Near-optimal distortion bounds for embedding doubling spaces into L_1

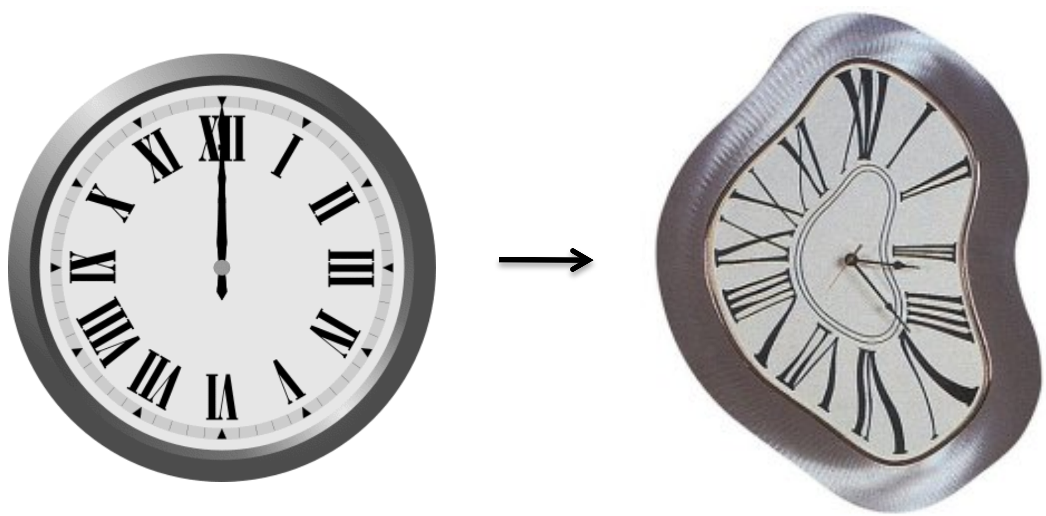
Anastasios Sidiropoulos (TTIC)
Joint work with James R. Lee (U. Washington)

Metric embeddings

- Given spaces $M=(X,d)$, $M'=(X',d')$
- Mapping $f : X \longrightarrow X'$
- Distortion c if:

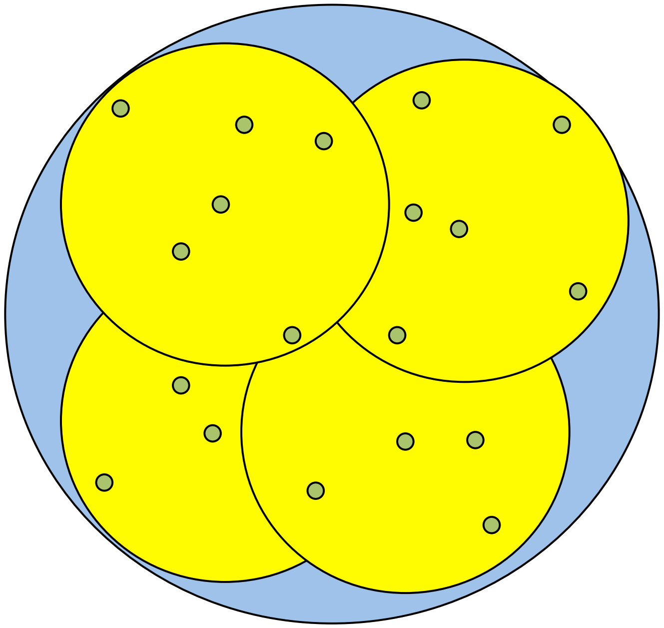
$$d(x_1,x_2) \leq d'(f(x_1),f(x_2)) \leq c \cdot d(x_1,x_2)$$

- $c_Y(X)$ = infimum distortion to embed X into Y



Doubling spaces

- A metric (X, d) is **doubling** if every ball of radius r can be covered by O(1) balls of radius r/2.
- Metric notion of “bounded dimension”



Distortion of L_1 embeddings

- n-point metrics: $O(\log n)$ [Bourgain '85]
- n-vertex expanders: $\Omega(\log n)$ [Linial,London,Rabinovich '95]
- Doubling metrics : $O(\log n)^{1/2}$ [Gupta,Krauthgamer,Lee '03]
- Doubling metrics : $\Omega(\log n)^{\delta}$, for some $\delta>0$ [Cheeger,Kleiner,Naor '09]

Our result

Theorem [Lee,S '11]

There exists an infinite family of uniformly doubling spaces that require distortion

$$\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$$

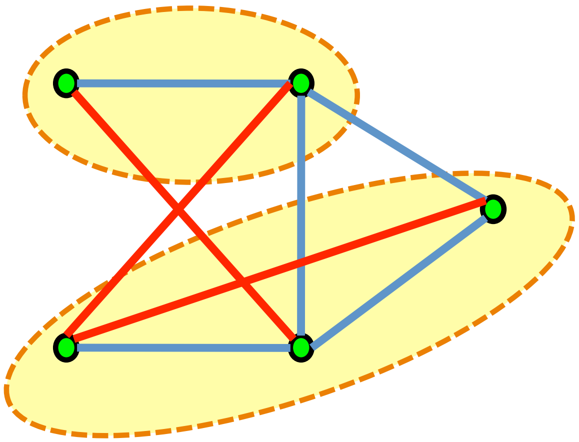
to be embedded into L_1.

I.e. matching the upper bound of Gupta-Krauthgamer-Lee up to a $O((\text{loglog } n)^{1/2})$ factor.

