How to Play any Unique Game

Eden Chlamtac * Konstantin Makarychev † Yury Makarychev ‡

Princeton University

{chlamtac, kmakaryc, ymakaryc}@cs.princeton.edu

Abstract

In this paper we present a new approximation algorithm for Unique Games. For a Unique Game with n vertices and k states (labels), if a $(1-\varepsilon)$ fraction of all constraints is satisfiable, the algorithm finds an assignment satisfying a

$$1 - O(\varepsilon\sqrt{\log n\,\log k})$$

fraction of all constraints. To this end, we introduce new embedding techniques for rounding semidefinite relaxations of problems with large domain size.

1 Introduction

The Unique Games Problem is a natural generalization of many constraint satisfaction problems. Particularly important special cases are MAX CUT and MAX 2-LIN mod p (systems of linear equations mod p with at most two variables in each equation). Formally, it is defined as follows:

Definition 1.1 (Unique Games Problem). *Given a constraint graph $G = (V, E)$ and a set of permutations π_{uv} on $[k] = \{1, \dots, k\}$ (for all edges (u, v)), assign a value (state) x_u from the set $[k]$ to each vertex u so as to satisfy the maximum number of constraints of the form $\pi_{uv}(x_u) = x_v$.*

In any instance of Unique Games if all constraints are satisfiable then it is easy to find a satisfying assignment. However, even if almost all constraints are satisfiable, it is NP-hard to find the optimal solution. Moreover, Khot [11] conjectured that for every positive ε and δ, there exists k such that it is NP-hard to distinguish whether a $(1-\varepsilon)$ fraction of all constraints is satisfiable, or only a δ fraction of all constraints is satisfiable. This conjecture, known as the Unique Games Conjecture, implies many inapproximability results for fundamental problems, which are not known to follow from more standard complexity assumptions. Thus it is interesting to determine what fraction of constraints can be satisfied for such instances as a function of ε, k and n (where n is the number of vertices). The recent algorithms of Charikar, Makarychev, and

*Supported by a Francis Upton fellowship.

†Supported by a Gordon Wu fellowship.

‡Supported by a Gordon Wu fellowship.

Makarychev [6] achieve near-optimal results as a function of ε and k: Khot, Kindler, Mossel, and O'Donnell [12] showed that any improvement in the guarantees will disprove the Unique Games Conjecture. However, the question of the best achievable results in terms of n is still wide open. Moreover, as our results indicate, there is a close connection between the results achievable as a function of n and the approximation factor for Sparsest Cut.

Algorithmically, Unique Games pose a challenge for the development of SDP based algorithms. While SDP approaches have been quite successful in dealing with binary constraint satisfaction problems, it is not clear how to extend these techniques to constraint satisfaction problems (CSPs) over large domains. Our study of Unique Games is motivated by attempting to develop an understanding of the SDP toolkit for CSPs over larger domains.

Let us begin by describing some known results. Using semidefinite programming Khot [11] constructed the first approximation algorithm for Unique Games in 2002. His algorithm satisfies a $\delta = 1 - O(k^2 \varepsilon^{1/5} \sqrt{\log(1/\varepsilon)})$ fraction of all constraints. Trevisan [14] developed an algorithm that satisfies a $1 - O(\sqrt[3]{\varepsilon \log n})$ fraction of all constraints.

Gupta and Talwar [7] suggested an algorithm based on linear programming that satisfies a $(1 - O(\varepsilon \log n))$ fraction of all constraints. They also raise the question of whether an algorithm based on Khot's semidefinite program can achieve a better performance guarantee. Finally, Charikar, Makarychev and Makarychev [6] developed two algorithms that satisfy roughly a $O(k^{-\frac{\varepsilon}{2-\varepsilon}})$ fraction and a $1 - O(\sqrt{\varepsilon \log k})$ fraction of all constraints, respectively. Note that the approximation guarantees of algorithms by Gupta and Talwar [7] and Charikar, Makarychev and Makarychev [7] are not always comparable. But these algorithms always have better approximation guarantee than algorithms by Khot [11] and Trevisan [14] respectively. We summarize the best known results in the following table (see the third column).

Technique	**MAX CUT:** $k = 2$		**Unique Games**	
SDP ($\varepsilon > 1/\log n$)	$1 - O(\sqrt{\varepsilon})$	[9]	$1 - O(\sqrt{\varepsilon \log k})$	[6]
SDP ($\varepsilon < 1/\log n$)	$1 - O(\varepsilon \sqrt{\log n})$	[1]	**?**	
LP ($\varepsilon < 1/\log n$)	$1 - O(\varepsilon \log n)$	[8]	$1 - O(\varepsilon \log n)$	[7]

It is instructive to compare these results with the best results for MAX CUT (in fact, all these results for MAX CUT apply to general Unique Games with only 2 states[1]). The Goemans–Williamson MAX CUT algorithm [9] satisfies a $(1 - O(\sqrt{\varepsilon}))$ fraction of all constraints. This approximation guarantee was generalized by Charikar, Makarychev, and Makarychev [6] for Unique Games with larger domain size k. However in the range $\varepsilon < 1/\log n$, the approximation guarantee for MAX CUT was significantly better than that for general Unique Games: the approximation algorithm by Agarwal, Charikar, Makarychev, and Makarychev [1] satisfies a $1 - O(\varepsilon \sqrt{\log n})$ fraction of all constraints for MAX CUT, but no analog of this was known for larger k. Compare this to the recent algorithm by Gupta and Talwar [7] which satisfies a $1 - O(\varepsilon \log n)$ fraction of all constraints[2]. Note that no algorithm with approximation guarantee $1 - O(\varepsilon \sqrt{\log n})$ was known even for $k = 3$ (or MAX 2-LIN mod 3). Our paper closes the gap between the Unique Games

[1]This problem is variously known as MAX RES CUT (see [9]) or MAX 2-LIN mod 2. Its complement is also called MIN UNCUT or MIN 2CNF≡ Deletion.

[2]It is interesting to note that the analog of Gupta and Talwar [7] for $k = 2$ was obtained earlier by Garg, Vazirani, and Yannakakis [8].

and the MAX CUT problems (thus giving a positive answer to the question raised by Gupta and Talwar [7]).

1.1. Our Results

The main result of this paper is as follows.

Theorem 1.2. *There exists a polynomial time algorithm that finds an assignment of values to vertices satisfying a $(1 - O(\varepsilon\sqrt{\log n \log k}))$ fraction of all constraints, for any instance of Unique Games for which a $(1 - \varepsilon)$ fraction of all constraints is satisfiable.*

Remark 1.1. *Since complexity reductions based on the Unique Games Conjecture use long codes, the parameter k is typically $O(\log n)$. For such k our algorithm satisfies $1 - O(\varepsilon\sqrt{\log n \log\log n})$ fraction of all constraints. For all values of $k = n^{o(1)}$, our algorithm has a better approximation guarantee than the algorithm of Gupta and Talwar [7].*

The SDP solution associates a collection of orthogonal vectors for every vertex, one such vector for every possible state of the vertex. The goal of the rounding algorithm is to pick one of these vectors for every vertex. As the main technical tool, we introduce a new type of random partitioning scheme, which we call an *m-orthogonal separator*. Specifically, we construct an algorithm that, given a set of vectors in an ℓ_2^2 space, produces random subsets S such that the probability that two orthogonal vectors belong to S is equals $1/m$ (we assume that $1/m$ is very small); and the distribution over corresponding cuts $(S, \bar{S})$ is a low distortion embedding from ℓ_2^2 to ℓ_1. In other words, for two orthogonal vectors u and v the events "$u \in S$" and "$v \in S$" are "almost" disjoint. This property is crucial for our algorithm: it essentially guarantees that we assign only one value to each vertex in a Unique Game. We stress that no known embedding satisfies this property.

