S'\Delta U) = S\Delta U$ for all $U \subseteq [k]$.

5. **Negative type property of the average**
For $v, w \in V$, and an integer $t > 0$, consider the distance function

$$d_t(v, w) := \left\| \frac{1}{\sqrt{N}} \sum_{S\subseteq[k]} \mathbf{v}_S^{\otimes t} - \frac{1}{\sqrt{N}} \sum_{T\subseteq[k]} \mathbf{w}_T^{\otimes t} \right\|_2^2.$$

Then, for all $u, v, w \in V$,

$$d_t(u, v) + d_t(v, w) \geq d_t(u, w).$$

Hence, d_t defines a negative type semi-metric (actually, ℓ_1 metric, as we will see). These vectors will be useful in the construction of the integrality gap instance for Maximum Cut.

Proof:

1. Notice that for any $f \in \mathcal{F}$, the set of vectors $\{\mathbf{u}_{f\chi_S}\}_{S\subseteq[k]}$ forms an orthonormal basis for the space $\mathbb{R}^{2^k}$.

2. Notice that for any $f, g, h \in \mathcal{F}$, $1 + \langle \mathbf{u}_f, \mathbf{u}_h \rangle \geq \langle \mathbf{u}_f, \mathbf{u}_g \rangle + \langle \mathbf{u}_g, \mathbf{u}_h \rangle$. This is because the set of vectors $\{\mathbf{u}_f\}_{f\in\mathcal{F}}$ are (normalized) truth-tables of $\{-1, 1\}$-valued functions.

3. Notice that for any $f, g \in \mathcal{F}$, and $T_1, T_2, S \subseteq [k]$,

$$\langle \mathbf{u}_{f\chi_{T_1}}, \mathbf{u}_{g\chi_{T_2}} \rangle = \langle \mathbf{u}_{f\chi_{T_1\triangle S}}, \mathbf{u}_{g\chi_{T_2\triangle S}} \rangle.$$

4. For any $f \in \mathcal{F}$ and $\boldsymbol{\mu} \in_\eta \mathcal{F}$, let $g := f\boldsymbol{\mu}$. Then, for $N \geq \Omega\left(\frac{1}{\eta}\log\frac{1}{\eta}\right)$, $\langle \mathbf{u}_f, \mathbf{u}_g \rangle \geq 1 - 4\eta$ with probability at-least $1 - \eta$.

5. See the discussion after Theorem 9.4.

■

Theorem 9.4 *[11, Theorem 8.2.5] Let $\langle G, \circ, \mathrm{id}\rangle$ be an abelian group of characteristic 2, i.e., $g \circ g = \mathrm{id}$ for all $g \in G$. Let $d(g, h)$ be a distance function[5] on G such that $d(g, h)$ depends only on $g \circ h$. Then, d is of negative type if and only if it is ℓ_1-embeddable.*

We view $\mathcal{F} = \{-1, 1\}^N$ as a group (denoted G) with the multiplicative operator $\circ$ being the point-wise multiplication between the boolean functions. The identity element id is the boolean function which is 1 at all points. The set $H := \{\chi_S\}_{S\subseteq[k]}$ forms a normal sub-group of G. The group and the distance we apply Theorem 9.4 on will be G/H and $d_t(\cdot, \cdot)$ respectively. It is easy to check that d_t is well defined. Further, G, and hence, H, has characteristic 2. We need to show that for $f, g \in G$, $d_t(fH, gH)$ depends only on fg. Since both $\frac{1}{\sqrt{N}}\sum_{S\subseteq[k]}\mathbf{u}_{f\chi_S}^{\otimes t}$ and $\frac{1}{\sqrt{N}}\sum_{T\subseteq[k]}\mathbf{u}_{g\chi_T}^{\otimes t}$ are unit vectors, it is sufficient to show that their inner product depends only on fg.

$$\begin{aligned}
\left\langle \frac{1}{\sqrt{N}}\sum_{S\subseteq[k]}\mathbf{u}_{f\chi_S}^{\otimes t}, \frac{1}{\sqrt{N}}\sum_{T\subseteq[k]}\mathbf{u}_{g\chi_T}^{\otimes t}\right\rangle &= \frac{1}{N}\sum_{S,T\subseteq[k]}\langle \mathbf{u}_{f\chi_S}, \mathbf{u}_{g\chi_T}\rangle^t \\
&= \frac{1}{N}\sum_{S,T\subseteq[k]}\left(\sum_{\mathbf{x}}\frac{f(\mathbf{x})g(\mathbf{x})\chi_{S\triangle T}(\mathbf{x})}{N}\right)^t \\
&= \frac{1}{N}\sum_{S,T\subseteq[k]}\left(\sum_{\mathbf{x}}\frac{(fg)(\mathbf{x})\chi_{S\triangle T}(\mathbf{x})}{N}\right)^t
\end{aligned}$$

Hence, Theorem 9.4 implies that d_t is ℓ_1-embaddable, and hence, defines a negative type metric.

[5]In general, a distance function may not satisfy the triangle inequality.

10 Integrality Gap Instances for Cut Problems

In this section, we describe the integrality gap instances for Balanced Separator, Maximum Cut and Minimum Uncut, along with their SDP solutions. These are constructed from the integrality gap instance of Unique Games, and its SDP solution (abstracted out in Section 9.3). Each construction essentially mimics the PCP reduction from Unique Games to the respective cut problem. These PCP reductions (in particular, the PCP soundness analysis that we need) are presented separately in Section 11.

10.1 The Graphs

The starting point of all the constructions is the (NO) instance of Unique Games described in Theorem 9.2. We recall the following notations which will be needed. For a permutation $\pi : [N] \mapsto [N]$ and a vector $\mathbf{x} \in \{-1,1\}^N$, the vector $\mathbf{x} \circ \pi$ is defined to be the vector with its j-th entry as $(\mathbf{x} \circ \pi)_j := \mathbf{x}_{\pi(j)}$. For $\varepsilon > 0$, the notation $\mathbf{x} \in_\varepsilon \{-1,1\}^N$ means that the vector $\mathbf{x}$ is a random $\{-1,1\}^N$ vector, with each of its bits independently set to -1 with probability ε, and set to 1 with probability $1-\varepsilon$. In this section we will also be using the notations from Section 8.2.

10.1.1 Balanced Separator

This instance has a parameter $\varepsilon > 0$ and we refer to it as $\mathcal{I}_\varepsilon^{\mathsf{BS}}(V^*, E^*)$. We start with the Unique Games instance $\mathcal{U}_\eta\left(G(V,E),[N],\{\pi_e\}_{e\in E},\mathbf{wt}\right)$ of Theorem 9.2. In $\mathcal{I}_\varepsilon^{\mathsf{BS}}$, each vertex $v \in V$ is replaced by a *block* of vertices denoted by $V^*[v]$. This block consists of vertices $(v,\mathbf{x})$ for each $\mathbf{x} \in \{-1,1\}^N$. Thus, the set of vertices for the Balanced Separator instance is

$$V^* := \{(v,\mathbf{x}) \mid v \in V,\ \mathbf{x} \in \{-1,1\}^N\} \qquad \text{and} \qquad V^* = \cup_{v\in V} V^*[v].$$

The edges in the Balanced Separator instance are defined as follows: For $e\{v,w\} \in E$, there is an edge e^* in $\mathcal{I}_\varepsilon^{\mathsf{BS}}$ between $(v,\mathbf{x})$ and $(w,\mathbf{y})$, with weight

$$\mathbf{wt}_{\mathsf{BS}}(e^*) := \mathbf{wt}(e) \cdot \mathbf{Pr}_{\substack{\mathbf{x}' \in_{1/2}\{-1,1\}^N \\ \boldsymbol{\mu} \in_\varepsilon \{-1,1\}^N}} \left[\left(\mathbf{x} = \mathbf{x}'\right) \wedge \left(\mathbf{y} = \mathbf{x}'\boldsymbol{\mu} \circ \pi_e^v\right)\right].$$

Notice that the size of $\mathcal{I}_\varepsilon^{\mathsf{BS}}$ is $|V^*| = |V| \cdot 2^N = n/\log n \cdot n = n^2/\log n$.

Theorem 10.1 *For every $t \in (\frac{1}{2}, 1)$, there exists a constant $c_t > 0$ such that the following holds: Let $\varepsilon > 0$ be sufficiently small and let $\mathcal{U}_\eta\left(G(V,E),[N],\{\pi_e\}_{e\in E},\mathbf{wt}\right)$ be an instance of* Unique Games *with $\mathsf{OPT}(\mathcal{U}_\eta) < 2^{-O(1/\varepsilon^2)}$. Let $\mathcal{I}_\varepsilon^{\mathsf{BS}}$ be the corresponding instance of* Balanced Separator *as defined above. Let $V^* = \cup_{v\in V} V^*[v]$ be the partition of its vertices as above. Then, any $\frac{5}{6}$-piecewise balanced cut $(A,\overline{A})$ in $\mathcal{I}_\varepsilon^{\mathsf{BS}}$ (in the sense of Definition 7.1) satisfies*

$$\sum_{e^* \in E^*(A,\overline{A})} \mathbf{wt}_{\mathsf{BS}}(e^*) \geq c_t \varepsilon^t.$$

This theorem is a direct corollary of Theorem 11.2. See Section 11.1 for more details.

10.1.2 Maximum Cut

This instance has a parameter $\rho \in (-1, 0)$ and we refer to it as $\mathcal{I}_\rho^{\mathsf{MC}}(V^*, E^*)$. We start with the Unique Games instance $\mathcal{U}_\eta\left(G(V,E), [N], \{\pi_e\}_{e\in E}, \mathbf{wt}\right)$ of Theorem 9.2. In $\mathcal{I}_\rho^{\mathsf{MC}}$, each vertex in $v \in V$ is replaced by a *block* of vertices $(v, \mathbf{x})$, where $\mathbf{x}$ varies over $\{-1,1\}^N$. For every pair of edges $e\{v,w\}, e'\{v,w'\} \in E$, there is an edge e^* in $\mathcal{I}_\rho^{\mathsf{MC}}$ between $(w, \mathbf{x})$ and $(w', \mathbf{y})$, with weight

$$\mathbf{wt}_{\mathsf{MC}}(e^*) := (\mathbf{wt}(e)\Psi_v(e')) \cdot \mathbf{Pr}_{\substack{\mathbf{x}'\in_{1/2}\{-1,1\}^N \\ \boldsymbol{\mu}\in_{\frac{1-\rho}{2}}\{-1,1\}^N}} \left[\left(\mathbf{x} = \mathbf{x}' \circ \pi_e^v\right) \wedge \left(\mathbf{y} = \mathbf{x}'\boldsymbol{\mu} \circ \pi_{e'}^v\right)\right].$$

Notice that the size of $\mathcal{I}_\rho^{\mathsf{MC}}$ is $n^2/\log n$.

Theorem 10.2 *For any constants $\rho \in (-1,0)$ and $\lambda > 0$, there is a constant $c(\rho,\lambda)$ such that the following holds: Let $\mathcal{U}_\eta\left(G(V,E), [N], \{\pi_e\}_{e\in E}, \mathbf{wt}\right)$ be an instance of* Unique Games *with $\mathsf{OPT}(\mathcal{U}_\eta) < c(\rho,\lambda)$. Let $\mathcal{I}_\rho^{\mathsf{MC}}$ be the corresponding instance of* Maximum Cut *as defined above. Then, any cut $(A, \overline{A})$ in $\mathcal{I}_\rho^{\mathsf{MC}}$ satisfies*

$$\sum_{e^*\in E^*(A,\overline{A})} \mathbf{wt}_{\mathsf{MC}}(e^*) \le \frac{1}{\pi}\arccos\rho + \lambda.$$

This theorem essentially follows from the results of [20] and [27]. See Section 11.2 for more details.

10.1.3 Minimum Uncut

This instance has a parameter $\varepsilon \in (0,1)$ and we refer to it as $\mathcal{I}_\varepsilon^{\mathsf{MUC}}(V^*, E^*)$. This is the same as $\mathcal{I}_{-1+2\varepsilon}^{\mathsf{MC}}$. More precisely, we start with the Unique Games instance $\mathcal{U}_\eta\left(G(V,E), [N], \{\pi_e\}_{e\in E}, \mathbf{wt}\right)$ of Theorem 9.2. In $\mathcal{I}_\varepsilon^{\mathsf{MUC}}$, each vertex in $v \in V$ is replaced by a *block* of vertices $(v, \mathbf{x})$, where $\mathbf{x}$ varies over $\{-1,1\}^N$. For every pair of edges $e\{v,w\}, e'\{v,w'\} \in E$, there is an edge e^* in $\mathcal{I}_\varepsilon^{\mathsf{MUC}}$ between $(w, \mathbf{x})$ and $(w', \mathbf{y})$, with weight

$$\mathbf{wt}_{\mathsf{MUC}}(e^*) := (\mathbf{wt}(e)\Psi_v(e')) \cdot \mathbf{Pr}_{\substack{\mathbf{x}'\in_{1/2}\{-1,1\}^N \\ \boldsymbol{\mu}\in_{1-\varepsilon}\{-1,1\}^N}} \left[\left(\mathbf{x} = \mathbf{x}' \circ \pi_e^v\right) \wedge \left(\mathbf{y} = \mathbf{x}'\boldsymbol{\mu} \circ \pi_{e'}^v\right)\right].$$

Notice that the size of $\mathcal{I}_\varepsilon^{\mathsf{MUC}}$ is $n^2/\log n$.