Sparsest-Cut

Instance:

- $G = (V,E)$
- cap : $V \times V \longrightarrow R$
- dem : $V \times V \longrightarrow R$



sparsity of a cut S = (capacity in S) / (demand crossing S)

Key step for a plethora of divide & conquer algorithms:

Crossing Number, Linear Arrangement, VLSI layout, Feedback Arc Set, Balanced Cut, Directed Cuts, Multi-way Cut, Scheduling, PRAM Emulation, Routing, Interval Graph Completion, Planar Edge Deletion, Pathwidth, Markov Chains, … [Leighton, Rao '99], …

Approximating the Sparsest-Cut

- O(log n)-approximation [Linial,London,Rabinovich'95], [Leighton,Rao'88]
- O($\log^{1/2}$ n loglog n)-approximation [Arora,Lee,Naor'05], [Arora,Rao,Vazirani'04]
- 1.001-hard [Ambuhl,Mastrolilli,Svensson'07]
- ω(1)-hard assuming Unique Games [Khot, Vishnoi '05], [Chawla,Krauthgamer,Kumar,Rabani,Sivakumar '05]

Negative type

(X,d) is in **NEG** if $c_2(X,d^{1/2}) = 1$

(X,d) is in **soft-NEG** if $c_2(X,d^{1/2}) = O(1)$

The geometry of graphs

SDP relaxation: $O(\log^{1/2} n \log\log n)$-approximation
[Arora,Lee,Naor'05], [Chawla,Gupta,Racke'05],
[Arora,Rao,Vazirani'04]

Theorem:

SDP integrality gap = min distortion to embed any n-point negative-type metric into L_1.

Theorem: [Arora,Lee,Naor'05]

Every n-point negative-type metric embeds into L_1 with distortion $O(\log^{1/2} n \log\log n)$.

NEG vs L_1

Major open question:

What is the integrality gap of the Sparsest-Cut SDP?

Equivalently:

What is the worst-case distortion required to embed a negative-type metric into L_1?

The Goemans-Linial conjecture

Conjecture [Goemans,Linial'94]

Every negative-type metric embeds into L_1 with distortion $O(1)$.

Theorem [Khot,Vishnoi'05]

There exist an n-point negative-type metric that requires distortion $\Omega(\log\log n)^c$ to embed into L_1.

(see also [Krauthgamer,Rabani], [Devanur,Khot,Saket,Vishnoi])

The Heisenberg group

Theorem [Lee,Naor'06]

The Heisenberg group $H^3(R)$, with the Carnot-Caratheodory metric is in NEG.

Theorem [Cheeger,Kleiner,Naor'09],[Cheeger,Kleiner'06]

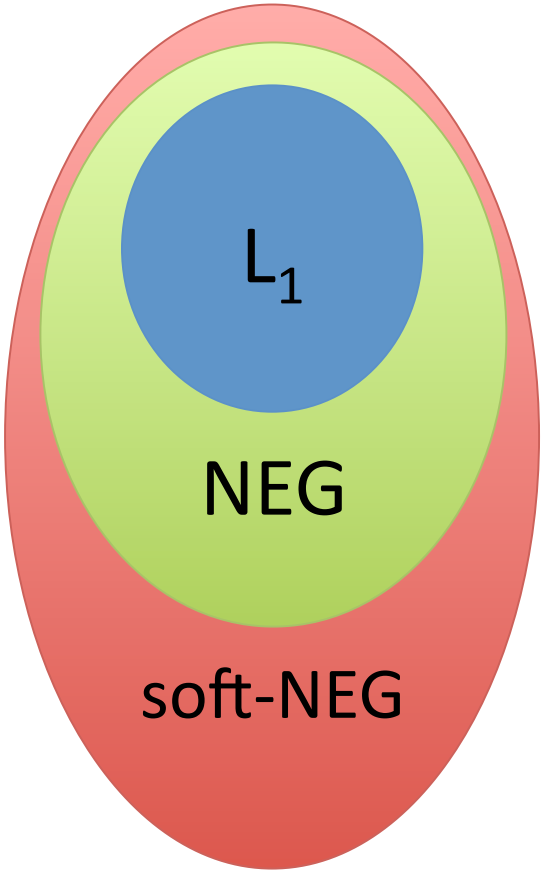
$H^3(R)$ requires distortion $\Omega((\log n)^c)$, for some $c>0$, to embed into L_1.

Corollary

The integrality gap of the Sparsest-Cut SDP is $\Omega((\log n)^c)$, for some $c>0$.

Soft negative-type

- All known algorithms for Sparsest-Cut require only soft-NEG
- This fact is essential for some fast algorithms [Sherman'09]



Our result

Theorem [Lee,S ‘11]
There exists a doubling space that requires distortion

$$\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$$

to be embedded into L_1.

Theorem [Assouad’83]
Every doubling space is in soft-NEG.

Corollary [Lee,S ‘11]
There exists a metric in soft-NEG that requires distortion

$$\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$$

to be embedded into L_1.

In other words…

Corollary [Lee, S '11]

Every known upper bound analysis of the Sparsest-Cut SDP, is tight up to $(\log\log n)^{O(1)}$ factors.

Main result

Sparsest-cut SDP:

$$\min\left\{\frac{\sum_{u,v} cap(u,v)\|x_u - x_v\|_2^2}{\sum_{u,v} dem(u,v)\|x_u - x_v\|_2^2} : (\{x_v\}_v, \|\cdot\|_2^2) \in NEG\right\}$$

Sparsest-cut **weak** SDP:

$$\min\left\{\frac{\sum_{u,v} cap(u,v)\|x_u - x_v\|_2^2}{\sum_{u,v} dem(u,v)\|x_u - x_v\|_2^2} : (\{x_v\}_v, \|\cdot\|_2^2) \in soft-NEG\right\}$$

Corollary [Lee,S]

The integrality gap of the weak SDP is $\Theta((\log n)^{1/2})$, up to $(\log\log n)^{O(1)}$ factors.

