Spectral Sparsification and Applications

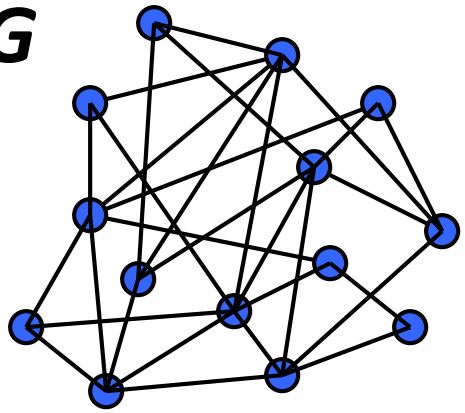
Nikhil Srivastava

Microsoft Research India

Part I: Graph Sparsification, Dimensionality Reduction, and Column Subset Selection

Graphs

G



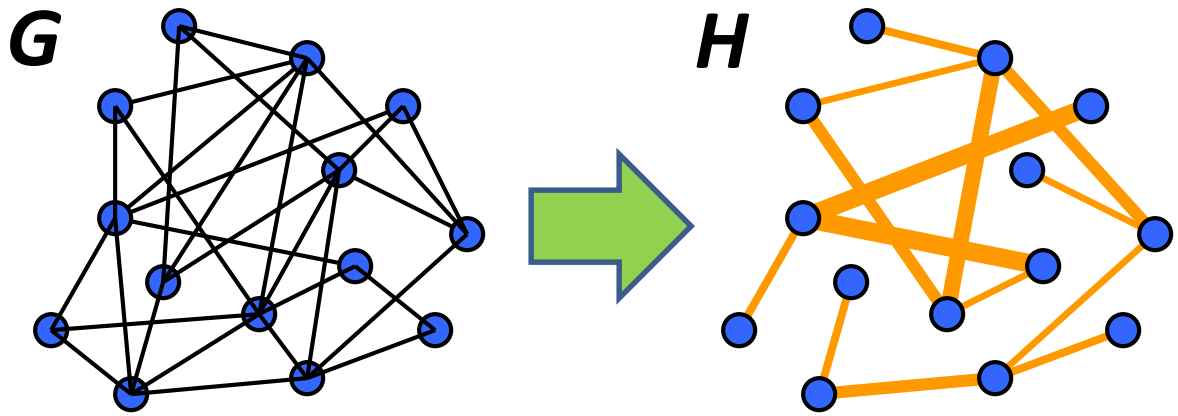
$\boldsymbol{G}=(V,E,w)$ *undirected*

$|V| = n$

$w: E \rightarrow \boldsymbol{R}_+$

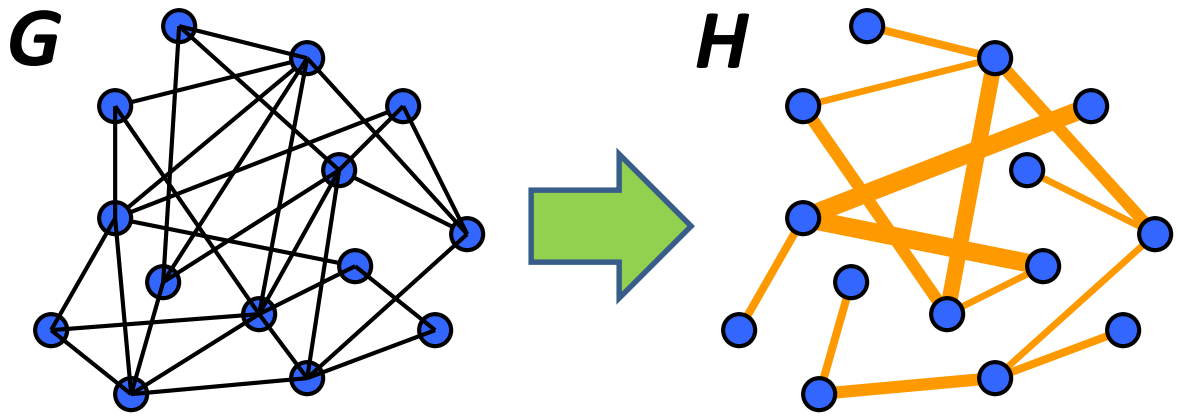
Sparsification

Approximate any graph ***G*** by a sparse graph ***H***.



Sparsification

Approximate any graph ***G*** by a sparse graph ***H***.



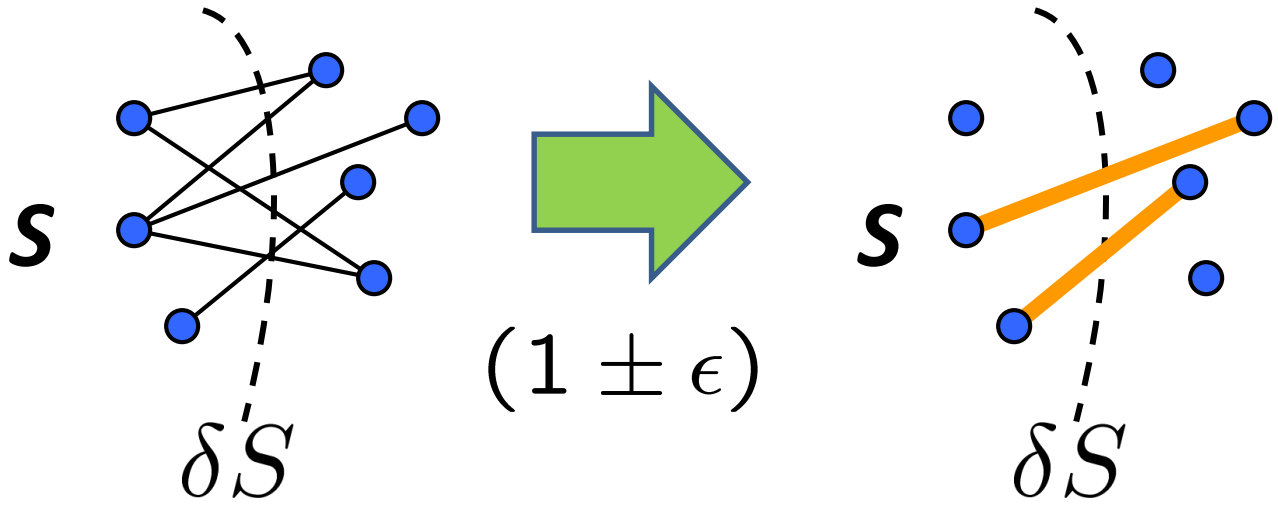
- Nontrivial statement about ***G***
- ***H*** is faster to compute with than ***G***

Preserving Cuts [Benczur-Karger'96]

H approximates ***G*** if

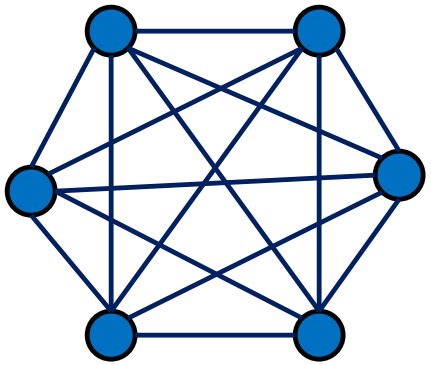
for every subset $S \subset V$

sum of weights of edges leaving ***S*** is preserved



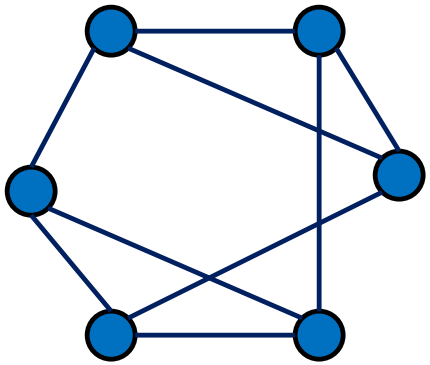
Example: The Complete Graph

G=K_n



$|E_G| = O(n^2)$

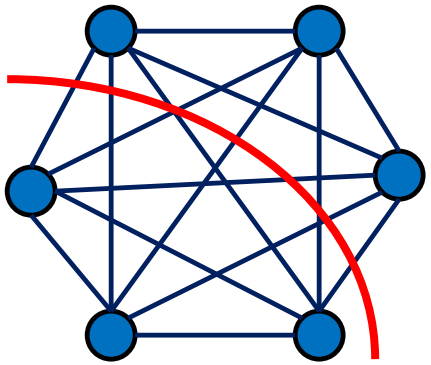
H = random d-regular



$|E_H| = O(dn)$

Example: The Complete Graph

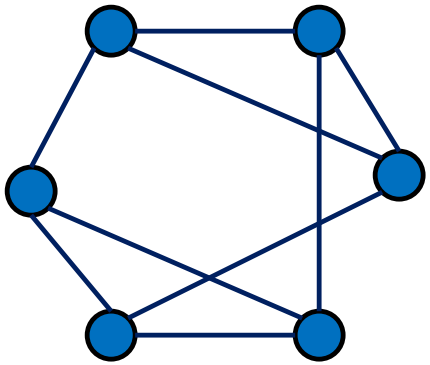
G=K_n



$|E_G| = O(n^2)$

$$wt_G(\delta S) = |S| \cdot |\overline{S}|$$

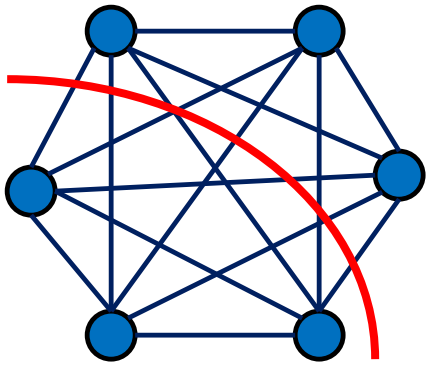
H = random d-regular



$|E_H| = O(dn)$

Example: The Complete Graph

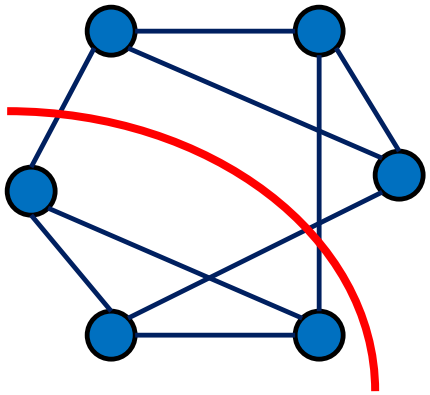
G=K_n



$|E_G| = O(n^2)$

$$wt_G(\delta S) = |S| \cdot |\overline{S}|$$

H = random d-regular



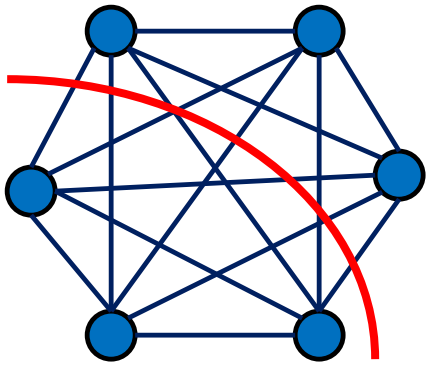
$|E_H| = O(dn)$

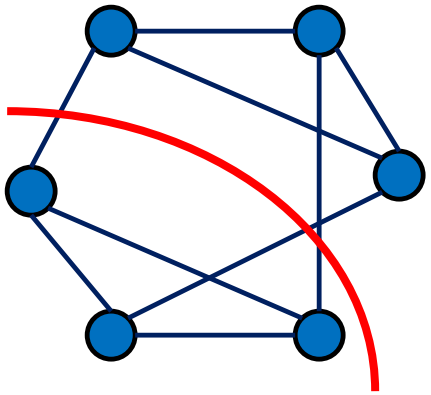
$$wt_H(\delta S) \simeq (d/n)|S| \cdot |\overline{S}|$$