Theorem 10.3 *For every $t \in (\frac{1}{2}, 1)$, there exists a constant $c_t > 0$ such that the following holds: Let $\varepsilon > 0$ be sufficiently small and let $\mathcal{U}_\eta\left(G(V,E), [N], \{\pi_e\}_{e\in E}, \mathbf{wt}\right)$ be an instance of* Unique Games *with $\mathsf{OPT}(\mathcal{U}_\eta) < 2^{-O(1/\varepsilon^2)}$. Let $\mathcal{I}_\varepsilon^{\mathsf{MUC}}$ be an instance of* Minimum Uncut *as defined above. Then, every cut in $\mathcal{I}_\varepsilon^{\mathsf{MUC}}$ has weight at most $1 - c_t\varepsilon^t$. In other words, every set of edges, removing which leaves $I_\varepsilon^{\mathsf{MUC}}$ bipartite, is of weight at-least $c_t\varepsilon^t$.*

This theorem is a direct corollary of Theorem 11.8, which is implicit in [19]. See Section 11.3 for more details.

10.2 SDP Solutions

Now we describe the SDP solutions for the Balanced Separator, Maximum Cut, and Minimum Uncut instances in Theorems 10.1, 10.2, and 10.3 respectively. Recall that $[N]$ is identified with $2^{[k]}$. Thus, $\mathbf{x} \in \{-1,1\}^N$ can be interpreted as a $\{+1,-1\}$-assignment to sets $T \subseteq [k]$.

Definition 10.4 *For $\mathbf{x} \in \{-1,1\}^N$, and $T \subseteq [k]$, let $\mathbf{x}(T)$ denote the j^{th} co-ordinate of $\mathbf{x}$ where $j \in [N]$ is the index identified with the set T.*

Definition 10.5 *For $\mathbf{x} \in \{-1,1\}^N$, and a bijection $\pi : 2^{[k]} \mapsto 2^{[k]}$, $\mathbf{x} \circ \pi$ is defined as*

$$(\mathbf{x} \circ \pi)(T) = \mathbf{x}(\pi(T)).$$

Notice that this definition is slightly abused: Previously, for $\mathbf{x} \in \{-1,1\}^N$, and $\pi : [N] \mapsto [N]$, $\mathbf{x} \circ \pi$ was defined to be the vector whose j-th co-ordinate is $\mathbf{x}_{\pi(j)}$. Upto the identification of $j \in [N]$ and $T \subseteq [k]$, this is consistent, and it will be clear from the context which definition we are referring to.

We start with the SDP solution of Theorem 9.3. Consider the following unit vectors, one for every pair $(v, \mathbf{x})$, where $v \in V$ and $\mathbf{x} \in \{-1,1\}^N$ (note that V is the set of vertices of the Unique Games instance of Theorem 9.2). Also, s is a positive integer (to be taken to be 4 in the rest of the paper).

$$\mathbf{V}_{v,s,\mathbf{x}} := \frac{1}{\sqrt{N}} \sum_{T \subseteq [k]} \mathbf{x}(T) \mathbf{v}_T^{\otimes 2s}. \tag{28}$$

Further, for $v \in V$, and an integer $t > 0$, consider the vector

$$\overline{\mathbf{V}}_{v,t} := \frac{1}{\sqrt{N}} \sum_{T \subseteq [k]} \mathbf{v}_T^{\otimes t}. \tag{29}$$

These vectors will be used as SDP solutions later on. In this section, we prove some of their generic properties.

Fact 10.6 *For every $v \in V$, $\mathbf{x} \in \{-1,1\}^N$ and every integer $s > 0$, $\|\mathbf{V}_{v,s,\mathbf{x}}\| = 1$. Further, for all integers $t > 0$, $\|\overline{\mathbf{V}}_{v,t}\| = 1$.*

Lemma 10.7 (Well Separated Lemma) *For any integer $s > 0$, and any odd integer $t > 0$,*

$$\frac{1}{2} \mathbf{E}_{\mathbf{x} \in_{1/2} \{-1,1\}^N,\ \mathbf{y} \in_{1/2} \{-1,1\}^N} \left[\|\mathbf{V}_{v,s,\mathbf{x}}^{\otimes t} - \mathbf{V}_{v,s,\mathbf{y}}^{\otimes t}\|^2 \right] = 1.$$

Lemma 10.8 *Let $0 < \eta < 1/2$, and assume that for $u, v \in V$ and $S, S' \subseteq [k]$, $\langle \mathbf{u}_S, \mathbf{v}_{S'} \rangle = 1 - \eta$. Let $\pi : 2^{[k]} \mapsto 2^{[k]}$ be defined as $\pi(S' \Delta U) := S \Delta U,\ \forall\ U \subseteq [k]$. Then*

- **Lower Bound:** $(1-\eta)^{2s}(1 - 2\mathbf{rdist}(\mathbf{x} \circ \pi, \mathbf{y})) - (2\eta)^s \leq \langle \mathbf{V}_{u,s,\mathbf{x}}, \mathbf{V}_{v,s,\mathbf{y}} \rangle$.
- **Upper Bound:** $\langle \mathbf{V}_{u,s,\mathbf{x}}, \mathbf{V}_{v,s,\mathbf{y}} \rangle \leq (1-\eta)^{2s}(1 - 2\mathbf{rdist}(\mathbf{x} \circ \pi, \mathbf{y})) + (2\eta)^s$.

Here, $\mathbf{rdist}(\mathbf{x}, \mathbf{y})$ denotes the fraction of points where $\mathbf{x}$ and $\mathbf{y}$ differ.

Corollary 10.9 *Let $0 < \eta < 1/2$, and assume that for $u, v \in V$ and $S, S' \subseteq [k]$, $\langle \mathbf{u}_S, \mathbf{v}_{S'} \rangle = 1 - \eta$. Then*

$$(1-\eta)^t - (2\eta)^{t/2} \leq \left\langle \overline{\mathbf{V}}_{u,t}, \overline{\mathbf{V}}_{v,t} \right\rangle \leq (1-\eta)^t + (2\eta)^{t/2}.$$

Theorem 10.10 *For $s = 4$ and $t = 2^{240} + 1$, the set of vectors $\{\mathbf{V}_{v,s,\mathbf{x}}^{\otimes t}\}_{v \in V, \mathbf{x} \in \{-1,1\}^N}$ give rise to a negative-type metric.*

Theorem 10.11 *For $s = 4$ and $t = 2^{241}$, the set of vectors $\{\mathbf{V}_{v,s,\mathbf{x}} \otimes \overline{\mathbf{V}}_{v,t}\}_{v \in V, \mathbf{x} \in \{-1,1\}^N}$ give rise to a negative-type metric.*

We differ the extremely technical proofs of these two theorems to Section 10.4.

10.2.1 Solution for the SDP of $\mathcal{I}_\varepsilon^{\mathsf{BS}}$

Now we present an SDP solution for $\mathcal{I}_\varepsilon^{\mathsf{BS}}(V^*, E^*, \mathbf{wt}_{\mathsf{BS}})$ that satisfies Properties (3), (4) and (5) of Theorem 7.2. This suffices to prove Theorem 7.3. For $(v, \mathbf{x}) \in V^*$, we associate the vector $\mathbf{V}_{v,4,\mathbf{x}}^{\otimes t}$, where $t = 2^{240} + 1$. Property (3) follows from Theorem 10.10, while Property (4) follows from Lemma 10.7. Property (5) follows from the following theorem.

Theorem 10.12 $\sum_{e^*=\{(v,\mathbf{x}),(w,\mathbf{y})\}\in E^*} \mathbf{wt}_{\mathsf{BS}}(e^*)\|\mathbf{V}_{v,4,\mathbf{x}}^{\otimes t} - \mathbf{V}_{w,4,\mathbf{y}}^{\otimes t}\|^2 \leq O(\eta + \varepsilon)$.

Proof: It is sufficient to prove that for an edge $e\{v, w\} \in E$ picked with probability $\mathbf{wt}(e)$ (from the UNIQUE GAMES instance $\mathcal{U}_\eta$), $\mathbf{x} \in_{1/2} \{-1, 1\}^N$, and $\boldsymbol{\mu} \in_\varepsilon \{-1, 1\}^N$,

$$\mathbf{E}_{e\{v,w\}}\left[\mathbf{E}_{\substack{\mathbf{x}\in_{1/2}\{-1,1\}^N\\ \boldsymbol{\mu}\in_\varepsilon\{-1,1\}^N}}\left\langle \mathbf{V}_{v,4,\mathbf{x}}^{\otimes t}, \mathbf{V}_{w,4,\mathbf{x}\boldsymbol{\mu}\circ\pi_e^v}^{\otimes t}\right\rangle\right] \geq 1 - O(t(\eta+\varepsilon)).$$

Since $e\{v, w\}$ is a random edge of $\mathcal{U}_\eta$, we know from the Closeness Property of Theorem 9.3, that w.h.p., there are $S, S' \subseteq [k]$ such that $\langle \mathbf{v}_S, \mathbf{w}_{S'}\rangle \geq 1-O(\eta)$. Moreover, $\pi_e^v(S'\Delta U) = S\Delta U$, $\forall\, U \subseteq [k]$. Further, using Chernoff Bounds, we see that except with probability ε, $\mathbf{rdist}(\mathbf{x}, \mathbf{x}\boldsymbol{\mu}) \leq 2\varepsilon$. Thus, using the lower bound estimate from Lemma 10.8, we get that

$$\left\langle \mathbf{V}_{v,4,\mathbf{x}}^{\otimes t}, \mathbf{V}_{w,4,\mathbf{x}\boldsymbol{\mu}\circ\pi_e^v}^{\otimes t}\right\rangle \geq 1 - O(t(\eta+\varepsilon)).$$

This completes the proof. ■

Using Theorem 10.12 along with Theorem 10.1, for the choices $\varepsilon = (\log\log n)^{-1/3}$, $\eta = O(\varepsilon)$, one gets the following theorem (note that $\mathsf{OPT}(\mathcal{U}_\eta) \leq \log^{-\eta} n$).

Theorem 10.13 *Non-uniform versions of* SPARSEST CUT *and* BALANCED SEPARATOR *have an integrality gap of at-least* $(\log\log n)^{1/6-\delta}$, *where* $\delta > 0$ *is arbitrary. The integrality gaps hold for standard* SDP*s with triangle inequality constraints.*

10.2.2 Solution for the SDP of $\mathcal{I}_\rho^{\mathsf{MC}}$

Now we present an SDP solution for $\mathcal{I}_\rho^{\mathsf{MC}}(V^*, E^*, \mathbf{wt}_{\mathsf{MC}})$ with an objective of at-least $\frac{1-\rho}{2} - O(\eta)$. For $(w, \mathbf{x}) \in V^*$, we associate the vector $\mathbf{V}_{w,4,\mathbf{x}} \otimes \overline{\mathbf{V}}_{w,t}$. The following theorem shows that this SDP solution to $\mathcal{I}_\rho^{\mathsf{MC}}$ has an objective at-least $\frac{1-\rho}{2} - O(\eta)$.

Theorem 10.14

$$\frac{1}{4}\sum_{e^*=\{(w,\mathbf{x}),(w',\mathbf{y})\}\in E^*} \mathbf{wt}_{\mathsf{MC}}(e^*)\|\mathbf{V}_{w,4,\mathbf{x}} \otimes \overline{\mathbf{V}}_{w,t} - \mathbf{V}_{w',4,\mathbf{y}} \otimes \overline{\mathbf{V}}_{w',t}\|^2 \geq \frac{1-\rho}{2} - O(\eta).$$

As a corollary to Theorem 10.2 and Theorem 10.14, we obtain an integrality gap of $\frac{\frac{1}{\pi}\arccos\rho+\lambda}{\frac{1-\rho}{2}-O(\eta)}$ for MAXIMUM CUT. Choosing the "critical" choice of ρ, and letting $\lambda, \eta \to 0$, we get the following theorem.

Theorem 10.15 *Let α_{GW} (≈ 0.878) be the approximation ratio obtained by Goemans-Williamson's algorithm for* Maximum Cut*. For every $\delta > 0$, the Goemans-Williamson's* SDP *has an integrality gap of at-least $\alpha_{\mathsf{GW}} + \delta$, even after including the triangle inequality constraints.*

Proof: [of Theorem 10.14] It is sufficient to prove that for an edge $e\{v,w\} \in E$ picked with probability $\mathbf{wt}(e)$, $e'\{v,w'\} \in \Gamma(v)$ picked with probability $\Psi_v(e')$, $\mathbf{x} \in_{1/2} \{-1,1\}^N$, and $\boldsymbol{\mu} \in_{\frac{1-\rho}{2}} \{-1,1\}^N$,

$$\mathbf{E}_{e\{v,w\},e'\{v,w'\}}\left[\mathbf{E}_{\substack{\mathbf{x}\in_{1/2}\{-1,1\}^N \\ \boldsymbol{\mu}\in_{\frac{1-\rho}{2}}\{-1,1\}^N}}\left\langle \mathbf{V}_{w,4,\mathbf{x}\circ\pi_e^v} \otimes \overline{\mathbf{V}}_{w,t}, \mathbf{V}_{w',4,\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^v} \otimes \overline{\mathbf{V}}_{w',t}\right\rangle\right] \leq \rho + O(\eta).$$

Notice that

$$\left\langle \mathbf{V}_{w,4,\mathbf{x}\circ\pi_e^v} \otimes \overline{\mathbf{V}}_{w,t}, \mathbf{V}_{w',4,\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^v} \otimes \overline{\mathbf{V}}_{w',t}\right\rangle = \left\langle \mathbf{V}_{w,4,\mathbf{x}\circ\pi_e^v}, \mathbf{V}_{w',4,\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^v}\right\rangle \left\langle \overline{\mathbf{V}}_{w,t}, \overline{\mathbf{V}}_{w',t}\right\rangle.$$

Since $e\{v,w\}$ and $e'\{v,w'\}$ are random (but not independent) edges of $\mathcal{U}_\eta$, we know from the Closeness Property, that w.h.p., there are $S,T,T' \subseteq [k]$ such that

$$\langle \mathbf{v}_S, \mathbf{w}_T\rangle \geq 1 - O(\eta), \quad \pi_e^v(T\Delta U) = S\Delta U, \ \forall\ U \subseteq [k]$$

$$\langle \mathbf{v}_S, \mathbf{w}'_{T'}\rangle \geq 1 - O(\eta), \quad \pi_{e'}^v(T'\Delta U) = S\Delta U, \ \forall\ U \subseteq [k].$$

This implies (via the triangle inequality) that $\langle \mathbf{w}_T, \mathbf{w}'_{T'}\rangle \geq 1 - O(\eta)$. Further, using Chernoff Bounds, we see that w.h.p., we have $\frac{1-\rho}{2} - \eta \leq \mathbf{rdist}(\mathbf{x}, \mathbf{x}\boldsymbol{\mu}) \leq \frac{1-\rho}{2} + \eta$. Now we use the estimate from Lemma 10.8 for $\langle \mathbf{V}_{w,4,\mathbf{x}\circ\pi_e^v}, \mathbf{V}_{w',4,\mathbf{x}\boldsymbol{\mu}\circ\pi_{e'}^v}\rangle$ and for $\langle \overline{\mathbf{V}}_{w,t}, \overline{\mathbf{V}}_{w',t}\rangle$. The former dot-product is $\rho \pm O(\eta)$ and the latter is $1 - O(\eta)$. ∎

10.2.3 Solution for the SDP of $\mathcal{I}_\varepsilon^{\mathsf{MUC}}$

Now we present an SDP solution for $\mathcal{I}_\varepsilon^{\mathsf{MUC}}(V^*, E^*, \mathbf{wt}_{\mathsf{MUC}})$ with an objective of at-most $O(\eta + \varepsilon)$. Set $t = 2^{240} + 1$. For $(w, \mathbf{x}) \in V^*$, we associate the vector $\mathbf{V}_{w,4,\mathbf{x}}^{\otimes t}$. The triangle inequality follows from Theorem 10.10. The following theorem shows that the SDP solution to $\mathcal{I}_\varepsilon^{\mathsf{MUC}}$ has a small objective.

