Approximating Sparsest Cut in Low Rank Graphs via Embeddings from Approximately Low Dimensional Spaces.

Yuval Rabani[*] Rakesh Venkat[†]

Abstract

We consider the problem of embedding a finite set of points $\{x_1, \ldots, x_n\} \in \mathbb{R}^d$ that satisfy ℓ_2^2 triangle inequalities into ℓ_1, when the points are *approximately* low-dimensional. Goemans (unpublished, appears in [19]) showed that such points residing in *exactly* d dimensions can be embedded into ℓ_1 with distortion at most $\sqrt{d}$. We prove the following robust analogue of this statement: if there exists a r-dimensional subspace Π such that the projections onto this subspace satisfy $\sum_{i,j \in [n]} \|\Pi x_i - \Pi x_j\|_2^2 \geq \Omega(1) \sum_{i,j \in [n]} \|x_i - x_j\|_2^2$, then there is an embedding of the points into ℓ_1 with $O(\sqrt{r})$ average distortion. A consequence of this result is that the integrality gap of the well-known Goemans-Linial SDP relaxation for the Uniform Sparsest Cut problem is $O(\sqrt{r})$ on graphs G whose r-th smallest normalized eigenvalue of the Laplacian satisfies $\lambda_r(G)/n \geq \Omega(1)\Phi_{SDP}(G)$. Our result improves upon the previously known bound of $O(r)$ on the average distortion, and the integrality gap of the Goemans-Linial SDP under the same preconditions, proven in [6, 5].

---

[*]The Hebrew University of Jerusalem, Israel. e-mail: `yrabani@cs.huji.ac.il`.

[†]The Hebrew University of Jerusalem, Israel. e-mail: `rakesh@cs.huji.ac.il`. Supported by an I-Core Algorithms Fellowship.

1 Introduction

A finite metric space consists of a pair $(\mathcal{X}, d)$, where $\mathcal{X}$ is a finite set of points, and $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$ is a distance function on pairs of points in $\mathcal{X}$. Many combinatorial optimization problems can be naturally formulated as a maximization or minimization problem over metric spaces $(\mathcal{X}, d)$ of some target class. However, since it might be computationally difficult to optimize over this class, one considers a *relaxation* that finds a solution $(\mathcal{Y}, d')$ amongst a class of computationally 'easy' metrics, and then looks to produce an *embedding* $\mathcal{Y} \hookrightarrow \mathcal{X}$ into the target space, while minimizing some measure of *distortion* between the distance functions d and d' incurred by the embedding. There has been much work that investigates various measures and costs of distortion incurred by embeddings between metric spaces, and applications thereof (see the surveys [11, 20, 17] and references therein).

In this work, we look at embeddings from ℓ_2^2 metrics to ℓ_1 metrics, motivated by applications to the Sparsest Cut problem. A ℓ_1 metric (or a ℓ_1 space) consists of a finite set of points represented in $\mathbb{R}^d$ with the distance given by the ℓ_1 distance between them. It is a natural target space that can be viewed as an non-negative combination of 'cut-metrics' on the underlying point set, and hence arises frequently in graph-cut based problems. A ℓ_2^2 space, on the other hand, is easy to optimize over, and consists of a finite set of points, say $X = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$, that satisfy triangle inequalities on the *squares* of distances:

$$\|x_i - x_j\|_2^2 + \|x_j - x_k\|_2^2 \geq \|x_i - x_k\|_2^2 \qquad \forall\, i, j, k \in [n]. \tag{1.1}$$

The Sparsest Cut problem is a fundamental NP-hard graph optimization problem that serves as a striking example of the utility of the metric embedding approach. In the (Uniform) Sparsest Cut problem, we are given a graph $G = (V, c)$, with a symmetric weight function c_{ij} on pairs $\{i, j\}$. The goal is to find a cut $(S, \overline{S})$ of minimum *sparsity* $\Phi(S)$, defined as follows (here, $\mathbb{I}_S(i)$ is 1, if $i \in S$, and 0 otherwise).

$$\Phi(S) := \frac{\sum_{i<j} c_{ij} \left|\mathbb{I}_S(i) - \mathbb{I}_S(j)\right|}{\sum_{i<j} \left|\mathbb{I}_S(i) - \mathbb{I}_S(j)\right|}$$

The best known approximation for the Sparsest Cut problem is due to Arora, Rao and Vazirani [3] (henceforth called the ARV algorithm), who considered the following semidefinite programming relaxation (SDP) introduced by Goemans and Linial (see [8] and [17]).

$$\textbf{SDP-1:} \quad \Phi_{SDP}(G) := \min_{\{x_i\}_{i \in [n]}} \frac{1}{n^2} \sum_{ij} c_{ij} \|x_i - x_j\|_2^2$$

$$\text{s.t} \quad \begin{cases} \|x_i - x_j\|_2^2 + \|x_j - x_k\|_2^2 \geq \|x_i - x_k\|_2^2 & \forall i, j, k \in [n]. \\ \sum_{kl} \|x_k - x_l\|_2^2 = n^2. \end{cases}$$

Clearly, $\Phi_{SDP}(G) \leq \Phi(G)$. Notice that any feasible solution to the above SDP constitutes a ℓ_2^2 space. The ARV algorithm works by producing an embedding of the solutions of the above SDP into a ℓ_1 space, with *average distortion* (see Section 2 for a definition) $O(\sqrt{\log n})$. It was shown in [18, 4] that producing an embedding of the SDP solutions into a ℓ_1 space with average distortion D suffices to get a $O(D)$ approximation to the Uniform Sparsest Cut problem.

Though the solutions to SDP-1 can lie in up to n dimensions, for certain graph classes, they are more structured. In particular, if the r-th smallest eigenvalue of the graph Laplacian satisfies $\lambda_r(G)/n \gg \Phi_{SDP}(G)$, then it turns out that the solutions are *approximately* r-dimensional (see Definition 1.2 and Section 3.4). Graphs whose r-th smallest eigenvalue is bounded away from 0 for a typically small r are called *low threshold-rank* graphs; note that spectral expanders are a special case of these for $r = 2$. The

work of Guruswami and Sinop [10] exploited higher levels of the Lasserre SDP hierarchy [15], along with the above structure, to give constant-factor guarantees for Sparsest Cut on these graphs. However, this involved partially solving a SDP of size $n^{O(r)}$[1], and did not say anything about the behaviour of the Goemans-Linial SDP on these graphs.

