

# Integrality Gaps for Sparsest Cut and Minimum Linear Arrangement Problems

Nikhil R. Devanur  
nikhil@cc.gatech.edu  
College of Computing,  
Georgia Tech.  
Atlanta, GA 30332.

Rishi Saket  
saket@cc.gatech.edu  
College of Computing,  
Georgia Tech.  
Atlanta, GA 30332.

Subhash A.Khot\*  
khot@cc.gatech.edu  
College of Computing,  
Georgia Tech.  
Atlanta, GA 30332.

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

## ABSTRACT

Arora, Rao and Vazirani [2] showed that the standard semi-definite programming (SDP) relaxation of the Sparsest Cut problem with the *triangle inequality* constraints has an integrality gap of  $O(\sqrt{\log n})$ . They conjectured that the gap is bounded from above by a constant. In this paper, we disprove this conjecture (referred to as the ARV-Conjecture) by constructing an  $\Omega(\log \log n)$  integrality gap instance. Khot and Vishnoi [16] had earlier disproved the *non-uniform* version of the ARV-Conjecture.

A simple “stretching” of the integrality gap instance for the Sparsest Cut problem serves as an  $\Omega(\log \log n)$  integrality gap instance for the SDP relaxation of the Minimum Linear Arrangement problem. This SDP relaxation was considered in [6, 11], where it was shown that its integrality gap is bounded from above by  $O(\sqrt{\log n} \log \log n)$ .

**Categories and Subject Descriptors:** F.2 [Analysis of Algorithms and Problem Complexity]: General;  
**General Terms:** Algorithms, Theory.

## 1. INTRODUCTION

Given an  $n$ -vertex graph  $G(V, E)$ , the *sparsity* of a cut  $(S, \bar{S})$  is defined as  $\frac{E(S, \bar{S})}{|S||\bar{S}|}$ , where  $E(S, \bar{S})$  denotes the set of edges crossing the cut. The Sparsest Cut problem is to find a cut with minimum sparsity. In the related problem of  $b$ -Balanced Separator, for some fixed constant  $0 < b \leq 1/2$ , the objective is to find a cut  $(S, \bar{S})$ , with  $|S|, |\bar{S}| \geq bn$ , which

\*This research was partially supported by the Microsoft New Faculty Fellowship.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

STOC'06, May 21–23, 2006, Seattle, Washington, USA.  
Copyright 2006 ACM 1-59593-134-1/06/0005 ...\$5.00.

minimizes the number of edges cut. It is well-known that a factor  $f(n)$  approximation algorithm for Sparsest Cut can be used iteratively to design a factor  $O(f(n))$  (pseudo-) approximation algorithm for Balanced Separator: Given a graph that has a  $(\frac{1}{2}, \frac{1}{2})$  partition cutting an  $\alpha$  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. A comprehensive survey of the applications of these two important problems in computer science can be found in [22].

The seminal work of Leighton and Rao [18] gave  $O(\log n)$  approximation algorithm for Sparsest Cut via an LP relaxation based on multicommodity flows. Aumann and Rabani [3] and Linial, London and Rabinovich [20] showed that  $O(\log n)$  approximation for Sparsest Cut also follows from a theorem of Bourgain [4] that gives embedding of any  $n$ -point metric into  $\ell_1$  with distortion  $O(\log n)$ . Specifically, they showed that the integrality gap for the *metric LP relaxation* of Sparsest Cut is bounded from above by  $O(\log n)$ .

Subsequently, it was realized that one could write an SDP relaxation of Sparsest Cut (see Fig. 1, Section 2.1), and enforce an additional condition, that the metric belongs to a special class of metrics, called the *negative type* or  $\ell_2^2$  metrics. Thus, a *better* embedding of  $\ell_2^2$  metrics into  $\ell_1$  would imply a *better* upper bound on the integrality gap of the Sparsest Cut SDP (and hence, a better approximation algorithm). Let us denote by  $g(n)$  the worst distortion needed to embed any  $n$ -point  $\ell_2^2$  metric into  $\ell_1$ , and let  $f(n)$  denote the worst integrality gap for the Sparsest Cut SDP. One can define a so-called *non-uniform* or *demands* version of the Sparsest Cut problem and consider a similar SDP relaxation for this problem. Let  $h(n)$  denote the worst integrality gap for this SDP. It is known that<sup>1</sup>

$$f(n) \leq h(n) \leq g(n).$$

<sup>1</sup>Among experts in the area, it is also known that  $h(n) \geq (1 - o(1))g(n)$ . In fact  $h(n) = g(n)$ , if the SDP relaxation for non-uniform Sparsest Cut does not require all vectors in the solution to have the same norm. We were unable to find an explicit reference.

The left hand side of this inequality holds because the non-uniform version of Sparsest Cut is more general, whereas, the right hand side of the inequality follows from results of [3, 20]. Bourgain’s Theorem shows that  $g(n) \leq O(\log n)$ , and it was an open problem whether  $\ell_2^2$  metrics admit a better embedding into  $\ell_1$ . In fact, Goemans [12] and Linial [19] conjectured that  $g(n) \leq C$  for some constant  $C$ . A breakthrough result of Arora, Rao and Vazirani [2] showed that  $f(n) \leq O(\sqrt{\log n})$ . They conjectured that  $f(n) \leq C$  for some constant  $C$  (henceforth, we refer to it as the ARV-Conjecture.) The *non-uniform version* of the ARV-Conjecture<sup>2</sup> would state that  $h(n) \leq C$  for some constant  $C$ .

Soon, Chawla, Gupta and Räcke [7] showed that  $g(n) \leq O(\log^{3/4} n)$ . This was improved by Arora, Lee and Naor [1] who showed that  $g(n) \leq O(\sqrt{\log n} \log \log n)$ . Both these results are actually stronger in the sense that they give embeddings of  $\ell_2^2$  into  $\ell_2$  (and  $\ell_2$  is known to embed isometrically into  $\ell_1$ ). From the lower bound side, progress was made by Khot and Vishnoi [16], showing that  $h(n) \geq (\log \log n)^{1/6-o(1)}$ . This disproved the Goemans-Linial Conjecture as well as the non-uniform ARV-Conjecture. The lower bound on  $h(n)$  was improved to  $\Omega(\log \log n)$  by Krauthgamer and Rabani [17].

We would like to point out that there is currently no hardness of approximation result for the Sparsest Cut (uniform or non-uniform) problem. It was shown in [8, 16] that, assuming the Unique Games Conjecture of Khot [15], non-uniform Sparsest Cut is hard to approximate within any constant factor.

## 1.1 Our Results

The main result of this paper is the disproof of the ARV-Conjecture. We construct an  $\Omega(\log \log n)$  integrality gap instance for Balanced Separator which implies the same gap for Sparsest Cut (i.e. we prove that  $f(n) \geq \Omega(\log \log n)$ ).

**THEOREM 1.1.** *The standard SDP relaxations of Sparsest Cut and Balanced Separator with the triangle inequality constraints have an integrality gap of at least  $\Omega(\log \log n)$ .*

Our result subsumes the results by Khot and Vishnoi [16], and Krauthgamer and Rabani [17]. As in [17, 8], our lower bound proof uses a Fourier analytic theorem of Kahn, Kalai and Linial [14] whereas Khot and Vishnoi use a theorem of Bourgain [5].

Our integrality gap instance easily extends to an integrality gap instance for an SDP relaxation of the Minimum Linear Arrangement problem. In this problem, given a graph  $G(V, E)$  on  $n$  vertices, the goal is to find a bijective assignment  $\pi : V \mapsto \{1, \dots, n\}$  which minimizes  $\sum_{e \in E} |\pi(i) - \pi(j)|$ . An approximation algorithm with ratio  $O(\log^2 n)$  for Minimum Linear Arrangement was given by Hansen [13], based on the work of Leighton and Rao [18]. It was improved to  $O(\log n \log \log n)$  by Even *et al.* [10], and to  $O(\log n)$  by Rao and Richa [21]. Recently, an SDP relaxation of Minimum Linear Arrangement was considered in [6, 11] where an  $O(\sqrt{\log n} \log \log n)$  integrality gap for the same was shown. We show that this SDP relaxation has an integrality gap of  $\Omega(\log \log n)$ .

<sup>2</sup>Arora, Rao and Vazirani did not make the conjecture about the non-uniform version, but we take the liberty to name it after them.

**THEOREM 1.2.** *The SDP relaxation of Minimum Linear Arrangement with the triangle inequality constraints has an integrality gap of at least  $\Omega(\log \log n)$ .*

## 1.2 Overview of the Paper

The formal statements of our integrality gap constructions appear in Sections 2.2 and 2.4. Section 2.2 also explains the basic idea behind our construction and how it differs from constructions in [16, 17]. The formal description of the construction of the integrality gap instance for Balanced Separator appears in Section 3. It has two parts: First, showing that the constructed graph has no small balanced cuts, and second, the constructed SDP solution satisfies the SDP constraints. The first part involves a simple application of the Kahn, Kalai and Linial Theorem [14], and is presented in Section 4.1. Construction of the SDP solution is rather technical and all the proofs are deferred to Section 4.

## 2. INTEGRALITY GAP CONSTRUCTIONS

In this section, we describe the SDP relaxations for Sparsest Cut, Balanced Separator and Minimum Linear Arrangement problems and give formal statements of our integrality gap constructions.

### 2.1 Sparsest Cut

**DEFINITION 2.1 (SPARSEST CUT).** *For a multi-graph  $G(V, E)$*   
*find  $\min_{\emptyset \neq S \subseteq V} \frac{\sum_{e \in E(S, \bar{S})} 1}{|S||\bar{S}|}$ .*

Figure 1 is an SDP relaxation for Sparsest Cut. To see that this is indeed a relaxation, for any cut  $(S, \bar{S})$ , consider the following vector assignment: Fix a unit vector  $\mathbf{w}$ . If  $i \in S$ , let  $\mathbf{v}_i := \mathbf{w}/\sqrt{|S||\bar{S}|}$  and if  $i \in \bar{S}$ , let  $\mathbf{v}_i := -\mathbf{w}/\sqrt{|S||\bar{S}|}$ . It is easy to check that this gives a valid SDP solution, and its objective value is equal to the sparsity of the cut.

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

Subject to

$$\begin{aligned} \forall i, j \in V & \quad \|\mathbf{v}_i\| = \|\mathbf{v}_j\| \\ \forall i, j, k \in V & \quad \|\mathbf{v}_i - \mathbf{v}_j\|^2 + \|\mathbf{v}_j - \mathbf{v}_k\|^2 \geq \|\mathbf{v}_i - \mathbf{v}_k\|^2 \\ & \quad \frac{1}{4} \sum_{i < j} \|\mathbf{v}_i - \mathbf{v}_j\|^2 = 1 \end{aligned}$$

Figure 1: SDP relaxation of Sparsest Cut

### 2.2 Balanced Separator

**DEFINITION 2.2 (BALANCED SEPARATOR).** *For a multi-graph  $G = (V, E)$ , and a balance parameter  $b \in (0, 1/2]$  (to be thought of as a fixed constant), the goal is to find a cut  $(S, \bar{S})$  that minimizes  $\sum_{e \in E(S, \bar{S})} 1$ , subject to  $\min\{|S|, |\bar{S}|\} \geq b|V|$ . The cuts that satisfy  $\min\{|S|, |\bar{S}|\} \geq b|V|$  are called  $(b, 1 - b)$  balanced cuts.*

Figure 2 is an SDP relaxation of Balanced Separator with parameter  $b$ . To see that this is indeed a relaxation, fix a

unit vector  $\mathbf{w}$  and let  $\mathbf{v}_i := \mathbf{w}$  or  $\mathbf{v}_i := -\mathbf{w}$  depending on which side of the cut vertex  $i$  belongs to.