Improves over the previous bound of $\Omega(\log n)^{1/4}$ [Lee,Moharrami'10]

Key ingredients of the proof

- A new topological construction of a hard space
- Discrete differentiation
- Discrete/approximate integral geometry in the plane

The Gupta-Newman-Rabinovich-Sinclair conjecture

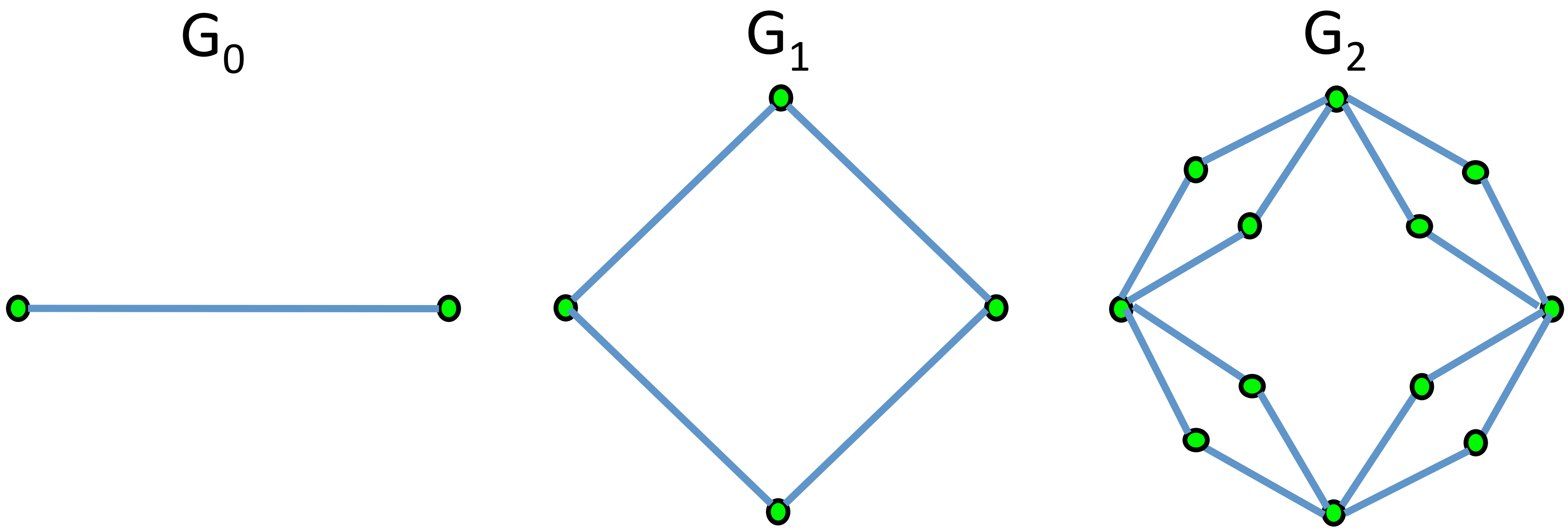
Conjecture [Gupta,Newman,Rabinovich,Sinclail '99]

Every minor-free family of graph embeds into L_1 with distortion O(1).

True for:

- Trees
- Series-parallel graphs [Gupta,Newman,Rabinovich,Sinclail '99]
- O(1)-Outerplanar graphs [Chekuri,Gupta,Newman,Rabinovich,Sinclair '2003]
- O(1)-pathwidth graphs [Lee, S '2009]
- (K_5\e)-free graphs [Chakrabarti, Jaffe, Lee, Vincent '2008]

The diamond graph



Embedding the diamond graph

Theorem [Rao ‘99], [Newman, Rabinovich ‘2002]

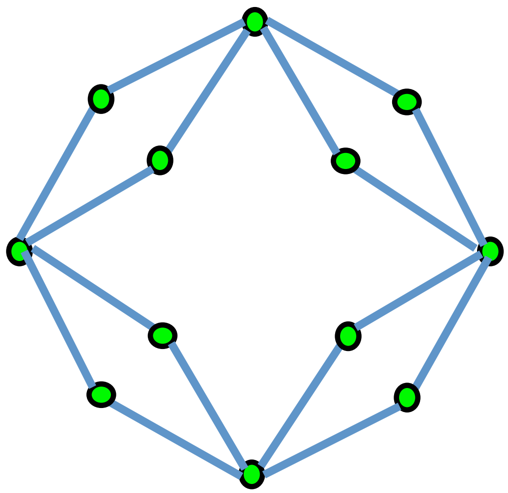
$$c_2(\text{diamond graph}) = \Theta(\sqrt{\log n})$$

Theorem [Gupta,Newman,Rabinovich,Sinclail ‘99], [Chakrabarti, Jaffe, Lee, Vincent ’2008]

$$c_1(\text{diamond graph}) \leq 2$$

Embedding the diamond graph

- In L_2, there is always a diagonal that incurs unbounded contraction. [Newman,Rabinovich'2002]

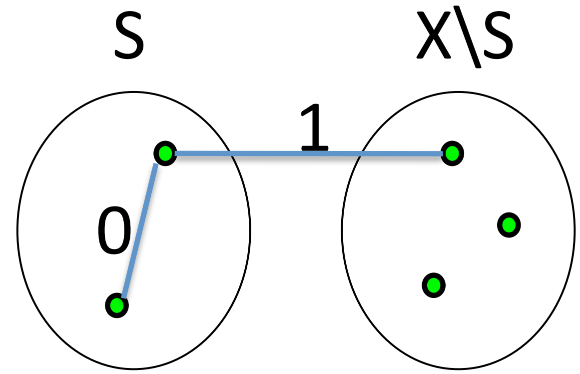


- Why not in L_1?

A combinatorial interpretation of L_1

An embedding into L_1 is a distribution over cuts

The cut cone



- For a finite set X, and $S \subseteq X$, let

  $d_S : X \times X \longrightarrow R$,

  $d_S(x,y) = |\mathbf{1}_S(x) - \mathbf{1}_S(y)|$

- A mapping $d : X \times X \longrightarrow R$ is in the **cut cone** if there exists a non-negative measure μ on 2^X, s.t.

$$\forall x, y \in X, d(x,y) = \int d_S(x,y) d\mu(S)$$

Fact:

A metric is isometrically embeddable into L_1, if and only if it is in the cut cone.