Goemans showed that if the points satisfying ℓ_2^2 triangle inequalities lie in d dimensions, then they can be embedded into ℓ_2 (and hence into ℓ_1, since there is an isometry from ℓ_2 to ℓ_1 [20]) with $\sqrt{d}$ distortion (unpublished, appears in [19], see also [5, Section 4] for an alternative proof).

Theorem 1.1 (Goemans [19, Appendix B]). *Let $x_1, x_2, \ldots, x_n \in \mathbb{R}^d$ be n points satisfying ℓ_2^2 triangle inequalities. Then there exists an embedding of these points into ℓ_2, $x_i \mapsto f(x_i)$, with distortion $\sqrt{d}$, that is,*

$$\frac{1}{\sqrt{d}} \, \|x_i - x_j\|_2^2 \leq \|f(x_i) - f(x_j)\|_2 \leq \|x_i - x_j\|_2^2, \quad \forall \, i, j \in V.$$

The immediate question that this raises is the following: can one reduce the dimension of ℓ_2^2 metrics, while preserving pairwise distances, *and* the ℓ_2^2 triangle inequalities? The Johnson-Lindenstrauss lemma [12] reduces the dimension to $O(\log n)$, while preserving pairwise distances approximately. However, this procedure does not preserve the ℓ_2^2 triangle inequalities, if the original points satisfied them. In fact, Magen and Moharammi [19] prove a strong lower bound against dimension reduction for ℓ_2^2 metrics.

It is interesting to note that the Johnson-Lindenstrauss lemma, while not preserving the ℓ_2^2 triangle inequalities exactly, does preserve them *approximately*, that is, every sequence of $k \leq n$ points $x_{i_1}, \ldots, x_{i_k}$ satisfies $\sum_{j=1}^{k-1} \left\| x_{i_j} - x_{i_{j+1}} \right\|_2^2 \geq \beta \cdot \|x_{i_1} - x_{i_k}\|_2^2$, for some $\beta = \Omega(1)$. An observation by Luca Trevisan (personal communication) shows that, in fact, Goemans' theorem is also true for points satisfying approximate triangle inequalities, but the proof uses the ARV machinery. However, even this does not yield anything better that $O(\sqrt{\log n})$, for approximately low-dimensional points.

The above discussion motivates one to ask if there is a more 'robust' analogue of Goemans' theorem that can be applied to low threshold-rank graphs. Deshpande, Harsha and Venkat [5] considered this question, and showed that one can prove a similar theorem for the case where the points are in approximately r dimensions, albeit giving a bound of $O(r)$ on the *average* distortion (which suffices for Sparsest Cut). One would expect an exact analogue to have a bound of $O(\sqrt{r})$, and it was left open if one could find such an embedding.

We show that there is, indeed, an embedding into ℓ_1 (in fact, into ℓ_2, since all our embeddings are one-dimensional) with $O(\sqrt{r})$ average distortion when the points are approximately r-dimensional.

1.1 Our Results

In order to state our main result, we use the following definition to quantify the notion of approximate rank of a set of points:

Definition 1.2. (η-Subspace rank) For any $\eta \in (0, 1]$, a set of points $X = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$ will be said to have η-subspace rank r, denoted by $\mathrm{ssr}_\eta(X) = r$, if there exists a subspace given by a projector $\Pi \in \mathbb{R}^{d \times d}$ with $\mathrm{rank}\,(\Pi) = r$ that satisfies:

$$\sum_{i,j \in [n]} \|\Pi x_i - \Pi x_j\|_2^2 \ \geq \ \eta \sum_{i,j \in [n]} \|x_i - x_j\|_2^2. \tag{1.2}$$

In this work, we will always consider $\eta = \Omega(1)$.

---

[1]In a separate work, Guruswami and Sinop [9] give an algorithm that solves the SDP partially, running in $2^{O(r)}\mathrm{poly}\,(n)$ time, and suffices for their algorithm.

Remark. *Since the subspace Π_r defined by the top-r left singular vectors of the matrix M with columns $\{x_i - x_j\}_{ij}$ satisfies $\|\Pi_r M\|_F^2 \geq \left\|\widetilde{\Pi} M\right\|_F^2$ for every $\widetilde{\Pi}$ with $\mathrm{rank}\left(\widetilde{\Pi}\right) \leq r$, we can always assume that $\Pi = \Pi_r(M)$ when we need to explicitly use the projections. Also, note that the subspace rank is independent of any scaling or shifting of the points, and is always at most the rank of the point set.*

Deshpande et al. [5] use a slightly different, stronger notion of approximate dimension, called the *stable-rank* of the point set, defined as $\mathrm{sr}(M) = \|M\|_F^2/\sigma_1(M)^2$, where σ_1 is the maximum singular value of the matrix M. Clearly, $\mathrm{sr}(M) \leq \mathrm{ssr}_\eta(X)/\eta$. For applications to the Sparsest Cut problem, the notion of subspace rank suffices and is natural (see Section 3.4). It would be interesting to see if other notions of approximate rank yield further applications or improvements, in Sparsest Cut, or elsewhere.

Our main result is the following:

Theorem 1.3. *Given a set of points $X = \{x_1, \ldots, x_n\} \in \mathbb{R}^d$ with $\mathrm{ssr}_\eta(X) = r$ that satisfy the ℓ_2^2 triangle inequalities, there is an embedding $X \hookrightarrow \ell_1$ with average distortion at most $O_\eta(\sqrt{r})$. That is, there is a constant $c(\eta)$ and a mapping $h : X \to \mathbb{R}^{d'}$ that satisfies:*

$$\|h(x_i) - h(x_j)\|_1 \leq \|x_i - x_j\|_2^2 \qquad \forall i, j \in [n] \tag{1.3}$$

$$\sum_{i,j\in[n]} \|h(x_i) - h(x_j)\|_1 \geq \frac{c(\eta)}{\sqrt{r}} \cdot \sum_{ij} \|x_i - x_j\|_2^2 \tag{1.4}$$

This matches Goemans' theorem in terms of the dependence on r, albeit for average-case distortion. Since the subspace rank is an *average* global condition on the point set, we cannot hope to prove a worst-case distortion guarantee like Goemans' theorem that depends only on the subspace rank (see Appendix A.1).