The result of Arora, Rao and Vazirani [2] established that the integrality gap of this SDP is at-most  $O(\sqrt{\log n})$ . They further conjectured that the integrality gap is  $O(1)$ . We disprove this conjecture by constructing  $\Omega(\log \log n)$  integrality gap instance for Balanced Separator which also implies the same gap for the Sparsest Cut SDP. The following theorem summarizes our construction.

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

Subject to

$$\forall i \in V \quad \|\mathbf{v}_i\|^2 = 1 \quad (2)$$

$$\forall i, j, k \in V \quad \|\mathbf{v}_i - \mathbf{v}_j\|^2 + \|\mathbf{v}_j - \mathbf{v}_k\|^2 \geq \|\mathbf{v}_i - \mathbf{v}_k\|^2 \quad (3)$$

$$\frac{1}{4} \sum_{i < j} \|\mathbf{v}_i - \mathbf{v}_j\|^2 \geq b \cdot (1 - b) \cdot |V|^2 \quad (4)$$

**Figure 2: SDP relaxation of b- Balanced Separator**

**THEOREM 2.3.** ( *$\Omega(\log \log n)$  Integrality Gap Instance for Balanced Separator*) *There are absolute constants  $c_1, c_2 > 0$  such that, for every large enough  $n$ , there exists a multi-graph  $G(V, E)$  on  $n$  vertices, and a vector assignment  $i \mapsto \mathbf{v}_i$  for every  $i \in V$  such that*

1. Every  $(\frac{1}{3}, \frac{2}{3})$  balanced cut must contain at least  $c_1 |E| \frac{\log \log n}{\log n}$  edges.
2. The vector assignment gives a **low** SDP objective value, i.e.,  $\frac{1}{4} \sum_{e \in \{i,j\}} \|\mathbf{v}_i - \mathbf{v}_j\|^2 \leq c_2 |E| \frac{1}{\log n}$ .
3. The vectors  $\{\mathbf{v}_i \mid i \in V\}$  are **well-separated**, i.e.,  $\frac{1}{4} \sum_{i < j \in V} \|\mathbf{v}_i - \mathbf{v}_j\|^2 \geq n^2$ .
4. The unit vectors  $\{\mathbf{v}_i \mid i \in V\}$  define a  $\ell_2^2$  metric, i.e., the following **triangle inequality** is satisfied:  $\forall i, j, k \in V, \|\mathbf{v}_i - \mathbf{v}_j\|^2 + \|\mathbf{v}_j - \mathbf{v}_k\|^2 \geq \|\mathbf{v}_i - \mathbf{v}_k\|^2$ .

### 2.3 Informal Description of the Construction in Theorem 2.3

We first highlight how the construction in Theorem 2.3 differs from the one in [16] (and also [17]). In these previous papers, the vertex set  $V$  is partitioned into sets  $V_1, V_2, \dots, V_l$ . Property (1) is no longer true, and in fact their graph does have “small” balanced cuts. For instance, there are small balanced cuts that place, for every  $1 \leq j \leq l$ , the entire set  $V_j$  on either side of the cut. This issue is handled by introducing the non-uniform version of the Balanced Separator problem. They define a “piece-wise balanced cut” as a cut that cuts “many” sets  $V_i$  in a balanced manner. Then they show that in their graph, there is no small piece-wise balanced cut. For the non-uniform version of the Balanced Separator problem, it suffices to worry only about the piece-wise balanced cuts.

We, however, need to construct a graph  $G(V, E)$  that has no small balanced cuts. Here is a simple approach: Start with a hypercube  $\mathcal{F} = \{-1, 1\}^N$  and a suitable group action on the  $N$  co-ordinates. The group naturally acts on the set of hypercube vertices and partitions it into “orbits”. We merge all vertices that fall into the same orbit. Call the

resulting multi-graph  $G(V, E)$ . Note that the hypercube has small balanced cuts, namely, the dimensionality cuts which cut  $1/N$  fraction of the edges. However, if the group is reasonable (e.g. transitive), then  $G(V, E)$  does not have small balanced cuts. A balanced cut in the graph  $G(V, E)$  corresponds to a balanced boolean function on the hypercube that is invariant on each orbit. Kahn, Kalai and Linial’s [14] result says that a balanced function must have a co-ordinate with “influence” at least  $\Omega(\log N/N)$ , and if the function is invariant under a transitive group action, all co-ordinates have the same influence. Thus, the sum of all influences is  $\Omega(\log N)$ . This is same as saying that every balanced cut in  $G(V, E)$  must cut  $\Omega(\log N/N)$  fraction of edges. Note that this lower bound is  $\Omega(\log N)$  factor larger than the size of the dimensionality cuts in the hypercube.

Thus, constructing graphs with no small cuts is quite easy (for that matter one could have taken an expander graph!). The tricky part is to construct a graph that also admits an SDP solution satisfying Properties (2), (3) and (4) of Theorem 2.3. The key idea here is to take the  $2N$ -dimensional hypercube  $\{-1, 1\}^N \times \{-1, 1\}^N = \mathcal{F}_1 \times \mathcal{F}_2$  and define a group action which in some sense “entangles” the two hypercubes  $\mathcal{F}_1$  and  $\mathcal{F}_2$  (We note that construction of [16] also uses this feature, but they do not use a group action. They use a more complicated way of partitioning the hypercube into orbits.) To be precise, let  $\Sigma$  be a group of cyclic shifts on the co-ordinates  $[1, 2, \dots, N]$ . Now consider the group action on  $[1, 2, \dots, N, N + 1, N + 2, \dots, 2N]$  which applies a cyclic shift  $\sigma \in \Sigma$  on the first  $N$  co-ordinates, and applies the *same* cyclic shift on the remaining  $N$  co-ordinates. Note that this group action is not transitive on the entire set of  $2N$  co-ordinates, but transitive separately on the first and the last  $N$  co-ordinates. Our graph is then obtained by merging vertices of  $\mathcal{F}_1 \times \mathcal{F}_2$  that fall into the same orbit. We can again show that the graph has no small balanced cuts via the theorem of Kahn, Kalai and Linial [14].

Now we outline the SDP solution. We want to assign one unit vector to each orbit. Here is the basic idea: An orbit consists of  $N$  elements of  $\mathcal{F}_1 \times \mathcal{F}_2$ . Pick one element in the orbit as a representative and call it  $(\mathbf{x}_1, \mathbf{y}_1)$ . Thus, all elements in the orbit are given by

$$(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_N, \mathbf{y}_N)$$

where  $\mathbf{x}_j$  ( $\mathbf{y}_j$  resp.) is  $(j - 1)^{th}$  cyclic shift of  $\mathbf{x}_1$  ( $\mathbf{y}_1$  resp.). We view  $\mathbf{x}_i$  as vectors with  $\pm 1$  co-ordinates and norm  $\sqrt{N}$ . Roughly speaking, the SDP solution assigns the following vector to the orbit:

$$\mathbf{V} := \frac{1}{N} \sum_{j=1}^N \mathbf{y}_1(j) \mathbf{x}_j$$

where  $\mathbf{y}_1(j)$  denotes the  $j^{th}$  co-ordinate of the bit-string  $\mathbf{y}_1$ . Couple of observations: (1) The vector  $\mathbf{V}$  does not depend on the choice of the representative  $(\mathbf{x}_1, \mathbf{y}_1)$ . This is because, in our group action, the *same* cyclic shift is applied to the first  $N$  and last  $N$  co-ordinates. (2) For a typical orbit, the vectors  $\mathbf{x}_j, 1 \leq j \leq N$  are almost orthogonal and therefore  $\mathbf{V}$  has norm close to 1.

This is only the basic idea and the actual SDP solution we construct (as well as the notation) is rather different (see Section 3.2).

## 2.4 Minimum Linear Arrangement

In this section we show that a simple *stretching* of the integrality gap instance for the SDP of Balanced Separator (see Fig. 2) leads to an integrality gap instance for the Minimum Linear Arrangement SDP.

**DEFINITION 2.4** (MINIMUM LINEAR ARRANGEMENT). *Given a multi-graph  $G = (V, E)$  on  $n$  vertices, the goal is to find a permutation  $\pi : V \rightarrow [n]$  that minimizes  $\text{obj}(\pi) := \sum_{e \in \{i,j\} \in E} |\pi(i) - \pi(j)|$ .*

Figure 3 is an SDP relaxation for the Minimum Linear Arrangement problem considered in [6, 11]. To see that this is indeed a relaxation, let  $\pi : V \mapsto [n]$  be any permutation (i.e. an integral solution). Define a metric  $d(i, j) = |\pi(i) - \pi(j)|$ . This is a *line metric* on integer points and hence satisfies the “spreading constraint”:

$$\forall i \in V, \forall S \subseteq V, \quad \sum_{j \in S} d(i, j) \geq (|S|^2 - 1)/4$$

Note that  $d(\cdot, \cdot)$  is an  $\ell_1$  metric, hence a negative type metric, and hence there exist vectors  $\mathbf{v}_i$  such that  $d(i, j) = \|\mathbf{v}_i - \mathbf{v}_j\|^2$ . This gives the SDP solution that achieves the same objective value as the integral solution.

$$\text{Minimize} \quad \sum_{e \in \{i,j\}} \|\mathbf{v}_i - \mathbf{v}_j\|^2 \quad (5)$$

Subject to

$$\begin{aligned} \forall i \in V, \forall S \subseteq V, \quad \sum_{j \in S} \|\mathbf{v}_i - \mathbf{v}_j\|^2 &\geq (|S|^2 - 1)/4 \\ \forall i, j, k \in V, \quad \|\mathbf{v}_i - \mathbf{v}_j\|^2 + \|\mathbf{v}_j - \mathbf{v}_k\|^2 &\geq \|\mathbf{v}_i - \mathbf{v}_k\|^2. \end{aligned}$$

**Figure 3: SDP relaxation of Minimum Linear Arrangement**

**THEOREM 2.5.** [6, 11] *The Integrality Gap of the SDP of Figure 3 is at-most  $O(\sqrt{\log n \log \log n})$ .*

The following theorem summarizes our integrality gap construction for the Minimum Linear Arrangement problem.

