Rounding the Sparsest-Cut SDP on Low Threshold-Rank Graphs

Rakesh Venkat

Tata Institute of Fundamental Research (TIFR), Mumbai

(Joint work with Amit Deshpande, MSR and Prahladh Harsha, TIFR)

SDP and Matrix Methods, NUS 2016

Outline

1 Introduction

2 Cheeger or Spectral Approach

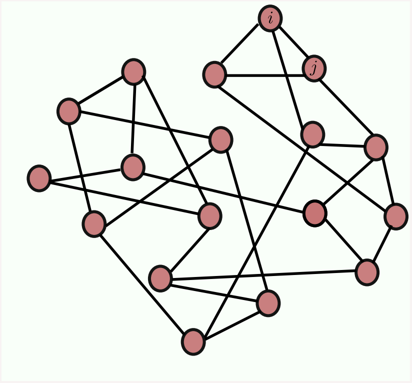
3 Rounding a stronger SDP
- Our Algorithm

4 Goemans' Theorem

5 Summary

The Sparsest Cut Problem

d-regular graph G



- Sparsity of a cut $(S, \bar{S})$ is

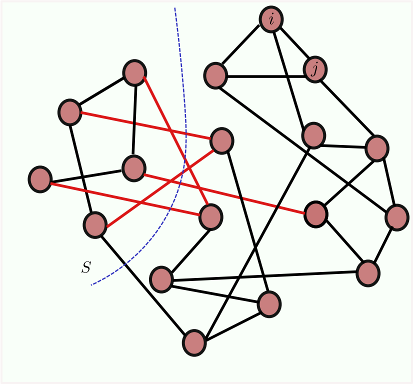
$$\Phi(S) = \frac{|E(S,\bar{S})|}{|S||\bar{S}|}$$

- Sparsest Cut in G is

$$\Phi(G) = \min_{S \subseteq V} \frac{|E(S,\bar{S})|}{|S||\bar{S}|}$$

The Sparsest Cut Problem

d-regular graph G



- Sparsity of a cut $(S, \bar{S})$ is

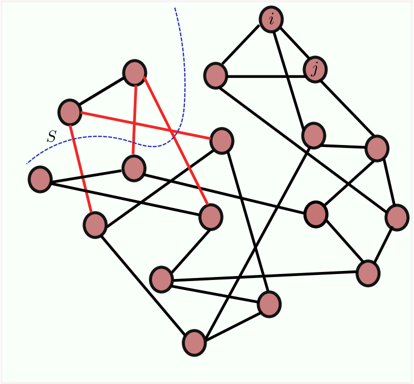
$$\Phi(S) = \frac{|E(S, \bar{S})|}{|S||\bar{S}|}$$

- Sparsest Cut in G is

$$\Phi(G) = \min_{S \subseteq V} \frac{|E(S, \bar{S})|}{|S||\bar{S}|}$$

The Sparsest Cut Problem

d-regular graph G



- Sparsity of a cut $(S, \bar{S})$ is

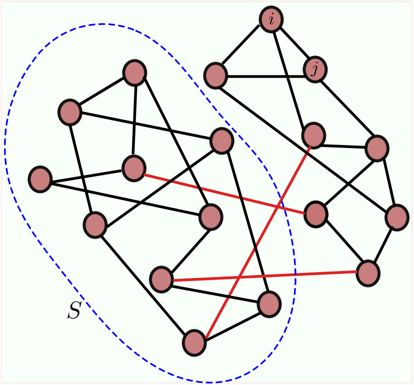
$$\Phi(S) = \frac{|E(S, \bar{S})|}{|S||\bar{S}|}$$

- Sparsest Cut in G is

$$\Phi(G) = \min_{S \subseteq V} \frac{|E(S, \bar{S})|}{|S||\bar{S}|}$$

The Sparsest Cut Problem

d-regular graph G



- Sparsity of a cut $(S, \bar{S})$ is

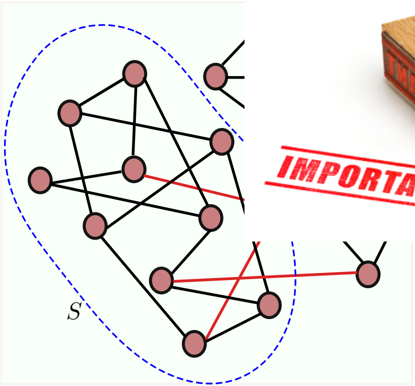
$$\Phi(S) = \frac{|E(S, \bar{S})|}{|S||\bar{S}|}$$

- Sparsest Cut in G is

$$\boxed{\Phi(G) = \min_{S \subseteq V} \frac{|E(S, \bar{S})|}{|S||\bar{S}|}}$$

The Sparsest Cut Problem

d-regular graph G



of a cut $(S, \bar{S})$ is

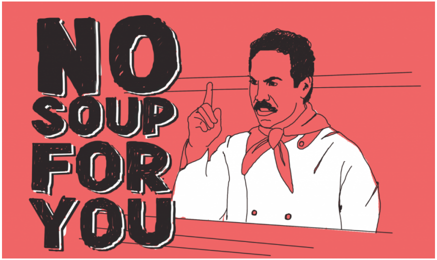
$$S) = \frac{|E(S, \bar{S})|}{|S||\bar{S}|}$$

Cut in G is

$$\Phi(G) = \min_{S \subseteq V} \frac{|E(S, \bar{S})|}{|S||\bar{S}|}$$

Complexity of Sparsest Cut: Lower bounds

- NP-hard to compute exactly
- Assuming the *Unique Games Conjecture*, it is NP-hard to *approximate* to any constant factor [CKKRS '05, Khot-Vishnoi '05]



Sparsest Cut Objective

$$\Phi_{OPT} = \min_S \frac{|E(S,\bar{S})|}{|S||\bar{S}|}$$

- Original Objective

$$\Phi_{OPT} = \min_{x_i \in \{0,1\}} \frac{\sum_{ij \in E}(x_i - x_j)^2}{\sum_{ij \in V \times V}(x_i - x_j)^2}$$

- Where $x_i = 1$ if $i \in S$, $x_i = 0$ if $i \in \bar{S}$

Sparsest Cut Objective

$$\Phi_{OPT} = \min_{S} \frac{|E(S,\bar{S})|}{|S||\bar{S}|}$$

- Original Objective

$$\Phi_{OPT} = \min_{x_i \in \{0,1\}} \frac{\sum_{ij \in E}(x_i - x_j)^2}{\sum_{ij \in V \times V}(x_i - x_j)^2}$$

- Where $x_i = 1$ if $i \in S$, $x_i = 0$ if $i \in \bar{S}$

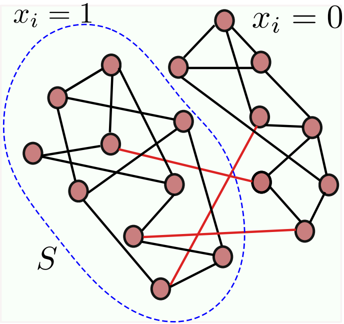
Sparsest Cut Objective

$$\Phi_{OPT} = \min_S \frac{|E(S, \bar{S})|}{|S||\bar{S}|}$$

- Original Objective

$$\Phi_{OPT} = \min_{x_i \in \{0,1\}} \frac{\sum_{ij \in E}(x_i - x_j)^2}{\sum_{ij \in V \times V}(x_i - x_j)^2}$$

