A Cheeger-Type Inequality on Simplicial Complexes

Scientific and Statistical Computing Seminar – U Chicago

Sayan Mukherjee

Departments of Statistical Science, Computer Science, Mathematics
Institute for Genome Sciences & Policy, Duke University
Joint work with:
J. Steenbergen (Duke), Carly Klivans (Brown)

November 9, 2012

Dimension reduction algorithm

- Examples include:
  1. Isomap - 2000
  2. Locally Linear Embedding (LLE) - 2000
  3. Hessian LLE - 2003
  4. Laplacian Eigenmaps - 2003
  5. Diffusion Maps - 2004
- Laplacian Eigenmaps is based directly on the graph Laplacian.

Laplacian eigenmaps

Given data points $x_1, x_2, \dots, x_n \in \mathbb{R}^p$ which we wish to map into $\mathbb{R}^k$, $k \ll p$,

1. Construct a graph (association matrix) A out of the data points.

Laplacian eigenmaps

Given data points $x_1, x_2, \ldots, x_n \in \mathbb{R}^p$ which we wish to map into $\mathbb{R}^k$, $k \ll p$,

1. Construct a graph (association matrix) A out of the data points.
2. Compute $L = D - A$ where $D_{ii} = \sum_j A_{ij}$

Laplacian eigenmaps

Given data points $x_1, x_2, \ldots, x_n \in \mathbb{R}^p$ which we wish to map into $\mathbb{R}^k$, $k \ll p$,

1. Construct a graph (association matrix) A out of the data points.
2. Compute $L = D - A$ where $D_{ii} = \sum_j A_{ij}$
3. Compute eigenvalues and eigenvectors of L

$$0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n, \quad f_0 = \mathbf{1}, ..., f_n$$

Laplacian eigenmaps

Given data points $x_1, x_2, \ldots, x_n \in \mathbb{R}^p$ which we wish to map into $\mathbb{R}^k$, $k \ll p$,

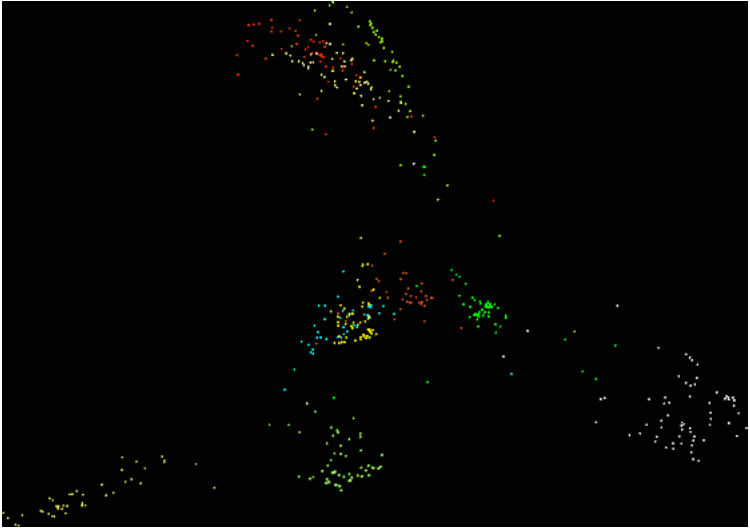
1. Construct a graph (association matrix) A out of the data points.
2. Compute $L = D - A$ where $D_{ii} = \sum_j A_{ij}$
3. Compute eigenvalues and eigenvectors of L

$$0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n, \quad f_0 = \mathbf{1}, \ldots, f_n$$

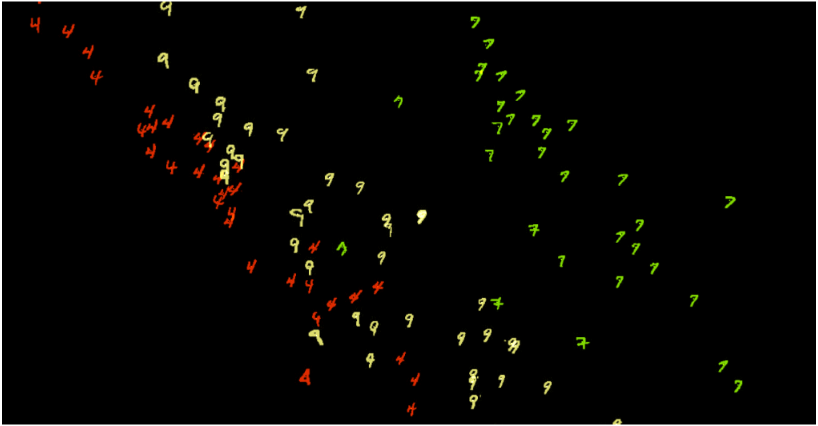
4. Map the data points into $\mathbb{R}^k$ by the map