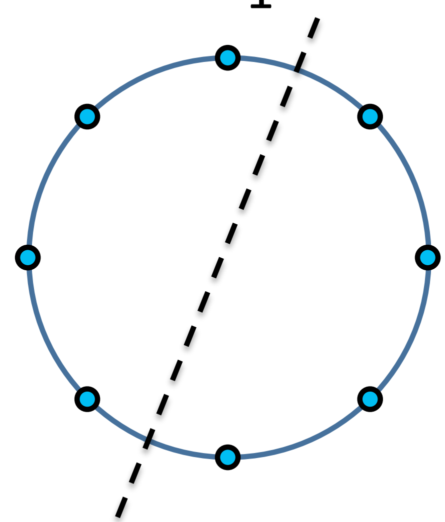
L_1 and the cut cone: example

- Embed the n-line into L_1

  Pick random x in {1,…,n-1}, and take the cut {1,…,x}

- Embed the n-cycle into L_1

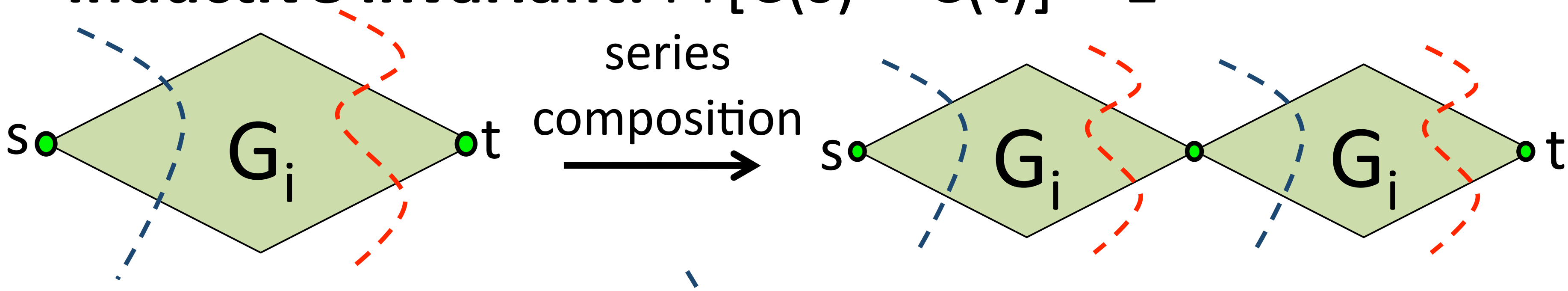
  Pick random angle

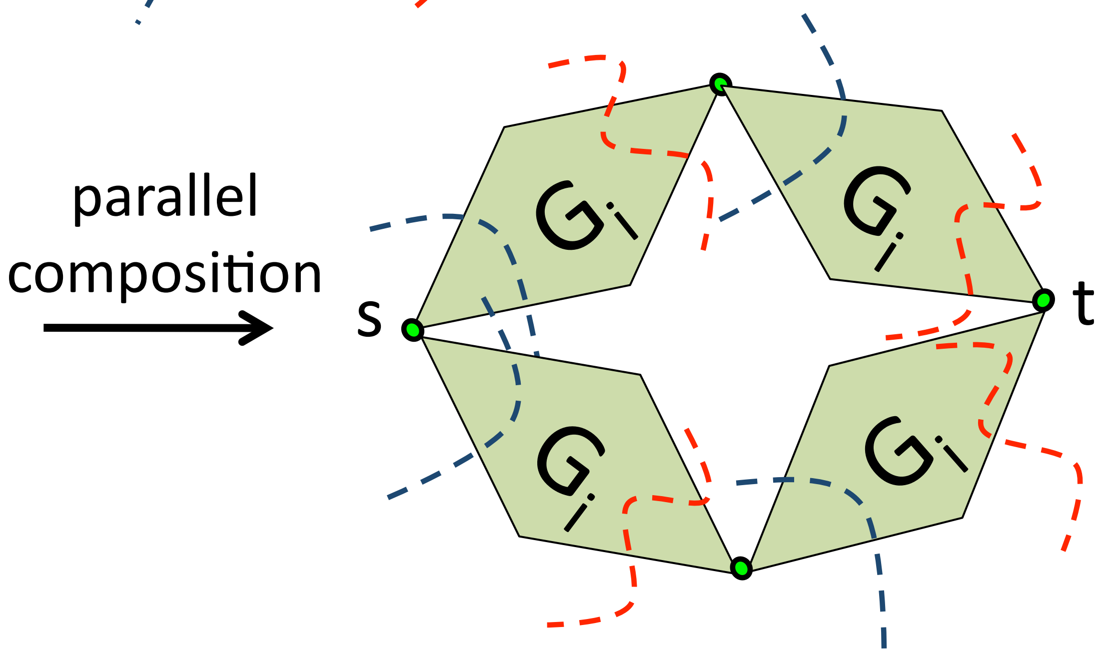


Embedding the diamond graph into L_1

[Gupta,Newman,Rabinovich,Sinclail '99]

Inductive invariant: $\Pr[C(s) \neq C(t)] = 1$





Key property:
The top and bottom copies are *independent*.

Towards a construction

Can we inductively construct a "simple" space s.t. the random cuts in smaller copies are not independent?

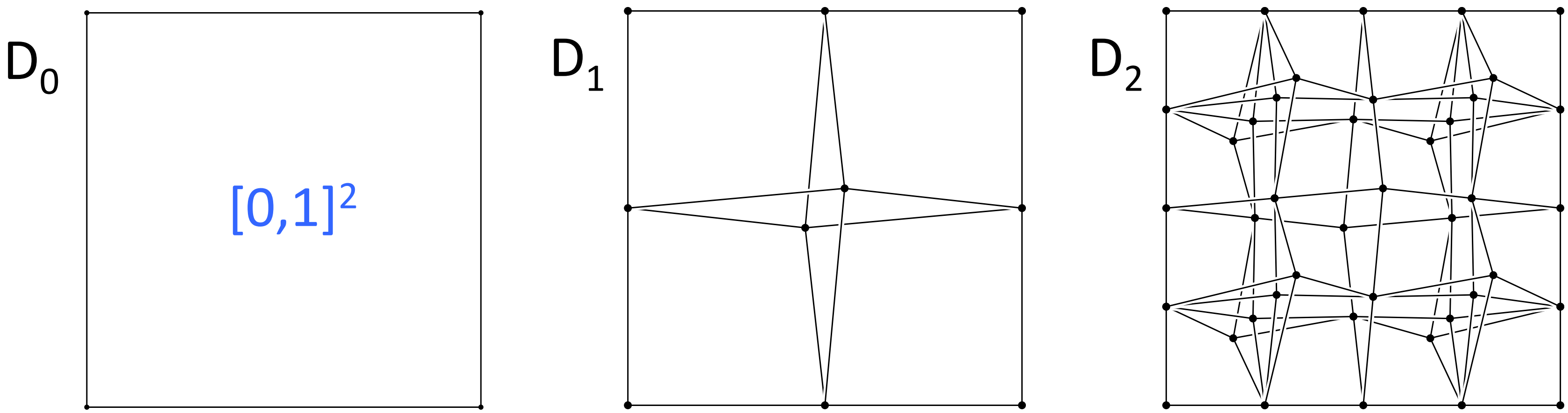
k-Sums Conjecture [Lee, S '09]

O(1)-Embeddability into L_1 is closed under k-sums.