Example: The Complete Graph

G=K_n

H = random d-regular





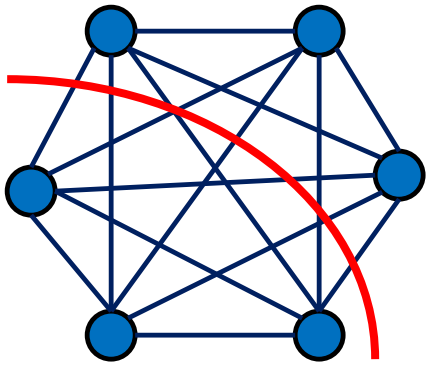
$|E_G| = O(n^2)$

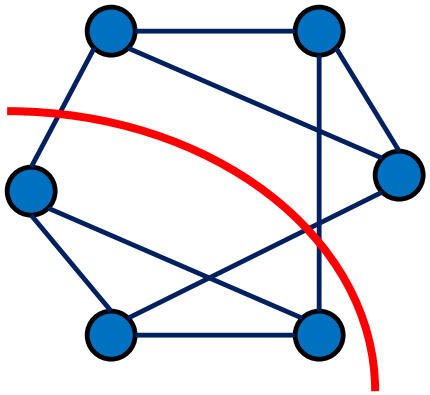
$|E_H| = O(dn)$

$$\forall S \subset V, \quad \frac{wt_G(\delta S)}{wt_H(\delta S)} \simeq (n/d)$$

Example: The Complete Graph

$G=K_n$ **H** = random d-regular x **(n/d)**





$|E_G| = O(n^2)$ $|E_H| = O(dn)$

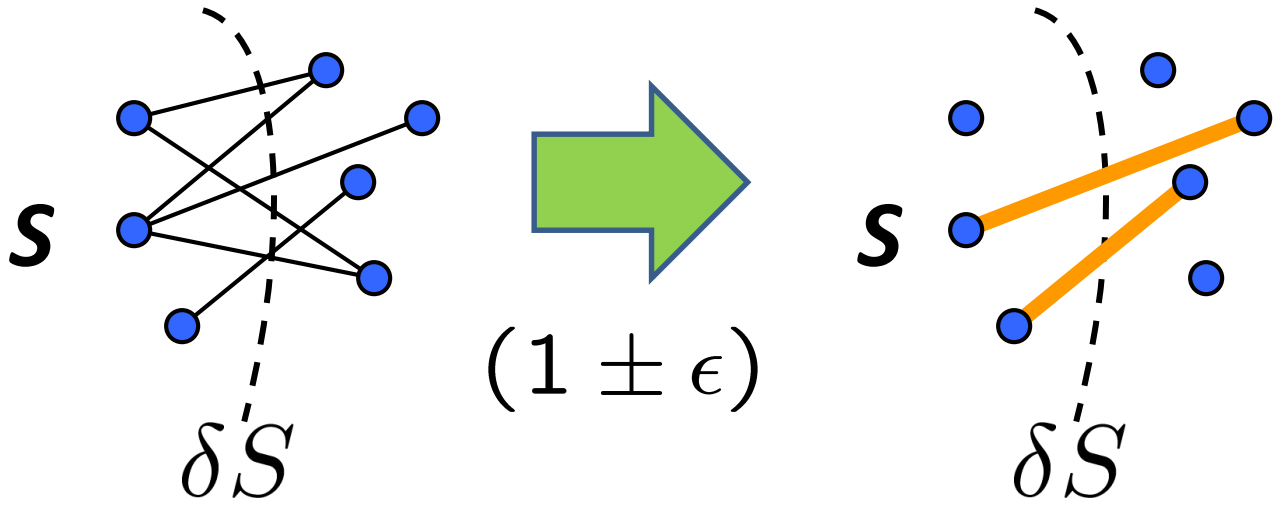
$$\forall S \subset V, \quad \frac{wt_G(\delta S)}{wt_{(n/d)H}(\delta S)} \simeq 1$$

Preserving Cuts [Benczur-Karger'96]

H approximates ***G*** if

for every subset $S \subset V$

sum of weights of edges leaving ***S*** is preserved



[Benczur-Karger'96]*:* For every ***G*** can quickly find ***H*** with **O(nlogn/ε^2)** edges.

A Stronger Notion

The Laplacian Matrix

$$L_G = D_G - A_G = \sum_{ijE} w_{ij}(\delta_i - \delta_j)(\delta_i - \delta_j)^T$$

The Laplacian Matrix

$$L_G = D_G - A_G = \sum_{ijE} w_{ij}(\delta_i - \delta_j)(\delta_i - \delta_j)^T$$

$$\begin{matrix} i \\ j \end{matrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{matrix} i & j \end{matrix}$$

The Laplacian Matrix

$$L_G = D_G - A_G = \sum_{ijE} w_{ij}(\delta_i - \delta_j)(\delta_i - \delta_j)^T$$

Quadratic form

$$x : V \to \mathbf{R}$$

$$x^T L_G x = \sum_{ij \in E} w_{ij}(x(i) - x(j))^2$$

The Laplacian Matrix

$$L_G = D_G - A_G = \sum_{ijE} w_{ij}(\delta_i - \delta_j)(\delta_i - \delta_j)^T$$

Quadratic form

$$x : V \to \mathbf{R}$$

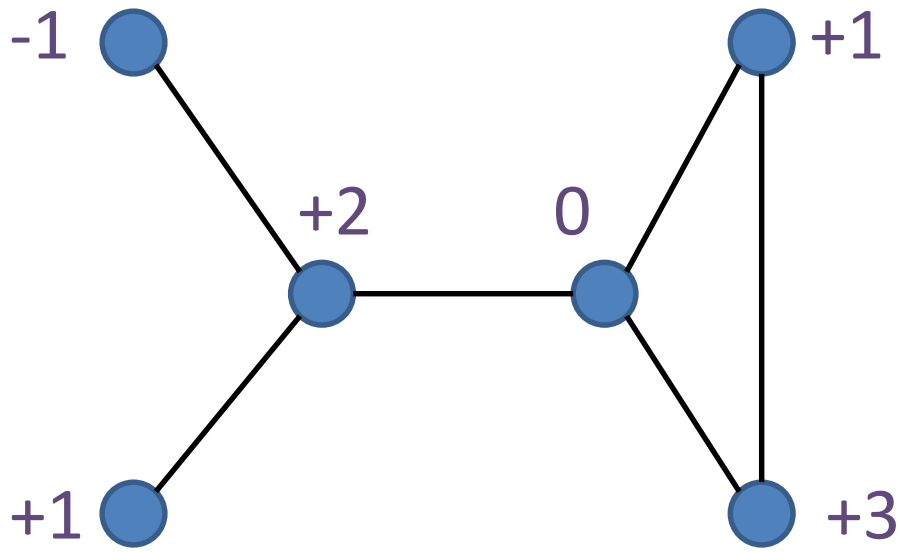
$$x^T L_G x = \sum_{ij \in E} w_{ij}(x(i) - x(j))^2$$

Positive semidefinite

Ker($\mathbf{L_G}$)=span(**1**) if ***G*** is connected

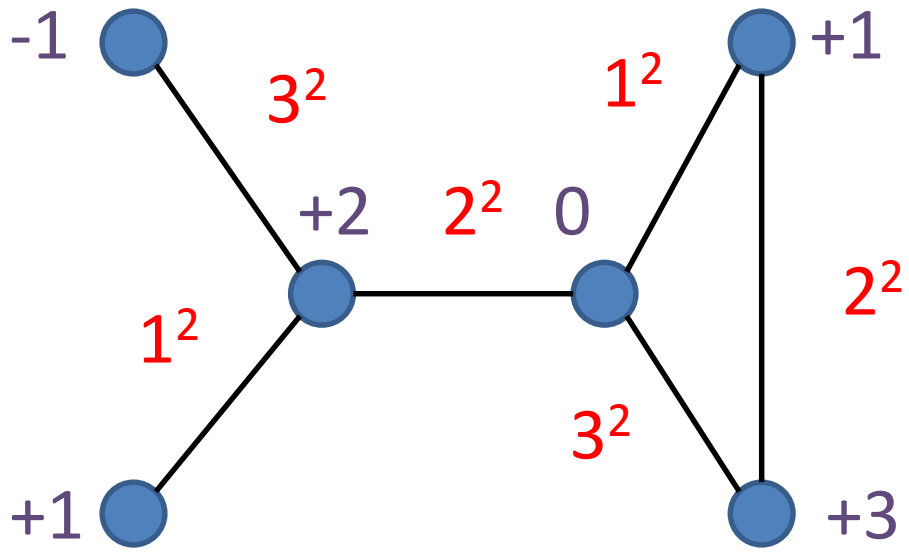
The Laplacian Quadratic Form

An example:



The Laplacian Quadratic Form

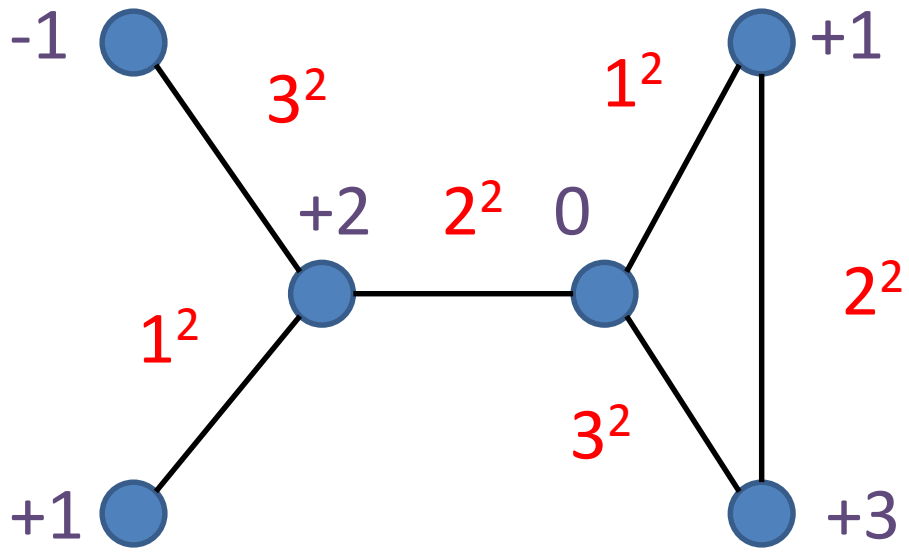
An example:



$$\mathbf{x}^T L\mathbf{x} = \sum_{i,j \in E} (\mathbf{x}(i)-\mathbf{x}(j))^2$$

The Laplacian Quadratic Form

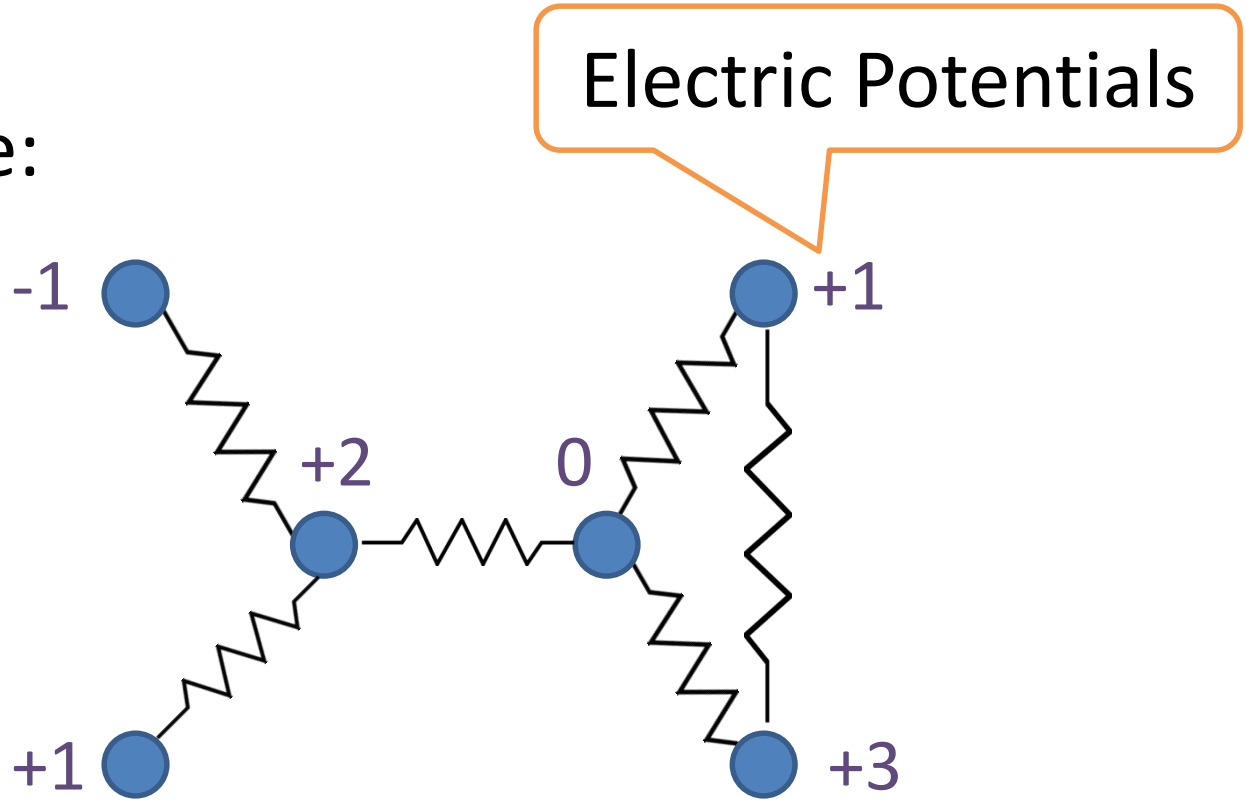
An example:



$$\mathbf{x}^T\boldsymbol{L}\mathbf{x} = \sum_{i,j \in E} (\mathbf{x}(i)-\mathbf{x}(j))^2 = 28$$