**THEOREM 2.6.** *There are absolute constants  $c_1, c_2 > 0$  such that, for every large enough  $n$ , there exists a multi-graph  $G'(V', E')$  on  $n$  vertices, and a vector<sup>3</sup> assignment  $i \mapsto \mathbf{v}'_i$  for every  $i \in V'$  such that*

1. *For every permutations  $\pi : V' \rightarrow [n]$ ,  $\text{obj}(\pi) \geq c_1 n |E'| \frac{\log \log n}{\log n}$ .*
2. *The vector assignment gives a **low** SDP objective value, i.e.,  $\sum_{e \in \{i,j\}} \|\mathbf{v}'_i - \mathbf{v}'_j\|^2 \leq c_2 n |E'| \frac{1}{\log n}$ .*
3. *The vectors  $\{\mathbf{v}'_i \mid i \in V'\}$  satisfy the **spreading constraints**, i.e.,  $\forall i \in V', \forall S \subseteq V', \sum_{j \in S} \|\mathbf{v}'_i - \mathbf{v}'_j\|^2 \geq (|S|^2 - 1)/4$ .*
4. *The vectors  $\{\mathbf{v}'_i \mid i \in V'\}$  satisfy the following triangle inequality:  $\forall i, j, k \in V', \|\mathbf{v}'_i - \mathbf{v}'_j\|^2 + \|\mathbf{v}'_j - \mathbf{v}'_k\|^2 \geq \|\mathbf{v}'_i - \mathbf{v}'_k\|^2$ .*

<sup>3</sup>They need not be unit vectors.

**PROOF.** The integrality gap instance for Minimum Linear Arrangement is the same as that for Balanced Separator described in Theorem 2.3.  $G'(V', E')$  is exactly the same as  $G(V, E)$  as in Theorem 2.3. The vector assignment is  $\mathbf{v}'_i := 2\sqrt{n}\mathbf{v}_i$ , where  $\mathbf{v}_i$  are as in Theorem 2.3. Properties (2) and (4) follow from Properties (2) and (4) in Theorem 2.3 respectively. Proof of Property (3) appear in Section 4.6. Property (1) follows from Property (1) in Theorem 2.3 and Lemma 2.7, which we prove next. The lemma says that a graph with no small balanced cut has no good linear arrangement either.  $\square$

**LEMMA 2.7.** *If a multi-graph  $G(V, E)$  has no  $(\frac{1}{3}, \frac{2}{3})$  balanced cut containing less than  $\mu$  edges, then for every  $\pi : V \rightarrow [n]$ ,  $\text{obj}(\pi) \geq \mu \frac{n}{3}$ .*

**PROOF.** For simplicity, for  $v_i, v_j \in V$ , let  $\text{wt}(v_i, v_j)$  denote the number of edges between  $v_i$  and  $v_j$ . Let  $\pi = v_1, v_2, \dots, v_n$  be the permutation of  $V$  defined as  $\pi(v_i) = i$ . Then,  $\text{obj}(\pi) = \sum_{1 \leq i < j \leq n} \text{wt}(v_i, v_j)(j-i) = \sum_{1 \leq i < j \leq n} \text{wt}(v_i, v_j) \left( \sum_{i \leq k < j} 1 \right) = \sum_{1 \leq k < n} \sum_{1 \leq i \leq k < j \leq n} \text{wt}(v_i, v_j) = \sum_{1 \leq k < n} \text{wt}(S_k, \bar{S}_k)$ , where  $S_k := \{v_1, v_2, \dots, v_k\}$  and  $\text{wt}(S_k, \bar{S}_k)$  is the weight of the cut  $(S_k, \bar{S}_k)$ . Note that for  $\frac{n}{3} \leq k \leq \frac{2n}{3}$ ,  $\text{wt}(S_k, \bar{S}_k) \geq \mu$ . Therefore,  $\text{obj}(\pi) \geq \mu \frac{n}{3}$ .  $\square$

## 3. INTEGRALITY GAP INSTANCE FOR Balanced Separator

In this section, we present the construction stated in Theorem 2.3. The correctness of this construction is proved in Section 4.

### 3.1 Constructing The Graph $G(V, E)$

Let  $N$  be an integer, which we assume to be prime for the rest of the paper.<sup>4</sup> Consider the hypercube  $\mathcal{F} = \{-1, 1\}^{2N}$ . The coordinates of an element in  $\mathcal{F}$  are divided into two blocks, each of which is of size  $N$ . For an element  $\mathbf{u} = (u_1, \dots, u_{2N}) \in \mathcal{F}$ ,  $\mathbf{u}^x := (u_1, \dots, u_N)$  and  $\mathbf{u}^y := (u_{N+1}, \dots, u_{2N})$ .  $\mathbf{u}^x$  and  $\mathbf{u}^y$  are referred to as the  $x$ -part and  $y$ -part of  $\mathbf{u}$  respectively. Let  $\sigma : \{-1, 1\}^N \mapsto \{-1, 1\}^N$  be the *rotation* operator defined as follows:

$$\sigma((u_1, u_2, \dots, u_{N-1}, u_N)) := (u_N, u_1, u_2, \dots, u_{N-1}).$$

Define  $\sigma^i$  recursively as follows:  $\sigma^1 := \sigma$ , and for all  $1 < i \leq N$ ,  $\sigma^i := \sigma \circ \sigma^{i-1}$ . This corresponds to applying the  $\sigma$  operator  $i$  times. The set of rotations  $\mathcal{H} := \{\sigma^i\}_{i=1}^N$  forms a group under composition.  $\mathcal{H}$  acts on  $\mathcal{F}$  as follows: For  $\mathbf{u} = (\mathbf{u}^x, \mathbf{u}^y) \in \mathcal{F}$  and  $\sigma^i \in \mathcal{H}$ ,

$$\sigma^i(\mathbf{u}) := (\sigma^i(\mathbf{u}^x), \sigma^i(\mathbf{u}^y)).$$

Hence,  $\mathcal{H}$  partitions  $\mathcal{F}$  into *orbits*  $\{\mathcal{O}_1, \dots, \mathcal{O}_m\}$ , for some  $m$ . Since  $N$  is a prime, all but four orbits have  $N$  elements each and hence  $m = 4 + (2^{2N} - 4)/N$ . We recall that for  $\mathbf{u}, \mathbf{v} \in \mathcal{F}$ , the inner product  $\langle \mathbf{u}, \mathbf{v} \rangle := \sum_{i=1}^{2N} u_i v_i$ . In terms of the  $x, y$ -parts, it is  $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}^x, \mathbf{v}^x \rangle + \langle \mathbf{u}^y, \mathbf{v}^y \rangle$ . We now identify certain orbits which have a particularly nice structure: the  $x$ -part of all the elements in it are *nearly orthogonal*.

**DEFINITION 3.1** (NEARLY ORTHOGONAL ORBIT). *An orbit  $\mathcal{O} \in \{\mathcal{O}_1, \dots, \mathcal{O}_m\}$  is said to be nearly orthogonal if it has*

<sup>4</sup>This assumption is not strictly necessary, but makes some of the proofs easier.

$N$  elements and for all  $\mathbf{u} \neq \mathbf{v} \in \mathcal{O}$

$$|\langle \mathbf{u}^x, \mathbf{v}^x \rangle| \leq 8\sqrt{N \log N}.$$

Without loss of generality, let  $\{\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_n\}$  be the set of orbits each of which is nearly orthogonal. The following lemma is a simple consequence of Chernoff Bounds. The proof appears in Section 4.2.

LEMMA 3.2 (MOST ORBITS ARE NEARLY ORTHOGONAL).

For every large enough  $N$ , the number  $n$  of nearly orthogonal orbits satisfies  $m \geq n \geq (1 - 4/N^2)m$ .

### The vertices of $G(V, E)$

There is a vertex (call it  $\mathcal{O}$ ) for every orbit  $\mathcal{O}$  which is nearly orthogonal, i.e., for every orbit in the set  $\{\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_n\}$ . Henceforth, we will only refer to nearly orthogonal orbits. We use the notation  $\mathcal{O} < \mathcal{O}'$  if orbit  $\mathcal{O}$  appears before the orbit  $\mathcal{O}'$  in the above canonical ordering.

### The edges of $G(V, E)$

Let  $\Delta(\cdot, \cdot)$  denote the Hamming distance. If there are  $\mathbf{u} \in \mathcal{O}$ ,  $\mathbf{v} \in \mathcal{O}'$  with  $\Delta(\mathbf{u}, \mathbf{v}) = 1$ , add an edge between  $\mathcal{O}$  and  $\mathcal{O}'$ . Note that if  $\Delta(\mathbf{u}, \mathbf{v}) = 1$ , then  $\Delta(\sigma^j(\mathbf{u}), \sigma^j(\mathbf{v})) = 1$  for every  $1 \leq j \leq N$ , and hence, there are exactly  $N$  edges between  $\mathcal{O}$  and  $\mathcal{O}'$ . Thus, edges in the multi-graph  $G(V, E)$  are in one-to-one correspondence with edges in the hypercube  $\{-1, 1\}^{2N}$ , except for the edges incident on  $\{-1, 1\}^{2N} \setminus \{\mathcal{O}_1, \dots, \mathcal{O}_n\}$ . Since almost all orbits are orthogonal,  $|E| = (1 - O(1/N^2))N \cdot 2^{2N}$ , where  $N \cdot 2^{2N}$  is the number of edges of the hypercube.

The following theorem, proved in Section 4.1, establishes that this graph has no small balanced cut. This is a consequence of a Fourier analytic result due to Kahn, Kalai and Linal [14].

THEOREM 3.3. *There is an absolute constant  $c > 0$ , such that every  $(\frac{1}{3}, \frac{2}{3})$  balanced cut in the graph  $G(V, E)$  cuts at least  $c \frac{\log \log n}{\log n}$  fraction of the edges.*

## 3.2 The SDP Solution

We now show how to associate vectors  $\mathcal{O} \mapsto \mathbf{V}_{\mathcal{O}}$  to vertices of  $G(V, E)$  so that Properties (2), (3), and (4) of Theorem 2.3 are satisfied. Fix integers  $r = 2^{12}$ ,  $s = 10$  and  $t = 2^{10^6} + 1$ . Write any orbit  $\mathcal{O}$ , as

$$\mathcal{O} = \{\mathbf{V}_{\mathcal{O},1}, \mathbf{V}_{\mathcal{O},2}, \dots, \mathbf{V}_{\mathcal{O},N}\}$$

where  $\mathbf{V}_{\mathcal{O},1}$  is fixed (arbitrarily) as the representative element of the orbit and  $\mathbf{V}_{\mathcal{O},j} = \sigma^{j-1}(\mathbf{V}_{\mathcal{O},1})$  for  $1 \leq j \leq N$ .