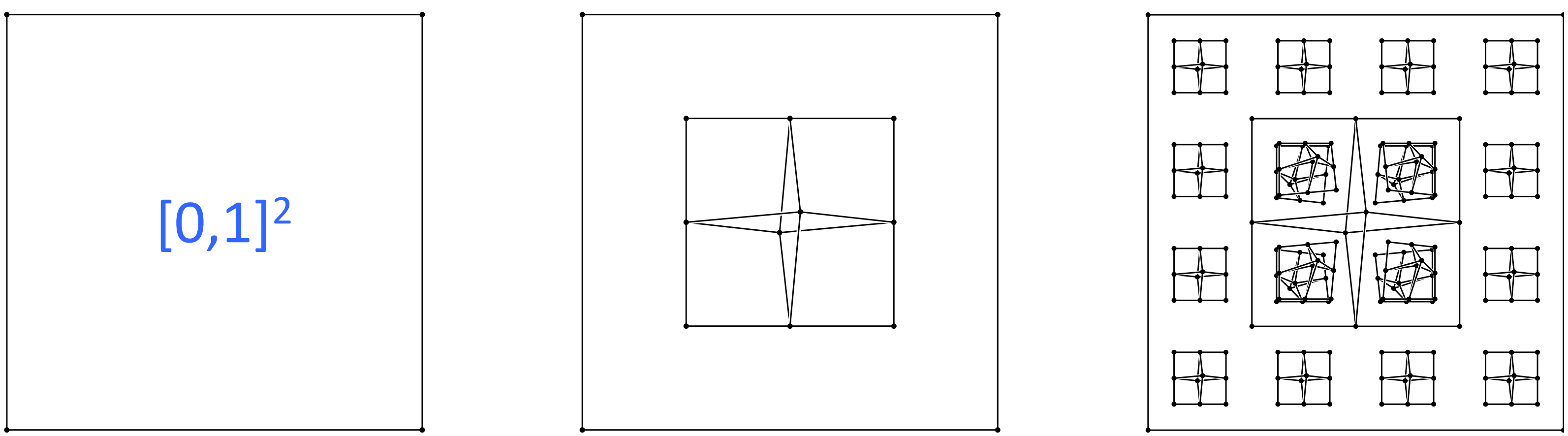
We need a *qualitatively* different inductive construction.

The new construction

The diamond-fold



The Laakso-fold



Differentiation of L_1-valued maps

- [Cheeger,Kleiner'06] develop a weak differentiation theory for maps into L_1.

- [Cheeger,Kleiner'09], [Lee,Raghaventra'07]
  Main idea: At a sufficiently small scale, almost all cuts are "well-structured".

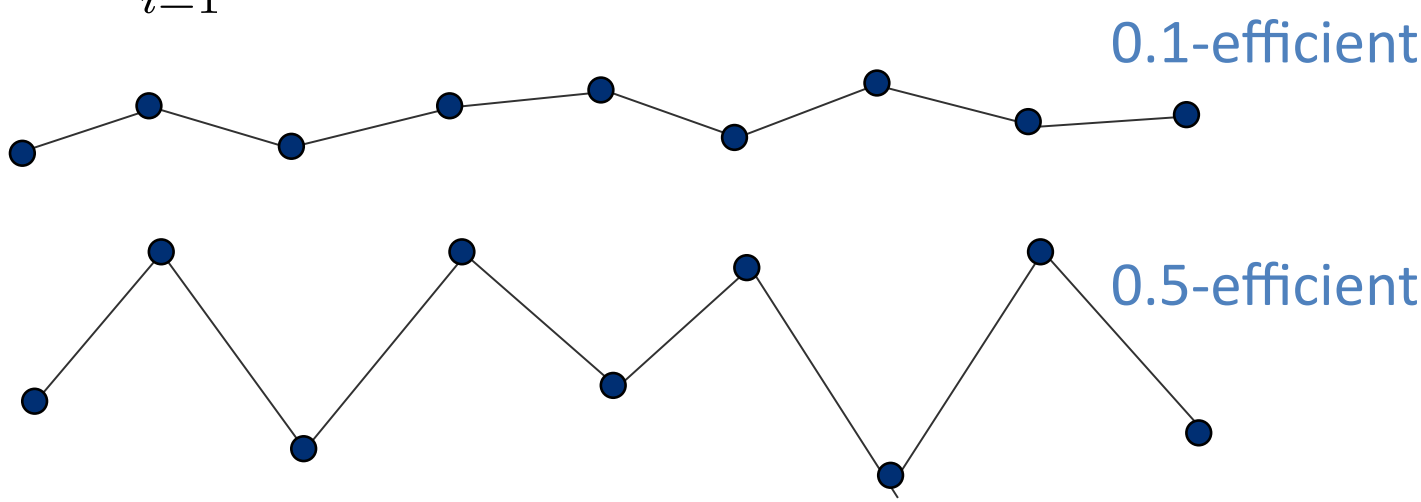
Coarse differentiation

[Matousek'99],[Eskin,Fisher,Whyte'06]

Let (Y,d) be any metric space, ε>0

$f:P_n \rightarrow Y$, is ε-efficient if

$$\sum_{i=1}^{n-1} d(f(x_i), f(x_{i+1})) \leq (1+\varepsilon) d(x_1, x_n)$$



Coarse differentiation (toy version)

Theorem [Matousek'99],[Eskin,Fisher,Whyte'06]

Let (Y,d) be any metric space, $D>0$.