Theorem 10.16

$$1 - \frac{1}{4} \sum_{e^*=\{(w,\mathbf{x}),(w',\mathbf{y})\}\in E^*} \mathbf{wt}_{\mathsf{MUC}}(e^*) \|\mathbf{V}_{w,4,\mathbf{x}}^{\otimes t} - \mathbf{V}_{w',4,\mathbf{y}}^{\otimes t}\|^2 \leq O(\eta + \varepsilon).$$

The proof of this theorem is similar to the one for Maximum Cut.

Using Theorem 10.16 along with Theorem 10.3, for the choices $\eta = O(\varepsilon)$ and $\varepsilon = (\log\log n)^{-1/3}$, one gets the following theorem (note that $\mathsf{OPT}(\mathcal{U}_\eta) \leq \log^{-\eta} n$).

Theorem 10.17 Minimum Uncut *has an integrality gap of at-least $(\log\log n)^{1/6-\delta}$, where $\delta > 0$ is arbitrary. The integrality gap holds for the standard* SDP *with the triangle inequality constraints.*

10.3 Proofs of Basic Properties of the SDP Solutions

Proof: [of Lemma 10.7]

$$\begin{aligned}
\frac{1}{2}\mathbf{E}_{\mathbf{x},\mathbf{y}}\left[\|\mathbf{V}_{v,s,\mathbf{x}}^{\otimes t}-\mathbf{V}_{v,s,\mathbf{y}}^{\otimes t}\|^2\right] &= \mathbf{E}_{\mathbf{x},\mathbf{y}}\left[1-\langle \mathbf{V}_{v,s,\mathbf{x}}^{\otimes t},\mathbf{V}_{v,s,\mathbf{y}}^{\otimes t}\rangle\right] \\
&= 1-\mathbf{E}_{\mathbf{x},\mathbf{y}}\left[\left(\frac{1}{N}\sum_{S,T\subseteq[k]}\mathbf{x}(T)\mathbf{y}(T)\langle \mathbf{v}_S,\mathbf{v}_T\rangle^{2s}\right)^t\right] \\
&= 1.
\end{aligned}$$

The last equality follows from the fact that the contribution of $(\mathbf{x},\mathbf{y})$ to the expectation is cancelled by that of $(\mathbf{x},-\mathbf{y})$. ■

Proof: [of Lemma 10.8]

$$\begin{aligned}
\langle \mathbf{V}_{u,s,\mathbf{x}},\mathbf{V}_{v,s,\mathbf{y}}\rangle &= \frac{1}{N}\sum_{T,T'\subseteq[k]}\mathbf{x}(T)\mathbf{y}(T')\langle \mathbf{u}_T,\mathbf{v}_{T'}\rangle^{2s} \\
&= \frac{1}{N}\sum_{T,T'\subseteq[k]}\mathbf{x}(S\triangle T)\mathbf{y}(S'\triangle T')\langle \mathbf{u}_{S\triangle T},\mathbf{v}_{S'\triangle T'}\rangle^{2s}.
\end{aligned}$$

We first show that in the above summation, terms with $T=T'$ dominate the summation. Since $\langle \mathbf{u}_S,\mathbf{v}_{S'}\rangle = 1-\eta$, the Matching Property implies that for all $T\subseteq[k]$, $\langle \mathbf{u}_{S\triangle T},\mathbf{v}_{S'\triangle T}\rangle = 1-\eta$. Further, since the vectors $\{\mathbf{v}_{T'}\}_{T'\subseteq[k]}$ form an orthonormal basis for $\mathbb{R}^N$, $\sum_{T'\subseteq[k]}\langle \mathbf{u}_{S\triangle T},\mathbf{v}_{S'\triangle T'}\rangle^2 = 1$. Hence,

$$\sum_{T'\subseteq[k],T'\neq T}\langle \mathbf{u}_{S\triangle T},\mathbf{v}_{S'\triangle T'}\rangle^{2s} \leq \left(1-(1-\eta)^2\right)^s \leq (2\eta-2\eta^2)^s = (2\eta)^s.$$

Now, $\langle \mathbf{V}_{u,s,\mathbf{x}},\mathbf{V}_{v,s,\mathbf{y}}\rangle$ is at-least

$$\frac{1}{N}\sum_{T\subseteq[k]}\mathbf{x}(S\triangle T)\mathbf{y}(S'\triangle T)\langle \mathbf{u}_{S\triangle T},\mathbf{v}_{S'\triangle T}\rangle^{2s} - \frac{1}{N}\sum_{\substack{T,T'\subseteq[k]\\T\neq T'}}\langle \mathbf{u}_{S\triangle T},\mathbf{v}_{S'\triangle T'}\rangle^{2s},$$

and at-most

$$\frac{1}{N}\sum_{T\subseteq[k]}\mathbf{x}(S\triangle T)\mathbf{y}(S'\triangle T)\langle \mathbf{u}_{S\triangle T},\mathbf{v}_{S'\triangle T}\rangle^{2s} + \frac{1}{N}\sum_{\substack{T,T'\subseteq[k]\\T\neq T'}}\langle \mathbf{u}_{S\triangle T},\mathbf{v}_{S'\triangle T'}\rangle^{2s}.$$

The first term in both these expressions is

$$\frac{1}{N}\sum_{T\subseteq[k]}\mathbf{x}(S\Delta T)\mathbf{y}(S'\Delta T)(1-2\eta)^{2s} = (1-2\mathbf{rdist}(\mathbf{x}\circ\pi,\mathbf{y}))(1-2\eta)^{2s}.$$

The second term is bounded by $(2\eta)^s$ as seen above. This completes the proof of the lemma. ■

10.4 Proving the Triangle Inequality

In this section, we prove Theorem 10.10 and Theorem 10.11. Unfortunately, the proof relies on heavy case-analysis and is not very illuminating.

10.4.1 Main Lemma

Lemma 10.18 *Let $\{\mathbf{u}_i\}_{i=1}^N, \{\mathbf{v}_i\}_{i=1}^N, \{\mathbf{w}_i\}_{i=1}^N$ be three sets of unit vectors in $\mathbb{R}^N$, such that the vectors in each set are mutually orthogonal. Assume that any three of these vectors satisfy the triangle inequality. Assume, moreover, that*

$$\langle \mathbf{u}_1, \mathbf{v}_1 \rangle = \langle \mathbf{u}_2, \mathbf{v}_2 \rangle = \cdots = \langle \mathbf{u}_N, \mathbf{v}_N \rangle, \tag{30}$$

$$\lambda := \langle \mathbf{u}_1, \mathbf{w}_1 \rangle = \langle \mathbf{u}_2, \mathbf{w}_2 \rangle = \cdots = \langle \mathbf{u}_N, \mathbf{w}_N \rangle \geq 0, \tag{31}$$

$$\forall 1 \leq i, j \leq N, \;\; |\langle \mathbf{u}_i, \mathbf{w}_j \rangle| \leq \lambda, \tag{32}$$

$$1 - \eta := \langle \mathbf{v}_1, \mathbf{w}_1 \rangle = \langle \mathbf{v}_2, \mathbf{w}_2 \rangle = \cdots = \langle \mathbf{v}_N, \mathbf{w}_N \rangle, \tag{33}$$

where $0 \leq \eta \leq 2^{-40s}$ and $s = 4$. Let $s_i, t_i, r_i \in \{-1, 1\}$ for $1 \leq i \leq N$. Define unit vectors

$$\mathbf{u} := \frac{1}{\sqrt{N}} \sum_{i=1}^N s_i \mathbf{u}_i^{\otimes 2s}, \quad \mathbf{v} := \frac{1}{\sqrt{N}} \sum_{i=1}^N t_i \mathbf{v}_i^{\otimes 2s} \quad \mathbf{w} := \frac{1}{\sqrt{N}} \sum_{i=1}^N r_i \mathbf{w}_i^{\otimes 2s}.$$

Then, the vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ satisfy the triangle inequality $1 + \langle \mathbf{u}, \mathbf{v} \rangle \geq \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$, i.e.,

$$N + \sum_{i,j=1}^N s_i t_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle^{2s} \geq \sum_{i,j=1}^N s_i r_j \langle \mathbf{u}_i, \mathbf{w}_j \rangle^{2s} + \sum_{i,j=1}^N t_i r_j \langle \mathbf{v}_i, \mathbf{w}_j \rangle^{2s}. \tag{34}$$

Proof: It suffices to show that for every $1 \leq j \leq N$,

$$1 + \sum_{i=1}^N s_i t_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle^{2s} \geq \sum_{i=1}^N s_i r_j \langle \mathbf{u}_i, \mathbf{w}_j \rangle^{2s} + t_j r_j \langle \mathbf{v}_j, \mathbf{w}_j \rangle^{2s} + \sum_{1 \leq i \leq N, i \neq j} \langle \mathbf{v}_i, \mathbf{w}_j \rangle^{2s}. \tag{35}$$

We consider four cases depending on value of λ.

(Case 1) $\lambda \leq \eta$: Since $\langle \mathbf{v}_j, \mathbf{w}_j \rangle = 1-\eta$, and $\sum_{1 \leq i \leq N} \langle \mathbf{v}_i, \mathbf{w}_j \rangle^2 = 1$, we have $\sum_{1 \leq i \leq N; i \neq j} \langle \mathbf{v}_i, \mathbf{w}_j \rangle^{2s} \leq (2\eta - \eta^2)^s$. Also, $\sum_{i=1}^N \langle \mathbf{u}_i, \mathbf{w}_j \rangle^{2s} \leq \lambda^{2s-2} \leq \eta^{2s-2}$. Moreover, for any $1 \leq i \leq N$, by the triangle inequality, $1 \pm \langle \mathbf{u}_i, \mathbf{v}_j \rangle \geq \langle \mathbf{v}_j, \mathbf{w}_j \rangle \pm \langle \mathbf{u}_i, \mathbf{w}_j \rangle \geq 1 - \eta - \lambda \geq 1 - 2\eta$, and therefore, $|\langle \mathbf{u}_i, \mathbf{v}_j \rangle| \leq 2\eta$. Therefore, $\sum_{i=1}^N \langle \mathbf{u}_i, \mathbf{v}_j \rangle^{2s} \leq (2\eta)^{2s-2}$. Thus, it suffices to prove that

$$1 \geq (2\eta)^{2s-2} + \eta^{2s-2} + (1-\eta)^{2s} + (2\eta - \eta^2)^s.$$

This is true when $\eta \leq 2^{-40s}$.

(Case 2) $\eta \leq \lambda \leq 1 - \sqrt{\eta}$: We will show that

$$1 + \sum_{i=1}^N s_i t_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle^{2s} \geq \sum_{i=1}^N s_i r_j \langle \mathbf{u}_i, \mathbf{w}_j \rangle^{2s} + t_j r_j (1-\eta)^{2s} + (2\eta - \eta^2)^s. \tag{36}$$

(Subcase i) $t_j \neq r_j$ **:** In this case it suffices to show that

$$1 + (1-\eta)^{2s} \geq \sum_{i=1}^{N} \langle \mathbf{u}_i, \mathbf{v}_j \rangle^{2s} + \sum_{i=1}^{N} \langle \mathbf{u}_i, \mathbf{w}_j \rangle^{2s} + (2\eta - \eta^2)^s.$$

Again, as before, we have that for every $1 \leq i \leq N$, $|\langle \mathbf{u}_i, \mathbf{w}_j \rangle| \leq \lambda \leq 1 - \sqrt{\eta}$, and $|\langle \mathbf{u}_i, \mathbf{v}_j \rangle| \leq \lambda + \eta \leq 1 - \sqrt{\eta} + \eta$. Thus, it suffices to prove that

$$1 + (1-\eta)^{2s} \geq (1 - \sqrt{\eta} + \eta)^{2s-2} + (1 - \sqrt{\eta})^{2s-2} + (2\eta - \eta^2)^s.$$

This also holds when $\eta \leq 2^{-40s}$.