The Laplacian Quadratic Form

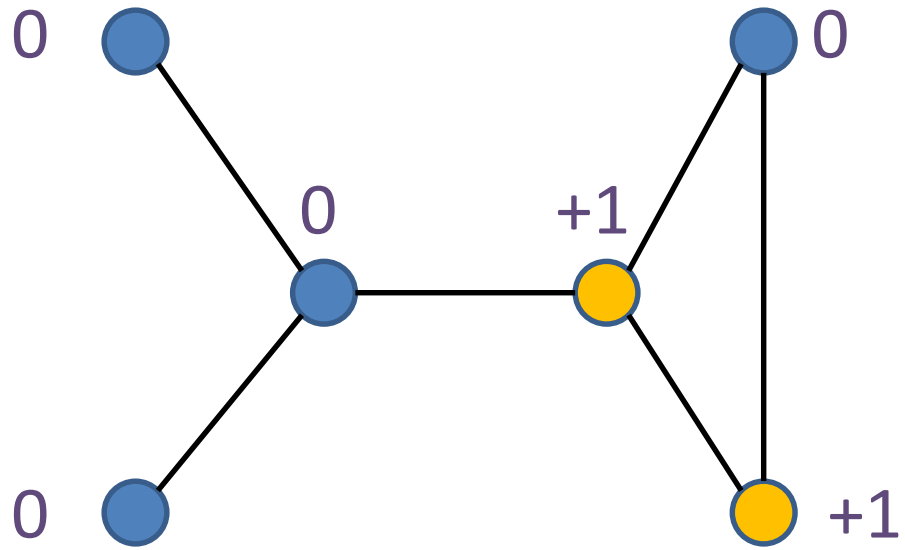
An example:



$$\mathbf{x}^T L\mathbf{x} = \sum_{i,j \in E} (\mathbf{x}(i)-\mathbf{x}(j))^2 = 28$$

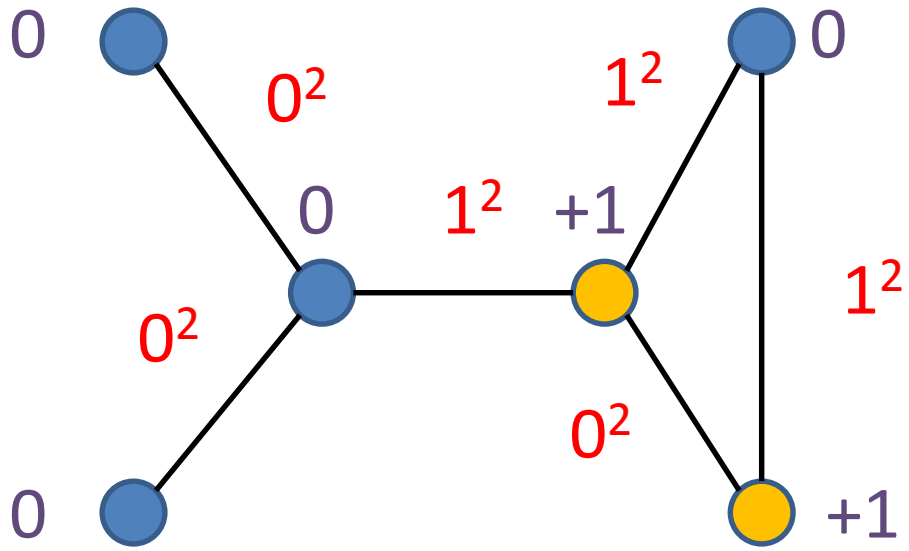
The Laplacian Quadratic Form

Another example:



The Laplacian Quadratic Form

Another example:

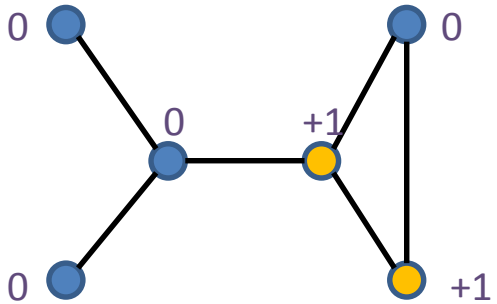


$$x^T L_G x = 3$$

Cuts and the Quadratic Form

For characteristic vector $x_S \in \{0,1\}^n$ of $S \subseteq V$

$$x_S^T L_G x_S = \sum_{ij \in E} w_{ij}(x(i) - x(j))^2$$



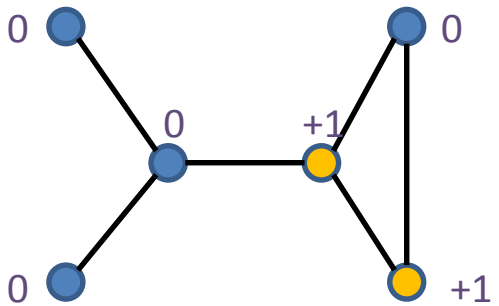
$$= \sum_{ij \in (S,\overline{S})} w_{ij}$$

$$= wt_G(S, \overline{S})$$

Cuts and the Quadratic Form

For characteristic vector $x_S \in \{0,1\}^n$ of $S \subseteq V$

$$x_S^T L_G x_S = \sum_{ij \in E} w_{ij}(x(i) - x(j))^2$$



$$= \sum_{ij \in (S,\overline{S})} w_{ij}$$

$$= wt_G(S, \overline{S})$$

So BK says:

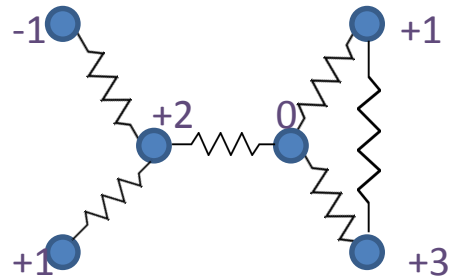
$$(1-\epsilon)x^T L_G x \leq x^T L_H x \leq (1+\epsilon)x^T L_G x$$

$$\forall x \in \{0,1\}^n$$

A Stronger Notion [ST'04]

For characteristic vector $x_S \in \{0,1\}^n$ of $S \subseteq V$

$$x_S^T L_G x_S = \sum_{ij \in E} w_{ij}(x(i) - x(j))^2$$



$$= \sum_{ij \in (S,\overline{S})} w_{ij}$$

$$= wt_G(S, \overline{S})$$

"Electrically Equivalent"

ST'04:

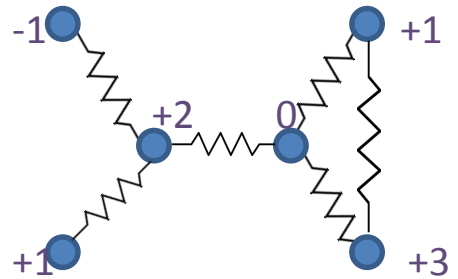
$$(1-\epsilon)x^T L_G x \leq x^T L_H x \leq (1+\epsilon)x^T L_G x$$

$$\forall x \in \mathbb{R}^n \qquad \forall x \in \{0,1\}^n$$

A Stronger Notion [ST'04]

For characteristic vector $x_S \in \{0,1\}^n$ of $S \subseteq V$

$$x_S^T L_G x_S = \sum_{ij \in E} w_{ij}(x(i) - x(j))^2$$



$$= \sum_{ij \in (S,\overline{S})} w_{ij}$$

$$= wt_G(S, \overline{S})$$

ST'04:

$$(1-\epsilon)L_G \preceq L_H \preceq L_G(1+\epsilon)$$

A Stronger Notion [ST'04]

$\{0,1\}^n$ of $S \subseteq V$

$- x(j))^2$

$= c(S)$

Thm[ST'04]

Can find H with $O(n \log^8 n/\epsilon^2)$ edges.

ST'04:

$$(1-\epsilon)L_G \preceq L_H \preceq L_G(1+\epsilon)$$

Why?

1. Solving $L_G x = b$ [ST'04]

$x^T L_G x \sim x^T L_H x$: can solve systems

in $\mathbf{L_G}$ by solving systems in $\mathbf{L_H}$.

Naïve:	$O(n^3)$
Fast Matrix Multiply:	$O(n^{2.37})$

1. Solving $L_G x = b$ [ST'04]

$\boldsymbol{x^T L_G x \sim x^T L_H x}$: can solve systems

in $\mathbf{L_G}$ by solving systems in $\mathbf{L_H}$.

Naïve:	$O(n^3)$
Fast Matrix Multiply:	$O(n^{2.37})$
ST'04	$O(m\log^{30} n)$
KMP'10	$O(m\log n)$

1. Solving $L_G x = b$ [ST'04]

$x^T L_G x \sim x^T L_H x$: can solve systems in L_G by solving systems in L_H.

Naïve:	$O(n^3)$
Fast Matrix Multiply:	$O(n^{2.37})$
ST'04	$O(m \log^{30} n)$
KMP'10	$O(m \log n)$

Thm [ST'04] $\forall G$ can find H with $O(n \log^8 n)$ edges.

2. Spectral Graph Theory

Courant-Fischer Thm: $\boldsymbol{x^T L_G x}$ determines $\lambda_i(L_G)$

$$\lambda_{max}(L) = \max \frac{x^T L x}{x^T x} \qquad \lambda_{min}(L) = \min \frac{x^T L x}{x^T x}$$

Thus for spectral sparsifier **H** of **G**:

$$(1-\epsilon)\lambda_i(G) \leq \lambda_i(H) \leq (1+\epsilon)\lambda_i(G)$$

Now **H** inherits many combinatorial properties:

random walks, colorings, spanning trees, etc.

3. Natural Setting

Geometric formulation more tractable:

$\boldsymbol{x^T L x}$ better behaved over $\mathbf{R}^n$ than $\{0,1\}^n$.

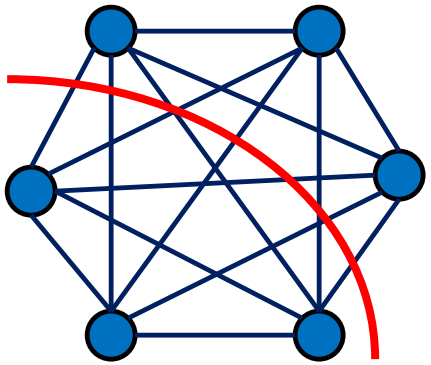
As we'll see, can get optimal O(n) sparsity

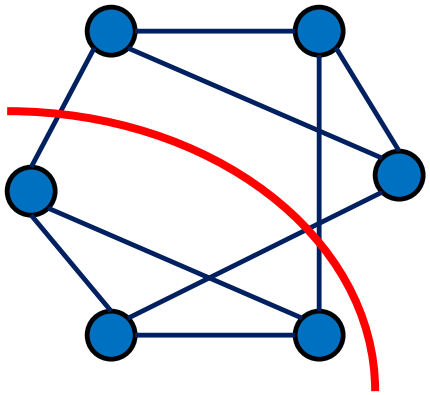
– **better** than O(nlogn) for cuts.