For any $\varepsilon>0$ (arbitrarily small),

there exists $n>0$, such that

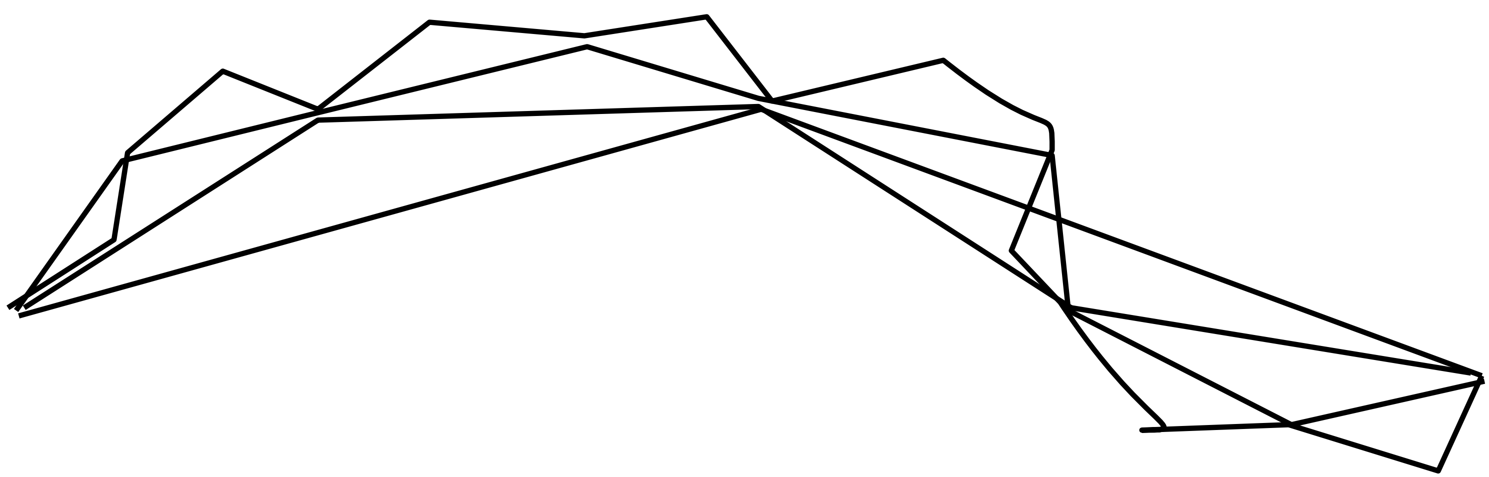
for any $f:P_n \longrightarrow Y$ with distortion D,

we can find an ε-efficient copy of P_3 in $f(P_n)$.

Coarse differentiation

Proof idea:

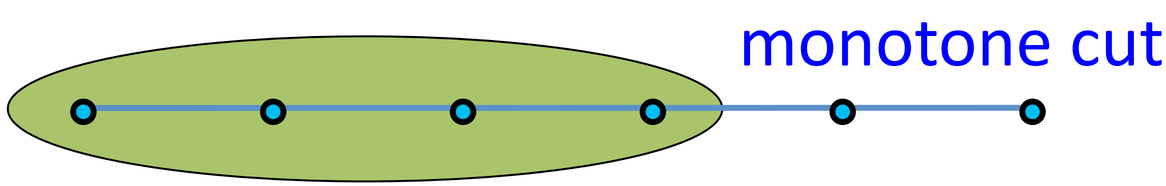
Suppose *no* scale is ε-efficient.

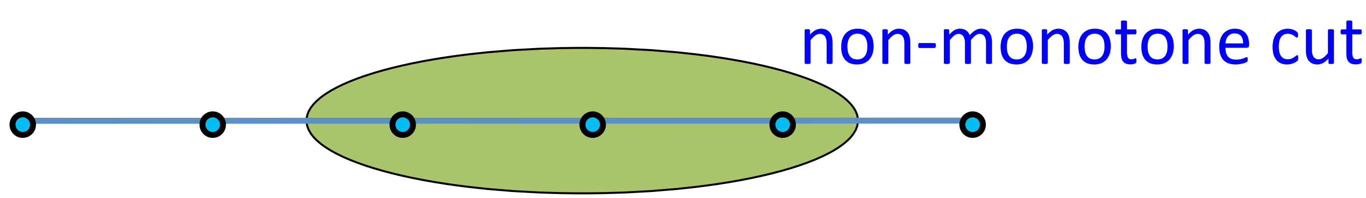


Differentiation in L_1

[Lee,Raghaventra'07], [Cheeger,Kleiner'09]