$$x_i \mapsto (f_1(x_i), f_2(i), \ldots, f_k(x_i))$$

Laplacian eigenmaps example

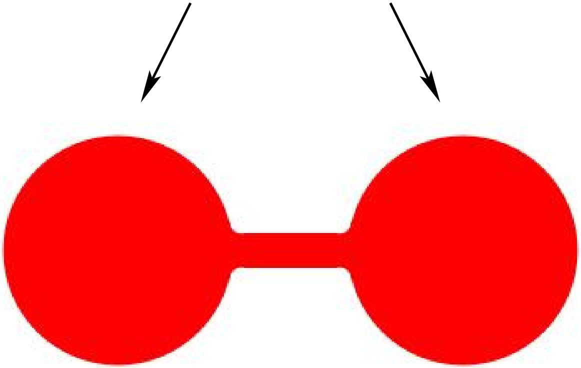


Laplacian eigenmaps example

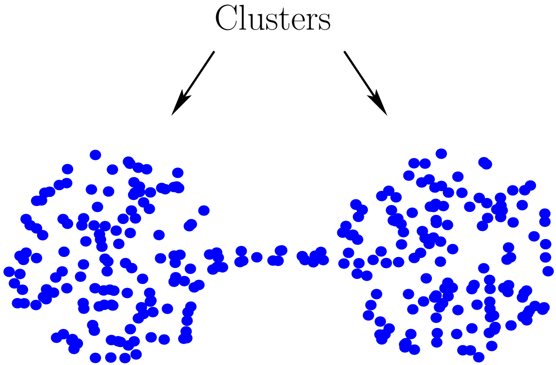


Near 0 homology

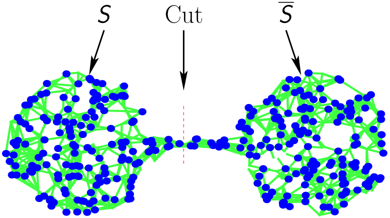
Near-Connected Components?



Clustering



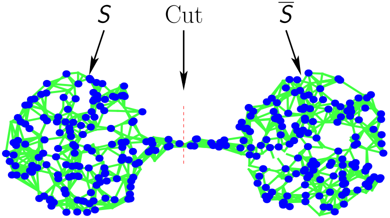
Near 0-homology for graphs



For a graph with vertex set V, the Cheeger number is defined to be

$$h = \min_{\emptyset \subsetneq S \subsetneq V} \frac{|\delta S|}{\min\left\{|S|, |\overline{S}|\right\}}.$$

Near 0-homology for graphs

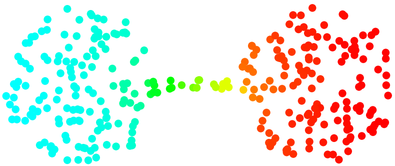


For a graph with vertex set V, the Cheeger number is defined to be

$$h = \min_{\emptyset \subsetneq S \subsetneq V} \frac{|\delta S|}{\min\left\{|S|, |\overline{S}|\right\}}.$$

$h = 0 \Leftrightarrow$ graph is disconnected.

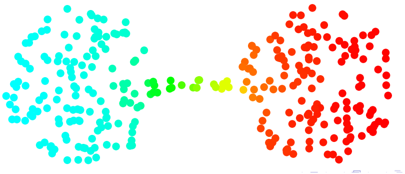
Near 0-homology for graphs



For a graph with vertex set V, the Fiedler number is defined to be

$$\lambda = \min_{\substack{f:V\to\mathbb{R}\\ f\perp 1}} \frac{\sum_{u\sim v}(f(u)-f(v))^2}{\sum_u (f(u))^2}.$$

Near 0-homology for graphs



For a graph with vertex set V, the Fiedler number is defined to be

$$\lambda = \min_{\substack{f:V\to\mathbb{R}\\ f\perp 1}} \frac{\sum_{u\sim v}(f(u)-f(v))^2}{\sum_u (f(u))^2}.$$

$\lambda = 0 \Leftrightarrow$ graph is disconnected.

Cheeger inequality for graphs

Theorem (Alon, Milman, Lawler & Sokal, Frieze & Kannan & Polson,...)

For any graph with Cheeger number h and Fiedler number λ

$$2h \geq \lambda_1 > \frac{h^2}{2M},$$

$M = \max_u u_d$, *maximum vertex degree.*

Edge expansion

Expander graphs are families of graphs that are sparse and strongly connected.

Edge expansion

Expander graphs are families of graphs that are sparse and strongly connected.

A family of expander graphs $\mathcal{G}$ has the property $h(G) > \epsilon > 0$ for all $G \in \mathcal{G}$.

Edge expansion

Expander graphs are families of graphs that are sparse and strongly connected.

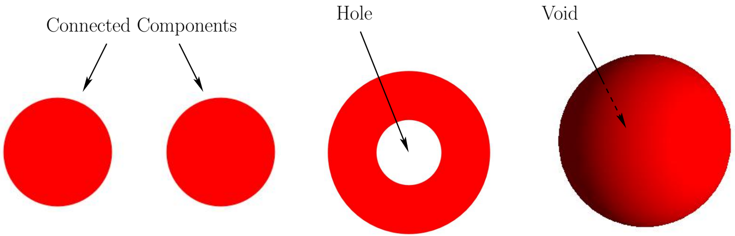
A family of expander graphs $\mathcal{G}$ has the property $h(G) > \epsilon > 0$ for all $G \in \mathcal{G}$.

