# Graphs in Machine Learning

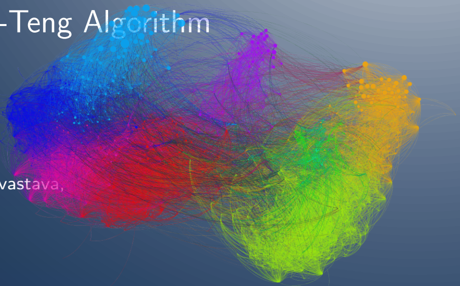
## Spectral Graph Sparsifiers: Theory

Effective Resistance and Spielman-Teng Algorithm

Michal Valko

*Inria & ENS Paris-Saclay, MVA*

Partially based on material by: Rob Fergus, Nikhil Srivastava,
Yiannis Koutis, Joshua Batson, Daniel Spielman



# Spectral Graph Sparsifiers

Rayleigh-Ritz gives:

# Spectral Graph Sparsifiers

Rayleigh-Ritz gives:

$$\lambda_{\min} = \min \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}} \quad \text{and} \quad \lambda_{\max} = \max \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}}$$

# Spectral Graph Sparsifiers

Rayleigh-Ritz gives:

$$\lambda_{\min} = \min \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}} \quad \text{and} \quad \lambda_{\max} = \max \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}}$$

What can we say about $\lambda_i(G)$ and $\lambda_i(H)$?

# Spectral Graph Sparsifiers

Rayleigh-Ritz gives:

$$\lambda_{\min} = \min \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}} \quad \text{and} \quad \lambda_{\max} = \max \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}}$$

What can we say about $\lambda_i(G)$ and $\lambda_i(H)$?

$$(1-\varepsilon)\mathbf{f}^\mathsf{T}\mathbf{L}_G\mathbf{f} \leq \mathbf{f}^\mathsf{T}\mathbf{L}_H\mathbf{f} \leq (1+\varepsilon)\mathbf{f}^\mathsf{T}\mathbf{L}_G\mathbf{f}$$

# Spectral Graph Sparsifiers

Rayleigh-Ritz gives:

$$\lambda_{\min} = \min \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}} \quad \text{and} \quad \lambda_{\max} = \max \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}}$$

What can we say about $\lambda_i(G)$ and $\lambda_i(H)$?

$$(1-\varepsilon)\mathbf{f}^\mathsf{T}\mathbf{L}_G\mathbf{f} \leq \mathbf{f}^\mathsf{T}\mathbf{L}_H\mathbf{f} \leq (1+\varepsilon)\mathbf{f}^\mathsf{T}\mathbf{L}_G\mathbf{f}$$

Eigenvalues are approximated well!

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

Using matrix ordering notation $(1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

# Spectral Graph Sparsifiers

Rayleigh-Ritz gives:

$$\lambda_{\min} = \min \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}} \quad \text{and} \quad \lambda_{\max} = \max \frac{\mathbf{x}^\mathsf{T}\mathbf{L}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}}$$

What can we say about $\lambda_i(G)$ and $\lambda_i(H)$?

$$(1-\varepsilon)\mathbf{f}^\mathsf{T}\mathbf{L}_G\mathbf{f} \le \mathbf{f}^\mathsf{T}\mathbf{L}_H\mathbf{f} \le (1+\varepsilon)\mathbf{f}^\mathsf{T}\mathbf{L}_G\mathbf{f}$$

Eigenvalues are approximated well!

$$(1-\varepsilon)\lambda_i(G) \le \lambda_i(H) \le (1+\varepsilon)\lambda_i(G)$$

Using matrix ordering notation $(1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

As a consequence, $\arg\min_{\mathbf{x}} \|\mathbf{L}_H\mathbf{x} - \mathbf{b}\| \approx \arg\min_{\mathbf{x}} \|\mathbf{L}_G\mathbf{x} - \mathbf{b}\|$