(Subcase ii) $t_j = r_j$ **:** We need to prove (36). It suffices to show that

$$1 - (1-\eta)^{2s} - (2\eta - \eta^2)^s \;\geq\; \sum_{i=1}^{N} |\langle \mathbf{u}_i, \mathbf{w}_j \rangle|^{2s} - \langle \mathbf{u}_i, \mathbf{v}_j \rangle^{2s}| = \sum_{i=1}^{N} |\theta_i^{2s} - \mu_i^{2s}|$$

where $\theta_i := |\langle \mathbf{u}_i, \mathbf{w}_j \rangle|$, $\mu_i := |\langle \mathbf{u}_i, \mathbf{v}_j \rangle|$. Clearly,

$$|\theta_i - \mu_i| \leq |\langle \mathbf{u}_i, \mathbf{v}_j \rangle - \langle \mathbf{u}_i, \mathbf{w}_j \rangle| \leq 1 - \langle \mathbf{v}_i, \mathbf{w}_j \rangle = \eta.$$

Here, we used the assumption that $(\mathbf{u}_i, \mathbf{v}_j, \mathbf{w}_j)$ satisfy the triangle inequality. Note also that $\max_{1 \leq i \leq N} \theta_i = \lambda$ and $\sum_{i=1}^{N} \theta_i^2 = 1$. Let $J := \{i \mid \theta_i \leq \eta\}$ and $I := \{i \mid \theta_i \geq \eta\}$. We have,

$$\begin{aligned}
\sum_{i=1}^{N} |\theta_i^{2s} - \mu_i^{2s}| &\leq \sum_{i \in J} (\theta_i^{2s} + \mu_i^{2s}) + \sum_{i \in I} ((\theta_i + \eta)^{2s} - \theta_i^{2s}) \\
&\leq (\eta)^{2s-2} + (2\eta)^{2s-2} + \sum_{i \in I} ((\theta_i + \eta)^{2s} - \theta_i^{2s}).
\end{aligned}$$

Lemma 10.19 implies that the summation on the last line above is bounded by

$$\sum_{l=1}^{2s-2} \binom{2s}{l} \lambda^{2s-l-2} \eta^l + (2s+1)\eta^{2s-2}.$$

Thus, it suffices to show that

$$1 - (1-\eta)^{2s} - (2\eta - \eta^2)^s \;\geq\; \sum_{l=1}^{2s-2} \binom{2s}{l} \lambda^{2s-l-2} \eta^l + (4\eta)^{2s-2}.$$

This is true if

$$2s\eta - \sum_{l=2}^{2s} \binom{2s}{l} \eta^l - (2\eta - \eta^2)^s \geq 2s\lambda^{2s-3}\eta + \sum_{l=2}^{2s} \binom{2s}{l} \eta^l + (4\eta)^{2s-2}.$$

This is true if $2s\eta(1 - \lambda^{2s-3}) \geq \eta^2(2^{2s} + 2^{2s} + 1 + 4^{2s})$. This is true if $2s\eta\sqrt{\eta} \geq \eta^2 \cdot 4^{2s+1}$, which holds when $\eta \leq 2^{-40k}$. Note that we used the fact that $\lambda \leq 1 - \sqrt{\eta}$.

(Case 3) $1-\sqrt{\eta} \le \lambda \le 1-\eta^2$ **:** We have $\langle \mathbf{v}_j, \mathbf{w}_j \rangle = 1-\eta$, $\langle \mathbf{u}_j, \mathbf{w}_j \rangle = \lambda =: 1-\zeta$. This implies that $\langle \mathbf{u}_j, \mathbf{v}_j \rangle = 1-\delta$, where by the triangle inequality

$$\eta \le \zeta + \delta, \quad \delta \le \eta + \zeta, \quad \zeta \le \eta + \delta.$$

Thus, to prove (35), it suffices to show that

$$1 + s_j t_j \langle \mathbf{u}_j, \mathbf{v}_j \rangle^{2s} \ge s_j r_j \langle \mathbf{u}_j, \mathbf{w}_j \rangle^{2s} + t_j r_j \langle \mathbf{v}_j, \mathbf{w}_j \rangle^{2s} + (2\eta-\eta^2)^s + (2\zeta-\zeta^2)^s + (2\delta-\delta^2)^s.$$

Depending on signs s_j, t_j, r_j, this reduces to proving one of the three cases:

$$1 + (1-\delta)^{2s} \ge (1-\zeta)^{2s} + (1-\eta)^{2s} + (2\eta-\eta^2)^s + (2\zeta-\zeta^2)^s + (2\delta-\delta^2)^s.$$

$$1 + (1-\eta)^{2s} \ge (1-\zeta)^{2s} + (1-\delta)^{2s} + (2\eta-\eta^2)^s + (2\zeta-\zeta^2)^s + (2\delta-\delta^2)^s.$$

$$1 + (1-\zeta)^{2s} \ge (1-\eta)^{2s} + (1-\delta)^{2s} + (2\eta-\eta^2)^s + (2\zeta-\zeta^2)^s + (2\delta-\delta^2)^s.$$

We will prove the first case, and the remaining two are proved in a similar fashion. We have that $1 + (1-\delta)^{2s} - (1-\zeta)^{2s} - (1-\eta)^{2s}$

$$\begin{aligned}
&\ge 1 + (1-(\zeta+\eta))^{2s} - (1-\zeta)^{2s} - (1-\eta)^{2s} \\
&\ge 2s(2s-1)\cdot\zeta\eta - \sum_{\substack{3\le i+j\le 2s \\ i\ge 1, j\ge 1}} \binom{2s}{i+j}\binom{i+j}{i}\zeta^i\eta^j \\
&\ge 2s(2s-1)\zeta\eta - \ 2^{8s}\zeta\eta\cdot\max\{\zeta,\eta,\delta\} \\
&\ge \min\{\zeta\eta, \eta\delta, \zeta\delta\},
\end{aligned}$$

provided that $2^{8s}\max\{\zeta,\eta,\delta\} \le 1$. Thus, it suffices to have

$$\min\{\zeta\eta, \eta\delta, \zeta\delta\} \ge (2\eta-\eta^2)^s + (2\zeta-\zeta^2)^s + (2\delta-\delta^2)^s.$$

This is clearly true, since, ζ, η, δ are within squares (or square-roots) of each other, and $\eta \le 2^{-40s}$.

(Case 4) $1-\eta^2 \le \lambda$ **:** This is essentially same as Case (2). Just interchange $1-\eta$ with λ and interchange $\mathbf{u}_i, \mathbf{v}_i$ for every i. This completes the proof of the Main Lemma. ∎

Lemma 10.19 *Let η, λ and $\{\theta_i\}_{i=1}^N$ be non-negative reals, such that $\sum_{i=1}^N \theta_i^2 \le 1$, and for all i, $\eta \le \theta_i \le \lambda$. Then*

$$\sum_{i=1}^N ((\theta_i+\eta)^{2s} - \theta_i^{2s}) \le \sum_{l=1}^{2s-2}\binom{2s}{l}\lambda^{2s-l-2}\eta^l + (2s+1)\eta^{2s-2}.$$

Proof: Clearly, $N \le 1/\eta^2$.

$$\begin{aligned}
\sum_{i=1}^{N} (\theta_i+\eta)^{2s} - \theta_i^{2s} &= \sum_{i=1}^{N}\sum_{l=1}^{2s} \binom{2s}{l} \theta_i^{2s-l}\eta^l \\
&= \sum_{l=1}^{2s-2} \binom{2s}{l} \sum_{i=1}^{N} \theta_i^{2s-l}\eta^l + 2s\cdot\left(\sum_{i=1}^{N}\theta_i\right)\eta^{2s-1} + N\eta^{2s} \\
&\le \sum_{l=1}^{2s-2} \binom{2s}{l} \lambda^{2s-l-2}\eta^l + 2s\cdot\sqrt{N}\eta^{2s-1} + N\eta^{2s} \\
&\le \sum_{l=1}^{2s-2} \binom{2s}{l} \lambda^{2s-l-2}\eta^l + (2s+1)\eta^{2s-2}.
\end{aligned}$$

∎

Lemma 10.20 *Let $a, b, c \in [-1, 1]$ such that $1 + a \ge b + c$. Then, $1 + a^t \ge b^t + c^t$ for every odd integer $t \ge 1$.*

Proof: First, we notice that it is sufficient to prove this inequality when $0 \le a, b, c \le 1$.

Suppose that $b < 0$ and $c < 0$, then $b^t + c^t < 0 \le 1 + a^t$. Hence, without loss of generality assume that $b \ge 0$. If $c < 0$ and $a \ge 0$, then $b^t + c^t < b^t \le 1 + a^t$. If $c < 0$ and $a < 0$, by hypothesis, $1 - c \ge b - a$, which is the same as $1 + |c| \ge b + |a|$, and proving $1 + a^t \ge b^t + c^t$ is equivalent to proving $1 + |c|^t \ge b^t + |a|^t$. Hence, we may assume that $c \ge 0$. If $a < 0$, then $1 + a^t = 1 - |a|^t \ge 1 - |a| = 1 + a \ge b + c \ge b^t + c^t$. Hence, we may assume that $0 \le a, b, c \le 1$.

Further, we may assume that $a < b \le c$. Since, if $a \ge b$, then $1 + a^t \ge c^t + b^t$. $1 + a \ge b + c$ implies that $1 - c \ge b - a$. Notice that both sides of this inequality are positive. It follows from the fact that $0 \le a < b \le c \le 1$, that $\sum_{i=0}^{t-1} c^i \ge \sum_{i=0}^{t-1} a^i b^{t-1-i}$. Multiplying these two inequalities, we obtain $1 - c^t \ge b^t - a^t$, which implies that $1 + a^t \ge b^t + c^t$. This completes the proof.

∎

10.4.2 Proof of Theorem 10.10

We will prove that any three vectors $\mathbf{V}_{u,s,\mathbf{x}}^{\otimes t}$, $\mathbf{V}_{v,s,\mathbf{y}}^{\otimes t}$ and $\mathbf{V}_{w,s,\mathbf{z}}^{\otimes t}$ satisfy the triangle inequality. For $T \subseteq [k]$, let $r_T := \mathbf{x}(T), s_T := \mathbf{y}(T), t_T := \mathbf{z}(T)$, so that

$$\mathbf{V}_{u,s,\mathbf{x}}^{\otimes t} = \left(\frac{1}{\sqrt{N}} \sum_{T\subseteq[k]} r_T \mathbf{u}_T^{\otimes 2s}\right)^{\otimes t} = \text{(say)} \ \ \mathbf{W}_u^{\otimes t},$$

$$\mathbf{V}_{v,s,\mathbf{y}}^{\otimes t} = \left(\frac{1}{\sqrt{N}} \sum_{T\subseteq[k]} s_T \ \mathbf{v}_T^{\otimes 2s}\right)^{\otimes t} = \text{(say)} \ \ \mathbf{W}_v^{\otimes t},$$

$$\mathbf{V}_{w,s,\mathbf{z}}^{\otimes t} = \left(\frac{1}{\sqrt{N}} \sum_{T\subseteq[k]} t_T \ \mathbf{w}_T^{\otimes 2s}\right)^{\otimes t} = \text{(say)} \ \ \mathbf{W}_w^{\otimes t}.$$

We need to show that

$$1 + \langle \mathbf{V}_{v,s,\mathbf{y}}^{\otimes t}, \mathbf{V}_{w,s,\mathbf{z}}^{\otimes t} \rangle \geq \langle \mathbf{V}_{u,s,\mathbf{x}}^{\otimes t}, \mathbf{V}_{v,s,\mathbf{y}}^{\otimes t} \rangle + \langle \mathbf{V}_{u,s,\mathbf{x}}^{\otimes t}, \mathbf{V}_{w,s,\mathbf{z}}^{\otimes t} \rangle.$$

We can assume that at-least one of the dot-products has magnitude at-least 1/3; otherwise, the inequality trivially holds. Assume, w.l.o.g., that $|\langle \mathbf{V}_{u,s,\mathbf{x}}^{\otimes t}, \mathbf{V}_{w,s,\mathbf{z}}^{\otimes t} \rangle| \geq 1/3$. This implies that $|\langle \mathbf{W}_u, \mathbf{W}_w \rangle|^t \geq 1/3$, and therefore, $|\langle \mathbf{W}_u, \mathbf{W}_w \rangle| = 1 - \eta'$, for some $\eta' = O(1/t)$. Hence, for some $S, T \subseteq [k]$, $|\langle \mathbf{u}_S, \mathbf{w}_T \rangle| = 1 - \eta$ for some $\eta \leq 2^{-40s}$. By relabeling, if necessary, we may assume that $\langle \mathbf{u}_\emptyset, \mathbf{w}_\emptyset \rangle = 1 - \eta$.