Viewed in this new framework, the random cuts generated by the algorithm for MAX CUT [1] may be seen as "∞-*antipodal* separators": the random set S never contains two *antipodal* vectors. (Recall that in constraint satisfaction problems with domain size $k = 2$, exclusive states of vertices are typically "encoded" by antipodal vectors; whereas if the domain size is larger, the exclusive states are encoded by orthogonal vectors.) Despite the similarity between orthogonal separators and antipodal separators, generating orthogonal separators seems much harder than generating antipodal separators. The reason is that any hyperplane cut separates antipodal vectors. However, there is no apparently simple way to separate orthogonal vectors.

In order to construct orthogonal separators we extend the methods of Charikar, Makarychev, and Makarychev [6] and combine them with powerful metric embedding techniques developed in the works of Arora, Rao, and Vazirani [4], of Lee [13], of Chawla, Gupta, and Räcke [5], and of Arora, Lee, and Naor [3]. We also introduce a special transformation of the space ℓ_2^2 , which we call "normalization". We present two algorithms: one using embeddings from ℓ_2^2 to ℓ_1 and the other using embeddings from ℓ_2^2 to ℓ_2. While the second algorithm gives a slightly better guarantee, the first algorithm would be improved even if better embedding techniques from ℓ_2^2 to ℓ_1 at one scale are found[3] (this cannot happen for embedding into ℓ_2, for which current guarantees are essentially tight).

[3]Previously, it was not known whether the MAX CUT algorithm of [1] exhibited this direct dependency.

Our new partitioning scheme is an analog of the structural theorem of Arora, Rao and Vazirani [4], which has found applications in several subsequent papers. We believe that our techniques will be useful for rounding SDP relaxations for problems involving more than two states.

In Section 2, we describe the semidefinite relaxation for Unique Games. Then in Section 3, we introduce the notion of embeddings separating orthogonal vectors and show how to use such embeddings to satisfy a $(1 - O(\varepsilon\sqrt{\log n\, \log k}))$ fraction of all constraints. Finally, in Section 4, we construct two embedding algorithms.

2 Semidefinite Relaxation

We use the vector relaxation of Khot [11] with additional triangle inequalities. For each vertex u and each state i we introduce a vector u_i. The intended integer solution is as follows. For every vector u_i set $u_i = 1$ if vertex u is assigned state i, otherwise let $u_i = 0$. Thus for a fixed u, only one u_i is not equal to zero. To model this proprty in the SDP we add constraints that u_i and u_j are orthogonal for $i \neq j$; and constraints $\|u_1\|^2 + \cdots + \|u_k\|^2 = 1$. We also add some triangle inequality constraints.

Notice, that if the constraint between u and v is satisfied, then $u_i = v_{\pi_{uv}(i)}$ for all $i \in [k]$. On the other hand if the constraint is not satisfied then the equality $u_i = v_{\pi_{uv}(i)}$ is violated for exactly two values of i. Thus the expression

$$\varepsilon_{uv} = \frac{1}{2}\sum_{i=1}^{k} \|u_i - v_{\pi_{uv}(i)}\|^2$$

is equal to 0, if the constraint is satisfied and 0, otherwise.

Using this observation we construct the following SDP:

$$\text{minimize } \frac{1}{2}\sum_{(u,v)\in E}\sum_{i=1}^{k} \|u_i - v_{\pi_{uv}(i)}\|^2$$

subject to

$$\begin{aligned}
\forall u \in V\ \forall i,j \in [k], i \neq j \qquad & \langle u_i, u_j\rangle = 0 && (1)\\
\forall u \in V \qquad & \sum_{i=1}^{k} \|u_i\|^2 = 1 && (2)\\
\forall u,v,w \in V\ \forall i,j,l \in [k] \qquad & \|u_i - w_l\|^2 \leq \|u_i - v_j\|^2 + \|v_j - w_l\|^2 && (3)\\
\forall u,v \in V\ \forall i,j \in [k] \qquad & \|u_i - v_j\|^2 \leq \|u_i\|^2 + \|v_j\|^2 && (4)\\
\forall u,v \in V\ \forall i,j \in [k] \qquad & \|u_i\|^2 \leq \|u_i - v_j\|^2 + \|v_j\|^2 && (5)
\end{aligned}$$

Note, that the objective function of the SDP measures how many constraints are *not satisfied.*

Remark 2.1. *The constraints* (4) *and* (5) *are ℓ_2^2 triangle inequalities with the zero vector:*

$$\begin{aligned}
\|u_i - v_j\|^2 &\leq \|u_i - 0\|^2 + \|v_j - 0\|^2;\\
\|u_i - 0\|^2 &\leq \|u_i - v_j\|^2 + \|v_j - 0\|^2.
\end{aligned}$$

A particularly important constraint is that the vectors u_i and u_j are orthogonal for $i \neq j$.

Remark 2.2. *The triangle inequalities also imply the "cycle constraints" introduced by Gupta and Talwar [7] in their LP. Thus the SDP relaxation is stronger than the LP relaxation. We do not use the "cycle constraints" specifically in our analysis.*

3 Overview of Techniques

In this section we describe the main technical tool of this paper. We introduce a new type of embeddings from ℓ_2^2 to ℓ_1: *embeddings separating orthogonal vectors.* Recall that a set of vectors X in $\mathbb{R}^d$ is an ℓ_2^2 space if it satisfies the following triangle inequalities:

$$\forall u, v, w \in X \; \|u-v\|^2 + \|v-w\|^2 \geq \|u-w\|^2.$$

Note that the vectors in any feasible solution to the SDP above, together with the zero vector, form an ℓ_2^2 space.

Definition 3.1. *Let X be an ℓ_2^2 space. We say that a random set $S \subset X$ is an m-orthogonal separator of X with distortion D and probability scale α if the following conditions hold:*

1. *For all u in X, $\Pr(u \in S) = \alpha \|u\|^2$.*
2. *For all orthogonal vectors u and v in X,*
$$\Pr(u \in S \text{ and } v \in S) \leq \frac{\min(\Pr(u \in S), \Pr(v \in S))}{m}.$$
Note that the right hand side is at most $\alpha/m \cdot \frac{\|u\|^2+\|v\|^2}{2}$.
3. *For all u and v in X, $\Pr(I_S(u) \neq I_S(v)) \leq \alpha D \|u-v\|^2$, where I_S is the indicator function of the set S.*

The novelty of Definition 3.1 is in property 2. It says that for every *orthogonal* vectors u and v the events "$u \in S$" and "$v \in S$" are almost disjoint. As always we can interpret a distribution over random cuts as an embedding into ℓ_1.