Examples

Example: The Complete Graph

$G=K_n$ **H** = random d-regular x **(n/d)**



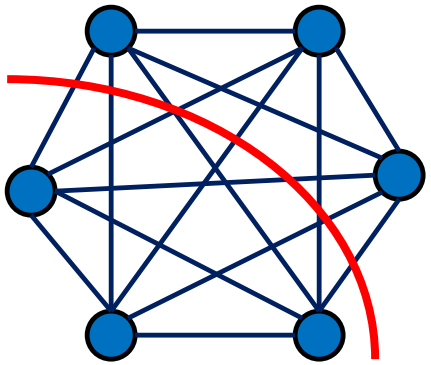


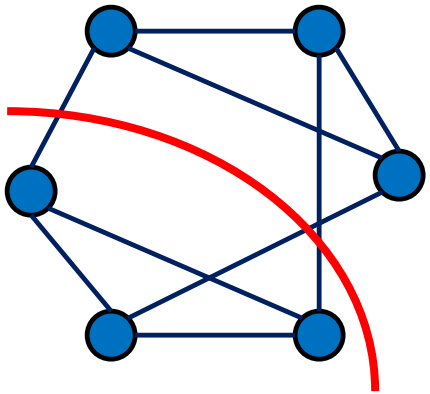
$|E_G| = O(n^2)$ $|E_H| = O(dn)$

$$\forall S \subset V, \quad \frac{wt_G(\delta S)}{wt_{(n/d)H}(\delta S)} \simeq 1 \pm \epsilon$$

Example: The Complete Graph

$G=K_n$ **H** = random d-regular x **(n/d)**



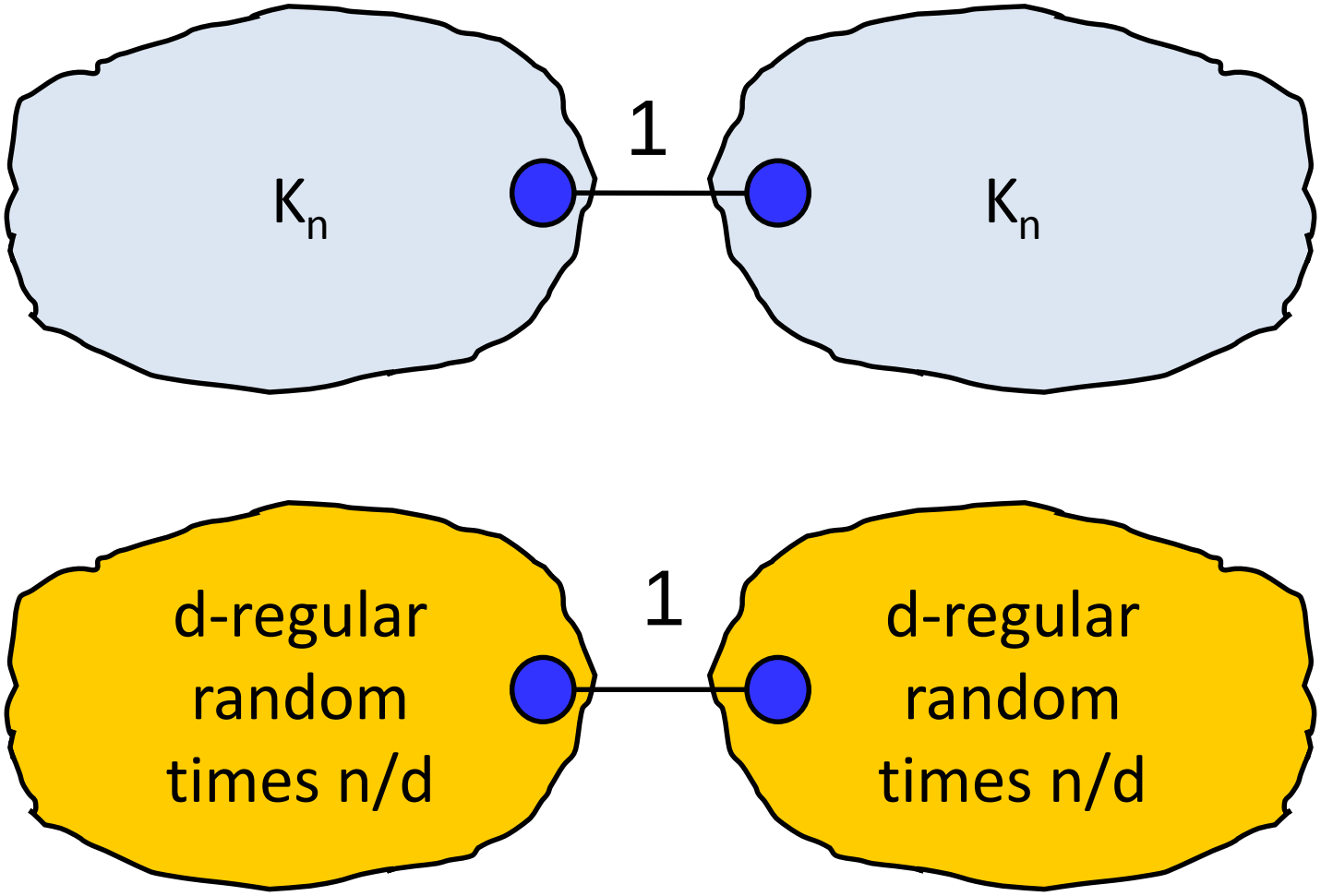


$|E_G| = O(n^2)$ $|E_H| = O(dn)$

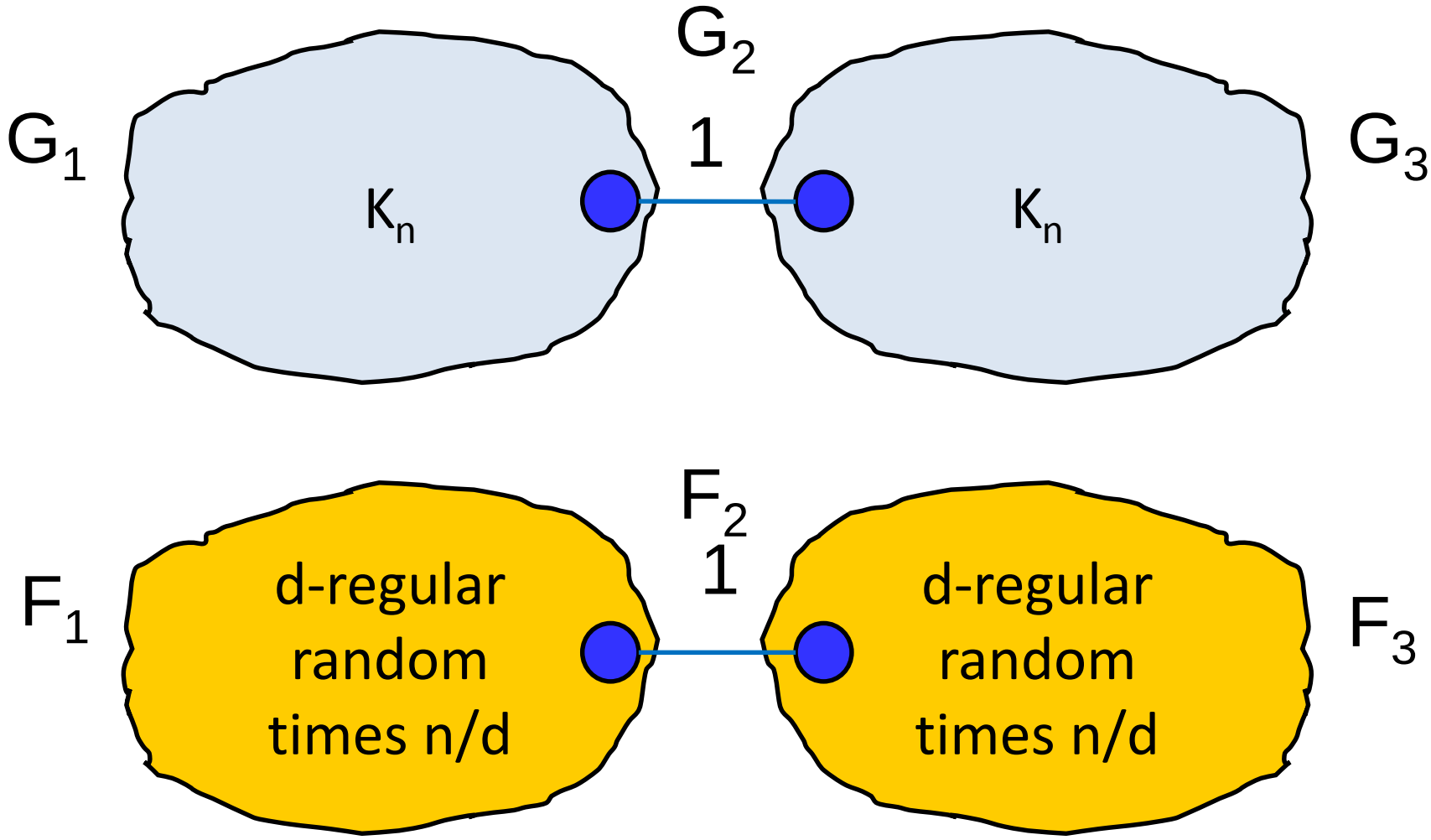
$$d = 1/\epsilon^2$$

$$\forall S \subset V, \quad \frac{wt_G(\delta S)}{wt_{(n/d)H}(\delta S)} \simeq 1 \pm \epsilon$$

Example: Dumbell



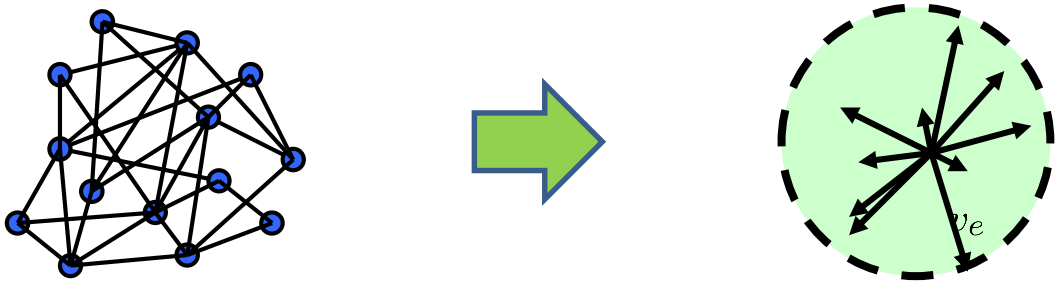
Example: Dumbell



$$G = G_1 + G_2 + G_3$$

$$x^T G x = x^T G_1 x + x^T G_2 x + x^T G_3 x$$

Will show how to do this for every graph…

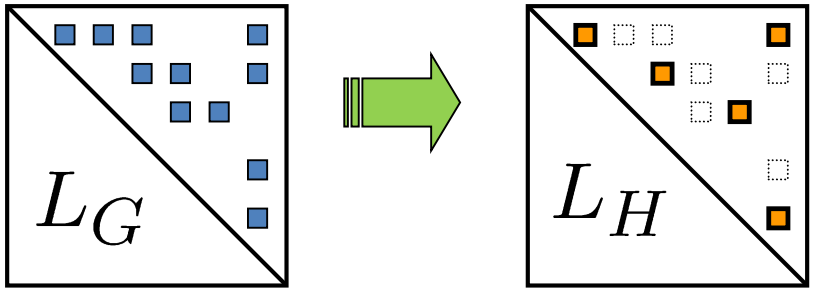


Step 1: Reduction to Geometric Problem

Original Goal

Given G

Find sparse H



satisfying $L_G \preceq L_H \preceq \kappa \cdot L_G$

Outer Product Expansion

Recall:

$$L_G = \sum_{ij \in E} (\delta_i - \delta_j)(\delta_i - \delta_j)^T = \sum_{e \in E} b_e b_e^T.$$

Outer Product Expansion

Recall:

$$L_G = \sum_{ij \in E} (\delta_i - \delta_j)(\delta_i - \delta_j)^T = \sum_{e \in E} b_e b_e^T.$$

For a weighted subgraph ***H***:

$$L_H = \sum_{e \in E} s_e b_e b_e^T$$

where s_e=wt(e) in H.