Cheeger inequality lets us use λ as criteria $\lambda(G) > 0$ for all $G \in \mathcal{G}$.

Higher-dimensional notions

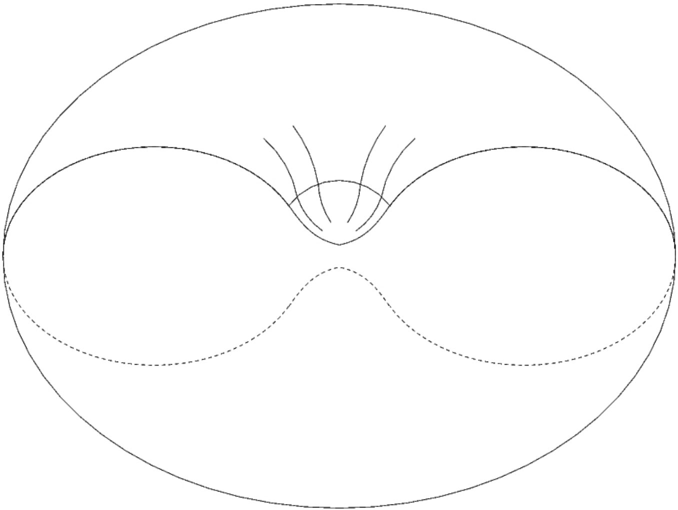
Homology

0-Homology | 1-Homology | 2-Homology

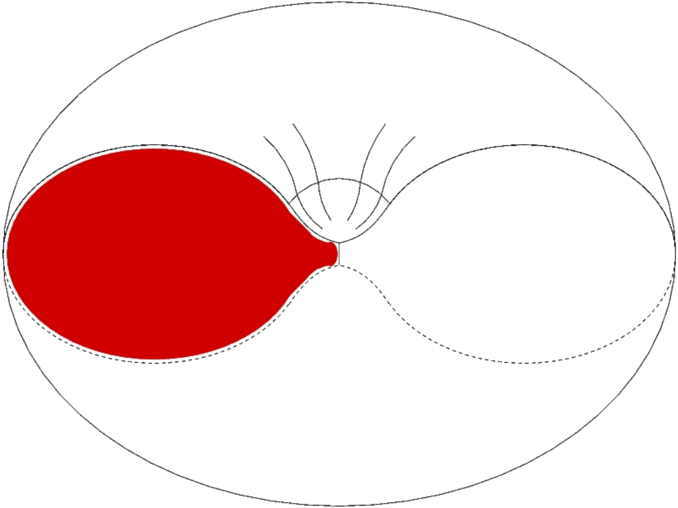


$\beta_0 = 2,\ \beta_1 = 0,\ \beta_2 = 0$ | $\beta_0 = 1,\ \beta_1 = 1,\ \beta_2 = 0$ | $\beta_0 = 1,\ \beta_1 = 0,\ \beta_2 = 1$

Near one homology

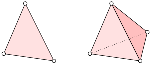


Near one homology



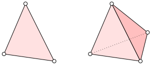
Simplicial complexes

A k-simplex is the convex hull of $k+1$ affinely independent points, $\sigma = \{u_0, ..., u_k\}$.



Simplicial complexes

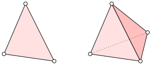
A k-simplex is the convex hull of $k+1$ affinely independent points, $\sigma = \{u_0, ..., u_k\}$.



A face τ of σ is the convex hull of a non-empty subset of the u_i, $\tau \leq \sigma$.

Simplicial complexes

A k-simplex is the convex hull of $k+1$ affinely independent points, $\sigma = \{u_0, ..., u_k\}$.



A face τ of σ is the convex hull of a non-empty subset of the u_i, $\tau \leq \sigma$.

A simplicial complex is a finite collection of simplices K such that $\sigma \in K$ and $\tau \leq \sigma$ implies $\tau \in K$, and $\sigma, \sigma_0 \in K$ implies $\sigma \cap \sigma_0$ is either empty or a face of both.