Write  $\mathbf{V}_{\mathcal{O},j} = (\mathbf{V}_{\mathcal{O},j}^x, \mathbf{V}_{\mathcal{O},j}^y)$ . Note that the set of vectors  $\{\frac{1}{\sqrt{N}}\mathbf{V}_{\mathcal{O},j}^x\}_{j=1}^N$  is a nearly orthogonal set of vectors, each with unit norm (every pairwise dot-product is bounded by  $O(\sqrt{\log N/N})$ ). We take high enough tensor powers of these vectors so that they become even more near-orthogonal. In particular, the vectors

$$\mathbf{T}_{\mathcal{O},j}^x := \left( \frac{1}{\sqrt{N}} \mathbf{V}_{\mathcal{O},j}^x \right)^{\otimes r} \quad 1 \leq j \leq N$$

are unit vectors with pairwise dot-products bounded by  $(O(\sqrt{\log N/N}))^r \leq 1/N^{r/3}$ . Now we apply Gram-Schmidt orthogonalization process to these vectors and obtain vectors  $\mathbf{W}_{\mathcal{O},j}^x, 1 \leq j \leq N$ . Since we apply Gram-Schmidt pro-

cess on vectors that are already nearly orthogonal, the resulting vectors are very close to the original ones. To be precise,  $\|\mathbf{W}_{\mathcal{O},j}^x - \mathbf{T}_{\mathcal{O},j}^x\| \leq 1/N^{r/10}$  (Lemma 4.6).

We are ready to assign a vector  $\mathbf{V}_{\mathcal{O}}$  to the orbit  $\mathcal{O}$ . Consider the representative element for the orbit  $\mathbf{V}_{\mathcal{O},1} = (\mathbf{V}_{\mathcal{O},1}^x, \mathbf{V}_{\mathcal{O},1}^y)$ . Let  $\mathbf{V}_{\mathcal{O},1}^y = (y_1, y_2, \dots, y_N) \in \{-1, 1\}^N$ . Define

$$\mathbf{V}_{\mathcal{O}} := \left( \frac{1}{\sqrt{N}} \sum_{j=1}^N y_j (\mathbf{W}_{\mathcal{O},j}^x)^{\otimes 2s} \right)^{\otimes t}.$$

Note that  $\mathbf{V}_{\mathcal{O}}$  is a unit vector because of orthonormality of vectors  $\mathbf{W}_{\mathcal{O},j}^x, 1 \leq j \leq N$ . The following lemmata establish Properties (2), (3) and (4) of Theorem 2.3. The proofs appear in Sections 4.3, 4.4 and 4.5 respectively. The most technical part is proving the triangle inequality. The proof is tedious and proceeds along similar lines as in [16]. We need to keep track of (the negligible) error terms introduced by the Gram-Schmidt orthogonalization process.

LEMMA 3.4 (LOW OBJECTIVE VALUE). *There is a fixed constant  $c > 0$  such that*

$$\frac{1}{4} \sum_{\substack{e \in E \\ \mathcal{O}, \mathcal{O}' \text{ are endpoints of } e}} \|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 \leq c \cdot |E| \frac{1}{\log n}.$$

LEMMA 3.5 (WELL-SEPARATEDNESS).

$$\frac{1}{4} \sum_{\mathcal{O} < \mathcal{O}' \in V} \|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 \geq n^2.$$

LEMMA 3.6 (TRIANGLE INEQUALITY).

$$\forall \mathcal{O}, \mathcal{O}', \mathcal{O}'' \in V, \|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 + \|\mathbf{V}_{\mathcal{O}'} - \mathbf{V}_{\mathcal{O}''}\|^2 \geq \|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}''}\|^2.$$

## 4. PROOFS

### 4.1 The Instance Has No Small Balanced Cuts

In this section we prove Theorem 3.3. Our proof relies on the following Fourier analytic result due to Kahn, Kalai and Linal [14]. First, we need a notion of the *influence*: Let  $f : \{-1, 1\}^K \rightarrow \{-1, 1\}$  be a boolean function. Let  $\mathbf{e}_j \in \{-1, 1\}^K$  be the vector containing  $-1$  at the  $j$ -th position and  $1$  at all other positions. Define  $\text{Inf}_j(f) := \Pr_{\mathbf{x} \in_R \{-1, 1\}^K} [f(\mathbf{x} \cdot \mathbf{e}_j) \neq f(\mathbf{x})]$ . In words, viewing  $f$  as a cut in the hypercube, influence of  $j^{\text{th}}$  co-ordinate equals the fraction of edges along the  $j^{\text{th}}$  dimension which are cut.

THEOREM 4.1. [14] *If  $f$  is a  $(\frac{1}{3}, \frac{2}{3})$  balanced boolean function on  $\{-1, 1\}^K$ , then there is a  $j \in [K]$  such that*

$$\text{Inf}_j(f) \geq c \frac{\log K}{K}$$

for some absolute constant  $c > 0$ .

PROOF OF THEOREM 3.3. Let  $C \subseteq V$  be a  $(\frac{1}{3}, \frac{2}{3})$  balanced cut in the instance graph. We need to lower bound the size of this cut. A cut  $C$  is viewed as a boolean function  $C : V \mapsto \{-1, 1\}$ . This naturally induces a cut  $C' : \{-1, 1\}^{2N} \mapsto \{-1, 1\}$  as follows: for  $\mathbf{u} \in \mathcal{O}_i, 1 \leq i \leq n$ , let  $C'(\mathbf{u}) := C(\mathcal{O}_i)$ . Without loss of generality assume that  $C'$  takes the value  $-1$  more often. For all points  $\mathbf{u} \in \{-1, 1\}^{2N} \setminus \{\mathcal{O}_1, \dots, \mathcal{O}_n\}$ , let  $C'(\mathbf{u}) = 1$ . This ensures that  $C'$  is also a  $(\frac{1}{3}, \frac{2}{3})$  balanced cut of  $\{-1, 1\}^{2N}$ . Note that  $C'$  is a boolean function invariant on each orbit.

For  $1 \leq i \leq 2N$ , let  $E_i$  denote the set of edges of dimension  $i$  in  $C'$ . Formally,  $E_i := \{\{\mathbf{x}, \mathbf{x} \cdot \mathbf{e}_i\} : \mathbf{x} \in \{-1, 1\}^{2N}\} \cap C'$ . Note that all the  $E_i$ 's are mutually disjoint. Hence,  $|C'| \geq \sum_{i=1}^{2N} |E_i|$ . By Theorem 4.1, there is a  $1 \leq j \leq 2N$ , such that  $|E_j| = \Omega\left(\frac{2^{2N} \log(2N)}{N}\right)$ . Without loss of generality, let  $1 \leq j \leq N$ . Since the cut  $C'$  is invariant on each orbit, the dimensions  $\{1, \dots, N\}$  should all have the same influence on  $C'$ , and hence,  $|E_i| = |E_j|$  for  $1 \leq i \leq N$ . Hence,  $|C'| \geq \Omega(2^{2N} \log(2N))$ .

Finally, we observe that the edge set  $E$  of the graph  $G(V, E)$  includes all but  $O(2^{2N}/N)$  edges of the hypercube  $\{-1, 1\}^{2N}$ . Thus

$$|C| \geq |C'| - O(2^{2N}/N) = \Omega\left(2^{2N} \log(2N)\right) = \Omega(|E| \frac{\log \log n}{\log n})$$

concluding the proof.  $\square$

## 4.2 Most Orbits are Orthogonal

In this section we give the proof of Lemma 3.2. Recall that an orbit  $O$  is nearly orthogonal if for all  $\mathbf{u}, \mathbf{v} \in O$  with distinct  $x$ -parts,

$$|\langle \mathbf{u}^x, \mathbf{v}^x \rangle| \leq 8\sqrt{N \log N}.$$

The following version of Chernoff Bound would be needed for the proof.

**THEOREM 4.2.** *If  $X_1, X_2, \dots, X_N$  are independent random variables where each  $X_i \in_{1/2} \{-1, 1\}$ , then for any  $\lambda > 0$*

$$\Pr \left[ |X_1 + X_2 + \dots + X_N| \geq \lambda \sqrt{N} \right] \leq 2 \exp(-\lambda^2/4).$$

Note that we have chosen  $N$  to be a large odd prime number greater than 3. This ensures that every orbit is of size  $N$ , except the ones containing  $\mathbf{1}$  and  $-\mathbf{1}$ . In particular we prove the following lemma.

**LEMMA 4.3.**

$$\Pr_{\mathbf{x} \in_{1/2} \{-1, 1\}^N} \left[ \exists 1 \leq l \leq N, |\langle \mathbf{x}, \sigma^l(\mathbf{x}) \rangle| \geq 8\sqrt{N \log N} \right] \leq 4/N^3.$$

**PROOF.** Let  $\mathbf{x} := (x_1 x_2 \dots x_N)$ . Then

$$\begin{aligned} \langle \mathbf{x}, \sigma(\mathbf{x}) \rangle &= x_1 x_2 + x_2 x_3 + \dots + x_{N-1} x_N + x_N x_1 \\ &= (x_1 x_2 + x_3 x_4 + \dots + x_{2i-1} x_{2i} + \dots + x_N x_1) \\ &\quad + (x_2 x_3 + x_4 x_5 + \dots + x_{2i} x_{2i+1} + \dots + x_{N-1} x_N) \\ &=: X + Y \end{aligned}$$

where we let  $X := (x_1 x_2 + x_3 x_4 + \dots + x_{2i-1} x_{2i} + \dots + x_{N-2} x_{N-1} + x_N x_1)$  and  $Y := (x_2 x_3 + x_4 x_5 + \dots + x_{2i} x_{2i+1} + \dots + x_{N-1} x_N)$ . For a randomly chosen  $\mathbf{x}$ ,  $X$  is the sum of  $(N+1)/2$  independent random variables, where each variable is 1 with probability  $1/2$ , and  $-1$  with probability  $1/2$ . Similarly,  $Y$  is the sum of  $(N-1)/2$  such independent random variables. We now analyse the probability that  $|\langle \mathbf{x}, \sigma(\mathbf{x}) \rangle|$  is greater than  $8\sqrt{N \log N}$ . All the probabilities are over  $\mathbf{x}$  chosen uniformly at random from  $\{-1, 1\}^N$ .

$$\begin{aligned} \Pr_{\mathbf{x}}[|\langle \mathbf{x}, \sigma(\mathbf{x}) \rangle| > 8\sqrt{N \log N}] &= \Pr_{\mathbf{x}}[|X + Y| > 8\sqrt{N \log N}] \\ &\leq \Pr_{\mathbf{x}}[|X| > 4\sqrt{N \log N}] \\ &\quad + \Pr_{\mathbf{x}}[|Y| > 4\sqrt{N \log N}] \\ &\leq 4/N^4. \end{aligned}$$

The last inequality follows from Theorem 4.2. Exactly the same analysis is true for  $\sigma^l$  instead of  $\sigma$ . The lemma follows by taking the union bound over all  $l$ 's.  $\square$

Note that this lemma actually implies that for all  $1 \leq l, l' \leq N$ ,  $\left| \langle \sigma^l(\mathbf{u}^x), \sigma^{l'}(\mathbf{v}^x) \rangle \right| \leq 8\sqrt{N \log N}$  where  $\mathbf{u}^x, \mathbf{v}^x \in \{-1, 1\}^N$  are random and independent. Therefore, except for  $4/N^3$  fraction of orbits, every orbit is nearly orthogonal.