# Spectral Graph Sparsifiers in ML

**Proposition** (Kyng et al., 2017; Spielman and Srivastava, 2011)

There exists an algorithm that can construct a spectral $\varepsilon$-sparsifier

# Spectral Graph Sparsifiers in ML

**Proposition** (Kyng et al., 2017; Spielman and Srivastava, 2011)

There exists an algorithm that can construct a spectral $\varepsilon$-sparsifier

- with only $\mathcal{O}(n \log(n)/\varepsilon^2)$ edges

# Spectral Graph Sparsifiers in ML

**Proposition** (Kyng et al., 2017; Spielman and Srivastava, 2011)

There exists an algorithm that can construct a spectral $\varepsilon$-sparsifier

- with only $\mathcal{O}(n \log(n)/\varepsilon^2)$ edges
- in $\mathcal{O}(m \log^2(n))$ time and $\mathcal{O}(n \log(n)/\varepsilon^2)$ space

# Spectral Graph Sparsifiers in ML

**Proposition** (Kyng et al., 2017; Spielman and Srivastava, 2011)

There exists an algorithm that can construct a spectral $\varepsilon$-sparsifier

- with only $\mathcal{O}(n \log(n)/\varepsilon^2)$ edges
- in $\mathcal{O}(m \log^2(n))$ time and $\mathcal{O}(n \log(n)/\varepsilon^2)$ space
- a single pass over the data

# Spectral Graph Sparsifiers in ML

Laplacian smoothing (denoising): given $\mathbf{y} \triangleq \mathbf{f}^\star + \xi$ and $G$ compute

$$\min_{\mathbf{f}\in\Re^n} (\mathbf{f}-\mathbf{y})^\mathsf{T}(\mathbf{f}-\mathbf{y}) + \lambda \mathbf{f}^\mathsf{T}\mathbf{L}_G\mathbf{f} \tag{1}$$

# Spectral Graph Sparsifiers in ML

Laplacian smoothing (denoising): given $\mathbf{y} \triangleq \mathbf{f}^\star + \xi$ and $G$ compute

$$\min_{\mathbf{f} \in \Re^n} (\mathbf{f} - \mathbf{y})^\mathsf{T}(\mathbf{f} - \mathbf{y}) + \lambda \mathbf{f}^\mathsf{T} \mathbf{L}_G \mathbf{f} \tag{1}$$

| | Preproc | Time | Space |
|---|---|---|---|
| $\widehat{\mathbf{f}} = (\lambda \mathbf{L}_G + \mathbf{I})^{-1}\mathbf{y}$ | $0$ | $\mathcal{O}(m \log(n))$ | $\mathcal{O}(m)$ |

# Spectral Graph Sparsifiers in ML

Laplacian smoothing (denoising): given $\mathbf{y} \triangleq \mathbf{f}^\star + \xi$ and $G$ compute

$$\min_{\mathbf{f} \in \Re^n} (\mathbf{f} - \mathbf{y})^\mathsf{T}(\mathbf{f} - \mathbf{y}) + \lambda \mathbf{f}^\mathsf{T} \mathbf{L}_G \mathbf{f} \tag{1}$$

| | Preproc | Time | Space |
|---|---|---|---|
| $\widehat{\mathbf{f}} = (\lambda \mathbf{L}_G + \mathbf{I})^{-1}\mathbf{y}$ | 0 | $\mathcal{O}(m \log(n))$ | $\mathcal{O}(m)$ |
| $\widetilde{\mathbf{f}} = (\lambda \mathbf{L}_H + \mathbf{I})^{-1}\mathbf{y}$ | $\mathcal{O}(m \log^2(n))$ | $\mathcal{O}(n \log^2(n))$ | $\mathcal{O}(n \log(n))$ |

# Spectral Graph Sparsifiers in ML

Laplacian smoothing (denoising): given $\mathbf{y} \triangleq \mathbf{f}^\star + \xi$ and $G$ compute

$$\min_{\mathbf{f} \in \Re^n} (\mathbf{f} - \mathbf{y})^\mathsf{T} (\mathbf{f} - \mathbf{y}) + \lambda \mathbf{f}^\mathsf{T} \mathbf{L}_G \mathbf{f} \tag{1}$$

| | Preproc | Time | Space |
|---|---|---|---|
| $\widehat{\mathbf{f}} = (\lambda \mathbf{L}_G + \mathbf{I})^{-1} \mathbf{y}$ | 0 | $\mathcal{O}(m \log(n))$ | $\mathcal{O}(m)$ |
| $\widetilde{\mathbf{f}} = (\lambda \mathbf{L}_H + \mathbf{I})^{-1} \mathbf{y}$ | $\mathcal{O}(m \log^2(n))$ | $\mathcal{O}(n \log^2(n))$ | $\mathcal{O}(n \log(n))$ |