- Where $x_i = 1$ if $i \in S$, $x_i = 0$ if $i \in \bar{S}$



Relaxing the Objective

- Assign a vector $x_i \in \mathbb{R}^m$ for each 0/1 variable
- Ideally, the vectors should be just scalars 0 or 1, i.e. one-dimensional
- Sparsest Cut objective relaxation (a SDP):

$$\Phi_1 = \min_{x_i \in \mathbb{R}^m} \quad \frac{\sum_{ij \in E} \|x_i - x_j\|_2^2}{\sum_{ij \in V \times V} \|x_i - x_j\|_2^2}$$

- $$\boxed{\Phi_1 \le \Phi_{OPT}}$$
- Can add in more constraints on vectors that 0/1 variables satisfy, e.g.

$$x_i^{\text{old}} \in [0,1] \text{ becomes } \|x_i\|^2 \le 1$$

Relaxing the Objective

- Assign a vector $x_i \in \mathbb{R}^m$ for each 0/1 variable
- Ideally, the vectors should be just scalars 0 or 1, i.e. one-dimensional
- Sparsest Cut objective relaxation (a SDP):

$$\Phi_1 = \min_{x_i \in \mathbb{R}^m} \quad \frac{\sum_{ij \in E} \|x_i - x_j\|_2^2}{\sum_{ij \in V \times V} \|x_i - x_j\|_2^2}$$

- $$\Phi_1 \leq \Phi_{OPT}$$

- Can add in more constraints on vectors that 0/1 variables satisfy, e.g.

$$x_i^{\text{old}} \in [0, 1] \text{ becomes } \|x_i\|^2 \leq 1$$

Relaxing the Objective

- Assign a vector $x_i \in \mathbb{R}^m$ for each 0/1 variable
- Ideally, the vectors should be just scalars 0 or 1, i.e. one-dimensional
- Sparsest Cut objective relaxation (a SDP):

$$\Phi_1 = \min_{x_i \in \mathbb{R}^m} \quad \frac{\sum_{ij \in E} \|x_i - x_j\|_2^2}{\sum_{ij \in V \times V} \|x_i - x_j\|_2^2}$$

- $$\boxed{\Phi_1 \leq \Phi_{OPT}}$$

- Can add in more constraints on vectors that 0/1 variables satisfy, e.g.

$$x_i^{\text{old}} \in [0,1] \text{ becomes } \|x_i\|^2 \leq 1$$

Relaxing the Objective

- Assign a vector $x_i \in \mathbb{R}^m$ for each 0/1 variable
- Ideally, the vectors should be just scalars 0 or 1, i.e. one-dimensional
- Sparsest Cut objective relaxation (a SDP):

$$\Phi_1 = \min_{x_i \in \mathbb{R}^m} \quad \frac{\sum_{ij \in E} \|x_i - x_j\|_2^2}{\sum_{ij \in V \times V} \|x_i - x_j\|_2^2}$$

- $$\Phi_1 \leq \Phi_{OPT}$$
- Can add in more constraints on vectors that 0/1 variables satisfy, e.g.

$$x_i^{\text{old}} \in [0, 1] \text{ becomes } \|x_i\|^2 \leq 1$$

Rounding the Relaxation

- Extract out a 1-dimensional 0/1 solution $\{y_i\}$ from the SDP solution $\{x_i\}$ in m dimensions
- We will lose in objective value since $\Phi_{\mathrm{ALG}} = \Phi(\{y_1, \ldots, y_n\}) \geq \Phi(\{x_1, \ldots, x_n\}) = \Phi_1$
- For Sparsest Cut, suffices to get an *embedding* into ℓ_1, rather than pure $\{0, 1\}$ solutions for y_i's.

Rounding the Relaxation

- Extract out a 1-dimensional 0/1 solution $\{y_i\}$ from the SDP solution $\{x_i\}$ in m dimensions
- We will lose in objective value since $\Phi_{\mathrm{ALG}} = \Phi(\{y_1, \ldots, y_n\}) \geq \Phi(\{x_1, \ldots, x_n\}) = \Phi_1$
- For Sparsest Cut, suffices to get an *embedding* into ℓ_1, rather than pure $\{0, 1\}$ solutions for y_i's.

Rounding the Relaxation

- Extract out a 1-dimensional 0/1 solution $\{y_i\}$ from the SDP solution $\{x_i\}$ in m dimensions
- We will lose in objective value since $\Phi_{\mathsf{ALG}} = \Phi(\{y_1, \ldots, y_n\}) \geq \Phi(\{x_1, \ldots, x_n\}) = \Phi_1$
- For Sparsest Cut, suffices to get an *embedding* into ℓ_1, rather than pure $\{0, 1\}$ solutions for y_i's.

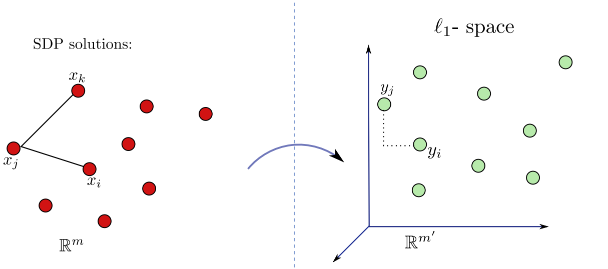
Connection to ℓ_1 embeddings [LLR '94, AR '94]

- Given a mapping of the points: $Y : V \to \mathbb{R}^{m'}$, we can produce a cut T of cost:

$$\Phi(T) \leq \frac{\sum_{ij \in V} \|y_i - y_j\|_1}{\sum_{kl \in V \times V} \|y_k - y_l\|_1}$$

- Sufficient to produce an *embedding* of the SDP solutions into ℓ_1-space

ℓ_1 embeddings of SDP solutions



$$OBJ = \frac{\sum_e \|x_i - x_j\|_2^2}{\sum_{i,j \in V} \|x_i - x_j\|_2^2} \approx OBJ(\ell_1) = \frac{\sum_e \|y_i - y_j\|_1}{\sum_{i,j \in V} \|y_i - y_j\|_1}$$

- Will compare $\|y_i - y_j\|_1$ to $\|x_i - x_j\|^2$

Average distortion and guarantee

If

- (Contraction):

$$\|y_i - y_j\|_1 \leq \|x_i - x_j\|^2, \text{ for every } i, j$$

- (Average Dilation)

$$\sum_{ij} \|y_i - y_j\|_1 \geq \frac{1}{D} \cdot \sum_{ij} \|x_i - x_j\|^2$$

then we have a $O(D)$ - approximation.

Average distortion and guarantee

If

- (Contraction):

$$\|y_i - y_j\|_1 \leq \|x_i - x_j\|^2, \text{ for every } i, j$$

- (Average Dilation)

$$\sum_{ij} \|y_i - y_j\|_1 \geq \frac{1}{D} \cdot \sum_{ij} \|x_i - x_j\|^2$$

then we have a $O(D)$ - approximation.

Average distortion and guarantee

If

- (Contraction):

$$\|y_i - y_j\|_1 \leq \|x_i - x_j\|^2, \text{ for every } i, j$$

- (Average Dilation)

$$\sum_{ij} \|y_i - y_j\|_1 \geq \frac{1}{D} \cdot \sum_{ij} \|x_i - x_j\|^2$$

then we have a $O(D)$ - approximation.