Note that we need to show that

$$1 + \langle \mathbf{W}_v, \mathbf{W}_w \rangle^t \geq \langle \mathbf{W}_u, \mathbf{W}_v \rangle^t + \langle \mathbf{W}_u, \mathbf{W}_w \rangle^t.$$

By Lemma 10.20, it suffices to show that

$$1 + \langle \mathbf{W}_v, \mathbf{W}_w \rangle \geq \langle \mathbf{W}_u, \mathbf{W}_v \rangle + \langle \mathbf{W}_u, \mathbf{W}_w \rangle.$$

For notational convenience, we replace indexing by $T \subseteq [k]$ with indexing by $1 \leq i \leq N$ where $N = 2^k$, and $\emptyset$ is identified with 1. Thus, we write s_i, t_i, r_i instead of s_T, t_T, r_T. We also write $\mathbf{u}_i$ instead of $\mathbf{u}_T$, $\mathbf{v}_i$ instead of $\mathbf{v}_T$ and $\mathbf{w}_i$ instead of $\mathbf{w}_T$. Thus, we need to show that

$$N + \sum_{i,j=1}^{N} s_i t_j \langle \mathbf{v}_i, \mathbf{w}_j \rangle^{2s} \geq \sum_{i,j=1}^{N} s_i r_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle^{2s} + \sum_{i,j=1}^{N} t_i r_j \langle \mathbf{u}_i, \mathbf{w}_j \rangle^{2s}.$$

As noted before, we may assume that $\langle \mathbf{u}_1, \mathbf{w}_1 \rangle = 1 - \eta$, and hence, by the Matching Property,

$$\langle \mathbf{u}_1, \mathbf{w}_1 \rangle = \langle \mathbf{u}_2, \mathbf{w}_2 \rangle = \cdots = \langle \mathbf{u}_N, \mathbf{w}_N \rangle = 1 - \eta.$$

Let $\lambda := \max_{1 \leq i,j \leq N} |\langle \mathbf{u}_i, \mathbf{v}_j \rangle|$. We may assume, w.l.o.g., that the maximum is achieved for $\mathbf{u}_1, \mathbf{v}_1$, and again by the Matching Property,

$$\langle \mathbf{u}_1, \mathbf{v}_1 \rangle = \langle \mathbf{u}_2, \mathbf{v}_2 \rangle = \cdots = \langle \mathbf{u}_N, \mathbf{v}_N \rangle = \lambda.$$

Now, the desired inequality follows from Lemma 10.18.

10.4.3 Proof of Theorem 10.11

The proof is similar to that of Theorem 10.10 that appears in Section 10.4.2. Recall that $s = 4$, and $t = 2^{241}$. First we need a simple lemma.

Lemma 10.21 *Let $a, b, c \in [-1, 1]$ such that $1 + a \geq b + c$. Let $a', b', c' \in [0, 1]$ such that $1 + a' \geq b' + c'$, $1 + b' \geq a' + c'$ and $1 + c' \geq a' + b'$. Then, $1 + aa' \geq bb' + cc'$.*

Proof: For fixed $a, b, c \in [-1, 1]$, consider the following LP: It can be verified that the only vertices of the (bounded) polytope for the above LP are: $(x, y, z) = (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1)$ and $(1, 1, 1)$. Hence, the minima has to occur at one of these vertices. The objective function at these 5 points is:

- At $(0, 0, 0)$, it is 1.

$$\text{Minimize} \quad 1 + ax - by - cz \tag{37}$$

Subject to

$$\begin{aligned} 0 \le x, y, z &\le 1 \\ x - y - z &\ge -1 \\ -x + y - z &\ge -1 \\ -x - y + z &\ge -1 \end{aligned}$$

- At $(1, 0, 0)$, it is $1 + a$, which is at-least 0 as $a \ge -1$.
- At $(0, 1, 0)$, it is $1 - b$, which is at-least 0 as $b \le 1$.
- At $(0, 0, 1)$, it is $1 - c$, which is at-least 0 as $c \le 1$.
- At $(1, 1, 1)$, it is $1 + a - b - c$, which is at-least 0 by hypothesis.

This shows that the objective is at-least 0, hence, $1 + ax - by - cz \ge 0$ for all feasible points. Since a', b', c' form a feasible point, the lemma follows. ■

We will prove that any three vectors $\mathbf{V}_{u,s,\mathbf{x}} \otimes \overline{\mathbf{V}}_{u,t}$, $\mathbf{V}_{v,s,\mathbf{y}} \otimes \overline{\mathbf{V}}_{v,t}$ and $\mathbf{V}_{w,s,\mathbf{z}} \otimes \overline{\mathbf{V}}_{w,t}$ satisfy the triangle inequality. For $T \subseteq [k]$, let $r_T := \mathbf{x}(T), s_T := \mathbf{y}(T), t_T := \mathbf{z}(T)$, so that

$$\mathbf{V}_{u,s,\mathbf{x}} \otimes \overline{\mathbf{V}}_{u,t} = \left(\frac{1}{\sqrt{N}} \sum_{T \subseteq [k]} r_T \mathbf{u}_T^{\otimes 2s} \right) \otimes \left(\frac{1}{\sqrt{N}} \sum_{T \subseteq [k]} \mathbf{u}_T^{\otimes t} \right) = \text{(say)} \;\; \mathbf{W}_u,$$

$$\mathbf{V}_{v,s,\mathbf{y}} \otimes \overline{\mathbf{V}}_{v,t} = \left(\frac{1}{\sqrt{N}} \sum_{T \subseteq [k]} s_T \mathbf{v}_T^{\otimes 2s} \right) \otimes \left(\frac{1}{\sqrt{N}} \sum_{T \subseteq [k]} \mathbf{v}_T^{\otimes t} \right) = \text{(say)} \;\; \mathbf{W}_v,$$

$$\mathbf{V}_{w,s,\mathbf{z}} \otimes \overline{\mathbf{V}}_{w,t} = \left(\frac{1}{\sqrt{N}} \sum_{T \subseteq [k]} t_T \mathbf{w}_T^{\otimes 2s} \right) \otimes \left(\frac{1}{\sqrt{N}} \sum_{T \subseteq [k]} \mathbf{w}_T^{\otimes t} \right) = \text{(say)} \;\; \mathbf{W}_w.$$

We need to show that

$$1 + \langle \mathbf{W}_v, \mathbf{W}_w \rangle \ge \langle \mathbf{W}_u, \mathbf{W}_v \rangle + \langle \mathbf{W}_u, \mathbf{W}_w \rangle.$$

We can assume that at-least one of the dot-products has magnitude at-least 1/3; otherwise, the inequality trivially holds. Assume, w.l.o.g., that $|\langle \mathbf{W}_u, \mathbf{W}_w \rangle| \ge 1/3$. Hence,

$$|\langle \mathbf{V}_{u,s,\mathbf{x}}, \mathbf{V}_{w,s,\mathbf{z}} \rangle| |\langle \overline{\mathbf{V}}_{u,t}, \overline{\mathbf{V}}_{w,t} \rangle| \ge 1/3.$$

Hence, $|\langle \overline{\mathbf{V}}_{u,t}, \overline{\mathbf{V}}_{w,t} \rangle| \ge 1/3$. Hence, for some $S, T \subseteq [k]$, $|\langle \mathbf{u}_S, \mathbf{w}_T \rangle| = 1 - \eta$, for some $\eta \le 2^{-40s}$. By relabeling, if necessary, we may assume that $\langle \mathbf{u}_\emptyset, \mathbf{w}_\emptyset \rangle = 1 - \eta$.

We know from Theorem 9.3 that

$$\begin{aligned} 1 + \langle \overline{\mathbf{V}}_{u,t}, \overline{\mathbf{V}}_{v,t} \rangle &\ge \langle \overline{\mathbf{V}}_{u,t}, \overline{\mathbf{V}}_{w,t} \rangle + \langle \overline{\mathbf{V}}_{v,t}, \overline{\mathbf{V}}_{w,t} \rangle \\ 1 + \langle \overline{\mathbf{V}}_{u,t}, \overline{\mathbf{V}}_{w,t} \rangle &\ge \langle \overline{\mathbf{V}}_{u,t}, \overline{\mathbf{V}}_{v,t} \rangle + \langle \overline{\mathbf{V}}_{v,t}, \overline{\mathbf{V}}_{w,t} \rangle \\ 1 + \langle \overline{\mathbf{V}}_{v,t}, \overline{\mathbf{V}}_{w,t} \rangle &\ge \langle \overline{\mathbf{V}}_{u,t}, \overline{\mathbf{V}}_{w,t} \rangle + \langle \overline{\mathbf{V}}_{u,t}, \overline{\mathbf{V}}_{v,t} \rangle \end{aligned}$$

Hence, by Lemma 10.21, it suffices to show that

$$1 + \langle \mathbf{V}_{v,s,\mathbf{y}}, \mathbf{V}_{w,s,\mathbf{z}} \rangle \geq \langle \mathbf{V}_{u,s,\mathbf{x}}, \mathbf{V}_{w,s,\mathbf{z}} \rangle + \langle \mathbf{V}_{u,s,\mathbf{x}}, \mathbf{V}_{v,s,\mathbf{y}} \rangle.$$

Rest is exactly as in the proof of Theorem 10.10, We restate it for completeness. For notational convenience, we replace indexing by $T \subseteq [k]$ with indexing by $1 \leq i \leq N$ where $N = 2^k$, and $\emptyset$ is identified with 1. Thus, we write s_i, t_i, r_i instead of s_T, t_T, r_T. We also write $\mathbf{u}_i$ instead of $\mathbf{u}_T$, $\mathbf{v}_i$ instead of $\mathbf{v}_T$ and $\mathbf{w}_i$ instead of $\mathbf{w}_T$. Thus, we need to show that

$$N + \sum_{i,j=1}^{N} s_i t_j \langle \mathbf{v}_i, \mathbf{w}_j \rangle^{2s} \geq \sum_{i,j=1}^{N} s_i r_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle^{2s} + \sum_{i,j=1}^{N} t_i r_j \langle \mathbf{u}_i, \mathbf{w}_j \rangle^{2s}.$$

As noted before, we may assume that $\langle \mathbf{u}_1, \mathbf{w}_1 \rangle = 1 - \eta$, and hence, by the Matching Property,

$$\langle \mathbf{u}_1, \mathbf{w}_1 \rangle = \langle \mathbf{u}_2, \mathbf{w}_2 \rangle = \cdots = \langle \mathbf{u}_N, \mathbf{w}_N \rangle = 1 - \eta.$$

Let $\lambda := \max_{1 \leq i,j \leq N} |\langle \mathbf{u}_i, \mathbf{v}_j \rangle|$. We may assume, w.l.o.g., that the maximum is achieved for $\mathbf{u}_1, \mathbf{v}_1$, and again by the Matching Property,

$$\langle \mathbf{u}_1, \mathbf{v}_1 \rangle = \langle \mathbf{u}_2, \mathbf{v}_2 \rangle = \cdots = \langle \mathbf{u}_N, \mathbf{v}_N \rangle = \lambda.$$

Now, the desired inequality follows from Lemma 10.18.

11 PCP Reductions

This section contains the proofs of Theorems 10.1, 10.2 and 10.3. In fact, we do more: We also establish, assuming the UGC, hardness of approximation results for non-uniform versions of Sparsest Cut and Balanced Separator, and for Minimum Uncut. Optimal hardness of approximation result for Maximum Cut (assuming the UGC) was already known, see [20]. The hardness results are proved via the standard paradigm of composing an inner PCP with the outer PCP provided by the UGC. We present a separate inner PCP for Balanced Separator, Maximum Cut and Minimum Uncut. The inner PCP for Maximum Cut was proposed by Khot *et al.* in [20], while that for Minimum Uncut is implicit in [19].

The reduction from a PCP to a graph theoretic problem is standard: Replace bits in the proof by vertices, and replace every (2-query) PCP test by an edge of the graph. The weight of the edge is equal to the probability that the test is performed by the PCP verifier. The task of proving Theorems 10.1, 10.2 and 10.3 corresponds exactly to the soundness analysis for the corresponding inner PCPs, i.e., proving Theorems 11.2, 11.5 and 11.8 respectively.

11.1 Inner PCP for Balanced Separator

In this section, we establish that, assuming the UGC, it is NP-hard to (pseudo)-approximate Balanced Separator to within any constant factor. By a standard reduction, this immediately implies, again assuming the UGC, that it is NP-hard to approximate Sparsest Cut to within any constant factor. First, we describe an inner PCP that reads two bits from the proof and accepts if and only if the two bits are equal. For $\varepsilon \in (0,1)$, the verifier $V_\varepsilon^{\mathsf{BS}}$ is given a Unique Games instance $\mathcal{U}(G(V,E),[N],\{\pi_e\}_{e\in E})$. The verifier expects, as a proof, the Long Code of the label of every vertex $v \in V$. Formally, a proof Π is $\{A^v\}_{v\in V}$, where each $A^v \in \{-1,1\}^N$ is the supposed Long Code of the label of v.

The Verifier $V_\varepsilon^{\mathsf{BS}}$ with Parameter $\varepsilon \in (0,1)$

1. Pick $e\{v,w\} \in E$ with probability $\mathbf{wt}(e)$.
2. Pick a random $\mathbf{x} \in_{1/2} \{-1,1\}^N$ and $\boldsymbol{\mu} \in_\varepsilon \{-1,1\}^N$.
3. Let $\pi := \pi_e : [N] \to [N]$ be the bijection corresponding to $e\{v,w\}$. Accept if and only if

$$A^v(\mathbf{x}) = A^w((\mathbf{x}\boldsymbol{\mu}) \circ \pi_e^v).$$

Theorem 11.1 (Completeness) *For every $\varepsilon \in (0,1)$, if $\mathsf{OPT}(\mathcal{U}) \geq 1-\eta$, then there is a proof Π such that $\mathbf{Pr}\left[V_\varepsilon^{\mathsf{BS}} \text{ accepts } \Pi\right] \geq (1-\eta)(1-\varepsilon)$. Moreover, every table A^v in Π is balanced, i.e., exactly half of its entries are $+1$ and the rest half are -1.*

Proof: Since $\mathsf{OPT}(\mathcal{U}) \geq 1-\eta$, there is a labeling λ for which the total weight of the edges satisfied is at-least $1-\eta$. Hence, if we pick an edge $e\{v,w\}$ with probability $\mathbf{wt}(e)$, with probability at-least $1-\eta$, we have $\lambda(v) = \pi_e^v(\lambda(w))$. Let the proof consist of Long Codes of the labels assigned by λ to the vertices. With probability $1-\varepsilon$, we have $\boldsymbol{\mu}_{\lambda(v)} = 1$. Hence, with probability at-least $(1-\eta)(1-\varepsilon)$,

$$A^v(\mathbf{x}) = \mathbf{x}_{\lambda(v)} = (\mathbf{x}\boldsymbol{\mu})_{\pi_e^v(\lambda(w))} = A^w((\mathbf{x}\boldsymbol{\mu}) \circ \pi_e^v).$$

Noting that a Long Code is balanced, this completes the proof. ∎

We say that a proof $\Pi = \{A^v\}_{v\in V}$ is θ-piecewise balanced if

$$\mathbf{E}_v\left[|\widehat{A^v_\emptyset}|\right] \leq \theta.$$

Here, the expectation is over the probability distribution p_v on the vertex set V.