Definition 3.2. *We say that a mapping f of an ℓ_2^2 space X to ℓ_1 is a distortion D embedding m-separating orthogonal vectors if the following conditions hold.*

1. *For all u in X, $\|f(u)\|_1 = \|u\|^2$.*
2. *For all orthogonal vectors u and v in X,*
$$\|f(u) - f(v)\|_1 \geq \|f(u)\|_1 + \|f(v)\|_1 - 2\,\frac{\min(\|f(u)\|_1, \|f(v)\|_1)}{m}.$$
3. *For all u and v in X, $\|f(u) - f(v)\|_1 \leq D\,\|u-v\|^2$.*

These definitions are equivalent in the following sense: if there exists an m-orthogonal separator of X then there exists an embedding m-separating orthogonal vectors with the same distortion and vice versa. We explain this connection in more detail in the full version of the paper. In this paper we work only with orthogonal separators i.e. Definition 3.1.

Let us state the main technical result.

Theorem 3.3. *There exists a randomized polynomial time algorithm that, given an ℓ_2^2 space X containing 0 and a parameter m, returns an m-orthogonal separator of X with distortion $D = O(\sqrt{\log |X| \log m})$ and probability scale $\alpha \geq 1/poly(m)$.*

In the next section we show how using this theorem we obtain an approximation algorithm for Unique Games. We shall prove Theorem 3.3 in Section 4.

3.1. Approximation Algorithm

Input: An instance of Unique Games.
Output: Assignment of states to vertices.

1. Solve the SDP.
2. Mark all vertices as unprocessed.
3. while (there are unprocessed vertices)
   (a) Produce an m-orthogonal separator S with distortion D and probability scale α as in Theorem 3.3, where $m = 4k$ and $D = O(\sqrt{\log n \log m})$.
   (b) For all unprocessed vertices u :
   - Let $S_u = \{i : u_i \in S\}$.
   - If S_u contains exactly one element i, then assign the state i to u, and mark the vertex u as processed.
4. If the algorithm performs more than n/α iterations, assign arbitrary values to any remaining vertices (note that $\alpha \geq 1/poly(k)$).

Lemma 3.4. *The algorithm satisfies the constraint between vertices u and v with probability $1 - O(D\varepsilon_{uv})$, where ε_{uv} is the SDP contribution of the term corresponding to the edge (u, v):* √

$$\varepsilon_{uv} = \frac{1}{2} \sum_{i=1}^{k} \|u_i - v_{\pi_{uv}(i)}\|^2.$$

Proof. If $D\varepsilon_{uv} \geq 1/4$, then the statement holds trivially, so we assume that $D\varepsilon_{uv} < 1/4$. For the sake of analysis we also assume that π_{uv} is the identity permutation (we can just rename the states of the vertex v, this clearly does not affect the execution of the algorithm). √

At the end of an iteration in which one of the vertices u or v assigned a value we mark the constraint as satisfied or not: the constraint is satisfied, if the the same state i is assigned to the vertices u and v; otherwise, the constraint is not satisfied (here we conservatively count the number of satisfied constraints: a constraint marked as not satisfied in the analysis may potentially be satisfied in the future).

Consider one iteration of the algorithm. There are three possible cases:

1. Both sets S_u and S_v are equal and contain only one element, then the constraint is satisfied.

2. The sets S_u and S_v are equal, but contain more than one or none elements, then no values are assigned at this iteration to u and v.

3. The sets S_u and S_v are not equal, then the constraint is not satisfied (a conservative assumption).

Let us estimate the probabilities of each of these cases. Using the fact that for all $i \neq j$ the vectors u_i and u_j are orthogonal, and the first and second properties of orthogonal separators we get (below α is the probability scale):

$$\begin{aligned}
\Pr(|S_u| = 1) &\geq \sum_{i\in[k]} \Pr(i \in S_u) - \sum_{\substack{i,j\in[k] \\ i\neq j}} \Pr(i \in S_u \text{ and } j \in S_u) \\
&= \sum_{i\in[k]} \Pr(u_i \in S) - \sum_{\substack{i,j\in[k] \\ i\neq j}} \Pr(u_i \in S \text{ and } u_j \in S) \\
&\geq \sum_{i\in[k]} \alpha \|u_i\|^2 - \frac{\alpha}{m} \sum_{i,j\in[k]} \frac{\|u_i\|^2 + \|u_j\|^2}{2} \\
&= \alpha - \alpha/4 = \frac{3}{4}\alpha.
\end{aligned}$$

The probability that the constraint is not satisfied is at most

$$\Pr(S_u \neq S_v) \leq \sum_{i\in[k]} \Pr(I_S(u_i) \neq I_S(v_i)) \leq \alpha D \sum_{i\in[k]} \|u_i - v_i\|^2 = \alpha D \varepsilon_{uv}.$$

Finally the probability of satisfying the constraint is at least

$$\Pr(|S_u| = 1 \text{ and } S_u = S_v) \geq \frac{3}{4}\alpha - \alpha D \varepsilon_{uv} \geq \frac{1}{2}\alpha.$$

Since the algorithm performs n/α iterations, the probability that it does not assign any value to u or v before step 4 is exponentially small. At each iteration the probability of failure is at most $O(D\varepsilon_{uv})$ times the probability of success, thus the probability that the constraint is not satisfied is $O(D\varepsilon_{uv})$. $\square$

We now show that the approximation algorithm satisfies $1 - O(\sqrt{\log n \, \log k}\, \varepsilon)$ fraction of all constraints.

Proof of Theorem 1.2. By Lemma 3.4, the expected number of unsatisfied constraints is equal to

$$\sum_{(u,v)\in E} O(D \times \varepsilon_{uv}) = O(\sqrt{\log n \, \log k}) \times SDP,$$

where SDP is the SDP value. Thus the expected fraction of unsatisfied constraints is $O(\sqrt{\log n \, \log k}) \times SDP$. Since $SDP \leq \varepsilon|E|$, the algorithm satisfies $1 - O(\sqrt{\log n \, \log k}\, \varepsilon)$ fraction of all constraints with high probability. $\square$

4 Producing orthogonal separators

In this section we present two algorithms that generate m-orthogonal separators with distortions $D_1 = O(\sqrt{\log |X| \log m})$ and $D_2 = O(\sqrt{\log |X| \log m})$. The main difference between the algorithms is that the first algorithm uses embeddings to ℓ_1 as an intermediate step, while the second one uses embeddings to ℓ_2. Thus any improvements in embeddings to ℓ_1 will result in a better distortion for the first algorithm. We also believe that the first algorithm is simpler than the second one.