Outline

1 Introduction

2 Cheeger or Spectral Approach

3 Rounding a stronger SDP
- Our Algorithm

4 Goemans' Theorem

5 Summary

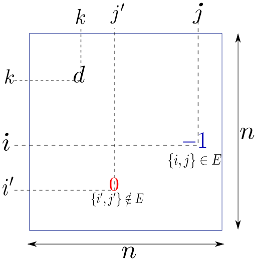
Cheeger Rounding (Alon-Milman '85)

$$\Phi_1 = \min_{x_i \in \mathbb{R}^m} \quad \frac{\sum_{ij \in E} \|x_i - x_j\|^2}{\sum_{ij \in V \times V} \|x_i - x_j\|^2}$$

- Given a SDP solution $\{x_i\}_i$ with objective value $\Phi_1 = \dfrac{\epsilon d}{n}$, can get a rounded solution (a cut) with value $O(\dfrac{\sqrt{\epsilon} d}{n})$.
- The ℓ_1 mapping is a simple one-dimensional embedding: find a specific co-ordinate t and set $y_i = x_i[t]$
- Works well for *expander* graphs

Spectral Graph Theory- Preliminaries

- Laplacian of a graph: $L = d\mathbb{I} - A$



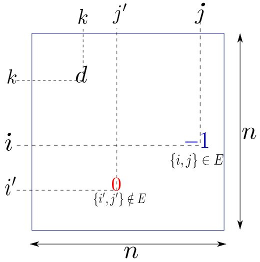
- Eigenvalues of the Laplacian

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \leq 2d$$

- Eigenvalues predict connectivity properties
  - $\lambda_2 = 0 \iff G$ is disconnected
  - $\lambda_n = 2d \iff G$ is bipartite

Spectral Graph Theory- Preliminaries

- Laplacian of a graph: $L = d\mathbb{I} - A$



- Eigenvalues of the Laplacian

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \leq 2d$$

- Eigenvalues predict connectivity properties
  - $\lambda_2 = 0 \iff G$ is disconnected
  - $\lambda_n = 2d \iff G$ is bipartite

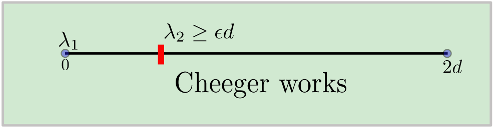
Cheeger Rounding Guarantee

- Can show the following:
  - $\Phi_1 = \dfrac{\lambda_2}{n}$
  - x_i's are in fact, one-dimensional. Furthermore, $x_i = u_2(i)$, where

$$Lu_2 = \lambda_2 u_2$$

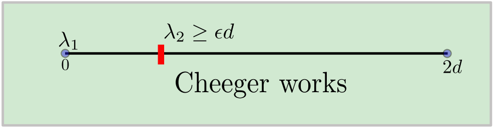
.

- $\Phi_{\text{ALG}} \leq O(\sqrt{\dfrac{d}{\lambda_2}})\ \Phi_{\text{OPT}}$
- This works when $\lambda_2 \geq \epsilon d$ (an *expander*)



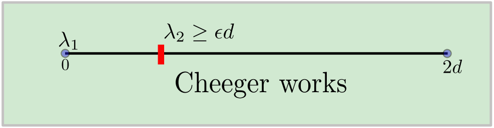
Cheeger Rounding Guarantee

- Can show the following:
  - $\Phi_1 = \dfrac{\lambda_2}{n}$
  - x_i's are in fact, one-dimensional. Furthermore, $x_i = u_2(i)$, where $$Lu_2 = \lambda_2 u_2$$.
- $\Phi_{\mathrm{ALG}} \leq O(\sqrt{\dfrac{d}{\lambda_2}})\ \Phi_{\mathrm{OPT}}$
- This works when $\lambda_2 \geq \epsilon d$ (an *expander*)



Cheeger Rounding Guarantee

- Can show the following:
  - $\Phi_1 = \dfrac{\lambda_2}{n}$
  - x_i's are in fact, one-dimensional. Furthermore, $x_i = u_2(i)$, where
$$Lu_2 = \lambda_2 u_2$$
.
- $\Phi_{\text{ALG}} \leq O(\sqrt{\dfrac{d}{\lambda_2}})\ \Phi_{\text{OPT}}$
- This works when $\lambda_2 \geq \epsilon d$ (an *expander*)

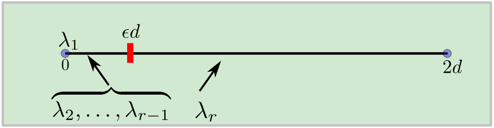


Improved Cheeger [KKLGT '13]

- An improved analysis by Kwok et al. gives the guarantee:

$$\Phi_{\mathsf{ALG}} \leq O(r)\sqrt{\frac{d}{\lambda_r}}\Phi_{\mathsf{OPT}}$$

- This is a $O(r)$ guarantee on graphs where $\lambda_r \geq \epsilon d$.
- Such graphs are said to have threshold rank r



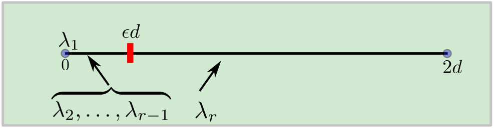
- Requires significantly more work than the original Cheeger analysis
- Dependence on r is tight

Improved Cheeger [KKLGT '13]

- An improved analysis by Kwok et al. gives the guarantee:

$$\Phi_{\mathsf{ALG}} \leq O(r)\sqrt{\frac{d}{\lambda_r}}\Phi_{\mathsf{OPT}}$$

- This is a $O(r)$ guarantee on graphs where $\lambda_r \geq \epsilon d$.
- Such graphs are said to have threshold rank r



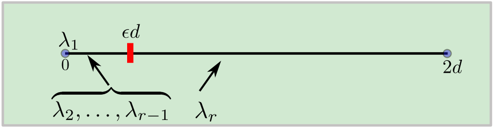
- Requires significantly more work than the original Cheeger analysis
- Dependence on r is tight

Improved Cheeger [KKLGT '13]

- An improved analysis by Kwok et al. gives the guarantee:

$$\Phi_{\text{ALG}} \leq O(r)\sqrt{\frac{d}{\lambda_r}}\Phi_{\text{OPT}}$$

- This is a $O(r)$ guarantee on graphs where $\lambda_r \geq \epsilon d$.
- Such graphs are said to have threshold rank r



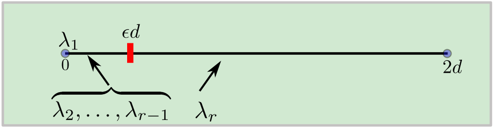
- Requires significantly more work than the original Cheeger analysis
- Dependence on r is tight

Improved Cheeger [KKLGT '13]

- An improved analysis by Kwok et al. gives the guarantee:

$$\Phi_{\mathrm{ALG}} \leq O(r)\sqrt{\frac{d}{\lambda_r}}\Phi_{\mathrm{OPT}}$$

- This is a $O(r)$ guarantee on graphs where $\lambda_r \geq \epsilon d$.
- Such graphs are said to have threshold rank r