Chains and cochains

- X = simplicial complex of dimension m, $(X) = m$.
- $C_k(\mathbb{F}) = \{\mathbb{F}$-linear combinations of oriented k-simplices$\}$
- $C^k(\mathbb{F}) = \{\mathbb{F}$-valued functions on oriented k-simplices$\}$

Chains and cochains

- X = simplicial complex of dimension m, $(X) = m$.
- $C_k(\mathbb{F}) = \{\mathbb{F}\text{-linear combinations of oriented } k\text{-simplices}\}$
- $C^k(\mathbb{F}) = \{\mathbb{F}\text{-valued functions on oriented } k\text{-simplices}\}$
- Chain Complex:

$$0 \longleftarrow C_0 \xleftarrow{\partial_1(\mathbb{F})} C_1 \xleftarrow{\partial_2(\mathbb{F})} \cdots \xleftarrow{\partial_m(\mathbb{F})} C_m \longleftarrow 0$$

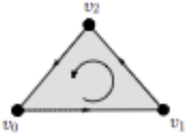
- Cochain Complex:

$$0 \longrightarrow C^0 \xrightarrow{\delta^0(\mathbb{F})} C^1 \xrightarrow{\delta^1(\mathbb{F})} \cdots \xrightarrow{\delta^{m-1}(\mathbb{F})} C^m \longrightarrow 0$$

- $\mathbb{F} = \mathbb{R}, \mathbb{Z}_2$

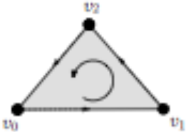
Chains and cochains

Given a simplicial complex $\sigma = \{v_0, ..., v_k\}$ an orientation of $[v_0, ..., v_k]$ is the equivalence class of even permutations of ordering.



Chains and cochains

Given a simplicial complex $\sigma = \{v_0, ..., v_k\}$ an orientation of $[v_0, ..., v_k]$ is the equivalence class of even permutations of ordering.

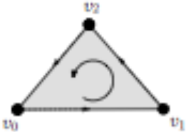


Boundary map $\partial_k(\mathbb{F}) : C_k(\mathbb{F}) \to C_{k-1}(\mathbb{F})$

$$\partial_k[v_0, ..., v_k] = \sum_{i=1}^{k}(-1)^i[v_0, ..., v_{i-1}, v_{i+1}, ..., v_k].$$

Chains and cochains

Given a simplicial complex $\sigma = \{v_0, ..., v_k\}$ an orientation of $[v_0, ..., v_k]$ is the equivalence class of even permutations of ordering.

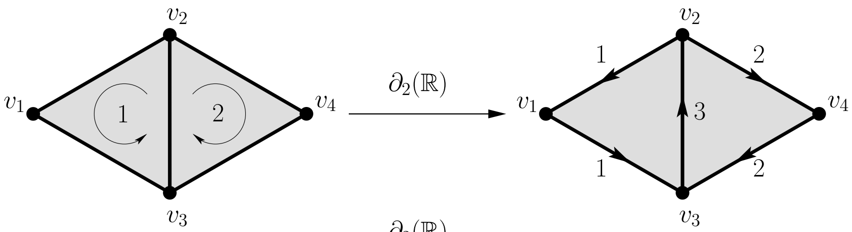


Boundary map $\partial_k(\mathbb{F}) : C_k(\mathbb{F}) \to C_{k-1}(\mathbb{F})$

$$\partial_k[v_0, ..., v_k] = \sum_{i=1}^{k}(-1)^i[v_0, ..., v_{i-1}, v_{i+1}, ..., v_k].$$

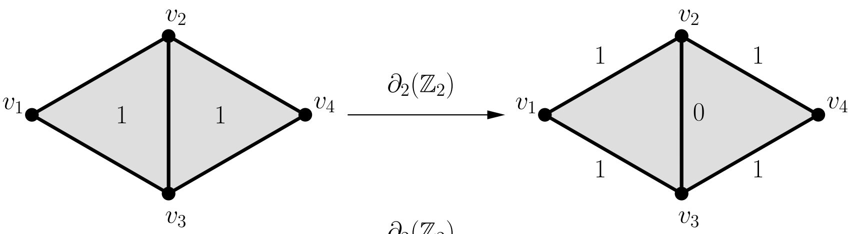
Coboundary map $\delta^{k-1}(\mathbb{F}) : C^{k-1}(\mathbb{F}) \to C^k(\mathbb{F})$ is the transpose of the boundary map.

$\partial(\mathbb{R})$



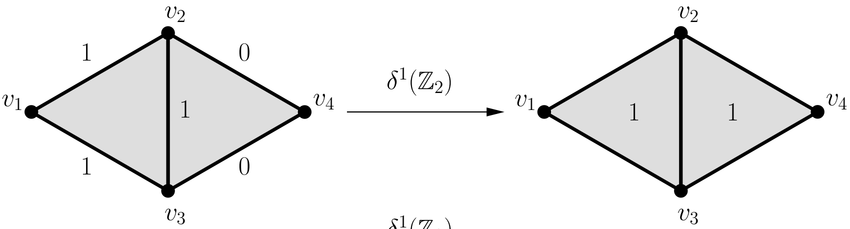
$$[v_1, v_3, v_2] + 2[v_2, v_4, v_3] \xrightarrow{\partial_2(\mathbb{R})} [v_1, v_3] + [v_2, v_1] + 3[v_3, v_2] + 2[v_2, v_4] + 2[v_4, v_3]$$