Original Goal

Given G

Find **sparse** H

satisfying

$$L_G \preceq L_H \preceq \kappa \cdot L_G$$

Original Goal

Given $L_G = \sum_{e \in G} b_e b_e^T \qquad b_{ij} = \delta_i - \delta_j$

Find **sparse**

$$s_e \geq 0$$

satisfying

$$L_G \preceq L_H = \sum_{e \in G} s_e b_e b_e^T \preceq \kappa \cdot L_G$$

More General Goal

Given $V = \sum_e v_e v_e^T$

Find **sparse**

$$s_e \geq 0$$

satisfying

$$V \preceq \sum_{e \in G} s_e v_e v_e^T \preceq \kappa \cdot V$$

Equivalent General Goal

Given $$I = \sum_e v_e v_e^T$$

Find **sparse**

$$s_e \geq 0$$

satisfying

$$I \preceq \sum_{e \in G} s_e v_e v_e^T \preceq \kappa \cdot I$$

Proof of Equivalence

Given $V = \sum_e v_e v_e^T$

Find **sparse**

$s_e \geq 0$

satisfying

$$V \preceq \sum_{e \in G} s_e v_e v_e^T \preceq \kappa \cdot V$$

Take

$v'_e = V^{-1/2} v_e.$

Proof of Equivalence

Given $V = \sum_e v_e v_e^T$

Find **sparse**

$s_e \geq 0$

satisfying

$V \preceq \sum_{e \in G} s_e v_e v$

Take

$v'_e = V^{-1/2} v_e.$

$\sum_e s_e v'_e v'^T_e \approx I$

$\sum_e v_e v_e^T \approx V.$

Core Problem

Given $$I = \sum_e v_e v_e^T$$

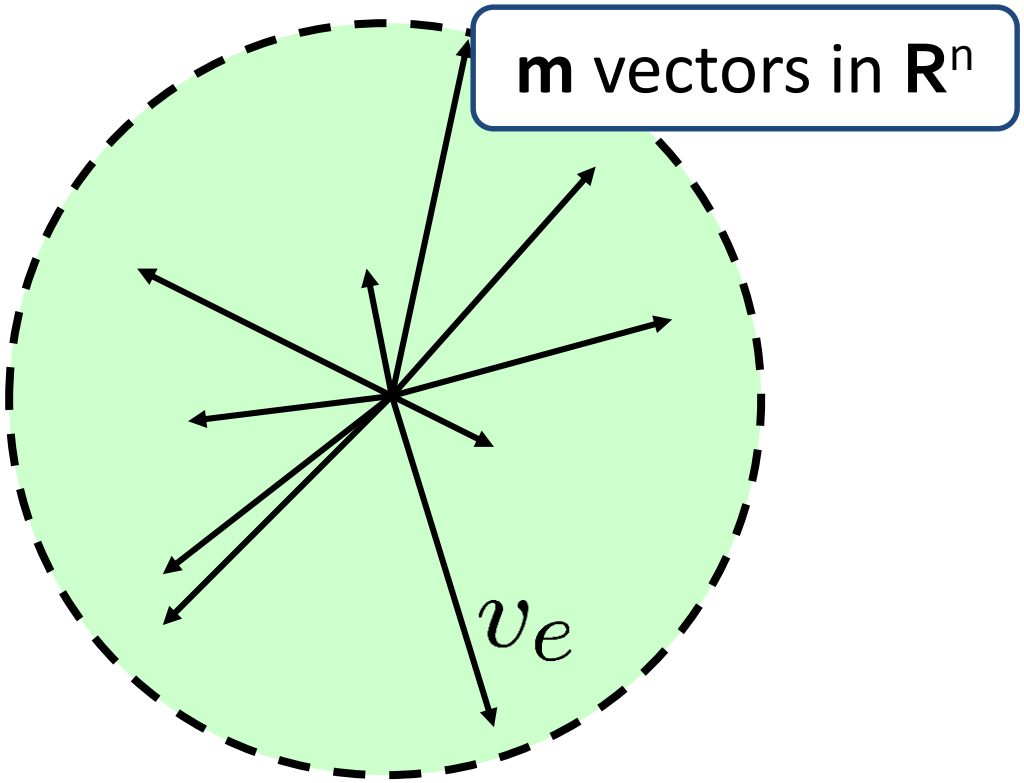
Find **sparse**

$$s_e \geq 0$$

satisfying

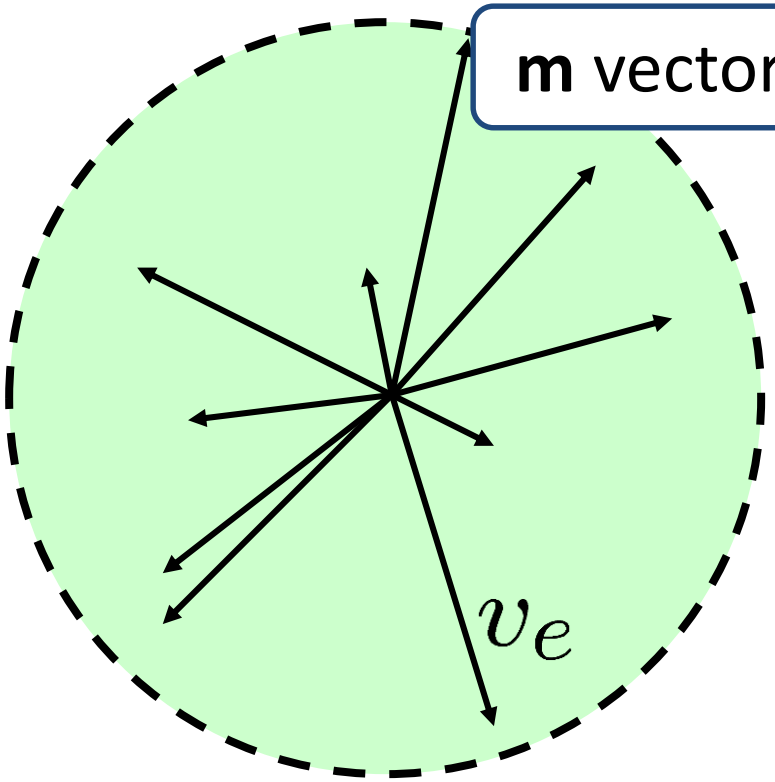
$$I \preceq \sum_{e \in G} s_e v_e v_e^T \preceq \kappa \cdot I$$