The above theorem holds even if the points satisfy the ℓ_2^2 triangle inequalities only *approximately*, since the steps in the analysis of the algorithm only need the points to satisfy an approximate version of the triangle inequalities[2]. Improving on the $\sqrt{r}$ bound above with any technique that works with approximate triangle inequalities would imply an improvement over the ARV algorithm's guarantee, since dimension reduction using the Johnson-Lindenstrauss [12] transform preserves pairwise distances (and hence the ℓ_2^2 inequalities) approximately, while reducing the dimension to $O(\log n)$.

Our main result immediately implies a $O(\sqrt{r})$ approximation algorithm for the Uniform Sparsest Cut problem on low threshold-rank graphs, using just the Goemans-Linial SDP.

Corollary 1.4. *Let $\epsilon \in (0, 1]$. Given a regular graph G with r-th smallest eigenvalue of the normalized Laplacian satisfying $\lambda_r(G) \geq \Phi_{SDP}(G)/(1-\epsilon)$, we can find a $O_\epsilon(\sqrt{r})$ approximation to the sparsest cut in the graph using SDP-1.*

This improves upon the previously known guarantee of $O(r/\epsilon)$ using the Goemans-Linial SDP in [5], under the same precondition.

Proof Techniques:
In order to prove our main result, we follow the generic approach of the ARV algorithm [3] that proceeds in two steps: If there is a dense cluster of the solution vectors, then a specific Fréchet embedding (see Section 2 for a definition) works. If not, then the solutions are 'well-spread', and one can always find two $\Omega(n)$-sized sets that are $O(1/\sqrt{\log n})$-apart in ℓ_2^2 distance, using a separating hyperplane algorithm.

---

[2]The points are said to satisfy *approximate ℓ_2^2* triangle inequalities, if every sequence of $k \leq n$ points $x_{i_1}, \ldots, x_{i_k}$ satisfies $\sum_{j=1}^{k-1} \|x_{i_j} - x_{i_{j+1}}\|_2^2 \geq \beta \cdot \|x_{i_1} - x_{i_k}\|_2^2$, for some $\beta = \Omega(1)$

This constitutes the core of the proof, and the analysis involves a 'chaining argument' which relies on the concentration of measure in high-dimensional spaces. These well-separated sets can then be used to construct a good Fréchet embedding into ℓ_1.

In our case, we would analogously like to find two large sets that are $\Omega(1/\sqrt{r})$-apart, and to do this, we need to work with the *projections* of the points. Note that the projections need not be in ℓ_2^2, while the ARV algorithm's analysis requires the use of ℓ_2^2 triangle inequalities at various points.

Thus, in order to prove Theorem 1.3, we follow and adapt the techniques in Naor, Rabani and Sinclair [21] (henceforth called the NRS analysis). Their work generalized the ARV algorithm's analysis to apply to the more general case of metrics *quasisymmetrically embeddable* into ℓ_2, which includes ℓ_2^2 as a special case. We do not need the complete machinery developed by them, though, and extend only a part of their analysis to our setting. In particular, the chaining argument in [21] works in Euclidean, rather than ℓ_2^2 space, making it useful in our case.

Our result, thus, also demonstrates the utility of isolating the chaining argument from the use of ℓ_2^2 triangle inequalities in the ARV algorithm's analysis.

1.2 Other related Work

We recall that the best known upper bound for the worst-case distortion of embedding $\ell_2^2 \hookrightarrow \ell_1$ is $O(\sqrt{\log n} \cdot \log \log n)$ by [2], building on the techniques in [3, 16]. The best known lower bound is $\Omega(\sqrt{\log n})$ for worst-case distortion [22], and $\exp(\Omega(\sqrt{\log \log n}))$ for average distortion [13]. On *low threshold-rank* graphs (where $\lambda_r \geq \Omega(1)\Phi_{SDP}$), an approximation guarantee of $O(1)$ for Sparsest Cut was obtained using $O(r)$ levels of the Lasserre hierarchy for SDPs [10]. In contrast, the works [6, 5] obtained a weaker $O(r)$ approximation, but using just the basic SDP relaxation. Oveis Gharan and Trevisan [7] also give a rounding algorithm for the basic SDP relaxation on low-threshold rank graphs, but require a stricter pre-condition on the eigenvalues ($\lambda_r \gg \log^{2.5} r \cdot \Phi(G)$), and leverage it to give a stronger $O(\sqrt{\log r})$-approximation guarantee. Their improvement comes from a new structure theorem on the SDP solutions of low threshold-rank graphs being clustered, and using the techniques in ARV for analysis.

Kwok et al. [14] showed that a better analysis of Cheeger's inequality gives a $O(r \cdot \sqrt{1/\lambda_r})$ approximation to the sparsest cut on regular graphs. In particular, when $\lambda_r(G) \geq \epsilon$, this gives a $O(r/\sqrt{\epsilon})$ approximation. Note that our result gives a better approximation in this setting (see Section 3.4).

2 Notation

We use $[n] = \{1, \ldots, n\}$. For a matrix $M \in \mathbb{R}^{d \times d}$, we say $M \succeq 0$ or M is positive-semidefinite (psd) if $y^T X y \geq 0$ for all $y \in \mathbb{R}^d$. The unit Euclidean Ball in $\mathbb{R}^d$ is denoted by B_2^d.

Graphs and Laplacians: All graphs will be defined on a vertex set $V = [n]$ of size n. The vertices will usually be referred to by indices $i, j, k, l \in [n]$. Given a graph with a symmetric weight function on pairs $W : V \times V \mapsto \mathbb{R}^+$, with $W(i, i) = 0 \, \forall i$, let $D(i) := \sum_j W(i, j)$ be the degree of vertex $i \in V$. The (normalized) graph Laplacian matrix is defined as:

$$L_W(i, j) := \begin{cases} -\frac{W(i,j)}{\sqrt{D(i)D(j)}} & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Note that $L_W \succeq 0$. We will denote the eigenvalues of (the Laplacian of) the graph G by $0 = \lambda_1(G) \leq \lambda_2(G) \ldots \leq \lambda_n(G)$, in *increasing* order. If the graph is *c-regular*, we have $D(i) = c$ for every $i \in V$. Note that c might be a fraction.