Theorem 11.2 *For every $t \in (\frac{1}{2},1)$, there exists a constant $b_t > 0$ such that the following holds: Let $\varepsilon > 0$ be sufficiently small and let $\mathcal{U}$ be an instance of* Unique Games *with $\mathsf{OPT}(\mathcal{U}) < 2^{-O(1/\varepsilon^2)}$. Then, for every $5/6$-piecewise balanced proof Π,*

$$\mathbf{Pr}\left[V_\varepsilon^{\mathsf{BS}} \text{ accepts } \Pi\right] < 1 - b_t\varepsilon^t.$$

Proof: The proof is by contradiction: We assume that there is a $5/6$-piecewise balanced proof Π, which the verifier accepts with probability at-least $1-b_t\varepsilon^t$, and deduce that $\mathsf{OPT}(\mathcal{U}) \geq 2^{-O(1/\varepsilon^2)}$. We let $b_t := \frac{1-e^{-2}}{96}c_t$, where c_t is the constant in Bourgain's Junta Theorem.
The probability of acceptance of the verifier is

$$\frac{1}{2} + \frac{1}{2}\mathbf{E}_{v,e\{v,w\},\mathbf{x},\boldsymbol{\mu}}\left[A^v(\mathbf{x})A^w(\mathbf{x}\boldsymbol{\mu} \circ \pi_e^v)\right].$$

Using the Fourier expansion $A^v = \sum_\alpha \widehat{A^v_\alpha}\chi_\alpha$ and $A^w = \sum_\beta \widehat{A^w_\beta}\chi_\beta$, and the orthonormality of characters, we get that this probability is

$$\frac{1}{2} + \frac{1}{2}\mathbf{E}_{v,e\{v,w\}}\left[\sum_\alpha \widehat{A^v_\alpha}\widehat{A^w_{\pi_e^w(\alpha)}}(1-2\varepsilon)^{|\alpha|}\right].$$

Here $\alpha \subseteq [N]$. Hence, the acceptance probability is

$$\frac{1}{2} + \frac{1}{2}\mathbf{E}_v \left[\sum_{\alpha} \widehat{A^v_\alpha} \mathbf{E}_{e\{v,w\}} \left[\widehat{A^w_{\pi^w_e(\alpha)}} \right] (1-2\varepsilon)^{|\alpha|} \right].$$

If this acceptance probability is at-least $1 - b_t\varepsilon^t$, then,

$$\mathbf{E}_v \left[\sum_{\alpha} \widehat{A^v_\alpha} \mathbf{E}_{e\{v,w\}} \left[\widehat{A^w_{\pi^w_e(\alpha)}} \right] (1-2\varepsilon)^{|\alpha|} \right] \geq 1 - 2b_t\varepsilon^t.$$

Hence, over the choice of v, with probability at-least $\frac{23}{24}$,

$$\sum_{\alpha} \widehat{A^v_\alpha} \mathbf{E}_{e\{v,w\}} \left[\widehat{A^w_{\pi^w_e(\alpha)}} \right] (1-2\varepsilon)^{|\alpha|} \geq 1 - 48b_t\varepsilon^t.$$

Call such vertices $v \in V$ *good*. Fix a good vertex v. Using the Cauchy-Schwarz Inequality we get,

$$\sum_{\alpha} \widehat{A^v_\alpha} \mathbf{E}_{e\{v,w\}} \left[\widehat{A^w_{\pi^w_e(\alpha)}} \right] (1-2\varepsilon)^{|\alpha|} \leq \sqrt{\sum_{\alpha} \widehat{A^v_\alpha}^2 (1-2\varepsilon)^{2|\alpha|} \sum_{\alpha} \mathbf{E}^2_{e\{v,w\}} \left[\widehat{A^w_{\pi^w_e(\alpha)}} \right]}.$$

Combining Jensen's Inequality and Parseval's Identity, we get that

$$\sum_{\alpha} \mathbf{E}^2_{e\{v,w\}} \left[\widehat{A^w_{\pi^w_e(\alpha)}} \right] \leq 1.$$

Hence,

$$1 - 96b_t\varepsilon^t \leq \sum_{\alpha} \widehat{A^v_\alpha}^2 (1-2\varepsilon)^{2|\alpha|}.$$

Now we combine Parseval's Identity with the fact that $1 - x \leq e^{-x}$ to obtain

$$\sum_{\alpha \,:\, |\alpha| > \frac{1}{\varepsilon}} \widehat{A^v_\alpha}^2 \leq \frac{96}{1 - e^{-2}} b_t\varepsilon^t = c_t\varepsilon^t.$$

Hence, by Bourgain's Junta Theorem

$$\sum_{\alpha \,:\, |\widehat{A^v_\alpha}| \leq \frac{1}{50}4^{-1/\varepsilon^2}} \widehat{A^v_\alpha}^2 \leq \frac{1}{2500}.$$

Call α *good* if $\alpha \subseteq [N]$ is nonempty, $|\alpha| \leq \varepsilon^{-1}$ and $|\widehat{A^v_\alpha}| \geq \frac{1}{50}4^{-1/\varepsilon^2}$.

Bounding the contribution due to large sets

Using the Cauchy-Schwarz Inequality, Parseval's Identity and Jensen's Inequality, we get

$$\left| \sum_{\alpha \,:\, |\alpha| > \frac{1}{\varepsilon}} \widehat{A^v_\alpha} \mathbf{E}_{e\{v,w\}} \left[\widehat{A^w_{\pi^w_e(\alpha)}} \right] (1-2\varepsilon)^{|\alpha|} \right| \leq \sqrt{\sum_{\alpha \,:\, |\alpha| > \varepsilon^{-1}} \widehat{A^v_\alpha}^2} < \sqrt{c_t\varepsilon^t}.$$

We can choose ε to be small enough so that the last term above is less than $1/50$.

Bounding the contribution due to small Fourier coefficients

Similarly, we use $\sum_{\alpha\,:\,|\widehat{A^v_\alpha}|\le\frac{1}{50}4^{-1/\varepsilon^2}} \widehat{A^v_\alpha}^2 \le \frac{1}{2500}$, and get

$$\left|\sum_{\alpha\,:\,|\widehat{A^v_\alpha}|\le\frac{1}{50}4^{-1/\varepsilon^2}} \widehat{A^v_\alpha}\mathbf{E}_{e\{v,w\}}\left[\widehat{A^w_{\pi^w_e(\alpha)}}\right](1-2\varepsilon)^{|\alpha|}\right| \le \frac{1}{50}.$$

Bounding the contribution due to the empty set

Since $\mathbf{E}_v\left[|\widehat{A^v_\emptyset}|\right] \le \frac{5}{6}$, $\mathbf{E}_v\left[\mathbf{E}_{e\{v,w\}}\left[|\widehat{A^v_\emptyset}\widehat{A^w_\emptyset}|\right]\right] \le \frac{5}{6}$. This is because each $|\widehat{A^v_\emptyset}| \le 1$. Hence, with probability at-least $\frac{1}{12}$ over the choice of v, $\mathbf{E}_{e\{v,w\}}\left[|\widehat{A^v_\emptyset}\widehat{A^w_\emptyset}|\right] \le \frac{10}{11}$. Hence, with probability at-least $\frac{1}{24}$ over the choice of v, v is good and $\mathbf{E}_{e\{v,w\}}\left[|\widehat{A^v_\emptyset}\widehat{A^w_\emptyset}|\right] \le \frac{10}{11}$. Call such a vertex *very good.*

Lower bound for a very good vertex with good sets

Hence, for a very good v,

$$\sum_{\alpha\text{ is good}} \widehat{A^v_\alpha}\mathbf{E}_{e\{v,w\}}\left[\widehat{A^w_{\pi^w_e(\alpha)}}\right](1-2\varepsilon)^{|\alpha|} \ge 1-\frac{1}{50}-\frac{1}{50}-\frac{10}{11} \ge \frac{1}{22}. \tag{38}$$

The labeling

Now we define a labeling for the Unique Games instance $\mathcal{U}$ as follows: For a vertex $v \in V$, pick α with probability $\widehat{A^v_\alpha}^2$, pick a random element of α and define it to be the label of v.

Let v be a very good vertex. It follows that the weight of the edges adjacent to v satisfied by this labeling is at-least

$$\mathbf{E}_{e\{v,w\}}\left[\sum_{\alpha\text{ is good}} \widehat{A^v_\alpha}^2\,\widehat{A^w_{\pi^w_e(\alpha)}}^2\frac{1}{|\alpha|}\right] \ge \varepsilon\ \mathbf{E}_{e\{v,w\}}\left[\sum_{\alpha\text{ is good}} \widehat{A^v_\alpha}^2\,\widehat{A^w_{\pi^w_e(\alpha)}}^2\right].$$

This is at-least

$$\varepsilon\ \frac{1}{2500}4^{-2/\varepsilon^2}\ \mathbf{E}_{e\{v,w\}}\left[\sum_{\alpha\text{ is good}} \widehat{A^w_{\pi^w_e(\alpha)}}^2\right],$$

which is at-least

$$\varepsilon\ \frac{1}{2500}4^{-2/\varepsilon^2}\ \mathbf{E}_{e\{v,w\}}\left[\sum_{\alpha\text{ is good}} \widehat{A^w_{\pi^w_e(\alpha)}}^2(1-2\varepsilon)^{|\alpha|}\right].$$

It follows from the Cauchy-Schwarz Inequality and Parseval's Identity that this is at-least

$$\varepsilon\ \frac{1}{2500}4^{-2/\varepsilon^2}\ \mathbf{E}_{e\{v,w\}}\left[\left|\sum_{\alpha\text{ is good}} \widehat{A^v_\alpha}\widehat{A^w_{\pi^w_e(\alpha)}}(1-2\varepsilon)^{|\alpha|}\right|^2\right].$$

Using Jensen's Inequality, we get that this is at-least

$$\varepsilon \ \frac{1}{2500} 4^{-2/\varepsilon^2} \ \mathbf{E}_{e\{v,w\}} \left| \left[\sum_{\alpha \text{ is good}} \widehat{A^v_\alpha} \widehat{A^w_{\pi^w_e(\alpha)}} (1-2\varepsilon)^{|\alpha|} \right] \right|^2 \geq \varepsilon \ \frac{1}{2500} 4^{-2/\varepsilon^2} \frac{1}{484}.$$

Here, the last inequality follows from our estimate in Equation (38). Since, with probability at-least $\frac{1}{24}$ over the choice of v, v is very good, our labeling satisfies edges with total weight at-least $\Omega\left(\varepsilon \ 4^{-2/\varepsilon^2}\right)$. This completes the proof of the theorem. ∎

Hardness of approximating Sparsest Cut *and* Balanced Separator

Assuming the UGC, for any $\eta, \zeta > 0$, it is NP-hard to determine whether an instance $\mathcal{U}$ of Unique Games has $\mathsf{OPT}(\mathcal{U}) \geq 1-\eta$ or $\mathsf{OPT}(\mathcal{U}) \leq \zeta$. We choose $\eta = \varepsilon$ and $\zeta \leq 2^{-O(1/\varepsilon^2)}$ so that (a) when $\mathsf{OPT}(\mathcal{U}) \geq 1-\eta$, there is a (piecewise balanced) proof that the verifier accepts with probability at-least $1-2\varepsilon$ and (b) when $\mathsf{OPT}(\mathcal{U}) \leq \zeta$, the verifier does not accept any 5/6-piecewise balanced proof with probability more than $1 - b_t\varepsilon^t$. Here b_t is as in the statement of Theorem 11.2.

Consider the instance of Balanced Separator, $\mathcal{I}^{\mathsf{BS}}_\varepsilon$, as described in Section 10.1.1. We recall it here: This instance has a parameter ε. Start with a Unique Games instance

$$\mathcal{U}\left(G(V,E), [N], \{\pi_e\}_{e\in E}, \mathbf{wt}\right),$$

and replace each vertex $v \in V$ by a *block* of vertices $(v, \mathbf{x})$ for each $\mathbf{x} \in \{-1,1\}^N$. For an edge $e\{v,w\} \in E$, there is an edge in $\mathcal{I}^{\mathsf{BS}}_\varepsilon$ between $(v,\mathbf{x})$ and $(w,\mathbf{y})$, with weight

$$\mathbf{wt}(e) \cdot \mathbf{Pr}_{\substack{\mathbf{x}' \in_{1/2} \{-1,1\}^N \\ \boldsymbol{\mu} \in_\varepsilon \{-1,1\}^N}} \left[(\mathbf{x} = \mathbf{x}') \wedge \left(\mathbf{y} = \mathbf{x}'\boldsymbol{\mu} \circ \pi^v_e\right) \right].$$

This is exactly the probability that $V^{\mathsf{BS}}_\varepsilon$ picks the edge $e\{v,w\}$, and decides to look at the $\mathbf{x}$-th (resp. $\mathbf{y}$-th) coordinate in the Long Code of the label of v (resp. w).

The demand function **dem** is 1 for any edge between vertices in the same block, and 0 otherwise. Let $B := \frac{1}{2}|V|\binom{2^N}{2}$, be half of the total demand.

Suppose that $\mathsf{OPT}(\mathcal{U}) \geq 1-\eta$. Let λ be a labeling that achieves the optimum. Consider the partition $(S, \overline{S})$ in $\mathcal{I}^{\mathsf{BS}}_\varepsilon$, where S consists of all vertices $(v,\mathbf{x})$ with the property that the Long Code of $\lambda(v)$ evaluated at $\mathbf{x}$ is 1. Clearly, the demand cut by this partition is exactly equal to B. Moreover, it follows from Theorem 11.1 that this partition cuts edges with weight at-most $\eta + \varepsilon = 2\varepsilon$.