Core Problem



$$I = \sum_e v_e v_e^T$$

Core Problem



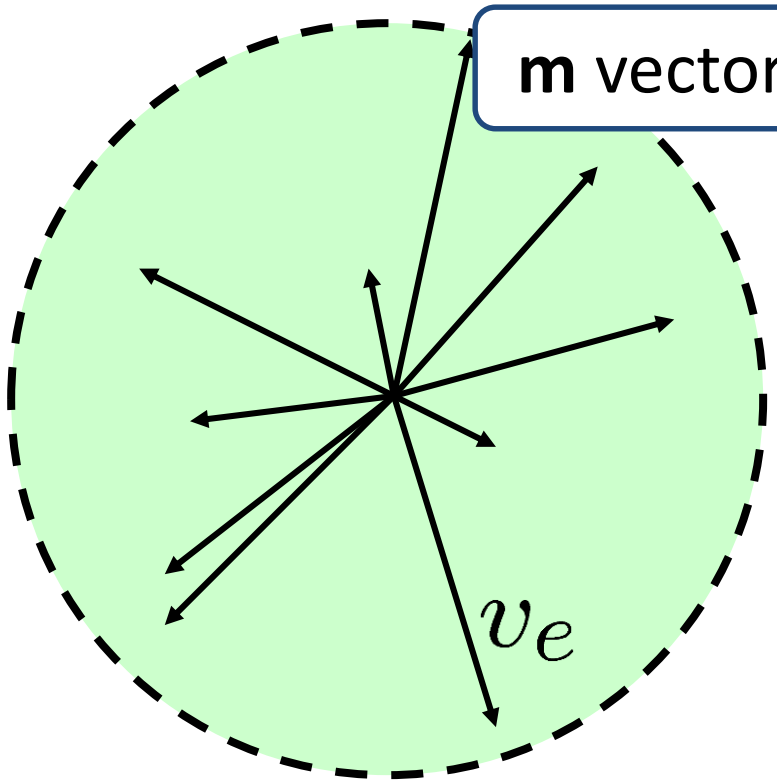
m vectors in $\mathbf{R}^n$

$$\forall u \quad \sum_e \langle u, v_e \rangle^2 = 1$$

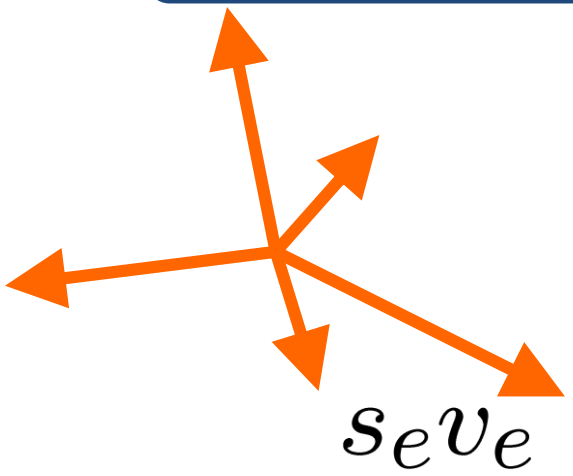
variance is the same in every direction

$$I = \sum_e v_e v_e^T$$

Core Problem



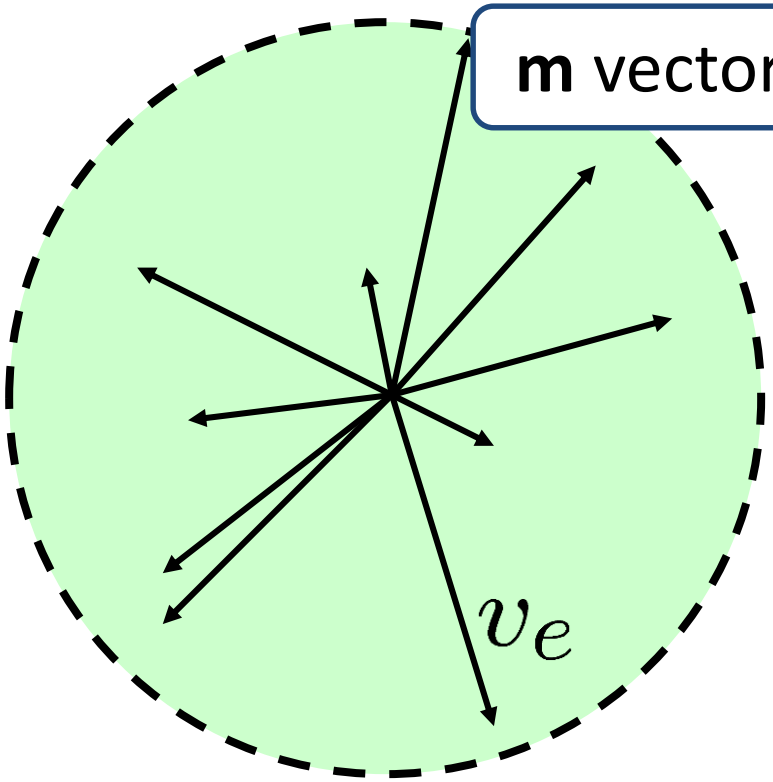
m vectors in $\mathbf{R}^n$

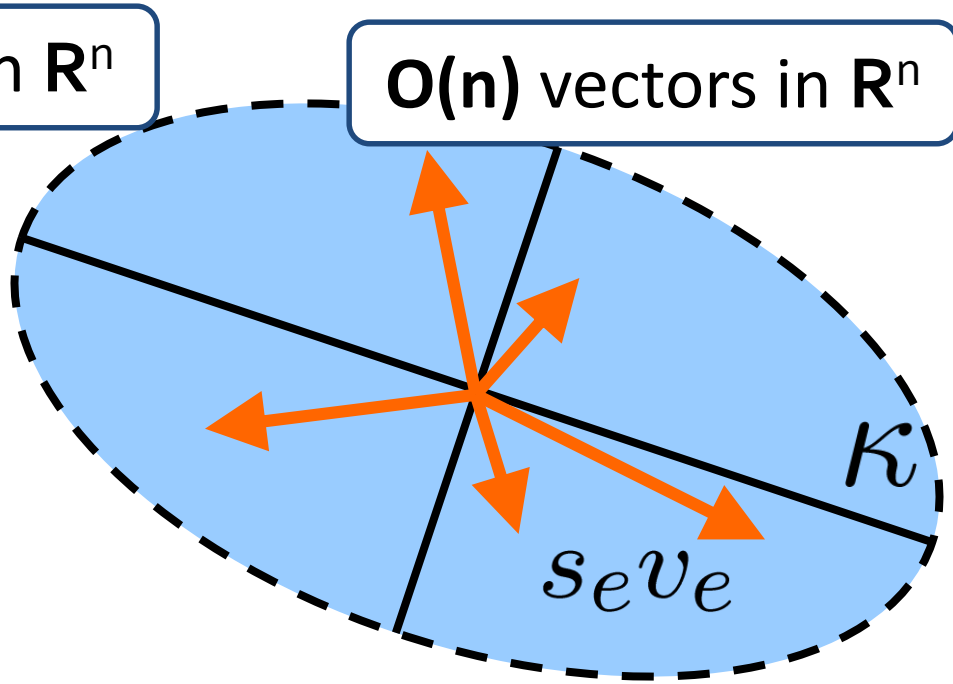


O(n) vectors in $\mathbf{R}^n$

$$I = \sum_e v_e v_e^T$$

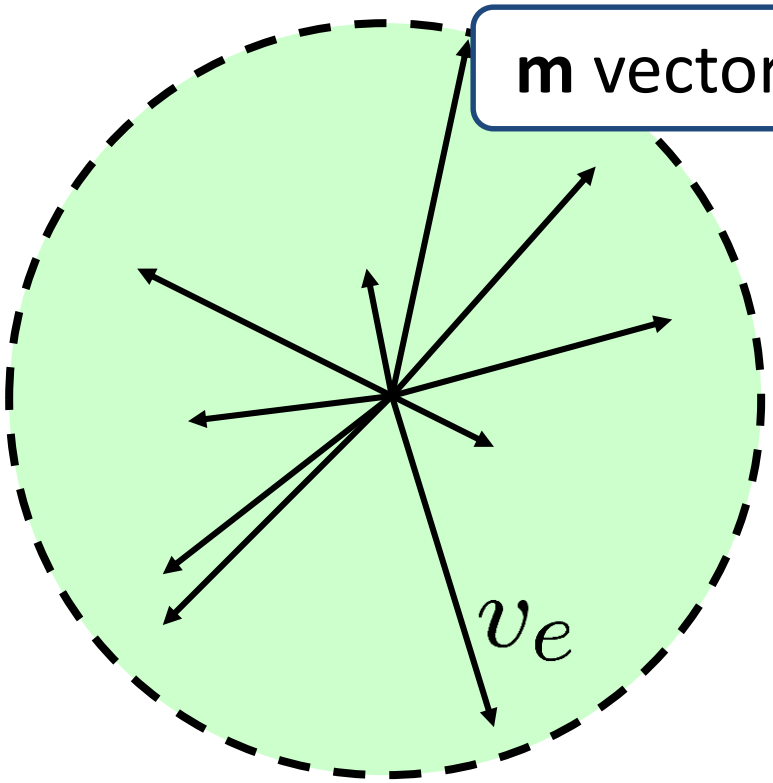
Core Problem

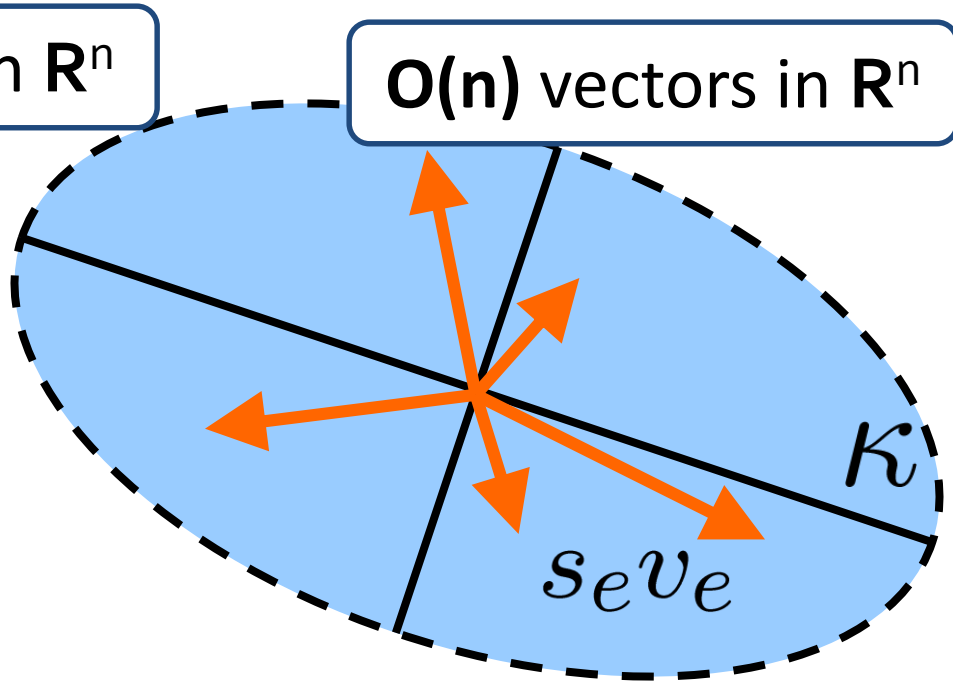




$$I = \sum_e v_e v_e^T \qquad I \preceq \sum_e s_e v_e v_e^T \preceq \kappa I$$

Core Problem





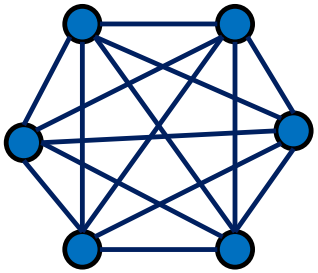
$$I = \sum_{i \leq m} v_i v_i^T \qquad I \preceq \sum_i s_e v_i v_i^T \preceq \kappa I$$