For nodes i, j in G, $d_G(i, j)$ is the shortest path between vertices i, j in G. For $S \subseteq [n]$, $G[S]$ is the subgraph induced by G on S. The *vertex expansion* of G, denoted by $h(G)$ is defined as the largest constant h such that for every set $S \subseteq V$ with $1 \geq |S| \geq |V|/2$, $|N_G(S)| \geq h|S|$ where $N_G(S) = \{j \in V : d_G(j, S) = 1\}$.

Embeddings and cuts: For our purposes, a (semi-)metric space (X, d) consists of a finite set of points $X = \{x_1, x_2, \ldots, x_n\}$ and a distance function $d : X \times X \mapsto \mathbb{R}_{\geq 0}$ satisfying the following three conditions:

1. $d(x, x) = 0, \forall x \in X$.

2. $d(x, y) = d(y, x)$.

3. (Triangle inequality) $d(x, y) + d(y, z) \geq d(x, z)$.

An *embedding* from a metric space (X, d) to a metric space (Y, d') is a mapping $f : X \to Y$. The embedding is called a *contraction*, if

$$d'(f(x_i), f(x_j)) \leq d(x_i, x_j), \qquad \forall x_i, x_j \in X.$$

For convenience, we will only deal with contractive mappings in this paper (this is without loss of generality). A contractive mapping is said to have (worst-case) distortion Δ, if: $\sup_{i,j} \frac{d(x_i, x_j)}{d'(f(x_i), f(x_j))} \leq \Delta$. It is said to have *average* distortion β, if $\frac{\sum_{i<j} d(x_i, x_j)}{\sum_{i<j} d'(f(x_i), f(x_j))} \leq \beta$.

Note that a mapping with worst-case distortion Δ also has average distortion Δ, but not necessarily vice-versa. *Fréchet* embeddings of (X, d) are a class of embeddings of $X \to \mathbb{R}^k$ into defined on the basis of distances to point sets: a co-ordinate of the embedding will be given by a map of the form $d(x_i, S) := \min_{j \in S} d(x_i, x_j)$ for some $S \subseteq X$. Note that Fréchet embeddings are always contractive in every co-ordinate.

When $X \subseteq \mathbb{R}^k$ is a ℓ_2^2 space, we will use $d(i, j) := \|x_i - x_j\|_2^2$, and $d(S, T) = \min_{i \in S, j \in T} d(i, j)$ for $S, T \subseteq [n]$. For $c \in \mathbb{R}$, $B(i, c) := \{j : d(i, j) \leq c\}$. We refer to the quantity $\frac{1}{n^2} \sum_{i,j} \|x_i - x_j\|_2^2$ as the *spread* of these points.

3 Proof of Main Theorem

3.1 Proof Outline

We prove Theorem 1.3 in two steps. First, we scale the points to lie within a ℓ_2 ball of radius 1; note that this would shrink the pairwise distances. Suppose that the points have *constant spread* after this scaling; i.e. they satisfy

$$\frac{1}{n^2} \sum_{i,j \in V} \|x_i - x_j\|_2^2 \geq \delta, \qquad \text{where } \delta = \Omega(1). \tag{3.1}$$

Since scaling does not affect the subspace rank, we continue to have $\mathrm{ssr}_\eta(X) = r$. In this case, we adapt the chaining argument from [21] to work on the *projections* $\{\Pi x_i\}_{i \in V}$ to conclude the existence of two large, Δ-separated sets for $\Delta = \Omega(1/\sqrt{r})$.

In the general case, we show that by appropriately utilizing the subspace criterion, we can either reduce it to the case of constant spread, or produce an $O(1)$ distortion Fréchet embedding by considering distances to an appropriate ℓ_2^2 ball centered at one of the points.

Let $V := [n]$. We will require the following definitions, following [3]:

Definition 3.1 (Largeness). A subset $A \subseteq V$ is β-large, if $|A| \geq \beta n$.

5

Definition 3.2 (Δ-separation). Subsets $L \subseteq V$ and $R \subseteq V$ are Δ-separated, if $d(L, R) \geq \Delta$

The following lemma, implicit in [3], gives a sufficient condition for the existence of a Fréchet embedding into ℓ_1 with low average distortion.

Lemma 3.3 (Sufficient condition). *If there is a set $S \subseteq [n]$ satisfying*

$$|S| \sum_{i \notin S} d(i, S) \geq c.n^2 \tag{3.2}$$

Then, there is an embedding of the points into ℓ_1 with average distortion $1/c$.

Proof. Consider the embedding $i \mapsto d(i, S)$. Clearly, this is a Fréchet embedding, and hence a contraction. Furthermore, we have:

$$\sum_{i,j \in V} |d(i, S) - d(j, S)| \geq \sum_{i \notin S, j \in S} |d(i, S) - 0|$$
$$= |S| \sum_{i \notin S} d(i, S) \geq cn^2$$

Thus, the average distortion of the map is at most $1/c$. $\qquad\square$

Note that the existence of two $\Omega(1)$-large, Δ-separated sets L, R would satisfy the above condition, with $S = L$ and $c = O(1/\Delta)$. The above can also be thought of as an embedding into ℓ_2, since it is one-dimensional.

3.2 The constant spread case

We will start by stating the following Proposition, which is a simple modification of Proposition 3.11 in [21]. Since the proof closely follows the original, requiring only a simple observation, we do not give it here.

Proposition 3.4 (From Proposition 3.11 in [21]). *Let $G = (V, E)$ be graph with vertex expansion $h(G) \geq 1/2$. Let $f : V \to B_2^d$ be a mapping that satisfies:*

$$\frac{1}{n^2} \sum_{i,j \in V} \|f(i) - f(j)\|_2 \geq \gamma \tag{3.3}$$

Then, there exists a pair $i, j \in V$, and constants $c_1(\gamma), c_2(\gamma)$ such that

$$\|f(i) - f(j)\|_2 \geq c_1(\gamma) \quad \text{and} \quad d_G(i, j) \leq c_2(\gamma)\sqrt{d} \tag{3.4}$$

Remark. *The modification only requires the observation that for any i, j with $\|f(i) - f(j)\|_2 \leq c_1(\gamma)$, and $u : \|u\|_2 = 1$, $\langle f(i) - f(j), u \rangle \leq c_1(\gamma)$. This avoids a union bound over the pairs of points in the last step of the proof, the rest of the steps being identical. Combined with the original statement of Proposition 3.11 in [21], the term $\sqrt{d}$ in the above can be replaced by $\min\left\{\sqrt{\log n}, \sqrt{d}\right\}$.*

We now proceed to prove a special case of Theorem 1.3 assuming condition (3.1).