## 4.3 Low SDP Optimum

In this section we show that the SDP solution that we constructed has a low optimum (Lemma 3.4). We show that if there is an edge between  $O$  and  $O'$ , then  $\|\mathbf{V}_O - \mathbf{V}_{O'}\|^2 \leq O(1/N)$ . Thus

$$\sum_{\substack{e \in E, \\ O, O' \text{ are endpoints of } e}} \|\mathbf{V}_O - \mathbf{V}_{O'}\|^2 \leq O\left(\frac{|E|}{\log n}\right).$$

Recall that there is an edge between  $O$  and  $O'$  if for some  $k$ ,  $\Delta(\mathbf{V}_{O,1}, \mathbf{V}_{O',k}) = 1$ . We consider two cases:

**Case 1:**  $\Delta(\mathbf{V}_{O,1}^y, \mathbf{V}_{O',k}^y) = 1$  and  $\mathbf{V}_{O,1}^x = \mathbf{V}_{O',k}^x$ .

We may assume that  $\mathbf{W}_{O,i}^x = \mathbf{W}_{O',i+k-1}^x$ . For notational convenience let  $\mathbf{u}_i := (\mathbf{W}_{O,i}^x)^{\otimes 2s}$ ,  $\mathbf{V}_{O,1}^y = (y_1, \dots, y_N)$ , and  $\mathbf{V}_{O',k}^y = (y'_1, \dots, y'_N)$ . Since  $\Delta(\mathbf{V}_{O,1}^y, \mathbf{V}_{O',k}^y) = 1$ ,  $\sum_{i=1}^N y_i y'_i = N - 2$ . Recall that  $\{\mathbf{u}_i\}_i$  is a set of orthonormal vectors. By definition  $\mathbf{V}_O = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y_i \mathbf{u}_i\right)^{\otimes t}$  and  $\mathbf{V}_{O'} = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y'_i \mathbf{u}_i\right)^{\otimes t}$ . Hence,

$$\begin{aligned} \langle \mathbf{V}_O, \mathbf{V}_{O'} \rangle^{1/t} &= \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y_i \mathbf{u}_i\right) \cdot \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y'_i \mathbf{u}_i\right) \\ &= \frac{1}{N} \sum_{i=1}^N y_i y'_i = \frac{N-2}{N}. \end{aligned}$$

This implies that  $\langle \mathbf{V}_O, \mathbf{V}_{O'} \rangle = (1 - 2/N)^t \geq 1 - 2t/N$ , and hence,

$$\|\mathbf{V}_O - \mathbf{V}_{O'}\|^2 = 2(1 - \langle \mathbf{V}_O, \mathbf{V}_{O'} \rangle) \leq \frac{4t}{N}.$$

**Case 2:**  $\Delta(\mathbf{V}_{O,1}^x, \mathbf{V}_{O',k}^x) = 1$  and  $\mathbf{V}_{O,1}^y = \mathbf{V}_{O',k}^y$ .

In this case let  $\mathbf{u}_i := (\mathbf{W}_{O,i}^x)^{\otimes 2s}$ ,  $\mathbf{v}_i := (\mathbf{W}_{O',i+k-1}^x)^{\otimes 2s}$  and  $\mathbf{V}_{O,1}^y = (y_1, \dots, y_N)$ . Again,  $\{\mathbf{u}_i\}_i$  and  $\{\mathbf{v}_i\}_i$  are sets of orthonormal vectors. Later in the proof we show that  $\langle \mathbf{u}_i, \mathbf{v}_i \rangle = 1 - O(1/N)$  for all  $i$ , and  $\langle \mathbf{u}_i, \mathbf{v}_j \rangle = o(1/N^2)$  for all  $i \neq j$ . Now by definition,  $\mathbf{V}_O = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y_i \mathbf{u}_i\right)^{\otimes t}$  and  $\mathbf{V}_{O'} = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y_i \mathbf{v}_i\right)^{\otimes t}$ . Hence,

$$\begin{aligned} \langle \mathbf{V}_O, \mathbf{V}_{O'} \rangle^{1/t} &= \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y_i \mathbf{u}_i\right) \cdot \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y_i \mathbf{v}_i\right) \\ &= \frac{1}{N} \left( \sum_{i=1}^N \langle \mathbf{u}_i, \mathbf{v}_i \rangle \right) + \frac{1}{N} \left( \sum_{i \neq j} y_i y_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle \right). \end{aligned}$$

This is at-least  $1 - O(1/N) - \frac{1}{N} N^2 o(1/N^2) \geq 1 - O(1/N)$ . This implies that  $\langle \mathbf{V}_O, \mathbf{V}_{O'} \rangle \geq 1 - O(1/N)$ , and hence,

$$\|\mathbf{V}_O - \mathbf{V}_{O'}\|^2 = 2(1 - \langle \mathbf{V}_O, \mathbf{V}_{O'} \rangle) \leq O\left(\frac{1}{N}\right).$$

We now show that  $\langle \mathbf{u}_i, \mathbf{v}_i \rangle = 1 - O(1/N)$  for all  $i$ . The fact that  $\Delta(\mathbf{V}_{O,1}^x, \mathbf{V}_{O',k}^x) = 1$  implies that  $\langle \mathbf{V}_{O,i}^x, \mathbf{V}_{O',i+k-1}^x \rangle =$

$N - 2$ . This implies that  $\langle \mathbf{T}_{\mathcal{O},i}^x, \mathbf{T}_{\mathcal{O}',i+k-1}^x \rangle \geq 1 - 2r/N$ . Hence,  $\langle \mathbf{W}_{\mathcal{O},i}^x, \mathbf{W}_{\mathcal{O}',i+k-1}^x \rangle \geq 1 - O(1/N)$ , and thus,  $\langle \mathbf{u}_i, \mathbf{v}_i \rangle \geq 1 - O(1/N)$ . The second last implication follows from Lemma 4.4. Now suppose that  $i \neq j$ . We need to show that  $\langle \mathbf{u}_i, \mathbf{v}_j \rangle = o(1/N^2)$ . Since the set of vectors  $\{\mathbf{W}_{\mathcal{O},i}^x\}_{\mathcal{O},i}$  satisfy the triangle inequality (see Lemma 4.8),  $1 + \langle \mathbf{W}_{\mathcal{O},i}^x, \mathbf{W}_{\mathcal{O},j}^x \rangle \geq \langle \mathbf{W}_{\mathcal{O},i}^x, \mathbf{W}_{\mathcal{O}',j+k-1}^x \rangle + \langle \mathbf{W}_{\mathcal{O},j}^x, \mathbf{W}_{\mathcal{O}',j+k-1}^x \rangle$ . Hence,  $\langle \mathbf{W}_{\mathcal{O},i}^x, \mathbf{W}_{\mathcal{O}',j+k-1}^x \rangle \leq 1 - \langle \mathbf{W}_{\mathcal{O},j}^x, \mathbf{W}_{\mathcal{O}',j+k-1}^x \rangle \leq O(1/N)$  because  $\langle \mathbf{W}_{\mathcal{O},i}^x, \mathbf{W}_{\mathcal{O},j}^x \rangle = 0$  and  $\langle \mathbf{W}_{\mathcal{O},j}^x, \mathbf{W}_{\mathcal{O}',j+k-1}^x \rangle \geq 1 - O(1/N)$ . Thus,  $|\langle \mathbf{u}_i, \mathbf{v}_j \rangle| = \langle \mathbf{W}_{\mathcal{O},i}^x, \mathbf{W}_{\mathcal{O}',j+k-1}^x \rangle^{2s} \leq o(1/N^2)$ .

LEMMA 4.4. *If the unit vectors  $\mathbf{W}, \mathbf{W}', \mathbf{T}, \mathbf{T}'$  are such that  $\|\mathbf{W} - \mathbf{T}\|^2, \|\mathbf{W}' - \mathbf{T}'\|^2 \leq O(1/N)$ , and  $\langle \mathbf{T}, \mathbf{T}' \rangle \geq 1 - O(1/N)$ , then  $\langle \mathbf{W}, \mathbf{W}' \rangle \geq 1 - O(1/N)$ .*

PROOF. It is easy to check that for  $a, b, c, d \geq 0$ , if  $a \leq b + c + d$ , then  $a^2 \leq 3(b^2 + c^2 + d^2)$ . Therefore, by the triangle inequality on the  $l_2$  norm of the vectors, it follows that

$$\|\mathbf{W} - \mathbf{W}'\|^2 \leq 3(\|\mathbf{W} - \mathbf{T}\|^2 + \|\mathbf{T} - \mathbf{T}'\|^2 + \|\mathbf{T}' - \mathbf{W}'\|^2).$$

Since each term in the RHS is  $O(1/N)$ , we get that  $1 - \langle \mathbf{W}, \mathbf{W}' \rangle = \frac{1}{2}\|\mathbf{W} - \mathbf{W}'\|^2 \leq O(1/N)$ .  $\square$

#### 4.4 The SDP Solution is ‘‘Well-Separated’’

In this section we prove Lemma 3.5. First we recall the SDP solution. For a nearly orthogonal orbit  $\mathcal{O}$ , the unit vector associated to it is

$$\mathbf{v}_{\mathcal{O}} := \left( \frac{1}{\sqrt{N}} \sum_{j=1}^N y_j (\mathbf{W}_{\mathcal{O},j}^x)^{\otimes 2s} \right)^{\otimes t}$$

where  $\mathbf{W}_{\mathcal{O},j}^x$ ,  $1 \leq j \leq N$ , are obtained by orthogonalizing the vectors,  $\mathbf{T}_{\mathcal{O},j}^x := \left( \frac{1}{\sqrt{N}} \mathbf{v}_{\mathcal{O},j}^y \right)^{\otimes r}$ , for  $1 \leq j \leq N$ . Observe that if an orbit  $\mathcal{O}$  has as its representative  $\mathbf{V}_{\mathcal{O},1} = (\mathbf{V}_{\mathcal{O},1}^x, \mathbf{V}_{\mathcal{O},1}^y)$ , there is a unique (and distinct from  $\mathcal{O}$ ) nearly orthogonal orbit  $\mathcal{O}''$  represented by  $\mathbf{V}_{\mathcal{O}'',1} = (\mathbf{V}_{\mathcal{O},1}^x, -\mathbf{V}_{\mathcal{O},1}^y)$ . Since the Gram-Schmidt orthogonalization is same for orbits identical in the  $x$ -part, and the fact that  $t$  is odd, we deduce that if  $\mathbf{V}_{\mathcal{O}}$  is part of our solution then so is  $-\mathbf{V}_{\mathcal{O}}$ . Now, we use the simple fact that if  $\mathbf{a}, \mathbf{b}$  are unit vectors then,  $\|\mathbf{a} - \mathbf{b}\|^2 + \|\mathbf{a} + \mathbf{b}\|^2 = 4$ . Hence, for any orbit  $\mathcal{O}$ ,  $\sum_{\mathcal{O}': \mathcal{O}' \neq \mathcal{O}} \|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 = 2n$ , where  $n$  is the number of orbits. Hence,  $\sum_{\mathcal{O}' < \mathcal{O}} \|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 = n^2$ . This proves the well-separatedness condition.