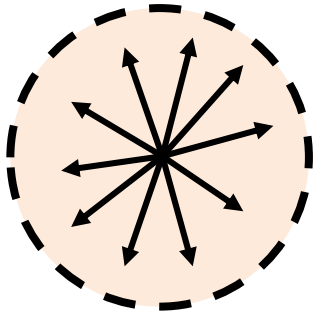
Examples of the Reduction

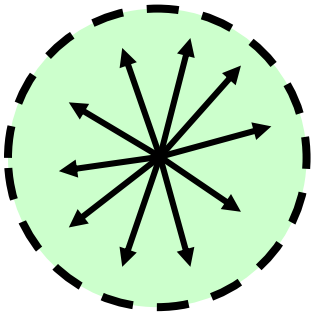
Graph

$L_G = \sum_e b_e b_e^T$

$I = \sum_e v_e v_e^T$







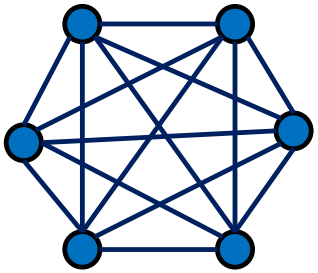
$v_e = L_G^{-1/2} b_e$

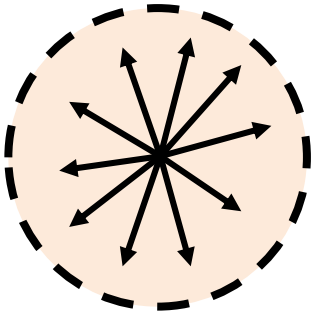
Examples of the Reduction

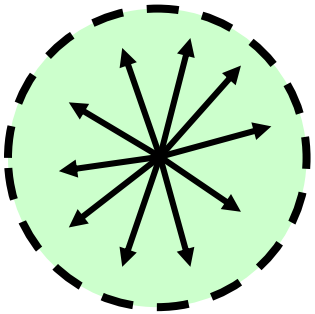
Graph

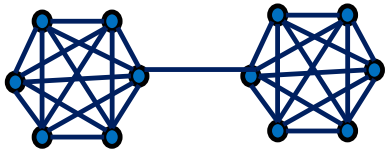
$L_G = \sum_e b_e b_e^T$

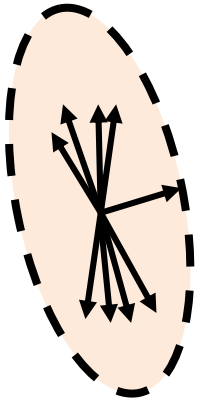
$I = \sum_e v_e v_e^T$

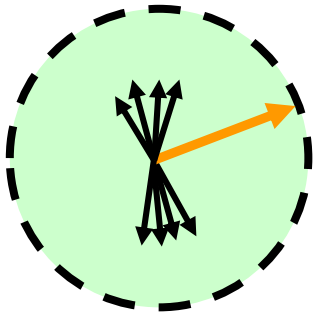






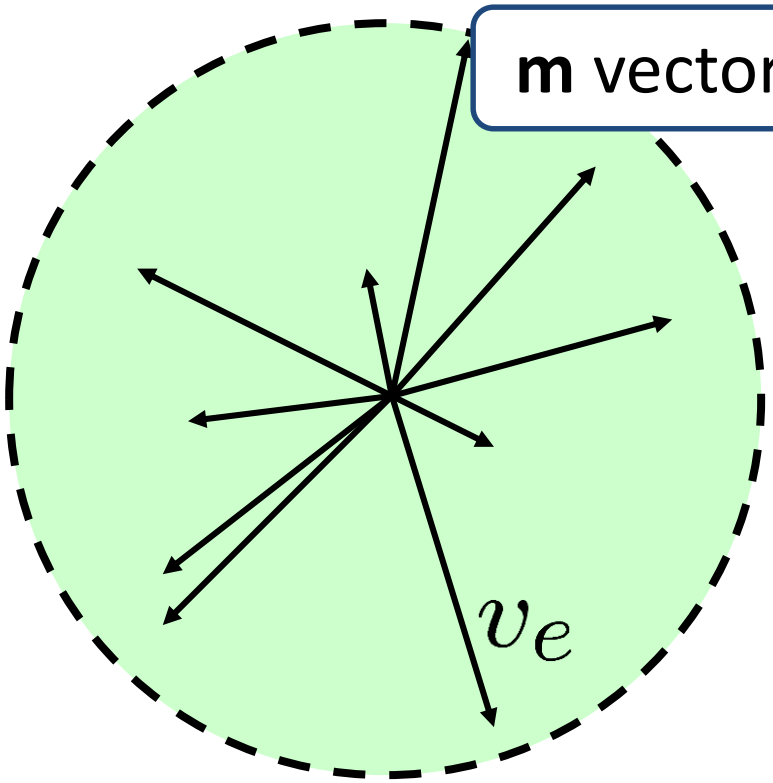


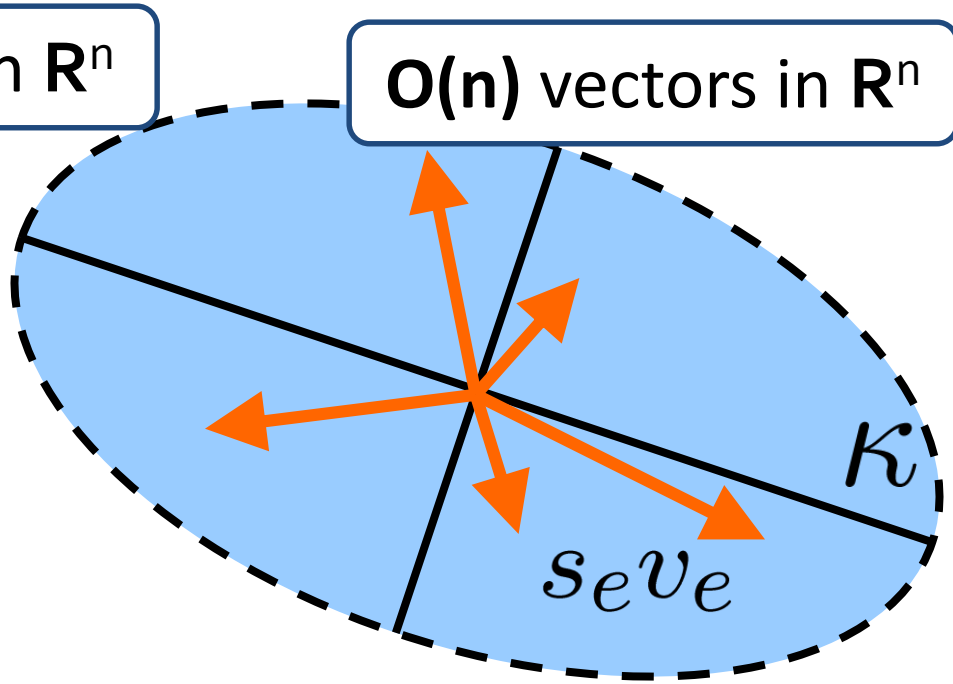




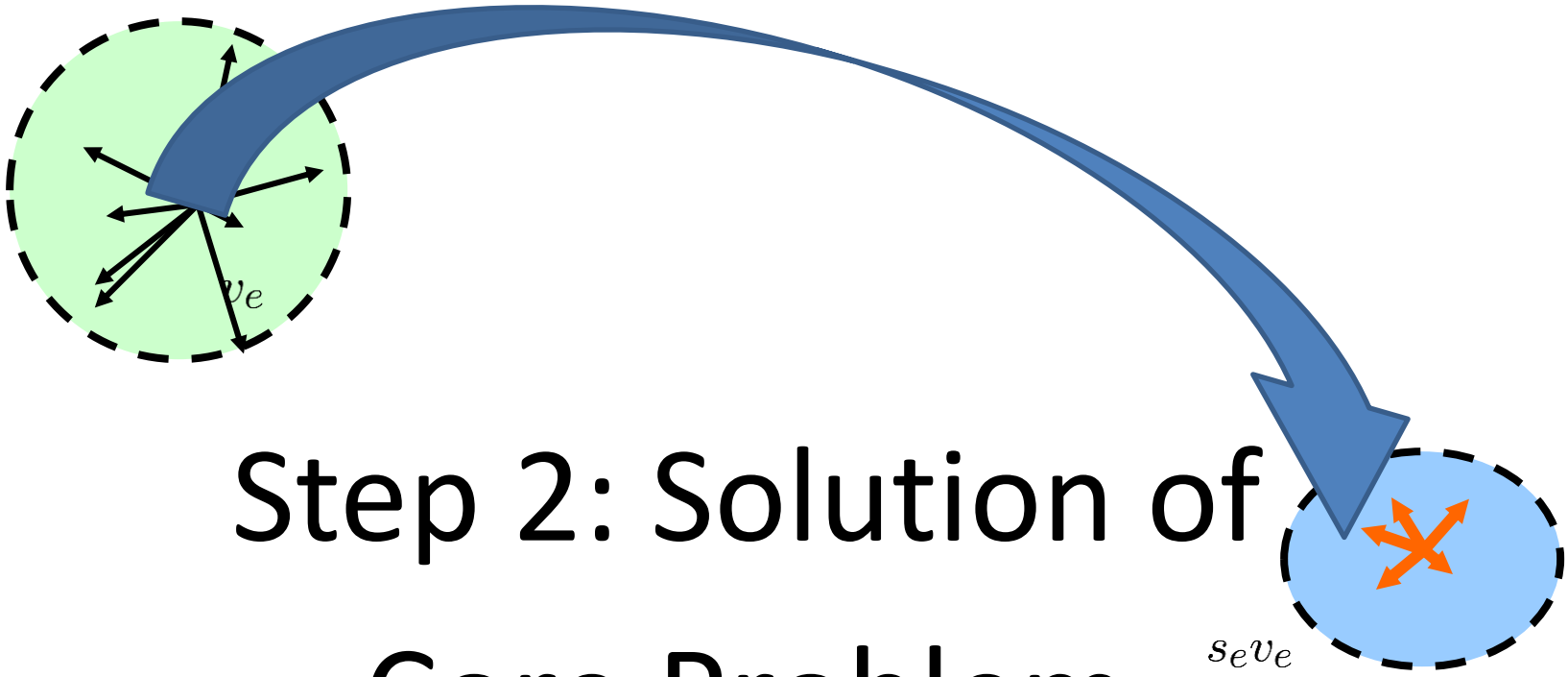
$v_e = L_G^{-1/2} b_e$

Core Problem

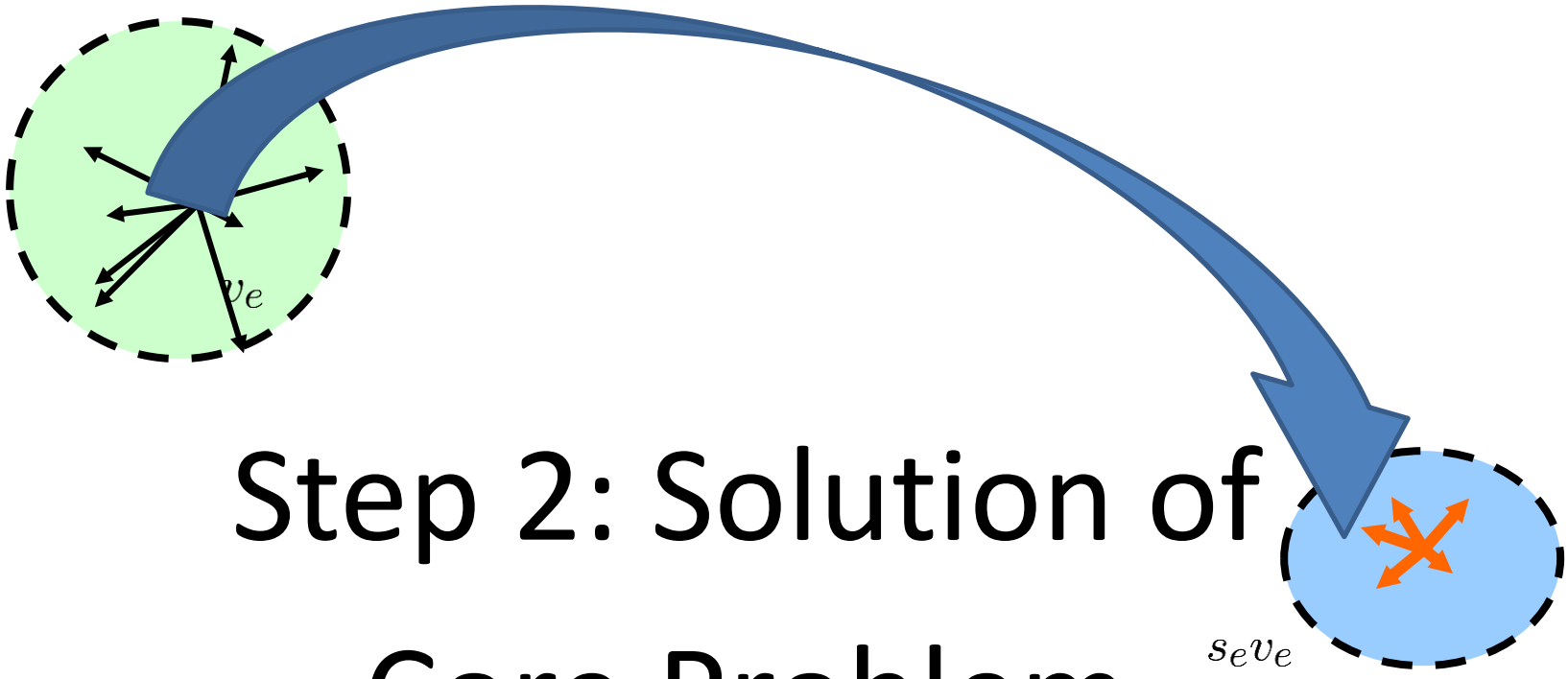




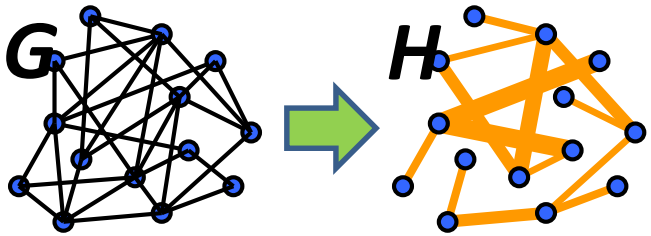
$$I = \sum_{i \leq m} v_i v_i^T \qquad I \preceq \sum_i s_e v_i v_i^T \preceq \kappa I$$



Step 2: Solution of Core Problem



Step 2: Solution of Core Problem

(implies 

)

Random Solution [Rudelson'97]

Given $\sum_i v_i v_i^T = I$

Sample $n \log n / \epsilon^2$ vectors randomly with replacement, by $p_i \propto \|v_i\|^2$.

Set $s_i = 1/p_i$ for chosen vectors.

Random Solution [Rudelson'97]

Given $\sum_i v_i v_i^T = I$

Sample $n \log n / \epsilon^2$ vectors randomly with replacement, by $p_i \propto \|v_i\|^2$.

Set $s_i = 1/p_i$ for chosen vectors.

Rudelson'97: This works whp:

$$1 - \epsilon \preceq \sum_i s_i v_i v_i^T \preceq 1 + \epsilon$$

Effective Resistance View [Spielman-S'08]

For a graph ***G***, the vectors are $v_e = L_G^{-1/2} b_e$

Lengths of vectors are:

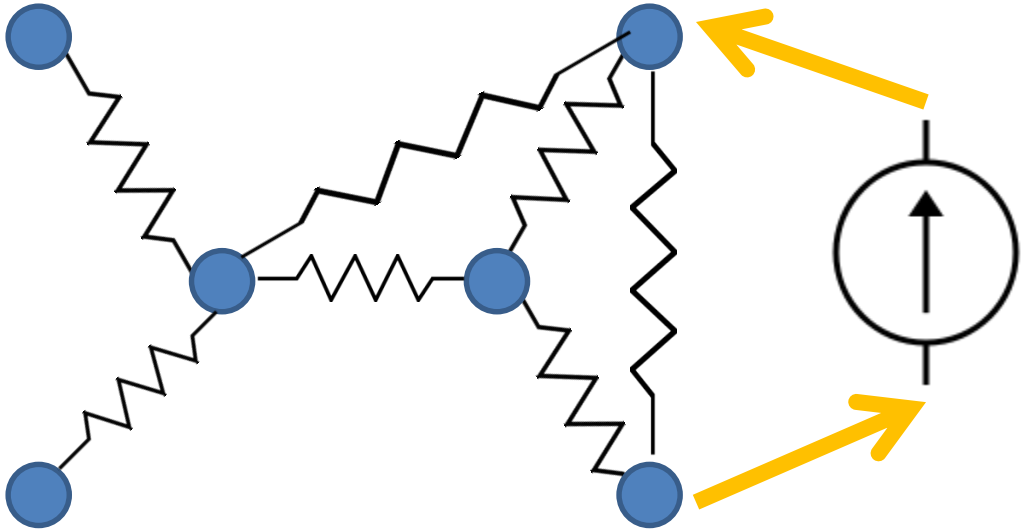
$$\|v_e\|^2 = \|L_G^{-1/2} b_e\|^2 = b_e^T L_G^{-1} b_e = \mathbf{Reff}_G(e)$$

Effective Resistance View [Spielman-S'08]

For a graph ***G***, the vectors are $v_e = L_G^{-1/2} b_e$

Lengths of vectors are:

$$\|v_e\|^2 = \|L_G^{-1/2} b_e\|^2 = b_e^T L_G^{-1} b_e = \mathbf{Reff}_G(e)$$



Random Solution [Rudelson'97]

Given $\sum_i v_i v_i^T = I$

Sample $n \log n / \epsilon^2$ vectors randomly with replacement, by $p_i \propto \|v_i\|^2$.

Set $s_i = 1/p_i$ for chosen vectors.

Rudelson'97: This works whp:

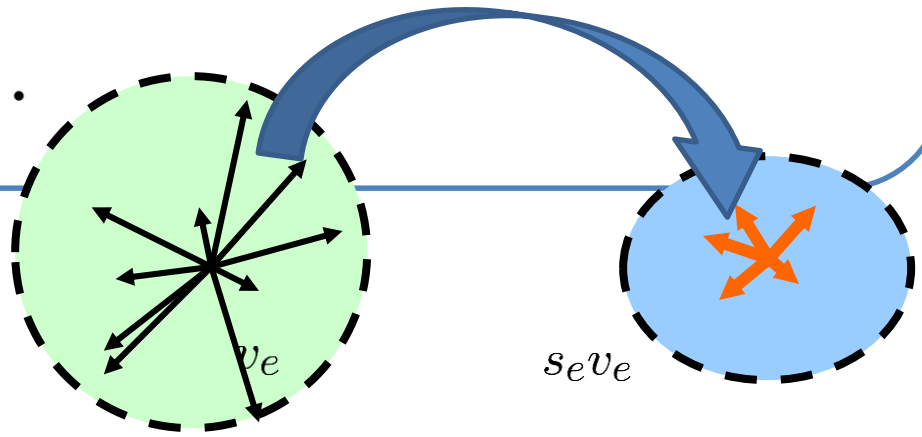
$$1 - \epsilon \preceq \sum_i s_i v_i v_i^T \preceq 1 + \epsilon$$

Deterministic Solution [Batson-Spielman-S'09]

Spectral Sparsification Theorem:

Given $\sum_{i \le m} v_i v_i^T = I_n$ there are $s_i \ge 0$ with:

- $(1-\epsilon) I \preceq \sum_i s_i v_i v_i^T \preceq (1+\epsilon) I$
- $\mathrm{supp}(s) \le 4n/\epsilon^2$.