Theorem 3.5. *Let* $X = \{x_1, \dots, x_n\}$ *satisfy* ℓ_2^2*-triangle inequalities, with* $X \subseteq B_2^d$ *and* $\mathrm{ssr}_\eta(X) = r$. *Furthermore, suppose that*

$$\frac{1}{n^2} \sum_{ij} \|x_i - x_j\|_2^2 \geq \delta, \qquad \text{where } \delta = \Omega(1).$$

Then there exist sets $A, B \subseteq X$, *with* $|A|, |B| \geq (\eta\delta/32)n$ *with* $d(A, B) \geq \Omega(1/\sqrt{r})$.

Proof. Let Π be the r-dimensional subspace containing an η fraction of the squared lengths of the difference vectors upon projection. Let $V = [n]$, and define $f : V \to B_2^r$ by

$$f(i) \triangleq \Pi x_i$$

Since the set X has η-subspace rank r, we have, by definition:

$$\frac{1}{n^2} \sum_{i,j \in V} \|f(i) - f(j)\|_2^2 \geq \eta\delta. \tag{3.5}$$

We will now follow the proof of Theorem 2.4 in [21], but switch to the projections where appropriate. Consider the graph $G = (V, E)$ with edges $E = \left\{ \{i, j\} : \|x_i - x_j\|_2^2 \leq \frac{\kappa}{\sqrt{r}} \right\}$, where $\kappa = \kappa(\eta, \delta)$ is a constant that we will set later.

Suppose, for the sake of contradiction, that every two sets $A, B \subseteq V$ with $|A|, |B| \geq (\eta\delta/32)n$ satisfy $d(A, B) \leq \kappa/\sqrt{r}$, which implies that $d_G(A, B) \leq 1$. We use the following lemma from [21]:

Lemma 3.6 (Lemma 2.3 in [21]). *Fix* $0 < \epsilon \leq \frac{1}{10}$, *and let* $G = (V, E)$ *be a graph such that for every* $X, Y \subseteq V$ *satisfying* $|X|, |Y| \geq \epsilon|V|$, $d_G(x, y) \leq 1$. *Then there is a* $U \subseteq V$ *with* $|U| \geq (1 - \epsilon)|V|$ *with* $h(G[U]) \geq \frac{1}{2}$.

Invoking Lemma 3.6 on G yields a subset $X' \subseteq V$, with $|X'| \geq (1 - \frac{\eta\delta}{32})n$ such that $h(G[X']) \geq \frac{1}{2}$. We claim the following:

$$\frac{1}{|X'|^2} \sum_{i,j \in X'} \|f(i) - f(j)\|_2 \geq \frac{(\eta\delta)^{3/2}}{32}. \tag{3.6}$$

To see this, note that $|X' \times X'| \geq (1 - \frac{\eta\delta}{16})n^2$. Let $D = \left\{ (i, j) \in V \times V : \|f(i) - f(j)\|_2^2 \geq \eta\delta/4 \right\}$. Since the diameter of the unit ball is 2, in order to satisfy (3.5), we should have $|D| \geq (\eta\delta/8)n^2$. Thus, $|D \cap (X' \times X')| \geq \frac{\eta\delta}{16}n^2$. This implies that the average ℓ_2-distance in $X' \times X'$ is at least:

$$\frac{1}{n^2} |D \cap (X' \times X')| \times \sqrt{\frac{\eta\delta}{4}} \geq \frac{(\eta\delta)^{3/2}}{32}. \tag{3.7}$$

This proves (3.6).

We can now apply Proposition 3.4 to $G[X']$, and the projections $\{f(i)\}_{i \in V}$, with $\gamma = (\eta\delta)^{3/2}/32$. We infer that there exists a path in G, of $k \leq c_2(\gamma)\sqrt{r} = a(\eta, \delta)\sqrt{r}$ vertices $i_1, i_2, \dots i_k \subseteq X'$ such that $\|f(i_1) - f(i_k)\|_2 \geq c_1(\gamma) = b(\eta, \delta)$, where $a(\eta, \delta)$ and $b(\eta, \delta)$ are constants depending on η and δ.

This implies that:

$$b^2(\eta, \delta) \overset{(a)}{\leq} \|f(i_1) - f(i_k)\|_2^2 \overset{(b)}{\leq} \|x_{i_1} - x_{i_k}\|_2^2 \overset{(c)}{\leq} \sum_{j=1}^{k-1} \|x_{i_j} - x_{i_{j+1}}\|_2^2 \overset{(d)}{\leq} a(\eta, \delta)\sqrt{r}\frac{\kappa}{\sqrt{r}}. \tag{3.8}$$

Above, (b) follows from the fact that projections can only decrease distances, (c) from the ℓ_2^2 property, and (d) from the definition of G. This is a contradiction, if we set $\kappa < \frac{b^2(\eta, \delta)}{a(\eta, \delta)}$. $\qquad \square$

Remark. *The last chain of inequalities above is the only place where the ℓ_2^2 triangle inequalities are invoked. Without them, we could still prove a weaker statement with $O(1/r)$ separation between the large sets, since (c) would hold with an additional multiplicative factor of k by convexity.*

3.3 The general case

We now extend our argument to the general case. Let us fix some notation before going to the proofs. We will take $V := [n]$, and $X = \{x_1, \ldots, x_n\}$ to satisfy the ℓ_2^2 triangle inequalities, with $\mathrm{ssr}_\eta(X) = r$. Let Π be the corresponding r-dimensional subspace. Let $f(i) := \Pi x_i$, as before. Define

$$d_f(i,j) := \|f(i) - f(j)\|_2^2$$

The terms $d_f(i,S), d_f(S,T)$ for $S, T \subseteq V$ are defined naturally, and denote $\mathrm{diam}_f(S) \triangleq \max_{i,j \in S} d_f(i,j)$. Note that $d_f(\cdot, \cdot)$ is *not necessarily* a distance, unlike $d(\cdot, \cdot)$. However, since f is a projection map, it satisfies:

$$d(i, S) \geq d_f(i, S) \qquad \forall i \in V, \forall S \subseteq V, \tag{3.9}$$

We will also assume that X is scaled to satisfy:

$$\frac{1}{n^2} \sum_{i,j \in V} \|x_i - x_j\|_2^2 = 1 \tag{3.10}$$

We first record a simple observation.