The algorithms generate orthogonal separators in three steps. First we *normalize* all vectors in a special way. Namely, we transform the set X into a set of functions in $L_2[0,\infty]$, so that the image of every non-zero vector is a function with L_2 norm 1; the images of orthogonal vectors are orthogonal; and the new configuration satisfies L_2^2 triangle inequalities. Then we embed the transformed set into ℓ_1 or ℓ_2. Finally, we produce orthogonal separators based on the ℓ_1 or ℓ_2 embedding and the original lengths of vectors.

4.1. Normalization: Embedding into $L_2[0,\infty]$

The space $L_2[0,\infty]$ is the space of square integrable functions $f : [0,\infty) \to \mathbb{R}^d$ equipped with the following inner product:

$$\langle f_1, f_2 \rangle = \int_0^{+\infty} \langle f_1(t), f_2(t) \rangle \, dt;$$

and norm:

$$\|f\|_2 = \sqrt{\langle f, f \rangle} \equiv \sqrt{\int_0^{+\infty} \|f(t)\|^2 \, dt}.$$

We construct a mapping φ from $\mathbb{R}^d$ into $L_2[0,\infty]$ as follows

$$\varphi(u)(t) = \begin{cases} u, & \text{if } t \le 1/\|u\|^2; \\ 0, & \text{otherwise.} \end{cases}$$

We map the zero vector to 0. Let us see what properties the embedding φ has.

Lemma 4.1. *Let $X \subset \mathbb{R}^d$ be an ℓ_2^2 metric space containing the zero vector. Then*

1. *The image $\varphi(X)$ satisfies triangle inequalities in L_2^2:*

$$\forall u, v, w \in X \;\; \|\varphi(u) - \varphi(v)\|_2^2 + \|\varphi(v) - \varphi(w)\|_2^2 \ge \|\varphi(u) - \varphi(w)\|_2^2.$$

2. *For all non-zero vectors u and v in X,*

$$\langle \varphi(u), \varphi(v) \rangle = \frac{\langle u, v \rangle}{\max(\|u\|^2, \|v\|^2)}.$$

3. *For all non-zero vectors u in X, $\|\varphi(u)\|_2^2 = 1$.*

4. *For all orthogonal u and v in X, the images $\varphi(u)$ and $\varphi(v)$ are also orthogonal.*

5. *For all non-zero vectors u and v in X,*

$$\|\varphi(v) - \varphi(u)\|_2^2 \leq \frac{2\,\|v-u\|^2}{\max(\|u\|^2, \|v\|^2)}.$$

Proof. 1. The triangle inequality for the functions $\varphi(u)$, $\varphi(v)$ and $\varphi(w)$ is equivalent to the following inequality:

$$\int_0^\infty \|\varphi(u)(t) - \varphi(v)(t)\|^2 + \|\varphi(v)(t) - \varphi(w)(t)\|^2 - \|\varphi(u)(t) - \varphi(w)(t)\|^2 \, dt \geq 0.$$

This inequality holds for every t, since the vectors $\varphi(u)(t)$, $\varphi(v)(t)$ and $\varphi(w)(t)$ lie in the set $\{0, u, v, w\} \subset X$ and vectors in X satisfy ℓ_2^2 triangle inequalities.
2. Without loss of generality assume that $\|u\| \leq \|v\|$, then

$$\langle \varphi(u), \varphi(v) \rangle \;=\; \int_0^\infty \langle \varphi(u)(t), \varphi(v)(t) \rangle \, dt = \int_0^{1/\|v\|^2} \langle u, v \rangle \, dt = \frac{\langle u, v \rangle}{\|v\|^2}.$$

Parts 3 and 4 follow from part 2. (The zero vector maps to 0, so $\varphi(0)$ is orthogonal to any function).
5. Assume without loss of generality that $\|u\| \leq \|v\|$, then

$$\begin{aligned}
\|\varphi(v) - \varphi(u)\|_2^2 &= \frac{1}{\|v\|^2} \cdot \|v-u\|^2 + \left(\frac{1}{\|u\|^2} - \frac{1}{\|v\|^2} \right) \cdot \|u\|^2 \\
&= \frac{1}{\|v\|^2} \cdot \left(\|v-u\|^2 + \|v\|^2 - \|u\|^2 \right) \\
&\leq \frac{2}{\|v\|^2} \cdot \left(\|v-u\|^2 \right).
\end{aligned}$$

Here we used the triangle inequality $\|v-0\|^2 \leq \|v-u\|^2 + \|u-0\|^2$. □

Remark 4.1. *How can we represent the embedding in $L_2^2[0,\infty]$ efficiently? Note that $L_2[0,\infty]$ is a Hilbert space, so the metric on every finite subset of $L_2^2[0,\infty]$ is uniquely determined by its Gram matrix[4]. Hence we just need to compute the Gram matrix for the vectors/functions from $\varphi(X)$. This can be done using the formula from Lemma 4.1.2. We get the following corollary.*

Corollary 4.2. *There exists a polynomial time algorithm that, given an ℓ_2^2 space X, computes the Gram matrix of the set of vectors $\varphi(X)$.*

4.2. Embedding into ℓ_1 and ℓ_2

We use the following theorem of Arora, Lee, and Naor [3], which is based on the results of Arora, Rao and Vazirani [4], Lee [13], and Chawla, Gupta and Räcke [5].

[4]The Gram matrix of a set of vectors is the matrix, where (ij)-th element is equal to the inner product of i-th and j-th vectors.

Theorem 4.3 ([3], Theorem 3.1). *There exist constants $C \geq 1$ and $0 < p < 1/2$ such that for every n-point ℓ_2^2 space X with distance $d(u,v) = \|u - v\|^2$ and every $\Delta > 0$, the following holds. There exists a distribution μ over subsets $U \subset X$ such that for every $u, v \in X$ with $d(u,v) \geq \Delta$,*

$$\mu \left\{ U : u \in U \text{ and } d(v, U) \geq \frac{\Delta}{C\sqrt{\log n}} \right\} \geq p.$$

Note that we can efficiently sample from the distribution μ. We need the following easy corollaries. We sketch their proofs in Appendix B.

Corollary 4.4. *There exists an efficient algorithm that, given an ℓ_2^2 space X, generates random subsets Y such that the following conditions hold.*

1. *For every u and v in X,*

$$\Pr\left(I_Y(u) \neq I_Y(v)\right) \leq D\,\|u - v\|^2.$$

2. *For every u and v s.t. $\|u - v\| \geq 1$,*

$$\Pr\left(I_Y(u) \neq I_Y(v)\right) \geq \beta,$$

where β is a universal constant, $D = O(\sqrt{\log |X|})$.

Corollary 4.5. *There exists an efficient algorithm, that constructs an embedding ψ of an ℓ_2^2 space X into ℓ_2 such that the following conditions hold.*

1. *For all u and v in X,*

$$\|\psi(u) - \psi(v)\| \leq D\,\|u - v\|^2.$$

2. *For every u and v s.t. $\|u - v\| \geq 1$,*

$$\|\psi(u) - \psi(v)\| \geq 2\gamma.$$

3. *The set $\psi(X)$ lies on the unit sphere: $\forall u \in X$*

$$\|\psi(u)\| = 1,$$

where γ is a universal constant; $D = O(\sqrt{\log |X|})$.