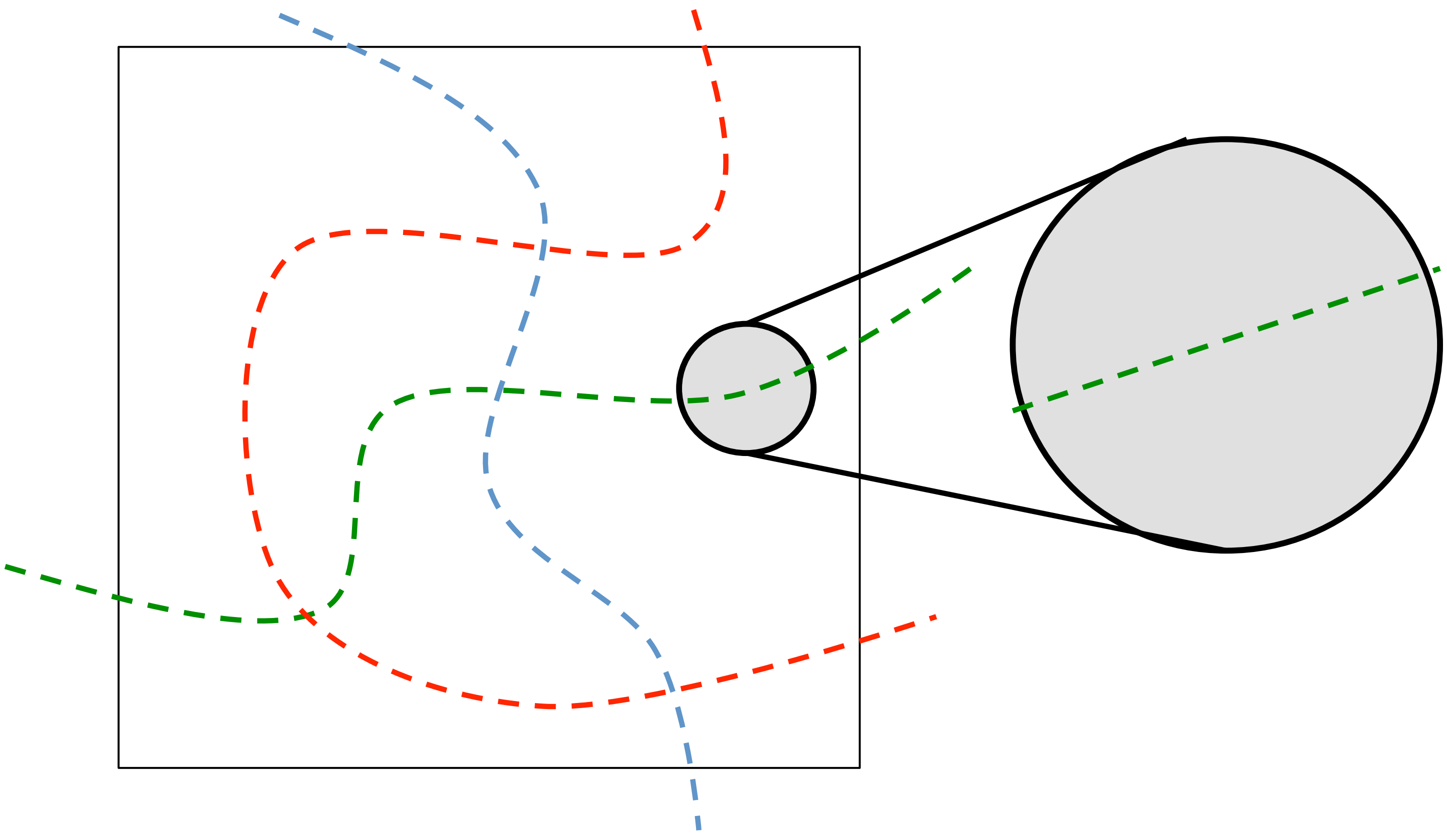
$f : P_n \rightarrow L_1$ is 0-efficient if and only if all cuts are half-lines.





Differentiation for maps $[0,1]^2 \to L_1$

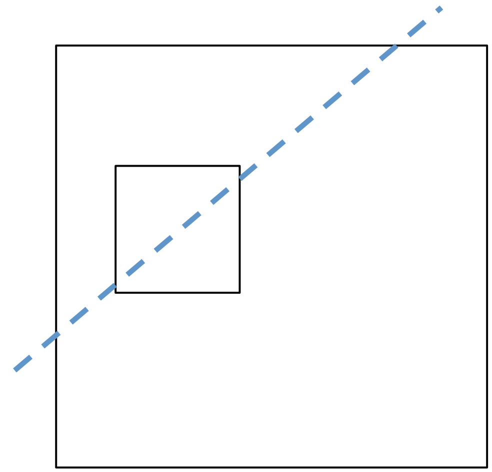
Locally, the distribution of cuts consists mostly of (near-)half-planes.

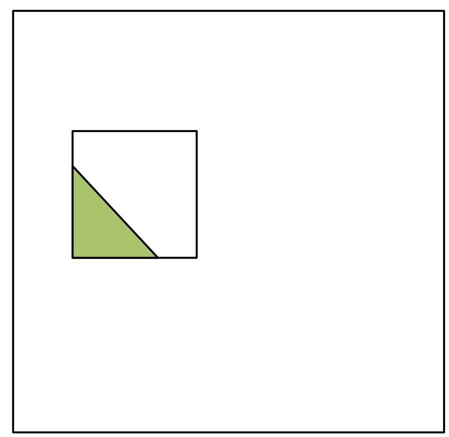


Differentiation and the diamond-fold

Main idea:

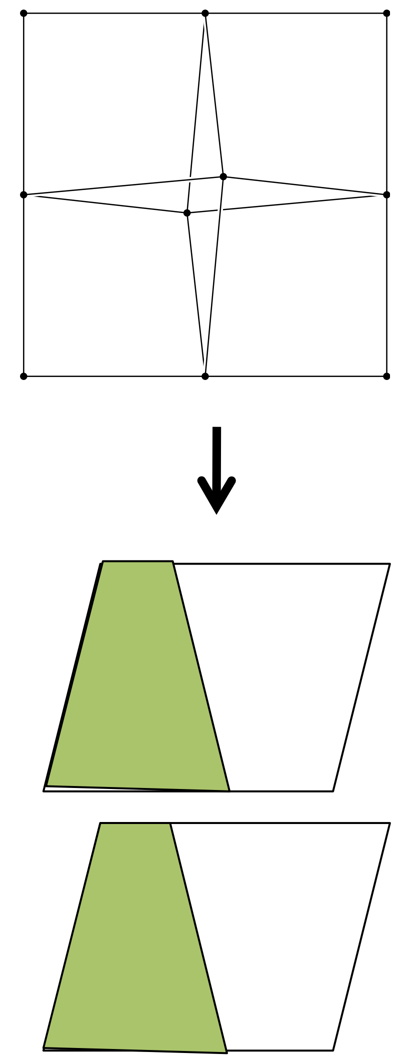
- Let $f : [0,1]^2 \longrightarrow L_1$
- Then, at a sufficiently small square X, for every line h intersecting X, almost all cuts restricted on h, are half-lines.
- Suppose that all cuts restricted to every line are half-lines. Then, all cuts are *half-planes*.