Now, suppose that $\mathsf{OPT}(\mathcal{U}) \leq \zeta$. Then, it follows from Theorem 11.2, that any B'-balanced partition, with $B/3 \leq B' \leq B$, cuts at-least $b_t\varepsilon^t$ fraction of the edges. This is due to the following: Any partition $(S, \overline{S})$ in $\mathcal{I}^{\mathsf{BS}}_\varepsilon$ corresponds to a proof Π in which we let the (supposed) Long Code of the label of v to be 1 at the point $\mathbf{x}$ if $(v,\mathbf{x}) \in S$, and -1 otherwise. Since $B/3 \leq B' \leq B$, as in the proof of Theorem 7.2, Π is 5/6 piecewise balanced and we apply Theorem 11.2.

Thus, we get hardness factor of $\Omega\left(1/\varepsilon^{1-t}\right)$ for Balanced Separator and hence, by Lemma 6.14, for Sparsest Cut as well.

Theorem 11.3 *Assuming the* UGC *(Conjecture 8.2), it is* NP*-hard to approximate (non-uniform versions of)* Balanced Separator *and* Sparsest Cut *to within any constant factor.*

11.2 Inner PCP for Maximum Cut

We describe an inner PCP that reads two bits from the proof and accepts if and only if the two bits are unequal. The verifier and the soundness analysis is due to Khot *et al.* [20]. The soundness analysis uses the Majority is Stablest Conjecture (see Theorem 6.23), which has been recently proved by Mossel *et al.* [27]. The hardness result for Maximum Cut follows from the standard reduction from the PCP to a graph.

The verifier V_ρ^{MC} is given a Unique Games instance $\mathcal{U}(G(V,E),[N],\{\pi_e\}_{e\in E})$. The verifier expects, as a proof, the Long Code of the label of every vertex $v \in V$. Formally, a proof Π is $\{A^v\}_{v\in V}$, where each $A^v \in \{-1,1\}^N$ is the supposed Long Code of the label of v. The verifier is parameterized by $\rho \in (-1,0)$.

The Verifier V_ρ^{MC} with Parameter $\rho \in (-1,0)$

1. Pick $v \in V$ with probability p_v. Then, pick two edges $e\{v,w\}, e'\{v,w'\}$, independent of each other, where e is picked with probability $\Psi_v(e)$ and e' with probability $\Psi_v(e')$.
2. Pick $\mathbf{x} \in_{1/2} \{-1,1\}^N$ and $\boldsymbol{\mu} \in_{(1-\rho)/2} \{-1,1\}^N$.
3. Let π_e and $\pi_{e'}$ be the bijections for edges e and e' respectively. Accept if and only if

$$A^w(\mathbf{x}\circ\pi_e^v) \neq A^{w'}((\mathbf{x}\boldsymbol{\mu})\circ\pi_{e'}^v).$$

Theorem 11.4 (Completeness) *For every $\rho \in (-1,0)$, if $\mathsf{OPT}(\mathcal{U}) \geq 1-\eta$, then there is a proof Π with $\mathbf{Pr}\left[V_\rho^{\mathsf{MC}} \text{ accepts } \Pi\right] \geq \frac{(1-2\eta)(1-\rho)}{2}$.*

Proof: Since $\mathsf{OPT}(\mathcal{U}) \geq 1-\eta$, there is a labeling λ for which the total weight of the edges satisfied is at-least $1-\eta$. Hence, with probability at-least $1-2\eta$ over the choice of edges $e\{v,w\}$ and $e'\{v,w'\}$, $\lambda(v) = \pi_e^v(\lambda(w)) = \pi_{e'}^v(\lambda(w'))$. Let the proof consist of Long Codes of the labels assigned by λ to the vertices. With probability $\frac{1-\rho}{2}$, $\boldsymbol{\mu}_{\lambda(v)} = -1$. Hence, with probability at-least $\frac{(1-2\eta)(1-\rho)}{2}$,

$$A^w(\mathbf{x}\circ\pi_e^v) = \mathbf{x}_{\pi_e^v(\lambda(w))} = \mathbf{x}_{\lambda(v)} \neq (\mathbf{x}\boldsymbol{\mu})_{\lambda(v)} = (\mathbf{x}\boldsymbol{\mu})_{\pi_{e'}^v(\lambda(w'))} = A^{w'}((\mathbf{x}\boldsymbol{\mu})\circ\pi_{e'}^v).$$

This completes the proof. ■

Theorem 11.5 (Soundness [20]) *For any constants $\rho \in (-1,0)$ and $\lambda > 0$, there is a constant $c(\rho,\lambda)$ such that the following holds: Let $\mathcal{U}$ be an instance of* Unique Games *with $\mathsf{OPT}(\mathcal{U}) < c(\rho,\lambda)$. Then, for any proof Π,*

$$\mathbf{Pr}\left[V_\rho^{\mathsf{MC}} \text{ accepts } \Pi\right] < \frac{1}{\pi}\arccos(\rho) + \lambda.$$

Proof: The proof is by contradiction. Assume that there is a proof $\Pi := \{A^v\}_{v\in V}$, where A^v is the supposed Long Code for the label assigned to v, which the verifier accepts with probability at-least $\frac{1}{\pi}\arccos\rho + \lambda$. Then, with probability at-least $\lambda/2$ over the choice of v,

$$\frac{1}{2} - \frac{1}{2}\mathbf{E}_{e\{v,w\},e'\{v,w'\}}\left[\sum_{\substack{\alpha,\beta \\ \pi_e^v(\alpha)=\pi_{e'}^v(\beta)}} \widehat{A^w_\alpha}\widehat{A^{w'}_\beta}\rho^{|\alpha|}\right] \geq \frac{1}{\pi}\arccos\rho + \frac{\lambda}{2}.$$

Call such a vertex v *good*. Fix a good vertex v. Note that $e\{v,w\}$ and $e'\{v,w'\}$ are identically distributed. Hence, the above can be written as

$$\sum_\alpha \left(\mathbf{E}_{e\{v,w\}}\left[\widehat{A^w_{\pi_e^w(\alpha)}}\right]\right)^2 \rho^{|\alpha|} \le 1 - \frac{2}{\pi}\arccos\rho - \lambda.$$

Let $C^v(\mathbf{x}) := \mathbf{E}_{e\{v,w\}}\left[A^w(\mathbf{x}\circ\pi_e^v)\right]$. Hence,

$$\sum_\alpha \widehat{C^v_\alpha}^2 \rho^{|\alpha|} \le 1 - \frac{2}{\pi}\arccos\rho - \lambda.$$

Applying the Majority is Stablest Theorem on C^v, we get that there is a coordinate, say i_v, such that $\mathrm{Inf}_{i_v}^{\le k}(C^v) \ge \zeta$. Here ζ, k are some constants depending on λ and ρ. Since $\mathrm{Inf}_{i_v}^{\le k}(C^v) \ge \zeta$,

$$\zeta \le \sum_{\substack{\alpha:\, i_v\in\alpha \\ |\alpha|\le k}} \mathbf{E}_{e\{v,w\}}\left[\widehat{A^w_{\pi_e^w(\alpha)}}\right]^2 \le \sum_{\substack{\alpha:\, i_v\in\alpha \\ |\alpha|\le k}} \mathbf{E}_{e\{v,w\}}\left[\widehat{A^w_{\pi_e^w(\alpha)}}^2\right].$$

The last inequality follows from Jensen's Inequality. Hence,

$$\zeta \le \mathbf{E}_{e\{v,w\}}\left[\mathrm{Inf}^{\le k}_{\pi_e^w(i_v)}(A^w)\right].$$

Hence, with probability at-least $\zeta/2$ over the choice of edges $e\{v,w\}$ adjacent to a good vertex v,

$$\frac{\zeta}{2} \le \mathrm{Inf}^{\le k}_{\pi_e^w(i_v)}(A^w). \tag{39}$$

The labeling

Now we define a labeling for the good vertices of the Unique Games instance $\mathcal{U}$ as follows: Let $\mathrm{Cand}_2[v] := \{j \in [N] : \mathrm{Inf}_j^{\le k}(C^v) \ge \zeta\}$, and let $\mathrm{Cand}_1[v] := \{j \in [N] : \mathrm{Inf}_j^{\le k}(A^v) \ge \zeta/2\}$. Then, let $\mathrm{Cand}[v] := \mathrm{Cand}_1[v] \cup \mathrm{Cand}_2[v]$. Since $\sum_j \mathrm{Inf}_j^{\le k}(A^v) \le k$, $|\mathrm{Cand}_1[v]| \le 2k/\zeta$. Similarly, $|\mathrm{Cand}_2[v]| \le k/\zeta$. Further, for a good vertex v, as noted above, $\mathrm{Cand}_2[v] \ne \emptyset$. The labeling is as follows: Pick a random i from $\mathrm{Cand}[v]$ and let it be the label of v. Hence, for a good vertex v, with probability at-least $\zeta/3k$, $i \in \mathrm{Cand}_2[v]$. Further, it follows from Equation (39), it follows that for $i \in \mathrm{Cand}_2[v]$, with probability at-least $\zeta/2$ over the choice of $e\{v,w\}$, $\pi_e^w(i) \in \mathrm{Cand}_1[w]$. This will be assigned to w with probability at-least $\zeta/3k$. Hence, this labeling satisfies at least $\frac{\lambda}{2}\frac{\zeta}{2}\left(\frac{\zeta}{3k}\right)^2$ fraction of the edges. Since k, ζ depend only on λ and ρ, we can let $c(\rho,\lambda) := \frac{\lambda}{2}\frac{\zeta}{2}\left(\frac{3k}{\zeta}\right)^2$ to complete the proof of the theorem.

∎

The following theorem, due to Khot *et al.*, follows immediately from Theorems 11.4 and 11.5.

Theorem 11.6 *[20] Assuming the* UGC *(Conjecture 8.2),* Maximum Cut *is* NP*-hard to approximate within any factor greater than* $\alpha_{\mathsf{GW}} \approx 0.878$.

11.3 Inner PCP for Minimum Uncut

In this section, we describe the verifier $V_\varepsilon^{\mathsf{MUC}}$. This is exactly the same as $V_{-1+2\varepsilon}^{\mathsf{MC}}$. We describe it explicitly for the sake of completeness. The verifier $V_\varepsilon^{\mathsf{MUC}}$ is given a Unique Games instance

$$\mathcal{U}(G(V,E),[N],\{\pi_e\}_{e\in E}).$$

The verifier expects, as a proof, the Long Code of the label of every vertex $v \in V$. Formally, a proof Π is $\{A^v\}_{v\in V}$, where each $A^v \in \{-1,1\}^N$ is the supposed Long Code of the label of v. The verifier is parameterized by $\varepsilon \in (0,1)$.

The Verifier $V_\varepsilon^{\mathsf{MUC}}$ with Parameter $\varepsilon \in (0,1)$

1. Pick $v \in V$ with probability p_v. Then, pick two edges $e\{v,w\}, e'\{v,w'\}$, independent of each other, where e is picked with probability $\Psi_v(e)$ and e' with probability $\Psi_v(e')$.
2. Pick $\mathbf{x} \in_{1/2} \{-1,1\}^N$ and $\boldsymbol{\mu} \in_{1-\varepsilon} \{-1,1\}^N$.
3. Let π_e and $\pi_{e'}$ be the bijections for edges e and e' respectively. Accept if and only if

$$A^w(\mathbf{x} \circ \pi_e^v) \neq A^{w'}((\mathbf{x}\boldsymbol{\mu}) \circ \pi_{e'}^v).$$

Theorem 11.7 (Completeness) *For every $\varepsilon \in (0,1)$, if $\mathsf{OPT}(\mathcal{U}) \geq 1-\eta$, then there is a proof Π with* $\mathbf{Pr}\left[V_\varepsilon^{\mathsf{MUC}} \text{ accepts } \Pi\right] \geq (1-2\eta)(1-\varepsilon)$.

Proof: Since $\mathsf{OPT}(\mathcal{U}) \geq 1-\eta$, there is a labeling λ for which the total weight of the edges satisfied is at-least $1-\eta$. Hence, with probability at-least $1-2\eta$ over the choice of edges $e\{v,w\}$ and $e'\{v,w'\}$, $\lambda(v) = \pi_e^v(\lambda(w)) = \pi_{e'}^v(\lambda(w'))$. Let the proof consist of Long Codes of the labels assigned by λ to the vertices. With probability $1-\varepsilon$, $\boldsymbol{\mu}_{\lambda(v)} = -1$. Hence, with probability at-least $(1-2\eta)(1-\varepsilon)$,

$$A^w(\mathbf{x} \circ \pi_e^v) = \mathbf{x}_{\pi_e^v(\lambda(w))} \neq (\mathbf{x}\boldsymbol{\mu})_{\pi_{e'}^v(\lambda(w'))} = A^{w'}((\mathbf{x}\boldsymbol{\mu}) \circ \pi_{e'}^v).$$

This completes the proof. ∎

Theorem 11.8 (Soundness) *For every $t \in (\frac{1}{2},1)$, there exists a constant $b_t > 0$ such that the following holds: Let $\varepsilon > 0$ be sufficiently small and let $\mathcal{U}$ be an instance of* Unique Games *with* $\mathsf{OPT}(\mathcal{U}) < 2^{-O(1/\varepsilon^2)}$. *Then, for every proof* Π,

$$\mathbf{Pr}\left[V_\varepsilon^{\mathsf{MUC}} \text{ accepts } \Pi\right] < 1 - b_t\varepsilon^t.$$

Proof: The proof is by contradiction: We assume that there is a proof $\Pi := \{A^v\}_{v\in V}$, where A^v is the supposed Long Code for the label assigned to v, which the verifier accepts with probability at-least $1-b_t\varepsilon^t$, and deduce that $\mathsf{OPT}(\mathcal{U}) \geq 2^{-O(1/\varepsilon^2)}$. We let $b_t := \frac{1-e^{-2}}{4}c_t$, where c_t is the constant in Bourgain's Junta Theorem.