Large computational improvement
↳ accuracy guarantees! Sadhanala et al., 2016

# Spectral Graph Sparsifiers in ML

Laplacian smoothing (denoising): given $\mathbf{y} \triangleq \mathbf{f}^\star + \xi$ and $G$ compute

$$\min_{\mathbf{f}\in\Re^n} (\mathbf{f}-\mathbf{y})^\mathsf{T}(\mathbf{f}-\mathbf{y}) + \lambda \mathbf{f}^\mathsf{T}\mathbf{L}_G\mathbf{f} \tag{1}$$

| | Preproc | Time | Space |
|---|---|---|---|
| $\widehat{\mathbf{f}} = (\lambda\mathbf{L}_G + \mathbf{I})^{-1}\mathbf{y}$ | 0 | $\mathcal{O}(m\log(n))$ | $\mathcal{O}(m)$ |
| $\widetilde{\mathbf{f}} = (\lambda\mathbf{L}_H + \mathbf{I})^{-1}\mathbf{y}$ | $\mathcal{O}(m\log^2(n))$ | $\mathcal{O}(n\log^2(n))$ | $\mathcal{O}(n\log(n))$ |

Large computational improvement
↳ accuracy guarantees! Sadhanala et al., 2016

Need to approximate spectrum only up to regularization level $\lambda$

# Ridge Spectral Graph Sparsifiers in ML

## Definition

An $(\varepsilon, \gamma)$-sparsifier of $G$ is a reweighted subgraph $H$ s.t.

$$(1-\varepsilon)\mathbf{L}_G - \varepsilon\gamma\mathbf{I} \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G + \varepsilon\gamma\mathbf{I} \tag{2}$$

# Ridge Spectral Graph Sparsifiers in ML

**Definition**

An $(\varepsilon, \gamma)$-sparsifier of $G$ is a reweighted subgraph $H$ s.t.

$$(1 - \varepsilon)\mathbf{L}_G - \varepsilon\gamma\mathbf{I} \preceq \mathbf{L}_H \preceq (1 + \varepsilon)\mathbf{L}_G + \varepsilon\gamma\mathbf{I} \quad (2)$$

Mixed multiplicative/additive error

- large (i.e. $\geq \gamma$) directions reconstructed accurately
- small (i.e. $\leq \gamma$) directions uniformly approximated ($\gamma\mathbf{I}$)

# Ridge Spectral Graph Sparsifiers in ML

## Definition

An $(\varepsilon, \gamma)$-sparsifier of $G$ is a reweighted subgraph $H$ s.t.

$$(1 - \varepsilon)\mathbf{L}_G - \varepsilon\gamma\mathbf{I} \preceq \mathbf{L}_H \preceq (1 + \varepsilon)\mathbf{L}_G + \varepsilon\gamma\mathbf{I} \quad (2)$$

Mixed multiplicative/additive error

- large (i.e. $\geq \gamma$) directions reconstructed accurately
- small (i.e. $\leq \gamma$) directions uniformly approximated ($\gamma\mathbf{I}$)

Adapted from Randomized Linear Algebra (RLA) community
↳ PSD matrix low-rank approx. Alaoui and Mahoney, 2015

# Ridge Spectral Graph Sparsifiers in ML

**Definition**

An $(\varepsilon, \gamma)$-sparsifier of $G$ is a reweighted subgraph $H$ s.t.

$$(1-\varepsilon)\mathbf{L}_G - \varepsilon\gamma\mathbf{I} \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G + \varepsilon\gamma\mathbf{I} \tag{2}$$

Mixed multiplicative/additive error

- large (i.e. $\geq \gamma$) directions reconstructed accurately
- small (i.e. $\leq \gamma$) directions uniformly approximated ($\gamma\mathbf{I}$)

Adapted from Randomized Linear Algebra (RLA) community
↳ PSD matrix low-rank approx. Alaoui and Mahoney, 2015

RLA → Graph: Improve over $\mathcal{O}(n \log n)$ exploiting regularization

Graph → RLA: Exploit $\mathbf{L}_G$ structure for fast $(\varepsilon, \gamma)$-sparsification

# Spectral Graph Sparsification: Intuition

Let us consider unweighted graphs: $w_{ij} \in \{0, 1\}$

$$\mathbf{L}_G = \sum_{ij} w_{ij} \mathbf{L}_{ij} = \sum_{ij \in E} \mathbf{L}_{ij}$$