#### 4.5 The Triangle Inequality

Consider any three orbits  $\mathcal{O}_1, \mathcal{O}_2$  and  $\mathcal{O}_3$ . We will prove that the vectors  $\mathbf{V}_{\mathcal{O}_1}, \mathbf{V}_{\mathcal{O}_2}$  and  $\mathbf{V}_{\mathcal{O}_3}$  satisfy the triangle inequality. Recall the definition of these vectors.

$$\mathbf{V}_{\mathcal{O}_1} = \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N y_{1i} \mathbf{W}_{\mathcal{O}_1,i}^x \right)^{\otimes t} =: \mathbf{U}_1^{\otimes t},$$

$$\mathbf{V}_{\mathcal{O}_2} = \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N y_{2i} \mathbf{W}_{\mathcal{O}_2,i}^x \right)^{\otimes t} =: \mathbf{U}_2^{\otimes t},$$

$$\mathbf{V}_{\mathcal{O}_3} = \left( \frac{1}{\sqrt{N}} \sum_{i=1}^N y_{3i} \mathbf{W}_{\mathcal{O}_3,i}^x \right)^{\otimes t} =: \mathbf{U}_3^{\otimes t}.$$

In this notation, we need to show that

$$1 + \langle \mathbf{U}_2^{\otimes t}, \mathbf{U}_3^{\otimes t} \rangle \geq \langle \mathbf{U}_1^{\otimes t}, \mathbf{U}_2^{\otimes t} \rangle + \langle \mathbf{U}_1^{\otimes t}, \mathbf{U}_3^{\otimes t} \rangle.$$

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  $|\langle \mathbf{U}_1^{\otimes t}, \mathbf{U}_3^{\otimes t} \rangle| \geq 1/3$ . This implies that  $|\langle \mathbf{U}_1, \mathbf{U}_3 \rangle|^t \geq 1/3$ , and therefore,  $|\langle \mathbf{U}_1, \mathbf{U}_3 \rangle| = 1 - \eta'$ , for some  $\eta' = O(1/t)$ . Hence,

$$\max_{1 \leq i, j \leq N} |\langle \mathbf{W}_{\mathcal{O}_1,i}^x, \mathbf{W}_{\mathcal{O}_3,j}^x \rangle| = 1 - \eta$$

for some  $\eta \leq 2^{-100s}$ . Since in our proof we consider  $y_1, y_2$  and  $y_3$  to take any in value  $\{-1, 1\}$ , we may relabel, if necessary, and assume that  $\langle \mathbf{W}_{\mathcal{O}_1,1}^x, \mathbf{W}_{\mathcal{O}_3,1}^x \rangle = 1 - \eta$ .

Note that we need to show that

$$1 + \langle \mathbf{U}_2, \mathbf{U}_3 \rangle^t \geq \langle \mathbf{U}_1, \mathbf{U}_2 \rangle^t + \langle \mathbf{U}_1, \mathbf{U}_3 \rangle^t.$$

By Lemma 4.9 it suffices to show that

$$1 + \langle \mathbf{U}_2, \mathbf{U}_3 \rangle \geq \langle \mathbf{U}_1, \mathbf{U}_2 \rangle + \langle \mathbf{U}_1, \mathbf{U}_3 \rangle.$$

We consider two cases:

CASE 1. Two of the three orbits have the same  $x$ -parts (after a suitable number of rotations). We may assume that  $\mathcal{O}_2$  and  $\mathcal{O}_3$  have the same  $x$ -part. Since the Gram Schmidt orthogonalization process yields the same set of vectors for orbits with identical  $x$ -parts, we may assume that  $\mathbf{W}_{\mathcal{O}_2,i}^x = \mathbf{W}_{\mathcal{O}_3,i}^x$  for all  $i \in [N]$ . After multiplying the triangle inequality by  $N$  and simplifying, we obtain the following inequality that needs to be proved:

$$N + \sum_{i=1}^N y_{2i} y_{3i} \geq \left\langle \sum_{j=1}^N (y_{2j} + y_{3j}) (\mathbf{W}_{\mathcal{O}_2,j}^x)^{\otimes 2s}, \sum_{j=1}^N y_{1j} (\mathbf{W}_{\mathcal{O}_1,j}^x)^{\otimes 2s} \right\rangle.$$

This is true since, for every  $i \in [N]$ ,  $1 + y_{2i} y_{3i} \geq (y_{2i} + y_{3i}) \sum_{j=1}^N y_{1j} \langle \mathbf{W}_{\mathcal{O}_2,i}^x, \mathbf{W}_{\mathcal{O}_1,j}^x \rangle^{2s}$ . Here we use the fact that  $\{\mathbf{W}_{\mathcal{O}_1,j}^x\}_j$  forms an orthonormal set of vectors. Hence, the triangle inequality for this special case is proved.

CASE 2. In this case no two orbits have the same  $x$ -parts. Therefore for any  $i \neq j$

$$|\langle \mathbf{T}_{\mathcal{O}_1,i}^x, \mathbf{T}_{\mathcal{O}_2,j}^x \rangle|, |\langle \mathbf{T}_{\mathcal{O}_1,i}^x, \mathbf{T}_{\mathcal{O}_3,j}^x \rangle|, |\langle \mathbf{T}_{\mathcal{O}_2,i}^x, \mathbf{T}_{\mathcal{O}_3,j}^x \rangle| \leq \left(1 - \frac{1}{N}\right)^r.$$

Since  $r$  is large enough, applying lemma 4.7 we obtain that

$$|\langle \mathbf{W}_{\mathcal{O}_1,i}^x, \mathbf{W}_{\mathcal{O}_2,j}^x \rangle|, |\langle \mathbf{W}_{\mathcal{O}_1,i}^x, \mathbf{W}_{\mathcal{O}_3,j}^x \rangle|, |\langle \mathbf{W}_{\mathcal{O}_2,i}^x, \mathbf{W}_{\mathcal{O}_3,j}^x \rangle| \leq 1 - \frac{1}{N}.$$

As noted before, we may assume that  $\langle \mathbf{W}_{\mathcal{O}_1,1}^x, \mathbf{W}_{\mathcal{O}_3,1}^x \rangle = 1 - \eta$ , and hence, by Lemma 4.5 and Lemma 4.7, for large enough  $N$ ,

$$1 - \eta - \frac{1}{N^{r/10-1}} \leq \langle \mathbf{W}_{\mathcal{O}_1,1}^x, \mathbf{W}_{\mathcal{O}_3,1}^x \rangle, \langle \mathbf{W}_{\mathcal{O}_1,2}^x, \mathbf{W}_{\mathcal{O}_3,2}^x \rangle, \dots \\ \dots, \langle \mathbf{W}_{\mathcal{O}_1,N}^x, \mathbf{W}_{\mathcal{O}_3,N}^x \rangle \leq 1 - \eta.$$

Let  $\alpha := \max_{1 \leq i, j \leq N} |\langle \mathbf{W}_{\mathcal{O}_1,i}^x, \mathbf{W}_{\mathcal{O}_2,j}^x \rangle|$ . We may assume, w.l.o.g., that the maximum is achieved for  $\mathbf{W}_{\mathcal{O}_1,1}^x, \mathbf{W}_{\mathcal{O}_2,1}^x$  and therefore,

$$\alpha - \frac{1}{N^{r/10-1}} \leq \langle \mathbf{W}_{\mathcal{O}_1,1}^x, \mathbf{W}_{\mathcal{O}_2,1}^x \rangle, \langle \mathbf{W}_{\mathcal{O}_1,2}^x, \mathbf{W}_{\mathcal{O}_2,2}^x \rangle, \dots \\ \dots, \langle \mathbf{W}_{\mathcal{O}_1,N}^x, \mathbf{W}_{\mathcal{O}_2,N}^x \rangle \leq \alpha.$$

Now, letting  $\mathbf{w}_i := \mathbf{W}_{\mathcal{O}_1, i}^x$ ,  $\mathbf{u}_i := \mathbf{W}_{\mathcal{O}_2, i}^x$ , and  $\mathbf{v}_i := \mathbf{W}_{\mathcal{O}_3, i}^x$  the desired inequality follows from Lemma 4.11 where Lemma 4.8 is used to make sure that the part of the hypothesis which requires that the set  $\{\mathbf{W}_{\mathcal{O}_i, i}^x\}_{\mathcal{O}_i}$  satisfies the triangle inequality.

LEMMA 4.5. *For any two orbits  $\mathcal{O}$ ,  $\mathcal{O}'$ , for any given  $j, k \in [N]$ ,*

$$\langle \mathbf{T}_{\mathcal{O}, j}^x, \mathbf{T}_{\mathcal{O}', k}^x \rangle = \langle \mathbf{T}_{\mathcal{O}, i+j}^x, \mathbf{T}_{\mathcal{O}', i+k}^x \rangle$$

for all  $i \in [N]$ .

PROOF. This follows from the fact that  $\mathbf{V}_{\mathcal{O}, j+k}^x = \sigma^k(\mathbf{V}_{\mathcal{O}, j}^x)$ , and therefore  $\langle \mathbf{V}_{\mathcal{O}, j}^x, \mathbf{V}_{\mathcal{O}', k}^x \rangle = \langle \mathbf{V}_{\mathcal{O}, i+j}^x, \mathbf{V}_{\mathcal{O}', i+k}^x \rangle$  for all  $i \in [N]$ .  $\square$

LEMMA 4.6. *For any orbit  $\mathcal{O}$ ,  $\|\mathbf{W}_{\mathcal{O}, i}^x - \mathbf{T}_{\mathcal{O}, i}^x\| \leq 1/N^{r/10}$  for all  $i \in [N]$ .*

PROOF. This follows from the fact that for any orbit  $\mathcal{O}$ , for  $i \neq j$ ,  $|\langle \mathbf{T}_{\mathcal{O}, i}^x, \mathbf{T}_{\mathcal{O}, j}^x \rangle| \leq 1/N^{r/3}$ . Therefore, applying the Gram-Schmidt orthogonalization process on the  $N$  vectors in an orbit does not change the norm of the vectors by more than  $1/N^{r/10}$ .