4.3. Generating orthogonal separators via ℓ_1

We present an algorithm to generate orthogonal separators with distortion $O(\sqrt{\log |X|}\log m)$. This result is not as strong as the one given in the next section, but is arguably simpler, and demonstrates a number of the same ideas. Using this algorithm, in conjunction with Lemma 3.4, implies the following result:

Theorem 4.6. *There exists a polynomial time algorithm that finds an assignment of values to vertices satisfying a $(1 - O(\varepsilon\sqrt{\log n}\log k))$ fraction of all constraints, for any instance of Unique Games for which a $(1 - \varepsilon)$ fraction of all constraints is satisfiable.*

The following algorithm generates orthogonal separators as specified above.

Input: An ℓ_2^2 set of vectors X (containing 0), a parameter m.
Output: A random set S.

1. Set $l = \lceil \ln m/\beta \rceil$ (where β is as in Corollary 4.4).
2. Obtain $\varphi(X)$, a normalization of X, as described in Section 4.1.
3. Apply the algorithm from Corollary 4.4 to the set $\varphi(X)$, to generate l random independent subsets $Y_1, \ldots, Y_l \subset \varphi(X)$.
4. For every vector $u \in X$, construct a word $W(u)$ of length l corresponding to inclusion or exclusion of $\varphi(u)$ from the sets Y_i:
$$W(u) = I_{Y_1}(\varphi(u)) \ldots I_{Y_l}(\varphi(u)).$$
5. Pick a random word W in $\{0,1\}^l$ s.t. the probability that $W = W(u)$ (for each u) equals $1/|X|$. This is feasible since the number of distinct words constructed in step 4 is at most $|X|$ (possibly we may pick a word not corresponding to any $W(u)$).
6. Pick a random uniform value r in the interval $(0,1)$.
7. Find all vectors u of ℓ_2^2-length at least r such that $W(u) = W$:
$$S = \left\{u \in X : \|u\|^2 \geq r \text{ and } W(u) = W\right\}.$$
8. Return S.

Lemma 4.7. *The algorithm generates an m-orthogonal separator of X with distortion $O(\sqrt{\log|X|}\log m)$ and probability scale $\alpha = 1/|X|$.*

Proof. Let us verify that all the conditions of Definition 3.1 hold.
1. Fix an arbitrary u. Conditional on the event $r \leq \|u\|^2$ the probability of picking u in S is equal to $1/|X|$. Thus
$$\Pr(u \in S) = \frac{1}{|X|} \cdot \Pr\left(r \leq \|u\|^2\right) = \frac{1}{|X|} \cdot \|u\|^2.$$
2. Fix orthogonal vectors u and v from X. By Lemma 4.1 (parts 3 and 4), $\|\varphi(u) - \varphi(v)\|_2^2 = 2$, hence by Corollary 4.4,
$$\Pr(I_{Y_i}(u) = I_{Y_i}(v)) \leq 1 - \beta.$$
Thus the probability that $W(u) = W(v)$ is at most $(1-\beta)^l \leq \frac{1}{m}$. The probability that u and v are in S is as follows:
$$\begin{aligned}
\Pr(u, v \in S) &= \Pr\left(W(u) = W(v) \text{ and } W = W(u) \text{ and } r \leq \min(\|u\|^2, \|v\|^2)\right) \\
&= \Pr(W(u) = W(v)) \cdot \Pr(W = W(u)) \cdot \Pr\left(r \leq \min(\|u\|^2, \|v\|^2)\right) \\
&\leq \frac{1}{|X|} \cdot \frac{\min(\|u\|^2, \|v\|^2)}{m} = \frac{\min(\Pr(u \in S), \Pr(v \in S))}{m}.
\end{aligned}$$

3. Fix u and v from X and assume $\|u\| \le \|v\|$. Similarly to part 2, we have

$$\begin{aligned}
\Pr\left(I_S(u) \neq I_S(v)\right) &= \Pr\big(W(u) \neq W(v) \text{ and } (W = W(u) \text{ or } W = W(v)) \text{ and } r \le \|u\|^2\big) \\
&\quad + \Pr\big(W = W(v) \text{ and } \|u\|^2 \le r \le \|v\|^2\big) \\
&\le \frac{2\|u\|^2}{|X|} \cdot \Pr\left(W(u) \neq W(v)\right) + \frac{1}{|X|}\left(\|v\|^2 - \|u\|^2\right).
\end{aligned}$$

Now, by Corollary 4.4.1 and Lemma 4.1.5,

$$\begin{aligned}
\Pr\left(W(u) \neq W(v)\right) &\le \sum_{i=1}^{l} \Pr\left(I_{Y_i}(\varphi(u)) \neq I_{Y_i}(\varphi(v))\right) \le l\sqrt{\log|X|} \cdot \|\varphi(u) - \varphi(v)\|_2^2 \\
&\le \frac{2l\sqrt{\log|X|} \cdot \|u - v\|^2}{\|v\|^2}.
\end{aligned}$$

Using ℓ_2^2 triangle inequality $\|v\|^2 - \|u\|^2 \le \|u - v\|^2$ we get

$$\begin{aligned}
\Pr\left(I_S(u) \neq I_S(v)\right) &\le \frac{1}{|X|}\left(\frac{4\|u\|^2}{\|v\|^2} \cdot l\sqrt{\log|X|} + 1\right)\|u - v\|^2 \\
&= \frac{1}{|X|}\, \|u - v\|^2 \cdot O(\sqrt{\log|X|}\ \log m).
\end{aligned}$$

□

4.4. Generating orthogonal separators via ℓ_2

In this section we prove Theorem 3.3, which in turn implies Theorem 1.2 (using Lemma 3.4). We present an algorithm to generate orthogonal separators using embeddings to ℓ_2. It uses ideas from the algorithm of Charikar, Makarychev and Makarychev [6].

Input: An ℓ_2^2 set of vectors X (containing 0), a parameter m.
Output: A random set S.

1. Fix $m' = p(m)$, where $p(x)$ is a polynomial we specify later.

2. Obtain $\varphi(X)$, a normalization of X, as described in Section 4.1.

3. Embed $\varphi(X)$ into the unit sphere in ℓ_2 (see Corollary 4.5). Denote the image of the vector $\varphi(u)$ by $\psi(u)$.

4. Generate a random Gaussian vector g with independent components distributed as $\mathcal{N}(0, 1)$.

5. Fix a threshold t s.t. $\Pr(\xi \ge t) = 1/m'$, where $\xi \sim \mathcal{N}(0, 1)$ (*i.e.* t is $(1 - 1/m')$-quantile of the standard normal distribution).