- Requires significantly more work than the original Cheeger analysis
- Dependence on r is tight

Outline

1 Introduction

2 Cheeger or Spectral Approach

3 Rounding a stronger SDP
- Our Algorithm

4 Goemans' Theorem

5 Summary

A stronger SDP relaxation

$$\Phi_\Delta = \min_{x_i \in \mathbb{R}^m} \quad \frac{\sum_{ij \in E} \|x_i - x_j\|^2}{\sum_{ij \in V \times V} \|x_i - x_j\|^2}$$

$$\text{s.t.} \quad \|x_i - x_j\|^2 + \|x_j - x_k\|^2 \geq \|x_i - x_k\|^2 \quad \forall i, j, k \in [n]$$

$(\ell_2^2$ inequality constraints$)$

- Constraints are triangle inequalities on *squares* of distances
  - Satisfied by $0, 1$ integral solutions

A stronger SDP relaxation

$$\Phi_\Delta = \min_{x_i \in \mathbb{R}^m} \quad \frac{\sum_{ij \in E} \|x_i - x_j\|^2}{\sum_{ij \in V \times V} \|x_i - x_j\|^2}$$

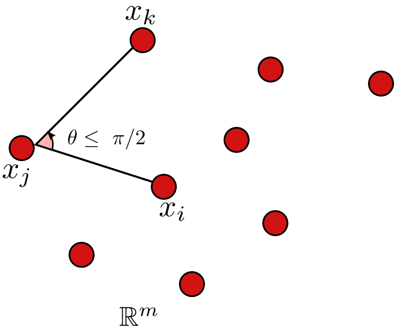
$$\text{s.t.} \quad \|x_i - x_j\|^2 + \|x_j - x_k\|^2 \geq \|x_i - x_k\|^2 \quad \forall i, j, k \in [n]$$

$$(\ell_2^2 \text{ inequality constraints})$$

- Here: A simple rounding algorithm for the above SDP with an $O(r)$ approximation when $\lambda_r \geq \epsilon d$.

ℓ_2^2 constraints

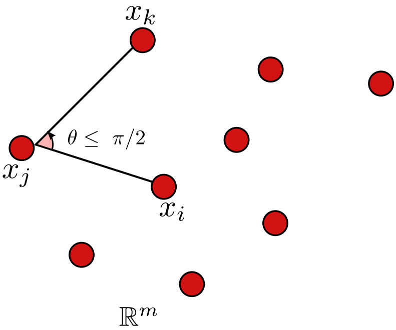
- $\|x_i - x_j\|^2 + \|x_j - x_k\|^2 \geq \|x_i - x_k\|^2$



- Equivalently: $\langle x_i - x_j, x_k - x_j \rangle \geq 0$
- One-dimensional solutions can't have three distinct points!

ℓ_2^2 constraints

- $\|x_i - x_j\|^2 + \|x_j - x_k\|^2 \geq \|x_i - x_k\|^2$



- Equivalently: $\langle x_i - x_j, x_k - x_j \rangle \geq 0$
- One-dimensional solutions can't have three distinct points!

Do ℓ_2^2 inequalities help?

- Best known unconditional guarantee for Sparsest Cut by Arora-Rao-Vazirani (ARV) rounds the above SDP to give

$$\boxed{\Phi_{\mathsf{ARV}} \leq O(\sqrt{\log n})\ \Phi_\Delta}$$

- Can we leverage them to do better on low threshold-rank graphs?
- Note: Can assume that $\Phi_\Delta \leq \dfrac{\epsilon d}{100n} \leq \dfrac{\lambda_r}{100n}$
  - Else, use Cheeger rounding to get a cut of sparsity $O(\frac{\sqrt{\epsilon} d}{n}) \leq \frac{1}{\sqrt{\epsilon}} \Phi_\Delta$

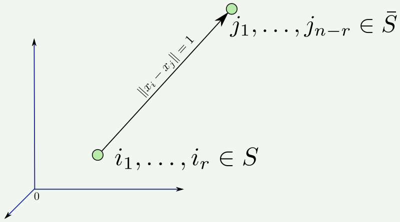
Do ℓ_2^2 inequalities help?

- Best known unconditional guarantee for Sparsest Cut by Arora-Rao-Vazirani (ARV) rounds the above SDP to give

$$\boxed{\Phi_{\mathrm{ARV}} \leq O(\sqrt{\log n})\ \Phi_\Delta}$$

- Can we leverage them to do better on low threshold-rank graphs?
- Note: Can assume that $\Phi_\Delta \leq \frac{\epsilon d}{100n} \leq \frac{\lambda_r}{100n}$
  - Else, use Cheeger rounding to get a cut of sparsity $O(\frac{\sqrt{\epsilon} d}{n}) \leq \frac{1}{\sqrt{\epsilon}} \Phi_\Delta$

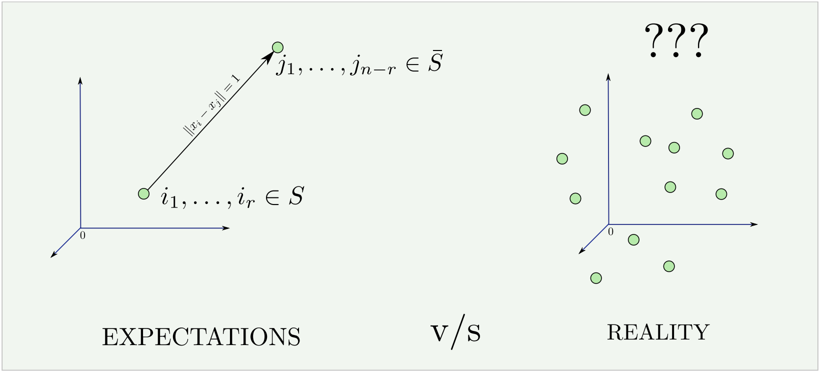
SDP Solutions



EXPECTATIONS

SDP Solutions

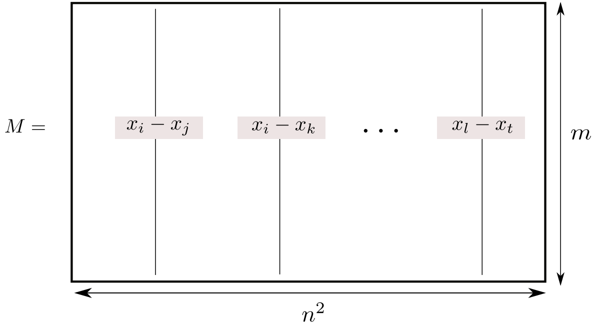
Can't beat NP-Hardness:



Life 101...