# Spectral Graph Sparsification: Intuition

Let us consider unweighted graphs: $w_{ij} \in \{0, 1\}$

$$\mathbf{L}_G = \sum_{ij} w_{ij}\mathbf{L}_{ij} = \sum_{ij \in E} \mathbf{L}_{ij} = \sum_{ij \in E} (\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)(\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)^\mathsf{T}$$

# Spectral Graph Sparsification: Intuition

Let us consider unweighted graphs: $w_{ij} \in \{0, 1\}$

$$\mathbf{L}_G = \sum_{ij} w_{ij} \mathbf{L}_{ij} = \sum_{ij \in E} \mathbf{L}_{ij} = \sum_{ij \in E} (\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)(\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)^\mathsf{T} = \sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T}$$

# Spectral Graph Sparsification: Intuition

Let us consider unweighted graphs: $w_{ij} \in \{0, 1\}$

$$\mathbf{L}_G = \sum_{ij} w_{ij}\mathbf{L}_{ij} = \sum_{ij \in E} \mathbf{L}_{ij} = \sum_{ij \in E} (\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)(\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)^\mathsf{T} = \sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T}$$

We look for a **subgraph** $H$

$$\mathbf{L}_H = \sum_{e \in E} s_e \mathbf{b}_e \mathbf{b}_e^\mathsf{T}$$

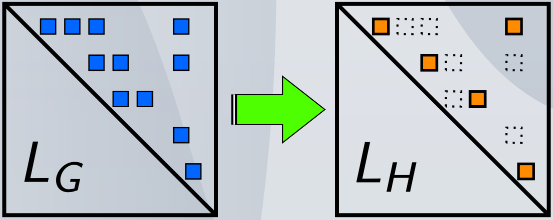
# Spectral Graph Sparsification: Intuition

Let us consider unweighted graphs: $w_{ij} \in \{0, 1\}$

$$\mathbf{L}_G = \sum_{ij} w_{ij}\mathbf{L}_{ij} = \sum_{ij \in E} \mathbf{L}_{ij} = \sum_{ij \in E} (\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)(\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)^\mathsf{T} = \sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T}$$

We look for a **subgraph** $H$

$$\mathbf{L}_H = \sum_{e \in E} s_e \mathbf{b}_e \mathbf{b}_e^\mathsf{T} \quad \text{where } s_e \text{ is a new weight of edge e}$$




https://math.berkeley.edu/~nikhil/

# Spectral Graph Sparsification: Intuition

We want $\quad (1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

# Spectral Graph Sparsification: Intuition

We want $\quad (1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

Equivalent, given $\mathbf{L}_G = \sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T}$

# Spectral Graph Sparsification: Intuition

We want $\quad (1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

Equivalent, given $\mathbf{L}_G = \sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{L}_G \preceq \sum_{e \in E} s_e \mathbf{b}_e \mathbf{b}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{L}_G$

# Spectral Graph Sparsification: Intuition

We want $\quad (1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

Equivalent, given $\mathbf{L}_G = \sum_{e\in E} \mathbf{b}_e\mathbf{b}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{L}_G \preceq \sum_{e\in E} s_e\mathbf{b}_e\mathbf{b}_e^\mathsf{T} \preceq \kappa\cdot\mathbf{L}_G$

Forget $\mathbf{L}$, given $\mathbf{A} = \sum_{e\in E} \mathbf{a}_e\mathbf{a}_e^\mathsf{T}$

# Spectral Graph Sparsification: Intuition

$$\text{We want} \quad (1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$$

$$\text{Equivalent, given } \mathbf{L}_G = \sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T} \text{ find } \mathbf{s}, \text{ s.t. } \mathbf{L}_G \preceq \sum_{e \in E} s_e \mathbf{b}_e \mathbf{b}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{L}_G$$

$$\text{Forget } \mathbf{L}, \text{ given } \mathbf{A} = \sum_{e \in E} \mathbf{a}_e \mathbf{a}_e^\mathsf{T} \text{ find } \mathbf{s}, \text{ s.t. } \mathbf{A} \preceq \sum_{e \in E} s_e \mathbf{a}_e \mathbf{a}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{A}$$