$\partial(\mathbb{Z}_2)$



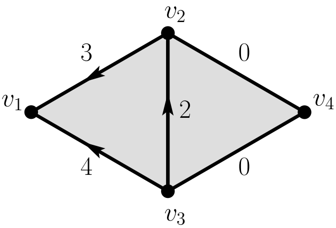
$$[v_1, v_2, v_3] + [v_2, v_3, v_4] \xrightarrow{\partial_2(\mathbb{Z}_2)} [v_1, v_2] + [v_1, v_3] + [v_2, v_4] + [v_3, v_4]$$

$\delta(\mathbb{Z}_2)$

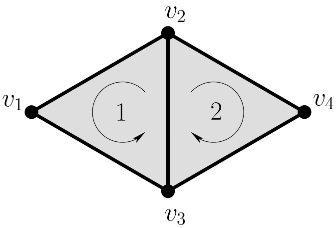


$$[v_1, v_2] + [v_1, v_3] + [v_2, v_3] \xrightarrow{\delta^1(\mathbb{Z}_2)} [v_1, v_2, v_3] + [v_2, v_3, v_4]$$

$\delta(\mathbb{R})$



$\xrightarrow{\delta^1(\mathbb{R})}$



$$3[v_2, v_1] + 4[v_3, v_1] + 2[v_3, v_2] \xrightarrow{\delta^1(\mathbb{R})} [v_1, v_3, v_2] + 2[v_2, v_4, v_3]$$

Cheeger numbers

$\mathbb{Z}_2$ homology and cohomology

$$H_k(\mathbb{Z}_2) = \frac{\ker\partial_k}{\mathrm{im}\partial_{k+1}}, \quad H^k(\mathbb{Z}_2) = \frac{\ker\delta^k}{\mathrm{im}\delta^{k+1}}.$$

Cheeger numbers

$\mathbb{Z}_2$ homology and cohomology

$$H_k(\mathbb{Z}_2) = \frac{\mathrm{ker}\partial_k}{\mathrm{im}\partial_{k+1}}, \quad H^k(\mathbb{Z}_2) = \frac{\mathrm{ker}\delta^k}{\mathrm{im}\delta^{k+1}}.$$

Coboundary and Boundary Cheeger numbers:

$$h^k \quad := \min_{\phi \in C^k(\mathbb{Z}_2) \backslash \mathrm{im}\delta} \left[\frac{|\delta\phi|}{\min_{\delta\phi=\delta\psi} |\psi|} \right] \frac{\leftarrow \text{distance from } \mathrm{ker}\delta^k}{\leftarrow \text{distance from } \mathrm{im}\delta^{k-1}}$$

Cheeger numbers

$\mathbb{Z}_2$ homology and cohomology

$$H_k(\mathbb{Z}_2) = \frac{\ker \partial_k}{\mathrm{im}\, \partial_{k+1}}, \quad H^k(\mathbb{Z}_2) = \frac{\ker \delta^k}{\mathrm{im}\, \delta^{k+1}}.$$

Coboundary and Boundary Cheeger numbers:

$$h^k := \min_{\phi \in C^k(\mathbb{Z}_2) \setminus \mathrm{im}\,\delta} \left[\frac{|\delta\phi|}{\min_{\delta\phi=\delta\psi} |\psi|} \right] \frac{\leftarrow \text{distance from } \ker\delta^k}{\leftarrow \text{distance from } \mathrm{im}\,\delta^{k-1}}$$

$$h_k := \min_{\phi \in C_k(\mathbb{Z}_2) \setminus \mathrm{im}\,\partial} \left[\frac{|\partial\phi|}{\min_{\partial\phi=\partial\psi} |\psi|} \right] \frac{\leftarrow \text{distance from } \ker\partial^k}{\leftarrow \text{distance from } \mathrm{im}\,\partial^{k+1}}.$$

Cheeger numbers

$\mathbb{Z}_2$ homology and cohomology

$$H_k(\mathbb{Z}_2) = \frac{\ker \partial_k}{\operatorname{im} \partial_{k+1}}, \quad H^k(\mathbb{Z}_2) = \frac{\ker \delta^k}{\operatorname{im} \delta^{k+1}}.$$

Coboundary and Boundary Cheeger numbers:

$$h^k := \min_{\phi \in C^k(\mathbb{Z}_2) \setminus \operatorname{im} \delta} \left[\frac{|\delta \phi|}{\min_{\delta\phi = \delta\psi} |\psi|} \right] \frac{\leftarrow \text{distance from } \ker \delta^k}{\leftarrow \text{distance from } \operatorname{im} \delta^{k-1}}$$

$$h_k := \min_{\phi \in C_k(\mathbb{Z}_2) \setminus \operatorname{im} \partial} \left[\frac{|\partial \phi|}{\min_{\partial\phi = \partial\psi} |\psi|} \right] \frac{\leftarrow \text{distance from } \ker \partial^k}{\leftarrow \text{distance from } \operatorname{im} \partial^{k+1}}.$$

$$H^k(\mathbb{Z}_2) \neq 0 \Leftrightarrow h^k = 0 \quad \text{and} \quad H_k(\mathbb{Z}_2) \neq 0 \Leftrightarrow h_k = 0.$$

h^0 is the Cheeger number for graphs. h^k defined by Dotterer and Kahle.