6. Pick a random uniform value r in the interval $(0, 1)$.

7. Find all vectors u of ℓ_2^2-length at least r such that $\langle \psi(u), g\rangle \geq t$:

$$S = \left\{u \in X : \|u\|^2 \geq r \text{ and } \langle \psi(u), g\rangle \geq t\right\}.$$

8. Return S.

Lemma 4.8. *The algorithm generates an m-orthogonal separator of X with distortion $O(\sqrt{\log |X| \log m'})$ and probability scale $\alpha = 1/m'$.*

Proof. Let us verify that all the conditions of Definition 3.1 hold.
1. Fix an arbitrary u. Conditional on the event $r \leq \|u\|^2$ the probability of picking u in S is equal to $1/m'$. Thus

$$\Pr(u \in S) = \frac{1}{m'} \cdot \Pr\left(r \leq \|u\|^2\right) = \frac{1}{m'} \cdot \|u\|^2.$$

2. Fix orthogonal vectors u and v from X. Simirlaly to Lemma 4.7 (part 2), we have

$$\begin{aligned}
\Pr(u \in S \text{ and } v \in S) &= \Pr\left(\langle \psi(u), g\rangle \geq t \text{ and } \langle \psi(v), g\rangle \geq t \text{ and } r \leq \min(\|u\|^2, \|v\|^2)\right) \\
&= \Pr(\langle \psi(u), g\rangle \geq t \text{ and } \langle \psi(v), g\rangle \geq t) \cdot \min(\|u\|^2, \|v\|^2) \\
&\leq \Pr(\langle (\psi(u) + \psi(v))/2, g\rangle \geq t) \cdot \min(\|u\|^2, \|v\|^2) \\
&= m' \Pr(\langle (\psi(u) + \psi(v))/2, g\rangle \geq t) \cdot \min(\Pr(u \in S), \Pr(v \in S)).
\end{aligned}$$

We need to show that

$$\Pr(\langle (\psi(u) + \psi(v))/2, g\rangle \geq t) \leq 1/(m \cdot m').$$

By Lemma 4.1 (parts 3 and 4), $\|\varphi(u) - \varphi(v)\|_2^2 = 2$. Thus by Corollary 4.5, $\|\psi(u) - \psi(v)\| \geq 2\gamma$, where γ is a positive constant. Hence

$$\mathrm{Var}\left[\langle (\psi(u) + \psi(v))/2, g\rangle\right] = \left\|\frac{\psi(u) + \psi(v)}{2}\right\|^2 \leq 1 - \gamma^2.$$

Now by Lemma A.1.2 from the Appendix,

$$\Pr\left(\left\langle \frac{\psi(u) + \psi(v)}{2}, g\right\rangle \geq t\right) \leq \tilde{\Phi}\left(\frac{t}{\sqrt{1-\gamma^2}}\right) \leq \frac{1}{t}\left(C \cdot \frac{t}{m'}\right)^{\frac{1}{1-\gamma^2}} = \frac{1}{m'} \cdot O\left(\left(\frac{\log m'}{m'}\right)^{\frac{1}{1-\gamma^2}-1}\right).$$

(here $\tilde{\Phi}(x)$ denotes the probability that a standard normal random variable is greater than x). Recall that we fixed m' to be $p(m)$, where $p(x)$ is a polynomial. It is easy to see that for an appropriate $p(x)$ (that depends only on the constant γ) the expression $O\left(\left(\frac{\log m'}{m'}\right)^{1/(1-\gamma^2)-1}\right)$ is less than $1/m$, therefore the value of the right hand side is less than $1/(m \cdot m')$.
3. For all u and v from X,

$$\begin{aligned}
\Pr(u \in S \text{ and } v \notin S) &= \Pr\left(\langle \psi(u), g\rangle \geq t \text{ and } \langle \psi(v), g\rangle \leq t \text{ and } r \leq \min(\|u\|^2, \|v\|^2)\right) \\
&\quad + \Pr\left(\langle \psi(u), g\rangle \geq t \text{ and } \|v\|^2 \leq r \leq \|u\|^2\right) \\
&\leq \Pr(\langle \psi(u), g\rangle \geq t \text{ and } \langle \psi(v), g\rangle \leq t) \cdot \min(\|u\|^2, \|v\|^2) \\
&\quad + 1/m' \cdot \left|\|u\|^2 - \|v\|^2\right|.
\end{aligned}$$

By Lemma A.2 from the Appendix,

$$\begin{aligned}\Pr\left(\langle\psi(u),g\rangle \ge t \text{ and } \langle\psi(v),g\rangle \le t\right) &= O(\|\psi(v)-\psi(u)\|\sqrt{\log m'}/m')\\ &\le O\left(\|\varphi(v)-\varphi(u)\|_2^2\cdot\sqrt{\log n}\cdot\sqrt{\log m'}/m'\right)\\ &\le O\left(\frac{\|v-u\|^2}{\max(\|u\|^2,\|v\|^2)}\cdot\sqrt{\log n}\cdot\sqrt{\log m'}/m'\right).\end{aligned}$$

Therefore,

$$\begin{aligned}\Pr(I_S(u)\ne I_S(v)) &= \Pr\left(u\in S \text{ and } v\notin S\right)+\Pr\left(u\notin S \text{ and } v\in S\right)\\ &\le O\left(\frac{\|v-u\|^2}{\max(\|u\|^2,\|v\|^2)}\cdot\sqrt{\log n}\cdot\sqrt{\log m'}/m'\right)\cdot\min(\|u\|^2,\|v\|^2)\\ &\quad+2/m'\cdot\left|\|u\|^2-\|v\|^2\right|\\ &\le O\left(\|v-u\|^2\sqrt{\log n}\cdot\sqrt{\log m'}/m'\right).\end{aligned}$$

This finishes the proof. □

5. Acknowledgements

We thank Moses Charikar for valuable discussions and comments.

References

[1] A. Agarwal, M. Charikar, K. Makarychev, and Y. Makarychev. *$O(\sqrt{\log n})$ approximation algorithms for Min UnCut, Min 2CNF Deletion, and directed cut problems.* In Proceedings of the 37th Annual ACM Symposium on Theory of Computing, pp. 573–581, 2005.

[2] G. Andersson, L. Engebretsen, and J. Håstad. *A new way of using semidefinite programming with applications to linear equations mod p.* Journal of Algorithms, vol. 39, issue 2, pp. 162–204, May 2001.

[3] S. Arora, J. R. Lee, and A. Naor. *Euclidean distortion and the sparsest cut.* In Proceedings of the 37th Annual ACM Symposium on Theory of Computing, pp. 553-562, 2005.

[4] S. Arora, S. Rao, and U. Vazirani. *Expander flows, geometric embeddings and graph partitioning.* In Proceedings of the 36th Annual ACM Symposium on Theory of Computing, pp. 222–231, 2004.

[5] S. Chawla, A. Gupta, and H. Räcke. *Approximations for generalized sparsest cut and embeddings of ℓ_2 into ℓ_1.* In Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 102–111, 2005.

[6] M. Charikar, K. Makarychev, and Y. Makarychev. *Near-Optimal Algorithms for Unique Games.* To appear in Proceedings of the 38th Annual ACM Symposium on Theory of Computing, 2006.

[7] A. Gupta and K. Talwar. *Approximating Unique Games.* In Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithms, pp. 99–106, 2006.

[8] N. Garg, V. V. Vazirani, and M. Yannakakis. *Approximate max-flow min-(multi)cut theorems and their applications.* In Proceedings of the 25th Annual ACM Symposium on Theory of Computing, pp. 698–707, 1993.