The probability that the verifier accepts Π is

$$\frac{1}{2} - \frac{1}{2}\mathbf{E}_{v,e\{v,w\},e'\{v,w'\},\mathbf{x},\boldsymbol{\mu}}\left[A^w(\mathbf{x} \circ \pi_e^v)A^{w'}(\mathbf{x}\boldsymbol{\mu} \circ \pi_{e'}^v)\right].$$

Using the Fourier expansion $A^w = \sum_\alpha \widehat{A^w_\alpha}\chi_\alpha$ and $A^{w'} = \sum_\beta \widehat{A^{w'}_\beta}\chi_\beta$, and the orthonormality of characters, we get that this probability is

$$\frac{1}{2} - \frac{1}{2}\mathbf{E}_{v,e\{v,w\},e'\{v,w'\}}\left[\sum_{\substack{\alpha,\beta \\ \pi^v_e(\alpha)=\pi^v_{e'}(\beta)}} \widehat{A^w_\alpha}\widehat{A^{w'}_\beta}(2\varepsilon-1)^{|\alpha|}\right].$$

Here $\alpha, \beta \subseteq [N]$. If the acceptance probability is at-least $1 - b_t\varepsilon^t$, then,

$$-\mathbf{E}_{v,e\{v,w\},e'\{v,w'\}}\left[\sum_{\substack{\alpha,\beta \\ \pi^v_e(\alpha)=\pi^v_{e'}(\beta)}} \widehat{A^w_\alpha}\widehat{A^{w'}_\beta}(2\varepsilon-1)^{|\alpha|}\right] \geq 1 - 2b_t\varepsilon^t.$$

Hence, over the choice of v, with probability at-least $\frac{1}{2}$,

$$-\mathbf{E}_{e\{v,w\},e'\{v,w'\}}\left[\sum_{\substack{\alpha,\beta \\ \pi^v_e(\alpha)=\pi^v_{e'}(\beta)}} \widehat{A^w_\alpha}\widehat{A^{w'}_\beta}(2\varepsilon-1)^{|\alpha|}\right] \geq 1 - 4b_t\varepsilon^t.$$

Call such vertices *good*. Fix a good vertex v. Now, $e\{v,w\}$ and $e'\{v,w'\}$ are identically distributed. Hence, the above can be written as

$$-\sum_\alpha \left(\mathbf{E}_{e\{v,w\}}\left[\widehat{A^w_{\pi^w_e(\alpha)}}\right]\right)^2 (2\varepsilon-1)^{|\alpha|} \geq 1 - 4b_t\varepsilon^t.$$

If we let $\varepsilon < 1/2$, the contribution of $\alpha = \emptyset$ will be negative. Hence, we may assume that $\widehat{A^w_\emptyset} = 0$ for every $w \in V$. Hence,

$$-\sum_\alpha (-1)^{|\alpha|}\left(\mathbf{E}_{e\{v,w\}}\left[\widehat{A^w_{\pi^w_e(\alpha)}}\right]\right)^2 (1-2\varepsilon)^{|\alpha|} \geq 1 - 4b_t\varepsilon^t.$$

Using Jensen' Inequality, it follows that

$$\mathbf{E}_{e\{v,w\}}\left[\sum_\alpha \widehat{A^w_{\pi^w_e(\alpha)}}^2 (1-2\varepsilon)^{|\alpha|}\right] \geq 1 - 4b_t\varepsilon^t.$$

Hence, as in the proof of Theorem 11.2,

$$\mathbf{E}_{e\{v,w\}}\left[\sum_{|\alpha|>\frac{1}{\varepsilon}} \widehat{A^w_{\pi^w_e(\alpha)}}^2\right] \leq \frac{4}{1-e^{-2}}b_t\varepsilon^t. \tag{40}$$

For every $v \in V$, define a function $C^v(\mathbf{x}) : \{-1,1\}^N \mapsto \mathbb{R}$ as $C^v(\mathbf{x}) := \mathbf{E}_{e\{v,w\}}\left[A^w(\mathbf{x} \circ \pi^v_e)\right]$. Then, $\widehat{C^v_\alpha} = \mathbf{E}_{e\{v,w\}}\left[\widehat{A^w_{\pi^w_e(\alpha)}}\right]$. Combining Equation (40) with Jensen's Inequality, for a good vertex v,

$$\sum_{|\alpha|>\frac{1}{\varepsilon}} \widehat{C^v_\alpha}^2 \leq \frac{4}{1-e^{-2}}b_t\varepsilon^t.$$

Hence, by Bourgain's Junta Theorem,

$$\sum_{\alpha:|\widehat{C^v_\alpha}|\leq \frac{1}{10}4^{-1/\varepsilon^2}} \widehat{C^v_\alpha}^2 \leq \frac{1}{100}. \tag{41}$$

Let α be such that $|\widehat{C^v_\alpha}| \geq \frac{1}{10}4^{-1/\varepsilon^2}$. Then, $\mathbf{E}_{e\{v,w\}}\left[\left|\widehat{A^w_{\pi^w_e(\alpha)}}\right|\right] \geq \frac{1}{10}4^{-1/\varepsilon^2}$. Hence,

$$\mathbf{Pr}_{e\{v,w\}}\left[\left|\widehat{A^w_{\pi^w_e(\alpha)}}\right| \geq \frac{1}{20}4^{-1/\varepsilon^2}\right] \geq \frac{1}{20}4^{-1/\varepsilon^2}. \tag{42}$$

The labeling

Now we define a labeling for the Unique Games instance $\mathcal{U}$ as follows: For a good vertex $v \in V$, define the set of candidate labels as follows: Let $\mathrm{Cand}_1[v] := \left\{\alpha | \alpha \neq \emptyset, |\alpha| \leq \frac{1}{\varepsilon}, |\widehat{A^v_\alpha}| \geq \frac{1}{20}4^{-1/\varepsilon^2}\right\}$ and $\mathrm{Cand}_2[v] := \left\{\alpha | \alpha \neq \emptyset, |\alpha| \leq \frac{1}{\varepsilon}, |\widehat{C^v_\alpha}| \geq \frac{1}{10}4^{-1/\varepsilon^2}\right\}$. Then, let $\mathrm{Cand}[v] := \mathrm{Cand}_1[v] \cup \mathrm{Cand}_2[v]$. It follows from Parseval's Identity that $|\mathrm{Cand}[v]| \leq |\mathrm{Cand}_1[v]| + |\mathrm{Cand}_2[v]| \leq 600 \cdot 4^{2/\varepsilon^2}$. Further, it follows from Equation (41) that $|\mathrm{Cand}_2[v]| \neq \emptyset$. The labeling is as follows: Pick a random α from $\mathrm{Cand}[v]$ and let a random element of α be the label of v. For a good vertex v, let $\alpha \neq \emptyset \in \mathrm{Cand}_2[v]$. It follows from Equation 42 that for at-least $\frac{1}{20}4^{-1/\varepsilon^2}$ fraction of neighbors w of v, $\left|\widehat{A^w_{\pi^w_e(\alpha)}}\right| \geq \frac{1}{20}4^{-1/\varepsilon^2}$. Since $0 < |\pi^w_e(\alpha)| \leq \frac{1}{\varepsilon}$, $\pi^w_e(\alpha) \in \mathrm{Cand}_1[w]$. Hence, $\mathrm{Cand}_2[v] \cap \mathrm{Cand}_1[w] \neq \emptyset$. Hence, this labeling satisfies at-least

$$\frac{1}{2} \cdot (20 \cdot 4^{1/\varepsilon^2})(600 \cdot 4^{2/\varepsilon^2})^2 \cdot \varepsilon$$

edges. This completes the proof of the theorem. ■

The following theorem now follows immediately:

Theorem 11.9 *Assuming the* UGC, *it is* NP*-hard to approximate* Minimum Uncut *to within any constant factor.*

12 Conclusion

We have presented a construction of an n-point negative type metric that requires distortion of $(\log\log n)^{1/6-\delta}$ to embed into ℓ_1. The best upper bound for embedding negative type metrics into ℓ_1 (rather, into $\ell_2 \subseteq \ell_1$) is $O(\sqrt{\log n}\log\log n)$ by Arora, Lee, and Naor [3]. It would be nice to close this gap. We believe that the connection between PCPs and metric embeddings, and the general approach of using PCP reductions to construct integrality gap instances would find several applications in future. In particular, combining the techniques in this paper with the PCP reduction for Vertex Cover by Khot and Regev [22] may lead to new integrality gap results for Vertex Cover.

It would be nice to give a more illuminating reason why our vectors satisfy the triangle inequality constraints. It may be true that our vectors satisfy the so-called *k-gonal inequalities* for small values of k (i.e. $k = 5, 7, \ldots$), and, perhaps, even for all values of k. This would imply that adding the k-gonal inequalities to the SDP relaxations does not increase their power. We leave this as an open problem.

The results in this paper seem to support the UGC, and resolving this conjecture remains a major open problem. The integrality gap instance for Unique Games (Theorem 2.6) might suggest approaches towards this goal.

Acknowledgments We are very thankful to Assaf Naor and James Lee for useful pointers and for ruling out some of our initial approaches. Many thanks to Sanjeev Arora, Moses Charikar, Umesh Vazirani, Ryan O'Donnell and Elchanan Mossel for insightful discussions at various junctures.

References

[1] Amit Agarwal, Moses Charikar, Konstantin Makarychev, and Yury Makarychev. $O(\sqrt{\log n})$ approximation algorithms for min-UnCut, min-2CNF deletion, and directed cut problems. In *Proceedings of the ACM Symposium on the Theory of Computing*, number 37, pages 573–581, 2005.

[2] Sanjeev Arora, Eli Berger, Elad Hazan, Guy Kindler, and Shmuel Safra. On non-approximability for quadratic programs. In *Annual Symposium on Foundations of Computer Science*, number 46, 2005. To appear.

[3] Sanjeev Arora, James R. Lee, and Assaf Naor. Euclidean distortion and the sparsest cut. In *Proceedings of the ACM Symposium on the Theory of Computing*, number 37, pages 553–562, 2005.

[4] Sanjeev Arora, László Lovász, Ilan Newman, Yuval Rabani, Yuri Rabinovich, and Santosh Vempala. Local versus global properties of metric spaces. Manuscript, 2005.

[5] Sanjeev Arora, Satish Rao, and Umesh V. Vazirani. Expander flows, geometric embeddings and graph partitioning. In *Proceedings of the ACM Symposium on the Theory of Computing*, number 36, pages 222–231, 2004.

[6] Yonatan Aumann and Yuval Rabani. An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm. *SIAM J. Comput.*, 27(1):291–301, 1998.

[7] Jean Bourgain. On lipschitz embeddings of finite metric spaces in hilbert space. *Israel Journal of Mathematics*, 52:46–52, 1985.

[8] Jean Bourgain. On the distribution of the fourier spectrum of boolean functions. *Israel Journal of Mathematics*, 131:269–276, 2002.

[9] Shuchi Chawla, Anupam Gupta, and Harald Räcke. Embeddings of negative-type metrics and an improved approximation to generalized sparsest cut. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 102–111, 2005.

[10] Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani, and D. Sivakumar. On the hardness of approximating multicut and sparsest-cut. In *Proceedings of the Annual IEEE Conference on Computational Complexity*, number 20, pages 144–153, 2005.

[11] M. Deza and Monique Laurent. *Geometry of cuts and metrics*. Springer-Verlag, New York, 1997.

[12] Peter Enflo. On the non-existence of uniform homeommorphism between L_p spaces. *Arkiv. Mat.*, 8:103–105, 1969.

[13] Uriel Feige, Mohammad Taghi Hajiaghayi, and James R. Lee. Improved approximation algorithms for minimum-weight vertex separators. In *Proceedings of the ACM Symposium on the Theory of Computing*, number 37, pages 563–572, 2005.

[14] Uriel Feige and László Lovász. Two-prover one-round proof systems, their power and their problems. In *Proceedings of the ACM Symposium on the Theory of Computing*, number 24, pages 733–744, 2002.

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

[16] Michel X. Goemans. Semidefinite programming in combinatorial optimization. *Math. Program.*, 79:143–161, 1997.

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

[18] George Karakostas. A better approximation ratio for the vertex cover problem. ECCC Report TR04-084, 2004.

[19] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the ACM Symposium on the Theory of Computing*, number 34, pages 767–775, 2002.

[20] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable CSPs? In *Annual Symposium on Foundations of Computer Science*, number 45, pages 146–154, 2004.

[21] Subhash Khot and Assaf Naor. Nonembeddability theorems via fourier analysis. In *Annual Symposium on Foundations of Computer Science*, number 46, 2005. To appear.

[22] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2-\varepsilon$. In *Proceedings of the Annual IEEE Conference on Computational Complexity*, number 18, pages 379–386, 2003.

[23] Frank Thomson Leighton and Satish Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *Journal of the ACM*, 46(6):787–832, 1999.

[24] Nathan Linial. Finite metric spaces: combinatorics, geometry and algorithms. In *Proceedings of the International Congress of Mathematicians*, number III, pages 573–586, 2002.

[25] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

[26] Jirí Matousek. *Lectures on discrete geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[27] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. In *Annual Symposium on Foundations of Computer Science*, number 46, 2005. To appear.

[28] Assaf Naor, Yuval Rabani, and Alistair Sinclair. Quasisymmetric embeddings, the observable diameter and expansion properties of graphs. Manuscript, 2005.

[29] R. O'Donnell. *Computational Aspects of Noise Sensitivity*. PhD thesis, 2004, MIT.

[30] Gideon Schechtman. *Handbook of the Geometry of Banach Spaces*, volume 2, chapter Concentration results and applications. North Holland, 2003.