A proof of this fact is presented in the full version [9].  $\square$

LEMMA 4.7. *Given any two orbits  $\mathcal{O}$  and  $\mathcal{O}'$ , for any  $i, j \in [N]$ ,  $|\langle \mathbf{W}_{\mathcal{O}, i}^x, \mathbf{W}_{\mathcal{O}', j}^x \rangle - \langle \mathbf{T}_{\mathcal{O}, i}^x, \mathbf{T}_{\mathcal{O}', j}^x \rangle| \leq 1/N^{r/10-1}$  for large enough  $N$ .*

PROOF. We can write  $|\langle \mathbf{W}_{\mathcal{O}, i}^x, \mathbf{W}_{\mathcal{O}', j}^x \rangle - \langle \mathbf{T}_{\mathcal{O}, i}^x, \mathbf{T}_{\mathcal{O}', j}^x \rangle|$  as  $|\langle \mathbf{T}_{\mathcal{O}, i}^x, \mathbf{W}_{\mathcal{O}', j}^x - \mathbf{T}_{\mathcal{O}', j}^x \rangle + \langle \mathbf{T}_{\mathcal{O}', j}^x, \mathbf{W}_{\mathcal{O}, i}^x - \mathbf{T}_{\mathcal{O}, i}^x \rangle + \langle \mathbf{W}_{\mathcal{O}', j}^x - \mathbf{T}_{\mathcal{O}', j}^x, \mathbf{W}_{\mathcal{O}, i}^x - \mathbf{T}_{\mathcal{O}, i}^x \rangle|$  (6)

and apply Lemma 4.6 to get the required bound.  $\square$

LEMMA 4.8. *For any  $\mathcal{O}, \mathcal{O}', \mathcal{O}''$  and  $i, j, k \in [N]$ ,  $\mathbf{W}_{\mathcal{O}, i}^x$ ,  $\mathbf{W}_{\mathcal{O}', j}^x$  and  $\mathbf{W}_{\mathcal{O}'', k}^x$  satisfy the triangle inequality.*

PROOF. Consider the vectors

$$\begin{aligned} \mathbf{T}_{\mathcal{O}, i}^x &= \left( \frac{1}{\sqrt{N}} \mathbf{V}_{\mathcal{O}, i}^x \right)^{\otimes r} \\ \mathbf{T}_{\mathcal{O}', j}^x &= \left( \frac{1}{\sqrt{N}} \mathbf{V}_{\mathcal{O}', j}^x \right)^{\otimes r} \\ \mathbf{T}_{\mathcal{O}'', k}^x &= \left( \frac{1}{\sqrt{N}} \mathbf{V}_{\mathcal{O}'', k}^x \right)^{\otimes r}. \end{aligned}$$

We may assume that no two of these are the same, otherwise the corresponding orthogonalized vectors would also be the same, and therefore the triangle inequality would be trivially valid. Applying lemma 4.10 to the set  $U := \{\mathbf{V}_{\mathcal{O}, i}^x, \mathbf{V}_{\mathcal{O}', j}^x, \mathbf{V}_{\mathcal{O}'', k}^x, -\mathbf{V}_{\mathcal{O}, i}^x, -\mathbf{V}_{\mathcal{O}', j}^x, -\mathbf{V}_{\mathcal{O}'', k}^x\}$  and  $D := N$ , and taking  $r$  to be  $2^{12}$  we obtain

$$\begin{aligned} N^r + \langle N^{r/2} \mathbf{T}_{\mathcal{O}', j}^x, N^{r/2} \mathbf{T}_{\mathcal{O}'', k}^x \rangle &\geq \langle N^{r/2} \mathbf{T}_{\mathcal{O}, i}^x, N^{r/2} \mathbf{T}_{\mathcal{O}', j}^x \rangle \\ &\quad + \langle N^{r/2} \mathbf{T}_{\mathcal{O}, i}^x, N^{r/2} \mathbf{T}_{\mathcal{O}'', k}^x \rangle \\ &\quad + N^{r-2}. \end{aligned}$$

Therefore  $1 + \langle \mathbf{T}_{\mathcal{O}', j}^x, \mathbf{T}_{\mathcal{O}'', k}^x \rangle \geq \langle \mathbf{T}_{\mathcal{O}, i}^x, \mathbf{T}_{\mathcal{O}', j}^x \rangle + \langle \mathbf{T}_{\mathcal{O}, i}^x, \mathbf{T}_{\mathcal{O}'', k}^x \rangle + \frac{1}{N^2}$ . Now using Lemma 4.7 and the fact that  $r$  is a large

number, for large enough  $N$  we get the desired triangle inequality,

$$1 + \langle \mathbf{W}_{\mathcal{O}', j}^x, \mathbf{W}_{\mathcal{O}'', k}^x \rangle \geq \langle \mathbf{W}_{\mathcal{O}, i}^x, \mathbf{W}_{\mathcal{O}', j}^x \rangle + \langle \mathbf{W}_{\mathcal{O}, i}^x, \mathbf{W}_{\mathcal{O}'', k}^x \rangle.$$

$\square$

LEMMA 4.9. [16] *Let  $a, b, c \in [-1, 1]$  such that  $1 + a \geq b + c$ . Then,  $1 + a^t \geq b^t + c^t$  for every odd integer  $t \geq 1$ .*

LEMMA 4.10. *Let  $U$  be a set of vectors in  $\mathbb{Z}^m$  that satisfy the following properties:*

1.  $\mathbf{u} \in U \Rightarrow -\mathbf{u} \in U$ .
2. There is a number  $D$  such that  $\|\mathbf{u}\|^2 \leq D$  for all  $\mathbf{u} \in U$ .
3. For every  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in U$ ,

$$D + \langle \mathbf{u}, \mathbf{v} \rangle \geq \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$$

Then, given any three vectors  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in U$  such that  $|\langle \mathbf{u}, \mathbf{v} \rangle|, |\langle \mathbf{u}, \mathbf{w} \rangle|, |\langle \mathbf{v}, \mathbf{w} \rangle| < D$ ,

$$D^{2^l} + \langle \mathbf{u}^{\otimes 2^l}, \mathbf{v}^{\otimes 2^l} \rangle \geq \langle \mathbf{u}^{\otimes 2^l}, \mathbf{w}^{\otimes 2^l} \rangle + \langle \mathbf{v}^{\otimes 2^l}, \mathbf{w}^{\otimes 2^l} \rangle + D^{2^l-2}$$

for all  $l \geq 1$ .

PROOF. The proof is by induction on  $l$ . Let  $\mathbf{u}, \mathbf{v}, \mathbf{w}$  be any three vectors in  $U$  such that they satisfy the condition that  $|\langle \mathbf{u}, \mathbf{v} \rangle|, |\langle \mathbf{u}, \mathbf{w} \rangle|, |\langle \mathbf{v}, \mathbf{w} \rangle| < D$ . Let  $x := \langle \mathbf{u}, \mathbf{v} \rangle$ ,  $y := \langle \mathbf{u}, \mathbf{w} \rangle$ ,  $z := \langle \mathbf{v}, \mathbf{w} \rangle$ .

BASE CASE:  $l = 1$ . We need to prove that,  $D^2 + x^2 \geq y^2 + z^2 + 1$ . It is sufficient to prove this when  $|x| \leq |y|$  and  $|x| \leq |z|$ . Hence, we may assume that  $|z| \geq |y| \geq |x|$ . Moreover, we may assume that  $z \geq 0$ . For if  $z < 0$ , then we argue about the vector  $-\mathbf{w}$  instead of  $\mathbf{w}$  to get  $z \geq 0$ . Consider two cases based on the sign of  $y$ :

1.  $y \geq 0$ . We have  $D + y \geq z + x$ . Since  $y \geq 0$ , we get that  $y \geq x$ . By hypothesis we know that  $D > z$ . Since all the numbers are integers we have that

$$D + y - 1 \geq z + x. \quad (7)$$

We also have that  $D + x \geq z + y$ , which implies that

$$D - y \geq z - x. \quad (8)$$

Since  $z + x \geq 0$  and  $z - x \geq 0$ , we can multiply inequalities (7), (8), and using the fact that  $D > y$  we get,  $D^2 - y^2 - D + y \geq z^2 - x^2$  which implies that  $D^2 + x^2 \geq y^2 + z^2 + D - y \geq y^2 + z^2 + 1$ .

2.  $y < 0$ . In this case  $-y \geq -x$ . Using this and the fact that  $D > z$  we deduce the following inequalities from (8):

$$\begin{aligned} D + y &\geq z + x \\ D - y - 1 &\geq z - x. \end{aligned}$$

Since  $z + x \geq 0$  and  $D + y > 0$ , multiplying the above two inequalities, we get  $D^2 - y^2 - D - y \geq z^2 - x^2$ , which implies that  $D^2 + x^2 \geq y^2 + z^2 + D + y \geq y^2 + z^2 + 1$ .

INDUCTIVE STEP: Assume that the lemma holds for some  $k \geq 1$ . We need to prove it for  $k+1$ . As before, we may assume that  $|x| \leq |y| \leq |z|$ , and hence, it is sufficient to show that  $D^{2^{k+1}} + x^{2^{k+1}} \geq y^{2^{k+1}} + z^{2^{k+1}} + D^{2^{k+1}-2}$ . By the induction hypothesis we have the following inequalities:

$$D^{2^k} + x^{2^k} \geq z^{2^k} + y^{2^k} + D^{2^k-2} \Rightarrow D^{2^k} - y^{2^k} - D^{2^k-2} \geq z^{2^k} - x^{2^k} \quad (9)$$

$$D^{2^k} + y^{2^k} \geq z^{2^k} + x^{2^k} \quad (10)$$

Observing that right hand sides of inequalities (9) and (10) are non-negative, we multiply both of them to get  $D^{2^{k+1}} - y^{2^{k+1}} - D^{2^{k+1}-2} - y^{2^k} D^{2^k-2} \geq z^{2^{k+1}} - x^{2^{k+1}}$ . This implies that  $D^{2^{k+1}} + x^{2^{k+1}} \geq y^{2^{k+1}} + z^{2^{k+1}} + D^{2^{k+1}-2}$  as desired.  $\square$

### 4.5.1 Main Lemma

LEMMA 4.11. Let  $\{\mathbf{u}_i\}_{i=1}^N, \{\mathbf{v}_i\}_{i=1}^N$  and  $\{\mathbf{w}_i\}_{i=1}^N$  be three sets of vectors, each set being an orthonormal set. Let  $s = 10$ . For some  $\gamma \geq 0$ , suppose these vectors satisfy:

1. **Mild Separation:** Dot-product of any two vectors is at most  $1 - \gamma$  in absolute value.
2. **Triangle Inequality:** Any three vectors satisfy the triangle inequality.
3. **Matching Property and Proper Indexing :** Let  $\mu := \max_{1 \leq i, j \leq N} |\langle \mathbf{v}_i, \mathbf{w}_j \rangle|$  and  $\lambda := \max_{1 \leq i, j \leq N} |\langle \mathbf{u}_i, \mathbf{w}_j \rangle|$ . Then

$$\mu - \gamma \leq \langle \mathbf{v}_i, \mathbf{w}_i \rangle \leq \mu \quad \forall 1 \leq i \leq N$$

$$\lambda - \gamma \leq \langle \mathbf{u}_i, \mathbf{w}_i \rangle \leq \lambda \quad \forall 1 \leq i \leq N$$

4. **Closeness:** Either  $\mu$  or  $\lambda$  is at least equal to  $1 - 2^{-100s}$ .

Let  $s_i, t_i, r_i \in \{-1, 1\}$  for  $1 \leq i \leq N$  and define

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

Then  $\mathbf{u}, \mathbf{v}, \mathbf{w}$  satisfy triangle inequality:  $1 + \langle \mathbf{u}, \mathbf{v} \rangle \geq \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ .

A proof of this lemma can be found in the full version [9].