# Spectral Graph Sparsification: Intuition

We want $(1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

Equivalent, given $\mathbf{L}_G = \sum_{e\in E} \mathbf{b}_e\mathbf{b}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{L}_G \preceq \sum_{e\in E} s_e\mathbf{b}_e\mathbf{b}_e^\mathsf{T} \preceq \kappa\cdot\mathbf{L}_G$

Forget $\mathbf{L}$, given $\mathbf{A} = \sum_{e\in E} \mathbf{a}_e\mathbf{a}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{A} \preceq \sum_{e\in E} s_e\mathbf{a}_e\mathbf{a}_e^\mathsf{T} \preceq \kappa\cdot\mathbf{A}$

Same as, given $\mathbf{I} = \sum_{e\in E} \mathbf{v}_e\mathbf{v}_e^\mathsf{T}$

# Spectral Graph Sparsification: Intuition

We want $(1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

Equivalent, given $\mathbf{L}_G = \sum_{e\in E} \mathbf{b}_e\mathbf{b}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{L}_G \preceq \sum_{e\in E} s_e\mathbf{b}_e\mathbf{b}_e^\mathsf{T} \preceq \kappa{\cdot}\mathbf{L}_G$

Forget $\mathbf{L}$, given $\mathbf{A} = \sum_{e\in E} \mathbf{a}_e\mathbf{a}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{A} \preceq \sum_{e\in E} s_e\mathbf{a}_e\mathbf{a}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{A}$

Same as, given $\mathbf{I} = \sum_{e\in E} \mathbf{v}_e\mathbf{v}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{I} \preceq \sum_{e\in E} s_e\mathbf{v}_e\mathbf{v}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{I}$

# Spectral Graph Sparsification: Intuition

We want $(1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

Equivalent, given $\mathbf{L}_G = \sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{L}_G \preceq \sum_{e \in E} s_e \mathbf{b}_e \mathbf{b}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{L}_G$

Forget $\mathbf{L}$, given $\mathbf{A} = \sum_{e \in E} \mathbf{a}_e \mathbf{a}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{A} \preceq \sum_{e \in E} s_e \mathbf{a}_e \mathbf{a}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{A}$

Same as, given $\mathbf{I} = \sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{I} \preceq \sum_{e \in E} s_e \mathbf{v}_e \mathbf{v}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{I}$

How to get it?

# Spectral Graph Sparsification: Intuition

We want $(1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

Equivalent, given $\mathbf{L}_G = \sum_{e\in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{L}_G \preceq \sum_{e\in E} s_e \mathbf{b}_e \mathbf{b}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{L}_G$

Forget $\mathbf{L}$, given $\mathbf{A} = \sum_{e\in E} \mathbf{a}_e \mathbf{a}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{A} \preceq \sum_{e\in E} s_e \mathbf{a}_e \mathbf{a}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{A}$

Same as, given $\mathbf{I} = \sum_{e\in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{I} \preceq \sum_{e\in E} s_e \mathbf{v}_e \mathbf{v}_e^\mathsf{T} \preceq \kappa \cdot \mathbf{I}$

How to get it? $\mathbf{v}_e \leftarrow \mathbf{A}^{-1/2}\mathbf{a}_e$

# Spectral Graph Sparsification: Intuition

We want $(1-\varepsilon)\mathbf{L}_G \preceq \mathbf{L}_H \preceq (1+\varepsilon)\mathbf{L}_G$

Equivalent, given $\mathbf{L}_G = \sum_{e\in E} \mathbf{b}_e\mathbf{b}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{L}_G \preceq \sum_{e\in E} s_e\mathbf{b}_e\mathbf{b}_e^\mathsf{T} \preceq \kappa\cdot\mathbf{L}_G$

Forget $\mathbf{L}$, given $\mathbf{A} = \sum_{e\in E} \mathbf{a}_e\mathbf{a}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{A} \preceq \sum_{e\in E} s_e\mathbf{a}_e\mathbf{a}_e^\mathsf{T} \preceq \kappa\cdot\mathbf{A}$

Same as, given $\mathbf{I} = \sum_{e\in E} \mathbf{v}_e\mathbf{v}_e^\mathsf{T}$ find $\mathbf{s}$, s.t. $\mathbf{I} \preceq \sum_{e\in E} s_e\mathbf{v}_e\mathbf{v}_e^\mathsf{T} \preceq \kappa\cdot\mathbf{I}$