Fiedler numbers

$\mathbb{R}$ homology and cohomology

$$H_k(\mathbb{R}) = \frac{\ker\partial_k}{\mathrm{im}\partial_{k+1}}, \quad H^k(\mathbb{R}) = \frac{\ker\delta^k}{\mathrm{im}\delta^{k+1}}.$$

Fiedler numbers

$\mathbb{R}$ homology and cohomology

$$H_k(\mathbb{R}) = \frac{\ker \partial_k}{\operatorname{im} \partial_{k+1}}, \quad H^k(\mathbb{R}) = \frac{\ker \delta^k}{\operatorname{im} \delta^{k+1}}.$$

Coboundary and Boundary Fiedler numbers:

$$\lambda^k \quad := \quad \min_{f \in C^k(\mathbb{R}) \setminus \operatorname{im} \delta} \left[\frac{\|\delta f\|_2^2}{\min_{\delta f = \delta g} \|g\|_2^2} \right] \; \frac{\leftarrow \text{distance from } \ker \delta^k}{\leftarrow \text{distance from } \operatorname{im} \delta^{k-1}}$$

Fiedler numbers

$\mathbb{R}$ homology and cohomology

$$H_k(\mathbb{R}) = \frac{\ker \partial_k}{\operatorname{im} \partial_{k+1}}, \quad H^k(\mathbb{R}) = \frac{\ker \delta^k}{\operatorname{im} \delta^{k+1}}.$$

Coboundary and Boundary Fiedler numbers:

$$\lambda^k \quad := \quad \min_{f \in C^k(\mathbb{R}) \setminus \operatorname{im} \delta} \left[\frac{\|\delta f\|_2^2}{\min_{\delta f = \delta g} \|g\|_2^2} \right] \frac{\leftarrow \text{distance from } \ker \delta^k}{\leftarrow \text{distance from } \operatorname{im} \delta^{k-1}}$$

$$\lambda_k \quad := \quad \min_{fi \in C_k(\mathbb{R}) \setminus \operatorname{im} \partial} \left[\frac{\|\partial f\|_2^2}{\min_{\partial f = \partial g} \|g\|_2^2} \right] \frac{\leftarrow \text{distance from } \ker \partial^k}{\leftarrow \text{distance from } \operatorname{im} \partial^{k+1}}.$$

Fiedler numbers

$\mathbb{R}$ homology and cohomology

$$H_k(\mathbb{R}) = \frac{\ker \partial_k}{\mathrm{im}\,\partial_{k+1}}, \quad H^k(\mathbb{R}) = \frac{\ker \delta^k}{\mathrm{im}\,\delta^{k+1}}.$$

Coboundary and Boundary Fiedler numbers:

$$\lambda^k \quad := \quad \min_{f \in C^k(\mathbb{R}) \setminus \mathrm{im}\,\delta} \left[\frac{\|\delta f\|_2^2}{\min_{\delta f = \delta g} \|g\|_2^2} \right] \; \frac{\leftarrow \text{distance from } \ker \delta^k}{\leftarrow \text{distance from } \mathrm{im}\,\delta^{k-1}}$$

$$\lambda_k \quad := \quad \min_{fi \in C_k(\mathbb{R}) \setminus \mathrm{im}\,\partial} \left[\frac{\|\partial f\|_2^2}{\min_{\partial f = \partial g} \|g\|_2^2} \right] \; \frac{\leftarrow \text{distance from } \ker \partial^k}{\leftarrow \text{distance from } \mathrm{im}\,\partial^{k+1}}.$$

$$H^k(\mathbb{R}) \neq 0 \Leftrightarrow \lambda^k = 0 \quad \text{and} \quad H_k(\mathbb{R}) \neq 0 \Leftrightarrow \lambda_k = 0.$$

λ^0 is the Fiedler number for graphs.

Combinatorial Laplacian

The k-th Laplacian of X

$$L_k := \partial_{k+1}\delta^k + \delta^{k-1}\partial_k$$

Combinatorial Laplacian

The k-th Laplacian of X

$$L_k := \partial_{k+1}\delta^k + \delta^{k-1}\partial_k$$

Eigenvalues of L_k measure near (co)homology

Theorem (Eckmann)

$$C^k = im(\partial_{k+1}) \bigoplus im(\delta^{k-1}) \bigoplus ker(L_k),$$

and

$$H_k(\mathbb{R}) \cong H^k(\mathbb{R}) \cong ker\, L_k.$$

Combinatorial Laplacian

$$C^1 = \mathrm{im}(\delta^0) \bigoplus \ker(L_1) \bigoplus \mathrm{im}(\partial_1).$$