## 4.6 Spreading Constraint for Minimum Linear Arrangement

In this section we show that the SDP solution to Minimum Linear Arrangement satisfies the *spreading* constraints. This would complete the proof of Theorem 2.6. Each vertex  $\mathcal{O}$  is associated with a unit vector  $\mathbf{V}_{\mathcal{O}}$ . The SDP solution to the Minimum Linear Arrangement instance in Theorem 2.6 is  $2\sqrt{n}\mathbf{V}_{\mathcal{O}}$ . We state the main theorem in terms of the vectors  $\mathbf{V}_{\mathcal{O}}$ . Let  $B(\mathcal{O}, r) := \{\mathcal{O}' \in V : \|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 \leq r\}$ .

LEMMA 4.12. 1.  $\forall \mathcal{O}, \forall r \geq 1/n, |B(\mathcal{O}, r)| \leq rn$ .

2.  $\forall \mathcal{O}, \forall S \subseteq V, \mathcal{O} \in S, \sum_{\mathcal{O}' \in S} \|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 \geq (|S|^2 - 1)/4n$ .

PROOF. The idea behind the proof is that from any given vertex, there are very few vertices within a  $(l_2^2)$  distance of 1, i.e.,  $|B(\mathcal{O}, 1)| \leq 4nN \exp(-\sqrt{N}/4)$  (see Lemma 4.15). The interesting radii are  $1/N^2 \leq r \leq 1$ . (If  $r > 1$  then

$|B(\mathcal{O}, r)| \leq n \leq rn$  holds trivially. If  $r \leq 1/N^2$ , then  $B(\mathcal{O}, r) = \{\mathcal{O}\}$  and the required inequality follows.) From Lemma 4.15,  $|B(\mathcal{O}, r)| \leq |B(\mathcal{O}, 1)| \leq 4nN \exp(-\sqrt{N}/4) \leq rn$ . Let  $d(\mathcal{O}, \mathcal{O}') := \|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2, S := \{\mathcal{O} = \mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_k\}$ , so that  $d(\mathcal{O}, \mathcal{O}_1) \leq d(\mathcal{O}, \mathcal{O}_2) \leq \dots \leq d(\mathcal{O}, \mathcal{O}_k)$ . Let  $r = d(\mathcal{O}, \mathcal{O}_{k/2})$ . Then  $|S|/2 \leq |B(\mathcal{O}, r)| \leq rn$ . Therefore  $\sum_{\mathcal{O}' \in S} d(\mathcal{O}, \mathcal{O}') \geq r|S|/2 \geq |S|^2/4n$ .  $\square$

We first show a necessary condition for two points to be at a distance lesser than 1. Next, we show that an exponentially small fraction of points satisfy this necessary condition.

LEMMA 4.13. Suppose  $\|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 \leq 1$ , then there is a  $1 \leq k \leq N$  such that

$$|\langle \mathbf{V}_{\mathcal{O},1}^x, \mathbf{V}_{\mathcal{O}',k}^x \rangle| \geq N^{3/4}.$$

PROOF. The proof is by contradiction. Suppose that for all  $1 \leq k \leq N, |\langle \mathbf{V}_{\mathcal{O},1}^x, \mathbf{V}_{\mathcal{O}',k}^x \rangle| \leq N^{3/4}$ . In fact, this implies that for all  $1 \leq i, j \leq N, |\langle \mathbf{V}_{\mathcal{O},i}^x, \mathbf{V}_{\mathcal{O}',j}^x \rangle| \leq N^{3/4}$ . This implies that  $|\langle \mathbf{T}_{\mathcal{O},i}^x, \mathbf{T}_{\mathcal{O}',j}^x \rangle| \leq 1/N^{r/4} = o(1/N^2)$ . Hence,  $|\langle \mathbf{W}_{\mathcal{O},i}^x, \mathbf{W}_{\mathcal{O}',j}^x \rangle| \leq o(1/N^2)$ . Thus,  $|\langle \mathbf{u}_i, \mathbf{v}_j \rangle| \leq o(1/N^2)$ , where  $\mathbf{u}_i := (\mathbf{W}_{\mathcal{O},i}^x)^{\otimes 2s}$  and  $\mathbf{v}_j := (\mathbf{W}_{\mathcal{O}',j}^x)^{\otimes 2s}$ . Let  $\mathbf{y} := \mathbf{V}_{\mathcal{O},1}^y$  and  $\mathbf{y}' := \mathbf{V}_{\mathcal{O}',1}^y$ . Recall that  $\mathbf{V}_{\mathcal{O}} := \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y_i \mathbf{u}_i\right)^{\otimes t}$ . Hence,

$$\begin{aligned} \langle \mathbf{V}_{\mathcal{O}}, \mathbf{V}_{\mathcal{O}'} \rangle^{1/t} &= \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y_i \mathbf{u}_i\right) \cdot \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N y'_i \mathbf{v}_i\right) \\ &\leq N^2 \cdot o(1/N^2) = o(1). \end{aligned}$$

Thus,  $\|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 = 2(1 - \langle \mathbf{V}_{\mathcal{O}}, \mathbf{V}_{\mathcal{O}'} \rangle) \geq 1$ .  $\square$

We now show that the event that there exists  $1 \leq k \leq N$  such that  $|\langle \mathbf{V}_{\mathcal{O},1}^x, \mathbf{V}_{\mathcal{O}',k}^x \rangle| \geq N^{3/4}$  occurs with low probability if  $\mathcal{O}'$  is picked uniformly at random.

LEMMA 4.14. For all  $\mathbf{x} \in \{-1, 1\}^N$ ,

$$\begin{aligned} \Pr_{\mathbf{x}' \in_{1/2}\{-1, 1\}^N} \left[ \exists 1 \leq l \leq N, |\mathbf{x} \cdot \sigma^l(\mathbf{x}')| \geq N^{3/4} \right] &\leq \\ &2 \exp(-\sqrt{N}/4)N. \end{aligned}$$

PROOF. This is a simple application of Chernoff bounds (Theorem 4.2). Proof follows from noting that for each  $l, \mathbf{x} \cdot \sigma^l(\mathbf{x}')$  is the sum of  $N$  random variables as required, and applying union bound over all  $l$ 's.  $\square$

We now put everything together to prove the main lemma.

LEMMA 4.15. For all  $\mathcal{O} \in V, \frac{|B(\mathcal{O}, 1)|}{n} \leq 4 \exp(-\sqrt{N}/4)N$ .

PROOF.

$$\begin{aligned} \frac{|B(\mathcal{O}, 1)|}{n} &\leq \Pr_{\mathcal{O}'} [\|\mathbf{V}_{\mathcal{O}} - \mathbf{V}_{\mathcal{O}'}\|^2 \leq 1] \\ &\leq 2 \Pr_{\mathbf{x}' \in_{1/2}\{-1, 1\}^N} \left[ \exists 1 \leq l \leq N, |\mathbf{x} \cdot \sigma^l(\mathbf{x}')| \geq N^{3/4} \right] \\ &\leq 4 \exp(-\sqrt{N}/4)N. \end{aligned}$$

where the inequality in the second line follows from Lemma 4.13 and the inequality in the third line follows from Lemma 4.14. Also, the probability in the first line is only over nearly orthogonal orbits, while the probability in the second line is over all strings. However, this introduces a factor of at most 2.  $\square$

<sup>5</sup>Strictly speaking, this proof works for only even-sized  $S$ . We barter this lack of accuracy for ease of presentation.

## 5. CONCLUSION

We have presented  $\Omega(\log \log n)$  integrality gap instance for Sparsest Cut problem. It would be interesting to close the gap between this lower bound and  $O(\sqrt{\log n})$  upper bound of Arora, Rao and Vazirani [2]. We do not know a simple and intuitive explanation why the SDP solution constructed in this paper (and also in [16]) satisfies the triangle inequality constraints. It would be nice to provide such an explanation and perhaps show that the SDP solution in fact satisfies higher  $k$ -gonal inequalities.

It remains a challenging open problem to prove a hardness of approximation result for Sparsest Cut (and Minimum Linear Arrangement) problem. Currently we only know a hardness result for non-uniform Sparsest Cut based on the Unique Games Conjecture (see [8, 16]).

## 6. REFERENCES

- [1] 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.
- [2] 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.
- [3] Yonatan Aumann and Yuval Rabani. An  $O(\log k)$  approximate min-cut max-flow theorem and approximation algorithm. *SIAM Journal of Computing*, 27(1):291–301, 1998.
- [4] Jean Bourgain. On lipschitz embeddings of finite metric spaces in hilbert space. *Israel Journal of Mathematics*, 52:46–52, 1985.
- [5] Jean Bourgain. On the distribution of the fourier spectrum of boolean functions. *Israel Journal of Mathematics*, 131:269–276, 2002.
- [6] Moses Charikar, Mohammad Hajiaghayi, Howard Karloff, and Satish Rao.  $l_2^2$  spreading metrics for vertex ordering problems. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms*, 2006. To appear.
- [7] 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.
- [8] 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.
- [9] Nikhil R Devanur, Subhash Khot, Rishi Saket, and Nisheeth Vishnoi. Integrality gaps for sparsest cut and minimum linear arrangement problems. Manuscript available from <http://www.cc.gatech.edu/~nikhil/>, 2006.
- [10] Guy Even, Joseph Naor, Satish Rao, and Baruch Schieber. Divide-and-conquer approximation algorithms via spreading metrics. *Journal of the ACM*, 47(4):585–616, 2000.
- [11] Uriel Feige and James Lee. An improved approximation ratio for the minimum linear arrangement problem. Manuscript, 2005.
- [12] Michel X. Goemans. Semidefinite programming in combinatorial optimization. *Math. Program.*, 79:143–161, 1997.
- [13] M. D. Hansen. Approximation algorithms for geometric embeddings in the plane with applications to parallel processing problems. In *Annual Symposium on Foundations of Computer Science*, number 30, pages 604–610, 1989.
- [14] Jeff Kahn, Gil Kalai, and Nati Linial. The influence of variables on boolean functions. In *Annual Symposium on Foundations of Computer Science*, number 29, pages 68–80, 1988.
- [15] 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.
- [16] Subhash Khot and Nisheeth K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into  $\ell_1$ . In *Annual Symposium on Foundations of Computer Science*, number 46, pages 53–62, 2005.
- [17] Robert Krauthgamer and Yuval Rabani. Improved lower bounds for embeddings into  $l_1$ . In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms*, 2006. To appear.
- [18] 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.
- [19] Nathan Linial. Finite metric spaces: combinatorics, geometry and algorithms. In *Proceedings of the International Congress of Mathematicians*, number III, pages 573–586, 2002.
- [20] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.
- [21] Satish Rao and Andréa W. Richa. New approximation techniques for some linear ordering problems. *SIAM Journal of Computing*, 34(2):388–404, 2004.
- [22] D. B. Shmoys. *Approximation Algorithms for NP-hard Problems*, chapter Approximation algorithms for Cut problems and their application to divide-and-conquer, pages 192–235. PWS, 1997.