Observation 3.7. *For any $i, j \in V$, and any $S \subseteq V$,*

$$d_f(i,j) \leq 3\left(d_f(i, S) + \mathrm{diam}_f(S) + d_f(j, S)\right).$$

Proof. Let $i^*, j^* \in S$ be such that $d_f(i, S) = d_f(i, i^*)$ and $d_f(j, S) = d_f(j, j^*)$. Since $\sqrt{d_f}$ obeys the triangle inequality, we have:

$$\left(\sqrt{d_f(i,j)}\right)^2 \leq \left(\sqrt{d_f(i, i^*)} + \sqrt{d_f(i^*, j^*)} + \sqrt{d_f(j, j^*)}\right)^2$$
$$\leq 3(d_f(i, S) + \mathrm{diam}_f(S) + d_f(j, S))$$

The last inequality follows from the convexity of the function $g(x) = x^2$, and the definition of diam_f. $\square$

We now consider various cases, and show that a low average-distortion embedding exists in each case.

Lemma 3.8 (Dense Ball). *If $\exists i \in V$, with $|B(i, 1/12)| \geq n/12$, then we can find an $O(1)$-average distortion embedding of X into ℓ_1.*

Proof. The proof follows the proof of a similar lemma in [3]. Let $i_0 \in V$ be such that $|B(i_0, 1/12)| \geq n/12$, and let $S = B(i_0, 1/12)$. Consider the embedding $i \mapsto d(i, S)$. This is a contraction. Since $\sum_{ij} \|x_i - x_j\|_2^2 = n^2$, we have :

$$n^2 = \sum_{i,j \in V} d(i,j)$$
$$\leq \sum_{i,j \in V} (d(i, S) + d(j, S)) \qquad \ldots \text{Using } \ell_2^2 \text{ triangle inequality}$$

8

$$= 2n \left(\sum_{i \notin S} d(i, S) \right)$$

This gives us that $\sum_{i \notin S} d(i, S) \geq n/12$. Since $|S| = \Omega(n)$, Lemma 3.3 applies, and proves that the above embedding has $O(1)$ average-distortion. [3] $\qquad \square$

Lemma 3.9 (Isolating a bounded ball). *If there is no $i \in V$ such that $|B(i, 1/12)| \geq n/12$, then there is a $j \in V$ such that $S = B(j, 12/9)$ satisfies $|S| \geq \frac{3}{12}n$, and*

$$\sum_{i,j \in S} d(i,j) \geq \left(\frac{2}{12} \right) \left(\frac{1}{12} \right) \frac{n^2}{12}$$

Proof. Suppose we had $|B(j, 12/9)| < (3n/12)$ for every $j \in V$. Then, for any $j \in V$, we would have $|\overline{B(j, 12/9)}| > 9n/12$, which gives us that $\sum_i d(j, i) > n$. Summing over $j \in V$ contradicts (3.10).
Now, let $j_0 := \arg\max_{j \in V} |B(j, 12/9)|$, and $S := B(j_0, 12/9)$. Define the set $A = B(j_0, 12/9) \setminus B(j_0, 1/12)$. From our assumption and the preceeding argument, $|A| \geq 2n/12$. Since $|B(i, 1/12)| \leq n/12$ for every $i \in A$, we have that $\left| \overline{B(i, 1/12)} \cap A \right| \geq n/12$. This gives us:

$$\sum_{i \in A, j \in A} d(i,j) \geq \frac{2n}{12} \times \frac{1}{12} \times \frac{n}{12}$$

$\qquad \square$

In next two lemmas, assume that the precondition of Lemma 3.9 holds, i.e., there is no $i \in V$ with $|B(i, 1/12)| \geq n/12$.

Lemma 3.10. *Let $j_0 = \arg\max_{j \in V} |B(j, 12/9)|$, and $S \triangleq B(j_0, 12/9)$. If S satisfies:*

$$\sum_{i,j \in S} d_f(i,j) \geq \frac{\eta}{600} |S|^2,$$

then there is an embedding of X into ℓ_1 with $O(\sqrt{r})$ average distortion.

Proof. Consider the map $g : V \to \mathbb{R}^d$ given by $g(i) \triangleq \sqrt{9/12} \cdot x_i$. This ensures that $g(i) \in B_2^d$ for every $i \in S$, and the mapping continues to obey the ℓ_2^2 triangle inequalities. Furthermore, from Lemma 3.9, the points in S satisfy:

$$\frac{1}{|S|^2} \sum_{i,j \in S} \|g(i) - g(j)\|_2^2 \geq \frac{9}{12} \times \frac{2}{12^3} = \Omega(1) \qquad (3.11)$$

From the assumption on S, we infer that:

$$\frac{1}{|S|^2} \sum_{i,j \in S} \|\Pi g(i) - \Pi g(j)\|_2^2 \geq \frac{9}{12} \times \frac{\eta}{600}$$

We can now invoke Theorem 3.5 on just the points in S to conclude that there exist sets $A, B \subseteq S$, such that $|A|, |B| \geq \Omega_\eta(n)$ with $d(A, B) \geq \Omega_\eta(1/\sqrt{r})$ (the scaling by a constant factor just shrinks some distances). As before, it is easy to see that A satisfies the conditions of Lemma 3.3 with $c = \Omega(1/\sqrt{r})$ and hence the mapping $h(i) \triangleq d(i, A)$ has average distortion $O(\sqrt{r})$. Note that by the ARV algorithm [3], the sets can be found with good probability by a random separating hyperplane through j_0. $\qquad \square$

---

[3] Strictly speaking, one could do without the ℓ_2^2 triangle inequality here by adjusting the constants appropriately, as we did in Observation 3.7.

Lemma 3.11. *Let $j_0 = \arg\max_{j \in V} |B(j, 12/9)|$, and $S \triangleq B(j_0, 12/9)$. If S satisfies:*

$$\sum_{ij \in S} d_f(i, j) \leq \frac{\eta}{600} |S|^2,$$

then we can find an embedding of X into ℓ_1 with $O(1)$ average distortion.