[9] M. Goemans and D. Williamson. *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming.* Journal of the ACM, vol. 42, no. 6, pp. 1115–1145, Nov. 1995.

[10] M. Goemans and D. Williamson. *Approximation algorithms for MAX-3-CUT and other problems via complex semidefinite programming.* Journal of Computer and System Sciences (Special Issue for STOC 2001), vol. 68, issue 2, pp. 442–470, Mar. 2004. Preliminary version in Proceedings of 33rd ACM Symposium on Theory of Computing, pp. 443–452, 2001.

[11] S. Khot. *On the power of unique 2-prover 1-round games.* In Proceedings of the 34th ACM Symposium on Theory of Computing, pp. 767–775, 2002.

[12] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. *Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?* ECCC Report TR05-101, 2005.

[13] J. R. Lee. *On distance scales, embeddings, and efficient relaxations of the cut cone.* In Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 92–101, 2005.

[14] L. Trevisan. *Approximation Algorithms for Unique Games.* In Proceedings of the 46th IEEE Symposium on Foundations of Computer Science, pp. 197–205, 2005.

A. Properties of Normal Distribution

For completeness we prove some standard results used in the paper. Denote the probability that a standard normal random variable is greater than $t \in \mathbb{R}$ by $\tilde{\Phi}(t)$, in other words

$$\tilde{\Phi}(t) \equiv 1 - \Phi_{0,1}(t) = \Phi_{0,1}(-t),$$

where $\Phi_{0,1}$ is the normal distribution function.

Lemma A.1. *1. For every $t > 0$,*

$$\frac{t}{\sqrt{2\pi}(t^2+1)} e^{-\frac{t^2}{2}} < \tilde{\Phi}(t) < \frac{1}{\sqrt{2\pi}t} e^{-\frac{t^2}{2}}.$$

2. There exist constants c_1, C_1, c_2, C_2 and C_3 such that for all $0 < p < 1/3$, $t \geq 0$ and $\rho \geq 1$ the following inequalities hold:

$$\frac{c_1}{\sqrt{2\pi}(t+1)} e^{-\frac{t^2}{2}} \leq \tilde{\Phi}(t) \leq \frac{C_1}{\sqrt{2\pi}(t+1)} e^{-\frac{t^2}{2}};$$

$$c_2\sqrt{\log(1/p)} \le \tilde{\Phi}^{-1}(p) \le C_2\sqrt{\log(1/p)};$$
$$\tilde{\Phi}(\rho t) \le \frac{1}{t}\,(Ct\tilde{\Phi}(t))^{\rho^2}.$$

Proof. 1. Observe, that in the limit $t \to \infty$ all three expressions are equal to 0. Hence the lemma follows from the following inequality on the derivatives:

$$\left(\frac{t}{\sqrt{2\pi}(t^2+1)}e^{-\frac{t^2}{2}}\right)' > \tilde{\Phi}(t)' \equiv -\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}} > \left(\frac{1}{\sqrt{2\pi}t}e^{-\frac{t^2}{2}}\right)'.$$

2. This trivially follows from (1).

□

Lemma A.2. *Let X and Y be two standard normal random variables with covariance $1-2\varepsilon^2$; and let $\tilde{\Phi}(t) = 1/m < 1/3$, $t > 1$. Then*

$$\Pr(X \ge t \text{ and } Y \le t) = O(\varepsilon\sqrt{\log m}/m).$$

Proof. If $\varepsilon t \ge 1$ or $\varepsilon \ge 1/2$, then we are done, since $\varepsilon\sqrt{\log m} = \Omega(\varepsilon t) = \Omega(1)$ and

$$\Pr(X \ge t \text{ and } Y \le t) \le \Pr(X \ge t) = \frac{1}{m}.$$

So assume that $\varepsilon t \le 1$ and $\varepsilon < 1/2$. Let

$$\xi = \frac{X+Y}{2\sqrt{1-\varepsilon^2}}; \qquad \eta = \frac{X-Y}{2\varepsilon}.$$

Then ξ and η are independent standard random variables. We have

$$\begin{aligned}
\Pr\Big(X \ge t \text{ and } Y \le t\Big) &= \Pr\left(\sqrt{1-\varepsilon^2}\,\xi + \varepsilon\eta \ge t \text{ and } \sqrt{1-\varepsilon^2}\,\xi - \varepsilon\eta \le t\right)\\
&= \int_0^\infty \Pr\left(\sqrt{1-\varepsilon^2}\,\xi + \varepsilon x \ge t \text{ and } \sqrt{1-\varepsilon^2}\,\xi - \varepsilon x \le t\right) dF_\eta(x)\\
&= \frac{1}{\sqrt{2\pi}}\int_0^\infty \Pr\left(t - \varepsilon x \le \sqrt{1-\varepsilon^2}\,\xi \le t + \varepsilon x\right) e^{-\frac{x^2}{2}}\,dx\\
&\le \frac{1}{\sqrt{2\pi}}\int_0^{t/\varepsilon} \Pr\left(t - \varepsilon x \le \sqrt{1-\varepsilon^2}\,\xi \le t + \varepsilon x\right) e^{-\frac{x^2}{2}}\,dx + \frac{1}{\sqrt{2\pi}}\int_{t/\varepsilon}^\infty e^{-\frac{x^2}{2}}\,dx.
\end{aligned}$$

The density of the random variable $\sqrt{1-\varepsilon^2}\,\xi$ on the interval $(t-\varepsilon x, t+\varepsilon x)$ for $x \in [0, t/\varepsilon]$ is at most

$$\frac{1}{\sqrt{2\pi(1-\varepsilon^2)}}e^{\frac{-(t-\varepsilon x)^2}{2(1-\varepsilon^2)}},$$

hence

$$\Pr\left(t - \varepsilon x \le \sqrt{1-\varepsilon^2}\,\xi \le t + \varepsilon x\right) \le \frac{2\varepsilon x}{\sqrt{2\pi(1-\varepsilon^2)}}\,e^{-\frac{(t-\varepsilon x)^2}{2(1-\varepsilon^2)}}.$$

Let us bound the first integral in the right hand side.