There is structure, though

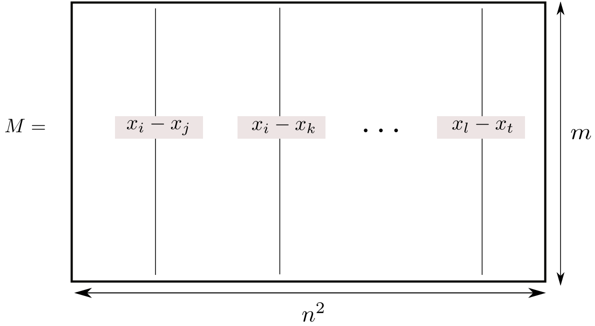
- The *Difference Matrix M* has some structure



- Let $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m$ be the singular values

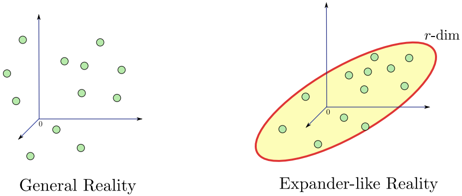
There is structure, though

- The *Difference Matrix* M has some structure



- Let $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m$ be the singular values

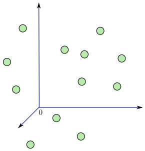
Tool 1 : Approximately Low-Dimensional solutions [GS'13]



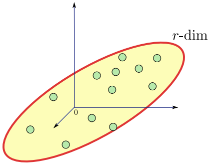
$$\boxed{\frac{\lambda_r}{n} \geq 100\Phi_\Delta \implies \sum_{i=1}^{r} \sigma_i^2 \geq 0.99 \sum_{i=1}^{n} \sigma_i^2}$$

- Constant fraction of the squared mass of the vectors $\{x_i - x_j\}_{ij}$ lies in a r-dimensional subspace
- Shift vectors x_i to have centroid as origin, above works with x_i
- *Stable Rank*: $\mathrm{sr}(M) \triangleq \frac{\|M\|_F^2}{\sigma_1(M)^2} \leq r/0.99$

Tool 1 : Approximately Low-Dimensional solutions [GS'13]



General Reality

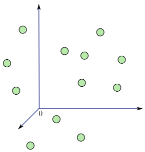


Expander-like Reality

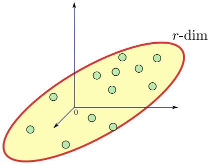
$$\frac{\lambda_r}{n} \geq 100\Phi_\Delta \implies \sum_{i=1}^{r} \sigma_i^2 \geq 0.99 \sum_{i=1}^{n} \sigma_i^2$$

- Constant fraction of the squared mass of the vectors $\{x_i - x_j\}_{ij}$ lies in a r-dimensional subspace
- Shift vectors x_i to have centroid as origin, above works with x_i
- *Stable Rank*: $\mathrm{sr}(M) \triangleq \dfrac{\|M\|_F^2}{\sigma_1(M)^2} \leq r/0.99$

Tool 1 : Approximately Low-Dimensional solutions [GS'13]



General Reality

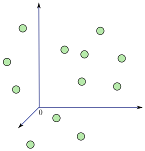


Expander-like Reality

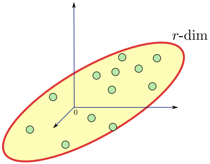
$$\boxed{\frac{\lambda_r}{n} \geq 100\Phi_\Delta \implies \sum_{i=1}^{r} \sigma_i^2 \geq 0.99 \sum_{i=1}^{n} \sigma_i^2}$$

- Constant fraction of the squared mass of the vectors $\{x_i - x_j\}_{ij}$ lies in a r-dimensional subspace
- Shift vectors x_i to have centroid as origin, above works with x_i
- *Stable Rank*: $\mathrm{sr}(M) \triangleq \frac{\|M\|_F^2}{\sigma_1(M)^2} \leq r/0.99$

Tool 1 : Approximately Low-Dimensional solutions [GS'13]



General Reality

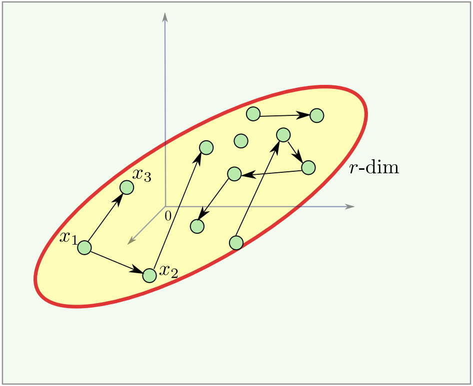


Expander-like Reality

$$\frac{\lambda_r}{n} \geq 100\Phi_\Delta \implies \sum_{i=1}^{r} \sigma_i^2 \geq 0.99 \sum_{i=1}^{n} \sigma_i^2$$

- Constant fraction of the squared mass of the vectors $\{x_i - x_j\}_{ij}$ lies in a r-dimensional subspace
- Shift vectors x_i to have centroid as origin, above works with x_i
- *Stable Rank*: $\mathrm{sr}(M) \triangleq \frac{\|M\|_F^2}{\sigma_1(M)^2} \leq r/0.99$

Tool 1 : Approximately Low-Dimensional solutions [GS'13]



Tool 2: Stronger Cauchy-Schwarz using ℓ_2^2 inequalities

Proposition

If x_i satisfy ℓ_2^2-inequalities, then $\forall i, j, k, l$, we have:

$$| \langle x_i - x_j, x_k - x_l \rangle | \leq \min \left\{ \|x_i - x_j\|^2, \|x_k - x_l\|^2 \right\}$$

Proof.

Left as easy exercise (see board). $\square$

- Note: Simple Cauchy Schwarz would give: $| \langle x_i - x_j, x_k - x_l \rangle | \leq \|x_i - x_j\| \|x_k - x_l\|$

Tool 2: Stronger Cauchy-Schwarz using ℓ_2^2 inequalities

Proposition

If x_i satisfy ℓ_2^2-inequalities, then $\forall i, j, k, l$, we have:

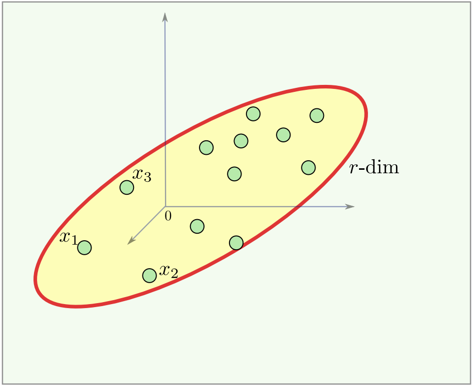
$$|\langle x_i - x_j, x_k - x_l \rangle| \leq \min\left\{ \|x_i - x_j\|^2, \|x_k - x_l\|^2 \right\}$$

Proof.

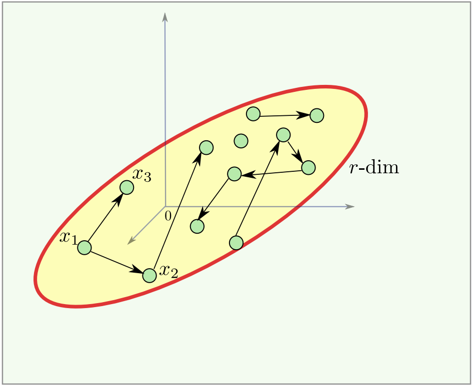
Left as easy exercise (see board). □

- Note: Simple Cauchy Schwarz would give: $|\langle x_i - x_j, x_k - x_l \rangle| \leq \|x_i - x_j\| \, \|x_k - x_l\|$