Proof. The proof will be similar to the proof of Lemma 3.8, except for the fact that we will work with projections instead of the original vectors. First, observe that there exists an $i_0 \in S$ such that $|B_f(i_0, \eta/24) \cap S| \geq 24|S|/25$. If not, then for every $i \in S$, we will have $\sum_{j \in S} d_f(i, j) > \frac{1}{25}|S| \times \eta/24 = \eta|S|/600$. Summing over $j \in S$ results in a contradiction to the precondition on S.

Let $T \triangleq B_f(i_0, \eta/24)$; from the preceding argument, we have $|T| = \Omega(n)$.

Claim 3.12. $\sum_{j \notin T} d_f(j, T) \geq \eta n/12$

Proof. We know that $\sum_{i,j \in V} \|f(i) - f(j)\|_2^2 = \sum_{i,j \in V} d_f(i, j) \geq \eta n^2$. Using Observation 3.7, we can infer:

$$\eta n^2 \leq \sum_{i,j \in V} d_f(i, j)$$

$$\leq 3 \sum_{i,j \in V} (d_f(i, T) + \operatorname{diam}_f(T) + d_f(j, T)) \qquad \ldots \text{Using Observation 3.7}$$

$$= 3 \left(2n \sum_{i \in V} d_f(i, T) + \frac{4\eta}{24} n^2 \right) \qquad \ldots \text{Since } \operatorname{diam}_f(T) \leq \frac{4\eta}{24}$$

This yields that $\sum_i d_f(i, T) \geq \frac{\eta}{12} n$, proving the claim. $\qquad \square$

Since $|T| = \Omega(n)$, and $d(i, T) \geq d_f(i, T)$, T satisfies the conditions of Lemma 3.3. This gives us an $O(1)$ average-distortion embedding of the points into ℓ_1. $\qquad \square$

We can now infer the proof of Theorem 1.3 by using the results above.

Proof of Theorem 1.3. The conditions covered in Lemmas 3.8, 3.9, 3.10 and 3.11 on the set of points $\{x_i\}_{i \in V}$ are exhaustive, and in each case yield an embedding with $O(\sqrt{r})$ average distortion. It is clear that each of these conditions can be easily checked, and the corresponding embeddings can be constructed efficiently. $\qquad \square$

Remark. *The Hamming Cube on N points, residing in $\log N$ dimensions, and having η-subspace rank $\Omega_\eta(\log N)$ by symmetry, has two $\Omega(N)$-sized sets that are $\Omega(1/\sqrt{\log N})$ apart, and shows that the above analysis is tight up to constants.*

3.4 Application to Sparsest Cut

The proof of Corollary 1.4 now follows easily, using the main result.

Proof of Corollary 1.4. Suppose $\lambda_r/n \geq \Phi_{SDP}/(1 - \epsilon)$. We invoke the following result of Guruswami and Sinop [10] (stated here for the special case of *Uniform* Sparsest Cut):

Proposition 3.13 (Von-Neumann inequality [10, Theorem 3.3])**.** *Let $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$ be the singular values of the matrix M with columns $\{(x_i - x_j)\}_{i<j}$. Then*

$$\frac{\sum_{t \geq r} \sigma_j^2}{\sum_{t=1}^{n} \sigma_j^2} \leq \frac{\Phi_{SDP}}{\lambda_r(G)/n}.$$

For every $l \leq n$, we know that $\sum_{i=1}^{l} \sigma_i^2 = \sum_{i<j} \|\Pi_l(x_i - x_j)\|_2^2$, where Π_l is the subspace defined by the the top l left singular vectors of M. This immediately gives us that $\mathrm{ssr}_\epsilon(X) = r-1$. Applying the main theorem gives us an $O(\sqrt{r})$ average distortion embedding into ℓ_1, and hence an $O_\epsilon(\sqrt{r})$ approximation to $\Phi(G)$ in this setting. □

Remark. *Under the same precondition, Guruswami and Sinop [10] give an $O(1/\epsilon)$ approximation, but by solving a SDP of size $n^{O(r)}$, using a partial solver that runs in time $2^{O(r)}\mathrm{poly}(n)$ [9]. They need to know r first, and set up the SDP and solver appropriately. The works [6, 5] give a $O(r/\epsilon^2)$ and $O(r/\epsilon)$ approximation respectively, using just the Goemans-Linial SDP; the rounding algorithms do not depend on r. Our algorithm too is independent of r, and we get a better guarantee of $O(\sqrt{r}/\epsilon)$ in this setting.*

Though the precondition of the corollary may seem involved, it can easily be related back to a simpler one, as the following corollary shows (proof in Appendix A.2).

Corollary 3.14. *If G is regular with $\lambda_r(G) \geq \epsilon$, then we can find a $O(\sqrt{r} + 1/\sqrt{\epsilon})$ approximation to the sparsest cut in G in $\mathrm{poly}(n)$ time.*

Remark. *It is clear that we get a $O(\sqrt{r})$ approximation for all graphs whose ℓ_2^2 representation always has subspace rank r. Graphs of low threshold-rank are one class of graphs that have this property.*

Acknowledgements

References

[1] Noga Alon and Vitali D Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *Journal of Combinatorial Theory, Series B*, 38(1):73–88, 1985.

[2] Sanjeev Arora, James R. Lee, and Assaf Naor. Euclidean distortion and the sparsest cut. *J. Amer. Math. Soc.*, 21:1–21, 2008. (Preliminary version in *37th STOC*, 2008). `doi:10.1090/S0894-0347-07-00573-5`.

[3] Sanjeev Arora, Satish Rao, and Umesh V. Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56(2), 2009. (Preliminary version in *36th STOC*, 2004). `doi:10.1145/1502793.1502794`.

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

[5] Amit Deshpande, Prahladh Harsha, and Rakesh Venkat. Embedding Approximately Low-Dimensional ℓ_2^2 Metrics into ℓ_1. In *36th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS), 2016*, LIPIcs, pages 10:1–10:13.

[6] Amit Deshpande and Rakesh Venkat. Guruswami-Sinop rounding without higher level Lasserre. In *Proc. 17th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*, volume 28 of *LIPIcs*, pages 105–114. Schloss Dagstuhl, 2014. `arXiv: 1406.7279, doi:10.4230/LIPIcs.APPROX-RANDOM.2014.105`.

[7] Shayan Oveis Gharan and Luca Trevisan. Improved ARV rounding in small-set expanders and graphs of bounded threshold rank. 2013. `arXiv:1304.2060`.