How to get it? $\mathbf{v}_e \leftarrow \mathbf{A}^{-1/2}\mathbf{a}_e$

Then $\sum_{e\in E} s_e\mathbf{v}_e\mathbf{v}_e^\mathsf{T} \approx \mathbf{I} \iff \sum_{e\in E} s_e\mathbf{a}_e\mathbf{a}_e^\mathsf{T} \approx \mathbf{A}$

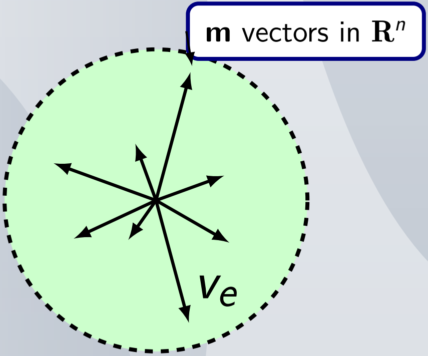
multiplying by $\mathbf{A}^{1/2}$ on both sides

# Spectral Graph Sparsification: Intuition

How does $\sum_{e\in E} \mathbf{v}_e\mathbf{v}_e^\mathsf{T} = \mathbf{I}$ look like geometrically?

# Spectral Graph Sparsification: Intuition

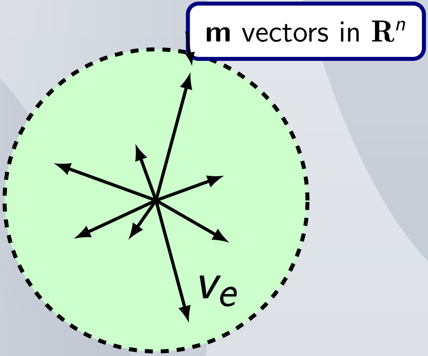
How does $\sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T} = \mathbf{I}$ look like geometrically?




Decomposition of identity: $\forall \mathbf{u}$ (unit vector): $\sum_{e \in E} (\mathbf{u}^\mathsf{T} \mathbf{v}_e)^2 = 1$

# Spectral Graph Sparsification: Intuition

How does $\sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T} = \mathbf{I}$ look like geometrically?




Decomposition of identity: $\forall \mathbf{u}$ (unit vector): $\sum_{e \in E} (\mathbf{u}^\mathsf{T} \mathbf{v}_e)^2 = 1$

moment ellipse is a sphere

https://math.berkeley.edu/~nikhil/

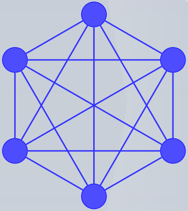
# Spectral Graph Sparsification: Intuition

Example: What happens with $K_n$?

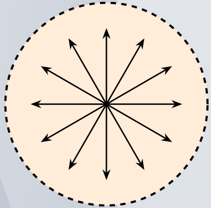
# Spectral Graph Sparsification: Intuition

Example: What happens with $K_n$?

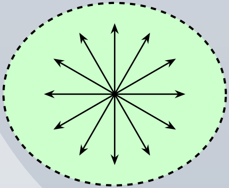
$K_n$ graph



$\sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T} = \mathbf{L}_G$



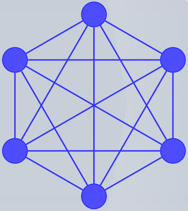
$\sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T} = \mathbf{I}$



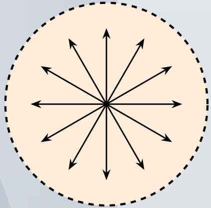
# Spectral Graph Sparsification: Intuition

Example: What happens with $K_n$?

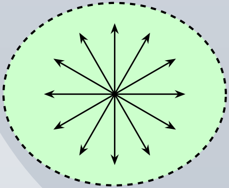
$K_n$ graph



$\sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T} = \mathbf{L}_G$



$\sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T} = \mathbf{I}$

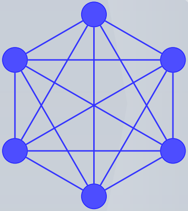


It is already isotropic! (looks like a sphere)

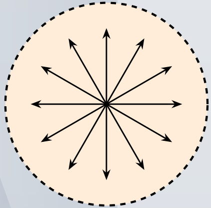
# Spectral Graph Sparsification: Intuition

Example: What happens with $K_n$?