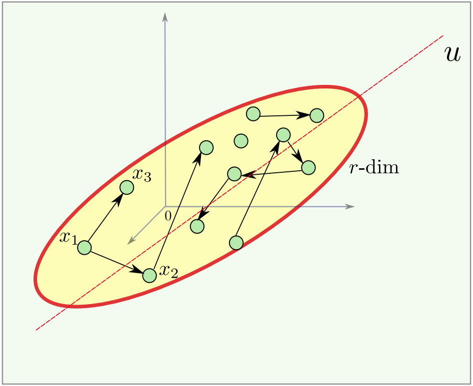
Algorithm



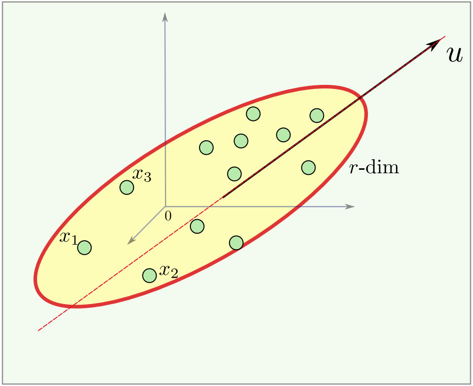
Algorithm



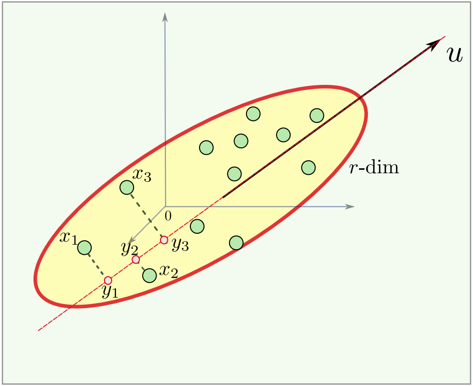
Algorithm



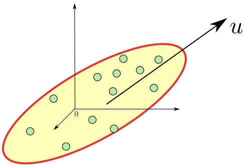
Algorithm



Algorithm



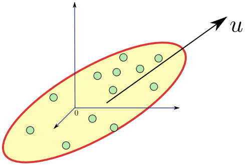
Algorithm



Expander-like Reality

1. Compute the top left-singular vector u, with singular value σ_1 of the matrix M
2. v is the top right-singular vector
3. 1-dimensional solutions are $y_i = \frac{\sigma_1}{\|v\|_1} \langle x_i, u \rangle$

Algorithm



Expander-like Reality

1. Compute the top left-singular vector u, with singular value σ_1 of the matrix M
2. v is the top right-singular vector
3. 1-dimensional solutions are $y_i = \frac{\sigma_1}{\|v\|_1} \langle x_i, u \rangle$

Analysis

$$\boxed{y_i = \frac{\sigma_1}{\|v\|_1} \langle x_i, u \rangle}$$

Will show:

- Contraction:

$$|y_i - y_j| \leq \|x_i - x_j\|^2 \quad \forall i, j$$

- (Average) Dilation:

$$\mathbb{E}_{ij}\left[|y_i - y_j|\right] \geq \frac{0.99}{r}\, \mathbb{E}_{ij}\left[\|x_i - x_j\|^2\right]$$

Analysis

$$\boxed{y_i = \frac{\sigma_1}{\|v\|_1} \langle x_i, u \rangle}$$

Will show:

- Contraction:

$$|y_i - y_j| \leq \|x_i - x_j\|^2 \quad \forall i, j$$

- (Average) Dilation:

$$\mathbb{E}_{ij}\left[|y_i - y_j|\right] \geq \frac{0.99}{r}\, \mathbb{E}_{ij}\left[\|x_i - x_j\|^2\right]$$

Analysis: Contraction

- $Mv = \sigma_1 u$, or equivalently, $\sigma_1 u = \sum_{kl} v_{kl}(x_k - x_l)$

Pick any (i,j). We have: $|y_i - y_j|_1 =$

$$
\begin{aligned}
\frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u\rangle| &= \frac{1}{\|v\|_1} \left| \left\langle x_i - x_j, \sum_{kl} v_{kl}(x_k - x_l) \right\rangle \right| \\
&\le \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, |\langle x_i - x_j, x_k - x_l\rangle| \\
&\le \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, \|x_i - x_j\|_2^2 \qquad \left[\text{By Stronger CS for } \ell_2^2\right] \\
&= \|x_i - x_j\|_2^2 .
\end{aligned}
$$

Analysis: Contraction

- $Mv = \sigma_1 u$, or equivalently, $\sigma_1 u = \sum_{kl} v_{kl}(x_k - x_l)$

Pick any (i,j). We have: $|y_i - y_j|_1 =$

$$\begin{aligned}
\frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u\rangle| &= \frac{1}{\|v\|_1} \left| \left\langle x_i - x_j, \sum_{kl} v_{kl}(x_k - x_l) \right\rangle \right| \\
&\le \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, |\langle x_i - x_j, x_k - x_l\rangle| \\
&\le \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, \|x_i - x_j\|_2^2 \quad \left[\text{By Stronger CS for } \ell_2^2\right] \\
&= \|x_i - x_j\|_2^2 .
\end{aligned}$$

Analysis: Contraction

- $Mv = \sigma_1 u$, or equivalently, $\sigma_1 u = \sum_{kl} v_{kl}(x_k - x_l)$

Pick any (i, j). We have: $|y_i - y_j|_1 =$

$$\frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u\rangle| = \frac{1}{\|v\|_1} \left| \left\langle x_i - x_j, \sum_{kl} v_{kl}(x_k - x_l) \right\rangle \right|$$

$$\leq \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, |\langle x_i - x_j, x_k - x_l\rangle|$$

$$\leq \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, \|x_i - x_j\|_2^2 \quad \left[\text{By Stronger CS for } \ell_2^2\right]$$

$$= \|x_i - x_j\|_2^2 .$$

Analysis: Contraction

- $Mv = \sigma_1 u$, or equivalently, $\sigma_1 u = \sum_{kl} v_{kl}(x_k - x_l)$

Pick any (i, j). We have: $|y_i - y_j|_1 =$

$$\frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u\rangle| = \frac{1}{\|v\|_1} \left| \left\langle x_i - x_j, \sum_{kl} v_{kl}(x_k - x_l) \right\rangle \right|$$

$$\leq \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, |\langle x_i - x_j, x_k - x_l\rangle|$$

$$\leq \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, \|x_i - x_j\|_2^2 \quad \left[\text{By Stronger CS for } \ell_2^2\right]$$

$$= \|x_i - x_j\|_2^2 .$$

Analysis: Contraction

- $Mv = \sigma_1 u$, or equivalently, $\sigma_1 u = \sum_{kl} v_{kl}(x_k - x_l)$

Pick any (i, j). We have: $|y_i - y_j|_1 =$

$$
\begin{aligned}
\frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u\rangle| &= \frac{1}{\|v\|_1} \left| \left\langle x_i - x_j, \sum_{kl} v_{kl}(x_k - x_l) \right\rangle \right| \\
&\leq \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, |\langle x_i - x_j, x_k - x_l\rangle| \\
&\leq \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, \|x_i - x_j\|_2^2 \quad \left[\text{By Stronger CS for } \ell_2^2\right] \\
&= \|x_i - x_j\|_2^2 .
\end{aligned}
$$