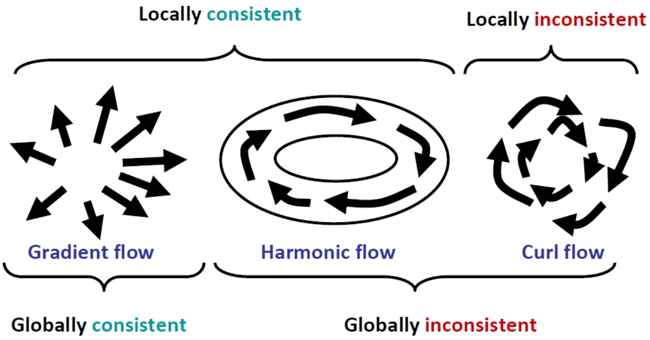
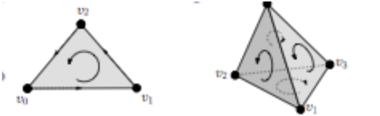


Figure: Courtesy of Pablo Parrilo

Orientability

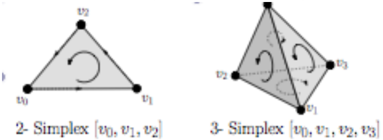
Two m-simplexes are lower adjacent if they share a common face



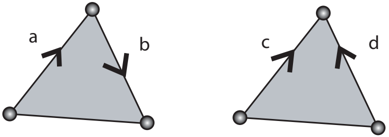
2- Simplex $[v_0, v_1, v_2]$ 3- Simplex $[v_0, v_1, v_2, v_3]$

Orientability

Two *m*-simplexes are lower adjacent if they share a common face

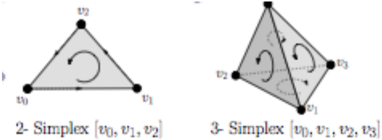


Two oriented lower adjacent *k*-simplexes τ and σ are dissimilar on a face ν if $\partial\tau$ and $\partial\sigma$ assign the same coefficient to ν.

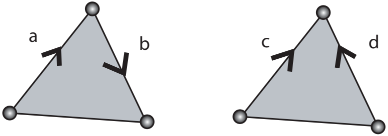


Orientability

Two *m*-simplexes are lower adjacent if they share a common face



Two oriented lower adjacent *k*-simplexes τ and σ are dissimilar on a face ν if $\partial\tau$ and $\partial\sigma$ assign the same coefficient to ν.



If X is a simplicial m-complex and all its *m*-simplices can be oriented similarly, then X is called orientable.

Pseudomanifold

A pseudomanifold is a combinatorial realization of a manifold with singularities.

Pseudomanifold

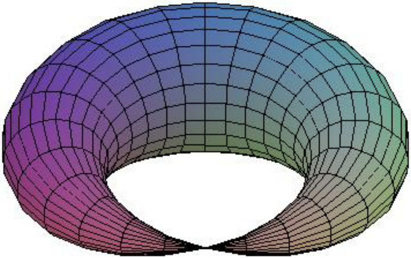
A pseudomanifold is a combinatorial realization of a manifold with singularities.
A topological space $\mathbb{X}$ endowed with a triangulation K is an m-dimensional pseudomanifold if the following conditions hold:

(1) $\mathbb{X} = |K|$ is the union of all n-simplices.

(2) Every $(m1)$-simplex is a face of exactly two m-simplices for $m > 1$.

(3) $\mathbb{X}$ is a strongly connected, there is a path between any pair of m-simplices in K.

Pseudomanifold



Positive result: chain complex

Proposition (Steenbergen & Klivans & M)

If X is an m-dimensional orientable pseudomanifold then

$$h_m \geq \lambda_m \geq \frac{h_m^2}{2(m+1)}.$$

Positive result: chain complex

Proposition (Steenbergen & Klivans & M)

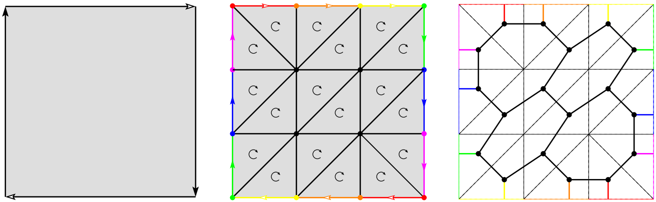
If X is an m-dimensional orientable pseudomanifold then

$$h_m \geq \lambda_m \geq \frac{h_m^2}{2(m+1)}.$$

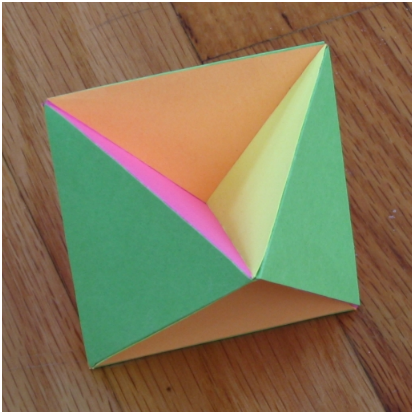
Discrete analog of the Cheeger inequality for manifolds with Dirichlet boundary condition, every $(m-1)$-simplex has at most two cofaces.

Orientation hypothesis

X is a triangulation of the real projective plane.
$H_2(\mathbb{Z}_2) \neq 0 \Rightarrow h_2 = 0$ and $H_2(\mathbb{R}) = 0 \Rightarrow \lambda_2 \neq 0$.