[8] Michel X. Goemans. Semidefinite programming in combinatorial optimization. *Mathematical Programming*, 79(1):143–161, 1997.

[9] Venkatesan Guruswami and Ali Kemal Sinop. Faster SDP Hierarchy Solvers for Local Rounding Algorithms. In *Proc. 53rd IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 197–206, 2012.

[10] Venkatesan Guruswami and Ali Kemal Sinop. Approximating non-uniform sparsest cut via generalized spectra. In *Proc. 24th Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 295–305, 2013. `arXiv:1112.4109, doi:10.1137/1.9781611973105.22`.

[11] Piotr Indyk and Jirí Matoušek. Low-distortion embeddings of finite metric spaces. In Jacob E. Goodman and Joseph O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, pages 177–196. Chapman and Hall/CRC, 2nd edition, 2004. `doi:10.1201/9781420035315.ch8`.

[12] William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference on Mondern Analysis and Probability*, volume 26 of *Contemporary Mathematics*, pages 189–206. Amer. Math. Soc., 1982. `doi:10.1090/conm/026`.

[13] Daniel M. Kane and Raghu Meka. A PRG for Lipschitz functions of polynomials with applications to sparsest cut. In *Proc. 45th ACM Symp. on Theory of Computing (STOC)*, pages 1–10, 2013. `arXiv:1211/1109, doi:10.1145/2488608.2488610`.

[14] Tsz Chiu Kwok, Lap Chi Lau, Yin Tat Lee, Shayan Oveis Gharan, and Luca Trevisan. Improved Cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap. In *Proc. 45th ACM Symp. on Theory of Computing (STOC)*, pages 11–20, 2013. `arXiv:1301. 5584, doi:10.1145/2488608.2488611`.

[15] Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[16] James R. Lee. On distance scales, embeddings, and efficient relaxations of the cut cone. In *Proc. of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, SODA, pages 92–101, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics. URL: `http://dl. acm.org/citation.cfm?id=1070432.1070446`.

[17] Nathan Linial. Finite metric spaces: combinatorics, geometry and algorithms. In *Proc. of the ICM, Beijing*, volume 3, pages 573–586, 2002. `arXiv:math/0304466`.

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

[19] Avner Magen and Mohammad Moharrami. On the nonexistence of dimension reduction for ℓ_2^2 metrics. In *Proc. 20th Annual Canadian Conf. on Comp. Geom.*, 2008. URL: `http://cccg.ca/ proceedings/2008/paper37full.pdf`.

[20] Jirí Matoušek. Embedding finite metric spaces into normed spaces. In *Lectures on Discrete Geometry*, Graduate Texts in Mathematics, chapter 5, pages 355–400. Springer, 2002. `doi: 10.1007/978-1-4613-0039-7_15`.

[21] Assaf Naor, Yuval Rabani, and Alistair Sinclair. Quasisymmetric embeddings, the observable diameter, and expansion properties of graphs. *Journal of Functional Analysis*, 227(2):273–303, 2005.

[22] Assaf Naor and Robert Young. The integrality gap of the Goemans–Linial SDP relaxation for Sparsest Cut is at least a constant multiple of $\Omega(\sqrt{\log n})$. To appear in *Proc.* 49*th ACM Symp. on Theory of Computing (STOC) 2017*. arXiv preprint: `http://arxiv.org/abs/1704.01200`.

[23] Luca Trevisan. Lecture notes on expansion, sparsest cut, and spectral graph theory, 2011. Available online. URL: `http://www.eecs.berkeley.edu/~luca/books/expanders.pdf`.

A   Appendix

A.1   Ruling out a worst-case distortion bound of $O(\sqrt{\mathrm{ssr}_\eta(X)})$.

We give a simple example of why one cannot hope to prove a worst-case distortion bound like Goemans' result, using the notion of subspace rank. Suppose that a certain point set X satisfies the ℓ_2^2 inequalities, and has *worst-case* distortion $\Omega(D)$ for embedding into ℓ_1. It is known that there exists such an X with $D = \Omega(\sqrt{\log n})$ [22]. Without loss of generality, let X be scaled to satisfy $\sum_{i,j} \|x_i - x_j\|_2^2 = n^2$, and $\|x_1 - x_2\|_2^2 = \max_{i,j} \|x_i - x_j\|_2^2$. Consider the set Y which has X, along with $C - 1$ additional copies of x_1 and x_2[4]. Clearly, Y satisfies the ℓ_2^2 triangle inequalities. Further, Y has η-subspace rank of 1 for a large enough C: the sum of all squared distances is at most $C + (C^2 - C) \|x_1 - x_2\|_2^2$, and the sum of squared distances along the direction $x_1 - x_2$ is at least $C^2 \|x_1 - x_2\|_2^2$. However, embedding Y with *worst*-case distortion $O(1)$ into ℓ_1 would contradict the lower bound on embedding X into ℓ_1.

A.2   Proof of Corollary 3.14

Proof (Of Corollary 3.14). The proof follows by using a combination of two algorithms, depending on how λ_r compares to $\Phi_{SDP}(G)$. Suppose that G is 1-regular by scaling the edge weights, without loss of generality, and let $X = \{x_1, \ldots, x_n\}$ be the optimal SDP solution. If $\Phi_{SDP} \geq \epsilon/100n$, then there is one co-ordinate of the SDP solution with objective value at least $\epsilon/100n$. In this case, running the Cheeger rounding algorithm [1, Lemma 2.1] (see also [23, Section 2.4] for an exposition) on this co-ordinate would output a cut of sparsity $O(\sqrt{\epsilon}/n) \leq O\left(\Phi_{SDP}(G)/\sqrt{\epsilon}\right)$.

If $\Phi_{SDP} \leq \epsilon/100n$ then we have $\lambda_r/n \geq 100\Phi_{SDP}$. Applying Corollary 1.4 with $\epsilon = 99/100$ gives us an $O(\sqrt{r})$ average-distortion embedding into ℓ_1, and hence an $O(\sqrt{r})$ approximation to $\Phi(G)$ in this setting. Thus, the best of the two cuts will be a $O(\sqrt{r} + 1/\sqrt{\epsilon})$ approximation to $\Phi(G)$. $\square$

---

[4]Technically, we are dealing with semi-metrics, and hence distinct points may overlap.