Analysis: Contraction

- $Mv = \sigma_1 u$, or equivalently, $\sigma_1 u = \sum_{kl} v_{kl}(x_k - x_l)$

Pick any (i,j). We have: $|y_i - y_j|_1 =$

$$
\begin{aligned}
\frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u\rangle| &= \frac{1}{\|v\|_1} \left| \left\langle x_i - x_j, \sum_{kl} v_{kl}(x_k - x_l) \right\rangle \right| \\
&\leq \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, |\langle x_i - x_j, x_k - x_l\rangle| \\
&\leq \frac{1}{\|v\|_1} \sum_{kl} |v_{kl}| \, \|x_i - x_j\|_2^2 \quad \left[\text{By Stronger CS for } \ell_2^2\right] \\
&= \|x_i - x_j\|_2^2 .
\end{aligned}
$$

Analysis: Dilation

$$
\begin{aligned}
\sum_{ij} |y_i - y_j|_1 &= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u\rangle| \\
&= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\sigma_1 v_{ij}| \quad [\text{Since } u^T M = \sigma_1 v^T] \\
&= \sigma_1^2 \\
&= \frac{\|M\|_F^2}{\mathrm{sr}(M)} \\
&= \frac{1}{\mathrm{sr}(M)} \sum_{ij} \|x_i - x_j\|_2^2 \\
&\geq \frac{0.99}{r} \sum_{ij} \|x_i - x_j\|_2^2 .
\end{aligned}
$$

Analysis: Dilation

$$
\begin{aligned}
\sum_{ij} |y_i - y_j|_1 &= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u \rangle| \\
&= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\sigma_1 v_{ij}| \quad \text{[Since } u^T M = \sigma_1 v^T \text{]} \\
&= \sigma_1^2 \\
&= \frac{\|M\|_F^2}{\mathrm{sr}(M)} \\
&= \frac{1}{\mathrm{sr}(M)} \sum_{ij} \|x_i - x_j\|_2^2 \\
&\geq \frac{0.99}{r} \sum_{ij} \|x_i - x_j\|_2^2 .
\end{aligned}
$$

Analysis: Dilation

$$\begin{aligned}
\sum_{ij} |y_i - y_j|_1 &= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u\rangle| \\
&= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\sigma_1 v_{ij}| \quad \text{[Since } u^T M = \sigma_1 v^T\text{]} \\
&= \sigma_1^2 \\
&= \frac{\|M\|_F^2}{\mathrm{sr}(M)} \\
&= \frac{1}{\mathrm{sr}(M)} \sum_{ij} \|x_i - x_j\|_2^2 \\
&\geq \frac{0.99}{r} \sum_{ij} \|x_i - x_j\|_2^2 .
\end{aligned}$$

Analysis: Dilation

$$\begin{aligned}
\sum_{ij} |y_i - y_j|_1 &= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u \rangle| \\
&= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\sigma_1 v_{ij}| \quad \text{[Since } u^T M = \sigma_1 v^T\text{]} \\
&= \sigma_1^2 \\
&= \frac{\|M\|_F^2}{\mathrm{sr}(M)} \\
&= \frac{1}{\mathrm{sr}(M)} \sum_{ij} \|x_i - x_j\|_2^2 \\
&\geq \frac{0.99}{r} \sum_{ij} \|x_i - x_j\|_2^2 .
\end{aligned}$$

Analysis: Dilation

$$
\begin{aligned}
\sum_{ij} |y_i - y_j|_1 &= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\langle x_i - x_j, u\rangle| \\
&= \sum_{ij} \frac{\sigma_1}{\|v\|_1} |\sigma_1 v_{ij}| \quad [\text{Since } u^T M = \sigma_1 v^T] \\
&= \sigma_1^2 \\
&= \frac{\|M\|_F^2}{\mathrm{sr}(M)} \\
&= \frac{1}{\mathrm{sr}(M)} \sum_{ij} \|x_i - x_j\|_2^2 \\
&\geq \frac{0.99}{r} \sum_{ij} \|x_i - x_j\|_2^2 .
\end{aligned}
$$

SoS hierarchies and comparison

- Algorithm by Guruswami and Sinop based on SoS hierarchy at level $O(r)$ gives a better result: $O(1)$ approximation
- GS algorithm runs in time $2^{O(r)}\texttt{poly}(n)$, but needs a specific solver

- Our algorithm does not need to know r a-priori
- Runs in time $\texttt{poly}(n)$, conceptually simpler

SoS hierarchies and comparison

- Algorithm by Guruswami and Sinop based on SoS hierarchy at level $O(r)$ gives a better result: $O(1)$ approximation
- GS algorithm runs in time $2^{O(r)}\texttt{poly}(n)$, but needs a specific solver

- Our algorithm does not need to know r a-priori
- Runs in time $\texttt{poly}(n)$, conceptually simpler
- Compared to improved Cheeger analysis, there is scope for improvement in dependence on r
- Gives a *projective* embedding, as against other known embeddings of ℓ_2^2 into ℓ_1 that are Frechet embeddings

SoS hierarchies and comparison

- Algorithm by Guruswami and Sinop based on SoS hierarchy at level $O(r)$ gives a better result: $O(1)$ approximation
- GS algorithm runs in time $2^{O(r)}\texttt{poly}(n)$, but needs a specific solver

- Our algorithm does not need to know r a-priori
- Runs in time $\texttt{poly}(n)$, conceptually simpler
- Compared to improved Cheeger analysis, there is scope for improvement in dependence on r
- Gives a *projective* embedding, as against other known embeddings of ℓ_2^2 into ℓ_1 that are Frechet embeddings

SoS hierarchies and comparison

- Algorithm by Guruswami and Sinop based on SoS hierarchy at level $O(r)$ gives a better result: $O(1)$ approximation
- GS algorithm runs in time $2^{O(r)}\mathtt{poly}(n)$, but needs a specific solver

- Our algorithm does not need to know r a-priori
- Runs in time $\mathtt{poly}(n)$, conceptually simpler
- Compared to improved Cheeger analysis, there is scope for improvement in dependence on r
- Gives a *projective* embedding, as against other known embeddings of ℓ_2^2 into ℓ_1 that are Frechet embeddings

Outline

1 Introduction

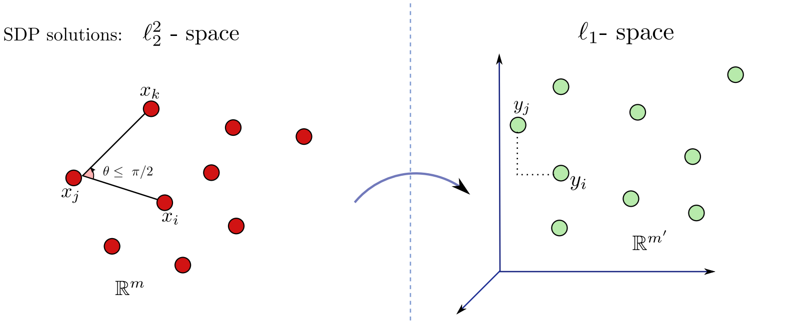
2 Cheeger or Spectral Approach

3 Rounding a stronger SDP
- Our Algorithm

4 Goemans' Theorem

5 Summary

Recap: Embedding



$$OBJ = \frac{\sum_e \|x_i - x_j\|_2^2}{\sum_{i,j \in V} \|x_i - x_j\|_2^2} \approx OBJ(\ell_1) = \frac{\sum_e \|y_i - y_j\|_1}{\sum_{i,j \in V} \|y_i - y_j\|_1}$$