Real projective plane



Orientation hypothesis

Theorem (Gundert & Wagner)

There is an infinite family of complexes that are not combinatorially expanding, $h = 0$, and whose spectral expansion is bounded away from zero, $\lambda > 0$.

Boundary hypothesis

Consider the graph below $h_1 = \frac{2}{3}$ and $\lambda_1 = \lambda^0 \leq \frac{2}{k+1}$.

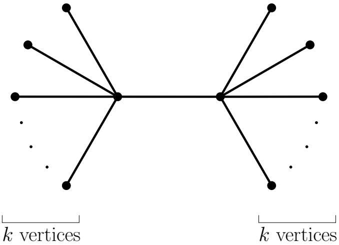


FIGURE 4. The family of graphs G_k.

Negative result: cochain complex

Proposition (Steenbergen & Klivans & M)

For every $m > 1$, there exist families of pseudomanifolds X_k and Y_k of dimension m such that

(1) $\lambda^{m-1}(X_k) \geq \frac{(m-1)^2}{2(m+1)}$ *for all* k *but* $h^{m-1}(X_k) \to 0$ *as* $k \to \infty$

Negative result: cochain complex

Proposition (Steenbergen & Klivans & M)

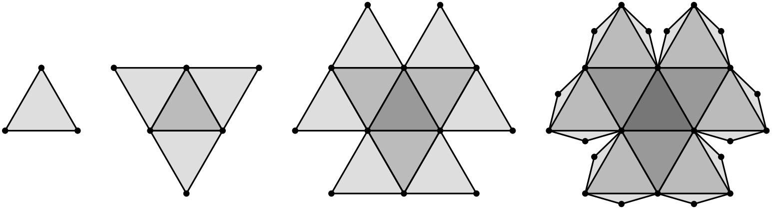
For every m >1, there exist families of pseudomanifolds X_k and Y_k of dimension m such that

(1) $\lambda^{m-1}(X_k) \geq \frac{(m-1)^2}{2(m+1)}$ *for all* k *but* $h^{m-1}(X_k) \to 0$ *as* $k \to \infty$

(2) $\lambda^{m-1}(Y_k) \leq \frac{1}{m^{k-1}}$ *for* $k > 1$ *but* $h^{m-1}(Y_k) \geq \frac{1}{k}$ *for all* k.

Buser part

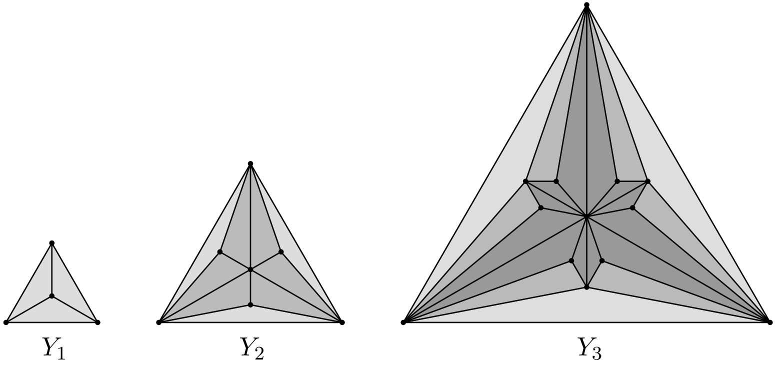
$$h^{m-1}(X_k) \not\lesssim \lambda^{m-1}(X_k)$$



X_1 X_2 X_3 X_4

Cheeger part

$$\lambda^{m-1}(Y_k) \not\lesssim h^{m-1}(Y_k)$$



Open problems

(1) Intermediate values of k – relation of h^k and λ^k or h_k and λ_k for $1 < k < m-1$.

Open problems

(1) Intermediate values of k – relation of h^k and λ^k or h_k and λ_k for $1 < k < m-1$.

(2) High-order eigenvalues – $\lambda^{k,j}$ and $h^{k,j}$ where $j > 1$ indexes the ordering of Fiedler/Cheeger numbers.

(3) Manifolds – The k-dimensional coboundary/boundary Cheeger numbers of a manifold M might be

$$\begin{aligned} h^k &= \inf_S \frac{\mathrm{Vol}_{m-k-1}(\partial S \setminus \partial M)}{\inf_{\partial T = \partial S} \mathrm{Vol}_{m-k}(T)}, \\ h_k &= \inf_S \frac{\mathrm{Vol}_{k-1}(\partial S)}{\inf_{\partial T = \partial S} \mathrm{Vol}_k(T)}. \end{aligned}$$

Acknowledgements

Thanks:
Matt Kahle, Shmuel Weinberger, Anna Gundert, Misha Belkin, Yusu Wang, Lek-Heng Lim, Yuan Yao, Bill Allard, Anil Hirani

Funding:

- Center for Systems Biology at Duke
- NSF DMS and CCF
- AFOSR
- NIH