Deterministic Solution [BSS'09]

Spectral Sparsification Theorem:

Given $\sum_{i \le m} v_i v_i^T = V$ there are $s_i \geq 0$ with:

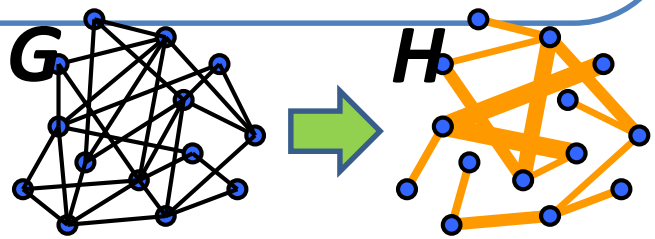
- $(1-\epsilon)V \preceq \sum_i s_i v_i v_i^T \preceq (1+\epsilon)V$
- $\mathrm{supp}(s) \leq 4\mathrm{rank}(V)/\epsilon^2$.

Deterministic Solution [BSS'09]

Spectral Sparsification Theorem:

Given $\sum_{i \le m} b_e b_e^T = L_G$ there are $s_e \geq 0$ with:

- $(1-\epsilon) L_G \preceq \sum_i s_i v_i v_i^T \preceq (1+\epsilon) L_G$
- $\mathrm{supp}(s) \leq 4(n-1)/\epsilon^2$.



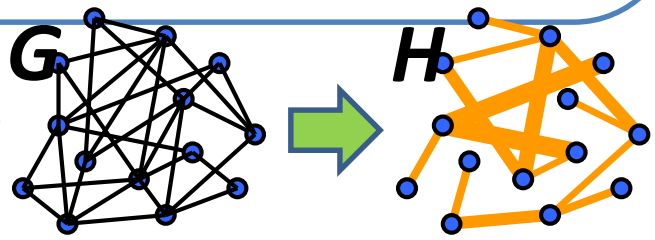
Deterministic Solution [BSS'09]

Spectral Sparsification Theorem:

Given $\sum_{i \le m} b_e b_e^T = L_G$ there are $s_e \ge 0$ with:

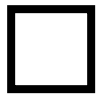
- $(1-\epsilon) L_G \preceq \sum_i s_i v_i v_i^T \preceq (1+\epsilon) L_G$
- $\mathrm{supp}(s) \le 4(n-1)/\epsilon^2$.

Deterministic,
"Twice-Ramanujan"



Proof of Theorem

[Part 2]



[Now]

Applications

Embedding Subsets of $\mathbf{L_1}$

$$x_1, \ldots, x_n \in \mathbf{R}^d \hookrightarrow y_1, \ldots, y_n \in \mathbf{R}^k$$

$$\|y_i - y_j\|_p \leq \|x_i - x_j\|_p \leq (1+\epsilon)\|y_i - y_j\|_p$$

How small can ***k*** be? Depends on ***p***.

Embedding Subsets of $\mathbf{L_1}$

$$x_1, \ldots, x_n \in \mathbf{R}^d \hookrightarrow y_1, \ldots, y_n \in \mathbf{R}^k$$

$$\|y_i - y_j\|_p \leq \|x_i - x_j\|_p \leq (1+\epsilon)\|y_i - y_j\|_p$$

How small can ***k*** be? Depends on ***p***.

p=2 [JL'84]: $k = O(\log n/\epsilon^2)$ tight.

Embedding Subsets of $\mathbf{L_1}$

$$x_1, \ldots, x_n \in \mathbf{R}^d \hookrightarrow y_1, \ldots, y_n \in \mathbf{R}^k$$

$$\|y_i - y_j\|_p \leq \|x_i - x_j\|_p \leq (1+\epsilon)\|y_i - y_j\|_p$$

How small can ***k*** be? Depends on ***p***.

p=2 [JL'84]: $k = O(\log n/\epsilon^2)$ tight.

p=1 [Sch'87] $k = O_\epsilon(n \log n)$

Embedding Subsets of $\mathbf{L_1}$

$$x_1, \ldots, x_n \in \mathbf{R}^d \hookrightarrow y_1, \ldots, y_n \in \mathbf{R}^k$$

$$\|y_i - y_j\|_p \leq \|x_i - x_j\|_p \leq (1+\epsilon)\|y_i - y_j\|_p$$

How small can ***k*** be? Depends on ***p***.

p=2 [JL'84]: $k = O(\log n/\epsilon^2)$ tight.

p=1 [Sch'87] $k = O_\epsilon(n \log n)$

[BC'04, LN'05, ASNN'11,R'11]: $k = \Omega(n^\alpha)$

[Newman-Rabinovich'10]: $k = O(n/\epsilon^2)$

$$\|x_i - x_j\|_1 = \sum_{S \subset [n]} \lambda_S |\mathbf{1}_S(i) - \mathbf{1}_S(j)|$$

2^n-dim embedding.

[Newman-Rabinovich'10]: $k = O(n/\epsilon^2)$

$$\|x_i - x_j\|_1 = \sum_{S \subset [n]} \lambda_S |\mathbf{1}_S(i) - \mathbf{1}_S(j)|$$

$$= \sum_{S \subset [n]} \langle e_i - e_j, \lambda_S^{1/2} \mathbf{1}_S \rangle^2$$

[Newman-Rabinovich'10]: $k = O(n/\epsilon^2)$

$$\|x_i - x_j\|_1 = \sum_{S \subset [n]} \lambda_S |\mathbf{1}_S(i) - \mathbf{1}_S(j)|$$

$$= \sum_{S \subset [n]} \langle e_i - e_j, \lambda_S^{1/2} \mathbf{1}_S \rangle^2$$

$$= (1 \pm \epsilon) \sum_{S \subset [n]} w_S \langle e_i - e_j, \lambda_S^{1/2} \mathbf{1}_S \rangle^2$$

$$\exists w_S \quad \forall i, j$$

[Newman-Rabinovich'10]: $k = O(n/\epsilon^2)$

$$\|x_i - x_j\|_1 = \sum_{S \subset [n]} \lambda_S |\mathbf{1}_S(i) - \mathbf{1}_S(j)|$$

$O(n/\epsilon^2)$ nonzero

$$= (1 \pm \epsilon) \sum_{S \subset [n]} w_S \langle e_i - e_j, \lambda_S^{1/2} \mathbf{1}_S \rangle^2$$

$\exists w_S \quad \forall i, j$

[Newman-Rabinovich'10]: $k = O(n/\epsilon^2)$

$$\begin{aligned}
\|x_i - x_j\|_1 &= \sum_{S \subset [n]} \lambda_S |\mathbf{1}_S(i) - \mathbf{1}_S(j)| \\
&= \sum_{S \subset [n]} \langle e_i - e_j, \lambda_S^{1/2} \mathbf{1}_S \rangle^2 \\
&= (1 \pm \epsilon) \sum_{S \subset [n]} w_S \langle e_i - e_j, \lambda_S^{1/2} \mathbf{1}_S \rangle^2 \\
&= (1 \pm \epsilon) \sum_{S : w_S \neq 0} w_S \lambda_S |\mathbf{1}_S(i) - \mathbf{1}_S(j)|.
\end{aligned}$$

[Newman-Rabinovich'10]: $k = O(n/\epsilon^2)$

$$\begin{aligned}
\|x_i - x_j\|_1 &= \sum_{S \subset [n]} \lambda_S |\mathbf{1}_S(i) - \mathbf{1}_S(j)| \\
&= \sum_{S \subset [n]} \langle e_i - e_j, \lambda_S^{1/2} \mathbf{1}_S \rangle^2 \\
&= (1 \pm \epsilon) \sum_{S \subset [n]} w_S \langle e_i - e_j, \lambda_S^{1/2} \mathbf{1}_S \rangle^2 \\
&= (1 \pm \epsilon) \sum_{S : w_S \neq 0} w_S \lambda_S |\mathbf{1}_S(i) - \mathbf{1}_S(j)|.
\end{aligned}$$

Take $y_i(S) = w_S \lambda_S \cdot x_i(S)$. □

Column Subset Selection [Boutsidis-Drineas-MagdonIsmail'11]

- **Input**: matrix **A** n x m with cols
- Best rank **k** approx: top k SVD:

$$\begin{aligned}\|A - P_k A\|_F^2 &= \mathrm{Tr}(P_k^{\perp} A A^T P_k^{\perp}) \\ &= \sum_{i \le m} \|P_k^{\perp} a_i\|^2 \\ &= \sum_{i > k} \sigma^2(A).\end{aligned}$$

- **Output**: subset **S** of **O(k)** cols of **A** with

$$\mathrm{Tr}(A_S^{\perp} A A^T A_S^{\perp}) \le 2\mathrm{Tr}(P_k^{\perp} A A^T P_k^{\perp})$$

Sketch [BDM+Schechtman'11]

- **Goal**: choose **S** which captures most of P_k
- **Idea:** Apply sparsification to

$$v_i = P_k a_i \oplus \|P_k^{\perp} a_i\| \in \mathrm{R}^{k+1}$$

$$\left[\sum_{i \le m} v_i v_i^T = P_k A A^T P_k \oplus \left(\sum_{i \le m} \|P_k^{\perp} a_i\|^2 \right) \right]$$

Sketch [BDM+Schechtman'11]

- **Goal**: choose **S** which captures most of P_k
- **Idea:** Apply sparsification to

$$v_i = P_k a_i \oplus \|P_k^{\perp} a_i\| \in \mathrm{R}^{k+1}$$

$$\left[\sum_{i \le m} v_i v_i^T = P_k A A^T P_k \oplus \left(\sum_{i \le m} \|P_k^{\perp} a_i\|^2 \right) \right]$$

Gives **O(k)** weights satisfying

$$\sum_{i \le m} s_i v_i v_i^T \simeq P_k A A^T P_k \oplus \mathrm{Tr}(P_k^{\perp} A A^T P_k^{\perp})$$

Sketch [BDM+Schechtman'11]

- **Goal**: choose **S** which captures most of P_k
- **Idea:** Apply sparsification to

$$v_i = P_k a_i \oplus \|P_k^{\perp} a_i\| \in \mathrm{R}^{k+1}$$

$$\left[\sum_{i \le m} v_i v_i^T \simeq P_k A A^T P_k \oplus \left(\sum_{i \le m} \|P_k^{\perp} a_i\|^2 \right) \right]$$

Behaves like SVD inside P_k

Gives **O(k)** weights satisfying

$$\sum_{i \le m} s_i v_i v_i^T \simeq P_k A A^T P_k \oplus \mathrm{Tr}(P_k^{\perp} A A^T P_k^{\perp})$$

Not much worse than SVD outside

Sketch

- Can show that taking support of **S** works

$$\mathrm{Tr}(A_S^{\perp} A A^T A_S^{\perp}) \leq 2\mathrm{Tr}(P_k^{\perp} A A^T P_k^{\perp})$$

- [BDM'11] Actually get 1+\epsilon approximation using other ideas related to volume sampling
- Improved by [Guruswami-Sinop'11] using volume sampling

Other Applications

- Embedding subspaces of ell_p

  [Schechtman'11]
- Sparsifying Unique Games and Min Uncut

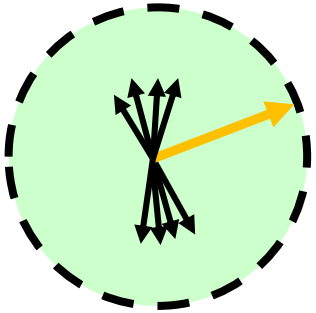
  [Spielman]
- Deterministic Matrix Completion

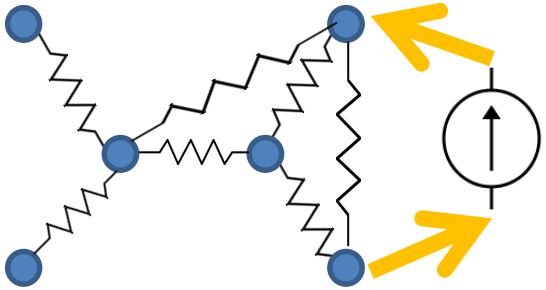
  [Heiman, Schechtman, Shraibman '11]
- Approximating Convex Sets by Polytopes

  [Barvinok'12]

Main Messages

1. It is useful to look at a graph as a circuit or a bunch of vectors





2. You can encode a lot of problems into quadratic forms of sums of rank one matrices

Next: Proof that sums can always be sparsified.