Goemans' Result

- Goemans' (unpublished) result:

Theorem (Goemans '00)

A set of points in $\mathbb{R}^m$ satisfying ℓ_2^2 triangle inequalities can be embedded into ℓ_1 with distortion $O(\sqrt{m})$

- Implies a $\sqrt{m}$ approximation to SPARSEST CUT on instances where solution has dimension m
- Does dimension reduction work in ℓ_2^2?
  - No. Very strong lower bounds [Magen-Moharammi '00].
  - Caveat: Only in worst-case distortion

Goemans' Result

- Goemans' (unpublished) result:

Theorem (Goemans '00)

A set of points in $\mathbb{R}^m$ satisfying ℓ_2^2 triangle inequalities can be embedded into ℓ_1 with distortion $O(\sqrt{m})$

- Implies a $\sqrt{m}$ approximation to SPARSEST CUT on instances where solution has dimension m
- Does dimension reduction work in ℓ_2^2?
  - No. Very strong lower bounds [Magen-Moharammi '00]. ☹
  - Caveat: Only in worst-case distortion

Goemans' Result

- Goemans' (unpublished) result:

Theorem (Goemans '00)

A set of points in $\mathbb{R}^m$ satisfying ℓ_2^2 triangle inequalities can be embedded into ℓ_1 with distortion $O(\sqrt{m})$

- Implies a $\sqrt{m}$ approximation to SPARSEST CUT on instances where solution has dimension m
- Does dimension reduction work in ℓ_2^2?
  - No. Very strong lower bounds [Magen-Moharammi '00]. ☹
  - Caveat: Only in worst-case distortion

Goemans' Result

- Goemans' (unpublished) result:

Theorem (Goemans '00)

A set of points in $\mathbb{R}^m$ satisfying ℓ_2^2 triangle inequalities can be embedded into ℓ_1 with distortion $O(\sqrt{m})$

- Implies a $\sqrt{m}$ approximation to SPARSEST CUT on instances where solution has dimension m
- Does dimension reduction work in ℓ_2^2?
  - No. Very strong lower bounds [Magen-Moharammi '00]. ☹
  - Caveat: Only in worst-case distortion

Comparison to our result

- Our rounding technique gives an embedding for ℓ_2^2 points with low *stable rank*: $\|M\|_F^2 \,/\, \|M\|^2$
- Stable rank is a well-known robust proxy for the rank
  - ML, column subset selection..
- Should be able to improve our bound to $O(\sqrt{\text{Stable Rank}})$.
- Is dimension reduction possible in terms of *stable rank*?
  - Only average distortion required
- Btw, our result also recovers Goemans' theorem using a (arguably) cleaner proof

Comparison to our result

- Our rounding technique gives an embedding for ℓ_2^2 points with low *stable rank*: $\|M\|_F^2 \,/\, \|M\|^2$
- Stable rank is a well-known robust proxy for the rank
  - ML, column subset selection..
- Should be able to improve our bound to $O(\sqrt{\text{Stable Rank}})$.
- Is dimension reduction possible in terms of *stable rank*?
  - Only average distortion required
- Btw, our result also recovers Goemans' theorem using a (arguably) cleaner proof

Comparison to our result

- Our rounding technique gives an embedding for ℓ_2^2 points with low *stable rank*: $\|M\|_F^2 \,/\, \|M\|^2$
- Stable rank is a well-known robust proxy for the rank
  - ML, column subset selection..
- Should be able to improve our bound to $O(\sqrt{\text{Stable Rank}})$.
- Is dimension reduction possible in terms of *stable rank*?
  - Only average distortion required
- Btw, our result also recovers Goemans' theorem using a (arguably) cleaner proof

Comparison to our result

- Our rounding technique gives an embedding for ℓ_2^2 points with low *stable rank*: $\|M\|_F^2 \,/\, \|M\|^2$
- Stable rank is a well-known robust proxy for the rank
  - ML, column subset selection..
- Should be able to improve our bound to $O(\sqrt{\text{Stable Rank}})$.
- Is dimension reduction possible in terms of *stable rank*?
  - Only average distortion required
- Btw, our result also recovers Goemans' theorem using a (arguably) cleaner proof

Comparison to our result

- Our rounding technique gives an embedding for ℓ_2^2 points with low *stable rank*: $\|M\|_F^2 \,/\, \|M\|^2$
- Stable rank is a well-known robust proxy for the rank
  - ML, column subset selection..
- Should be able to improve our bound to $O(\sqrt{\text{Stable Rank}})$.
- Is dimension reduction possible in terms of *stable rank*?
  - Only average distortion required
- Btw, our result also recovers Goemans' theorem using a (arguably) cleaner proof

Dimension Reduction

- Standard Johnson-Lindenstrauss dimension reduction preserves ℓ_2^2 triangle inequalities *approximately* (in $O(\log n/\epsilon^2)$ dimensions)

$$\|z_i - z_j\|^2 + \|z_k - z_j\|^2 \geq (1 - O(\epsilon)) \|z_i - z_k\|^2$$

- Goemans' theorem is true with approximate ℓ_2^2 inequalities, but requires ARV analysis [Trevisan]
- Can we modify our algorithm to work with approximate triangle inequalities?
  - Or 'fix' ℓ_2^2 inequalities without blowing up approximate dimension

Dimension Reduction

- Standard Johnson-Lindenstrauss dimension reduction preserves ℓ_2^2 triangle inequalities *approximately* (in $O(\log n/\epsilon^2)$ dimensions)

$$\|z_i - z_j\|^2 + \|z_k - z_j\|^2 \geq (1 - O(\epsilon))\,\|z_i - z_k\|^2$$

- Goemans' theorem is true with approximate ℓ_2^2 inequalities, but requires ARV analysis [Trevisan]
- Can we modify our algorithm to work with approximate triangle inequalities?
  - Or 'fix' ℓ_2^2 inequalities without blowing up approximate dimension

Dimension Reduction

- Standard Johnson-Lindenstrauss dimension reduction preserves ℓ_2^2 triangle inequalities *approximately* (in $O(\log n/\epsilon^2)$ dimensions)

$$\|z_i - z_j\|^2 + \|z_k - z_j\|^2 \geq (1 - O(\epsilon)) \|z_i - z_k\|^2$$

- Goemans' theorem is true with approximate ℓ_2^2 inequalities, but requires ARV analysis [Trevisan]
- Can we modify our algorithm to work with approximate triangle inequalities?
  - Or 'fix' ℓ_2^2 inequalities without blowing up approximate dimension

Outline

1 Introduction

2 Cheeger or Spectral Approach

3 Rounding a stronger SDP
- Our Algorithm

4 Goemans' Theorem

5 Summary

Summary and Future directions

- A simple SDP algorithm that gives non-trivial guarantees, using ℓ_2^2 inequalities
- Unconditional guarantees?
- Dimension reduction techniques to get ARV-like guarantees?

Summary and Future directions

- A simple SDP algorithm that gives non-trivial guarantees, using ℓ_2^2 inequalities
- Unconditional guarantees?
- Dimension reduction techniques to get ARV-like guarantees?

Summary and Future directions

- A simple SDP algorithm that gives non-trivial guarantees, using ℓ_2^2 inequalities
- Unconditional guarantees?
- Dimension reduction techniques to get ARV-like guarantees?

Thank you.