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} &\int_0^{t/\varepsilon} \Pr\left(t-\varepsilon x \le \sqrt{1-\varepsilon^2}\,\xi \le t+\varepsilon x\right) e^{-\frac{x^2}{2}}\,dx \\
&\le \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{2\varepsilon x}{\sqrt{2\pi(1-\varepsilon^2)}}\, e^{-\frac{(t-\varepsilon x)^2}{2(1-\varepsilon^2)}}\, e^{-\frac{x^2}{2}}\,dx \\
&\le \frac{\varepsilon}{\pi\sqrt{(1-\varepsilon^2)}} \int_0^{\infty} e^{-\frac{(t-\varepsilon x)^2}{2}}\, e^{-\frac{x^2}{2}}\,x\,dx \\
&\le \frac{2\varepsilon e^{-t^2/2}}{\pi} \int_0^{\infty} e^{\varepsilon t x}\, e^{-\frac{x^2}{2}}\,x\,dx \\
&\le \frac{2\varepsilon e^{-t^2/2}}{\pi} \int_0^{\infty} e^{x}\, e^{-\frac{x^2}{2}}\,x\,dx \\
&= O\left(\varepsilon e^{-t^2/2}\right) = O\left(\varepsilon t/m\right) = O\left(\varepsilon\sqrt{\log m}/m\right).
\end{aligned}$$

Now we estimate the second integral.

$$\frac{1}{\sqrt{2\pi}} \int_{t/\varepsilon}^{\infty} e^{-\frac{x^2}{2}}\,dx = \tilde{\Phi}(t/\varepsilon) = O\left(\frac{\varepsilon\, e^{-\frac{t^2}{2\varepsilon^2}}}{t}\right) = O\left(\frac{\varepsilon\, e^{-\frac{t^2}{2}}}{t}\right) = O(\varepsilon/m).$$

$\square$

B. Details of Section 4.2

We sketch the proofs of Corollary 4.4 and Corollary 4.5.

Corollary 4.4. *There exists an efficient algorithm that, given an ℓ_2^2 space X, generates random subsets Y such that the following conditions hold.*

1. *For every u and v in X,*

$$\Pr\left(I_Y(u) \ne I_Y(v)\right) \le D\,\|u-v\|^2.$$

2. *For every u and v s.t. $\|u-v\| \ge 1$,*

$$\Pr\left(I_Y(u) \ne I_Y(v)\right) \ge \beta,$$

where β is a universal constant, $D = O(\sqrt{\log |X|})$.

Sketch of the proof. We apply Theorem 4.3 to the space X with the distance function $d(u,v) = \|u-v\|^2$ and $\Delta = 1$. Let r be a random variable uniformly distributed in $[0, \frac{1}{C\sqrt{\log n}}]$, where C is the constant from Theorem 4.3. Let Y be the r-neighborhood of U. Then

$$\begin{aligned}
\Pr\left(I_Y(u) \ne I_Y(v)\right) &= \Pr\left(d(u,U) < r \le d(v,U) \text{ or } d(v,U) < r \le d(u,U)\right) \\
&\le C\sqrt{\log n}\cdot \mathbb{E}\left[|d(u,U) - d(v,U)|\right] \le C\sqrt{\log n}\cdot \|u-v\|^2.
\end{aligned}$$

We verified condition 1 for $D = C\sqrt{\log n}$. Now if $\|u - v\|^2 \geq 1$ by Theorem 4.3 we have

$$\Pr\left(u \in Y, v \notin Y\right) \geq \Pr\left(u \in U \text{ and } d(v, U) \geq \frac{\Delta}{C\sqrt{\log n}}\right) \geq p.$$

Therefore, $\Pr\left(I_Y(u) \neq I_Y(v)\right) = \Pr\left(u \in Y, v \notin Y\right) + \Pr\left(u \notin Y, v \in Y\right) \geq 2p$. We proved that condition 2 holds for $\beta = 2p$. $\square$

Corollary B.1 (*cf.* [3], Lemma 3.5). *There exists an efficient algorithm, that constructs an embedding h of an ℓ_2^2 space X into L_2 such that the following conditions hold.*

1. *For all u and v in X,*
$$\|h(u) - h(v)\| \leq D\,\|u - v\|^2.$$
2. *For every u and v s.t. $\|u - v\| \geq 1$,*
$$\|h(u) - h(v)\| \geq 2\gamma.$$
3. *The set $h(X)$ lies in the unit ball:*
$$\forall u \in X\ \|h(u)\| \leq 1.$$

where γ is a universal constant; $D = O(\sqrt{\log |X|})$.

Sketch of the proof. We apply Theorem 4.3 to the space X with the distance function $d(u, v) = \|u - v\|^2$ and $\Delta = 1$. Define an embedding h of X into $L_2(\mu)$ as follows:

$$h(u) = \min(C\sqrt{\log n} \cdot d(u, U), 1).$$

We verify that all conditions 1–3 are satisfied.
1. We prove that the expansion of h is at most $D \equiv C\sqrt{\log n}$.

$$\begin{aligned}\|h(u) - h(v)\|^2_{L_2(\mu)} &= \mathbb{E}\left[|h(u) - h(v)|^2\right] = \mathbb{E}\left[|\min(D \cdot d(u, U), 1) - \min(D \cdot d(v, U), 1)|^2\right] \\ &\leq \mathbb{E}\left[\left(D \cdot \|u - v\|^2\right)^2\right] = (D \cdot \|u - v\|^2)^2.\end{aligned}$$

2. Now if $\|u - v\|^2 \geq 1$ by Theorem 4.3 we have

$$\Pr\left(u \in U \text{ and } d(v, U) \geq \frac{\Delta}{C\sqrt{\log n}}\right) \geq p.$$

Therefore, $\Pr\left(h(u) = 0, h(v) = 1\right) \geq p$. Hence $\|h(u) - h(v)\|^2_{L_2(\mu)} \geq \Pr\left(h(u) = 0, h(v) = 1\right) + \Pr\left(h(u) = 1, h(v) = 0\right) \geq 2p$. We verified condition 2 for $\gamma = \sqrt{p/2}$.
3. We have

$$\|h(u)\|^2 = \mathbb{E}\left[\min(C\sqrt{\log n} \cdot d(u, U), 1)^2\right] \leq 1.$$

$\square$

Corollary 4.5. *There exists an efficient algorithm, that constructs an embedding ψ of an ℓ_2^2 space X into ℓ_2 such that the following conditions hold.*

1. *For all u and v in X, $\|\psi(u) - \psi(v)\| \leq D\,\|u - v\|^2$.*
2. *For every u and v s.t. $\|u - v\| \geq 1$, $\|\psi(u) - \psi(v)\| \geq 2\gamma$.*
3. *The set $\psi(X)$ lies on the unit sphere: $\forall u \in X$ $\|\psi(u)\| = 1$,*

where γ is a universal constant; $D = O(\sqrt{\log|X|})$.

Proof. Construct an embedding $h(u)$ from Corollary B.1. Standard arguments show that we can assume that $h(u)$ is an embedding into ℓ_2 (which is isometric to L_2). Define a new embedding as follows:

$$\psi(u) = h(u)/2 + \sqrt{1 - \|h(u)\|^2/4} \cdot e,$$

where e is a unit vector orthogonal to all vectors in $h(X)$. It is easy to see that the embedding ψ satisfies conditions 2 and 3. Let us check condition 1:

$$\begin{aligned}
\|\psi(u) - \psi(v)\| &\leq \|h(u) - h(v)\|/2 + |\sqrt{1 - \|h(u)\|^2/4} - \sqrt{1 - \|h(v)\|^2/4}| \\
&\leq C_1\|h(u) - h(v)\| \leq C_1\, D\|u - v\|^2,
\end{aligned}$$

since $0 \leq \|h(u)\|/2 \leq 1/2$, and the function $\sqrt{1 - x^2}$ is a Lipschitz function on the interval $[0, 1/2]$. $\square$