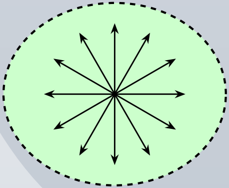
$K_n$ graph



$\sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T} = \mathbf{L}_G$



$\sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T} = \mathbf{I}$



It is already isotropic! (looks like a sphere)

rescaling $\mathbf{v}_e = \mathbf{L}^{-1/2}\mathbf{b}_e$ does not change the shape

https://math.berkeley.edu/~nikhil/

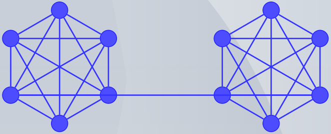
# Spectral Graph Sparsification: Intuition

Example: What happens with a dumbbell?

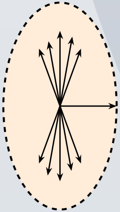
# Spectral Graph Sparsification: Intuition

**Example: What happens with a dumbbell?**

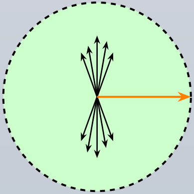
Dumbbell



$\sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T} = \mathbf{L}_G$



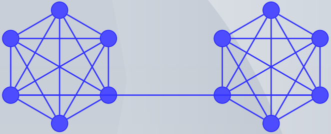
$\sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T} = \mathbf{I}$



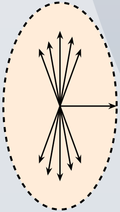
# Spectral Graph Sparsification: Intuition

Example: What happens with a dumbbell?

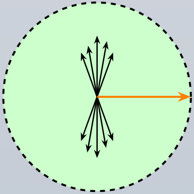
Dumbbell



$\sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T} = \mathbf{L}_G$



$\sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T} = \mathbf{I}$

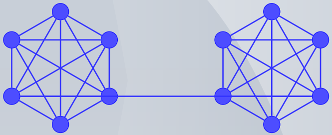


The vector corresponding to the link gets stretched!

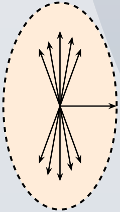
# Spectral Graph Sparsification: Intuition

Example: What happens with a dumbbell?

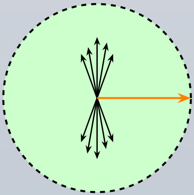
Dumbbell



$\sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^{\mathsf{T}} = \mathbf{L}_G$



$\sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^{\mathsf{T}} = \mathbf{I}$



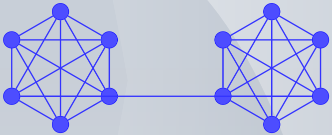
The vector corresponding to the link gets stretched!

because this transformation makes all the directions important

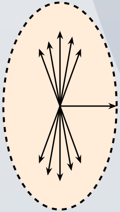
# Spectral Graph Sparsification: Intuition

Example: What happens with a dumbbell?

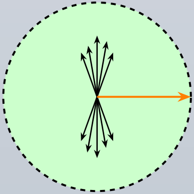
Dumbbell



$\sum_{e \in E} \mathbf{b}_e \mathbf{b}_e^\mathsf{T} = \mathbf{L}_G$



$\sum_{e \in E} \mathbf{v}_e \mathbf{v}_e^\mathsf{T} = \mathbf{I}$



The vector corresponding to the link gets stretched!

because this transformation makes all the directions important

rescaling reveals the vectors that are critical

https://math.berkeley.edu/~nikhil/

# References I


- Alaoui, A. E., & Mahoney, M. W. (2015). Fast randomized kernel methods with statistical guarantees. *Neural Information Processing Systems*.
- Kyng, R., Pachocki, J., Peng, R., & Sachdeva, S. (2017). A framework for analyzing resparsification algorithms. *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2032–2043.
- Sadhanala, V., Wang, Y.-X., & Tibshirani, R. J. (2016). Graph sparsification approaches for Laplacian smoothing. *International Conference on Artificial Intelligence and Statistics*, 1250–1259.
- Spielman, D. A., & Srivastava, N. (2011). Graph sparsification by effective resistances. *Journal on Computing*, *40*(6).

# Michal Valko

michal.valko@inria.fr

Inria & ENS Paris-Saclay, MVA


https://misovalko.github.io/mva-ml-graphs.html