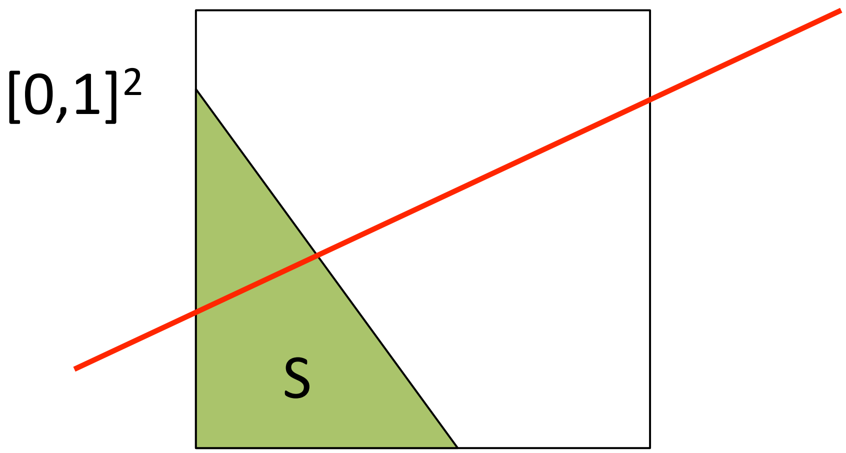


Differentiation and the diamond-fold (cont.)

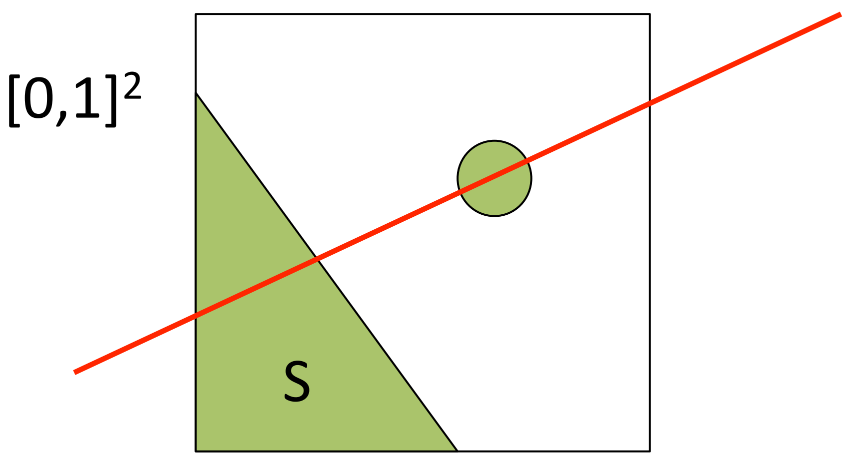
- It follows that there exists a copy of D_1, such that in both copies of $[0,1]^2$, all cuts are half-planes.
- But then the half-planes must be **identical** in both sheets.
- Thus, the two sheets are **collapsed**.



Differentiation and the diamond-fold



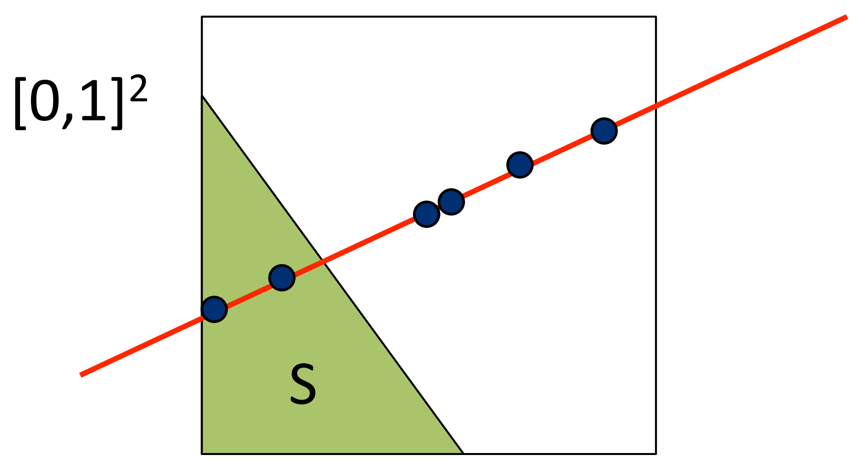
A map is 0-efficient if and only if every cut is a **halfplane**



Obstacle: An ε-efficient map might have **no** halfplane cuts

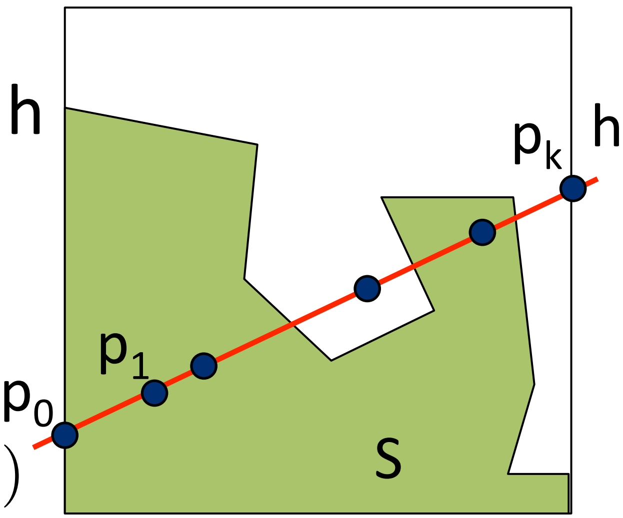
The quantitative bound

- We define efficiency w.r.t. **random** lines in the unit square.
- Avoid periodicities: define efficiency w.r.to a **random** subset of points in every line.



Taming ε-efficient maps

- Pick random line h
- Pick random set P of $k=1/\varepsilon^{O(1)}$ points in h
- p_0, p_k are on the boundary
- "Complexity" of a set:



$$C^*(S) = \int \mathbb{E}_P \sum_j |\mathbf{1}_S(p_j) - \mathbf{1}_S(p_{j+1})| d\mu(h)$$

$$C(S) = \int \mathbb{E}_P |\mathbf{1}_S(p_0) - \mathbf{1}_S(p_k)| d\mu(h)$$

Fact: C*(S) = C(S) iff S is a half-plane

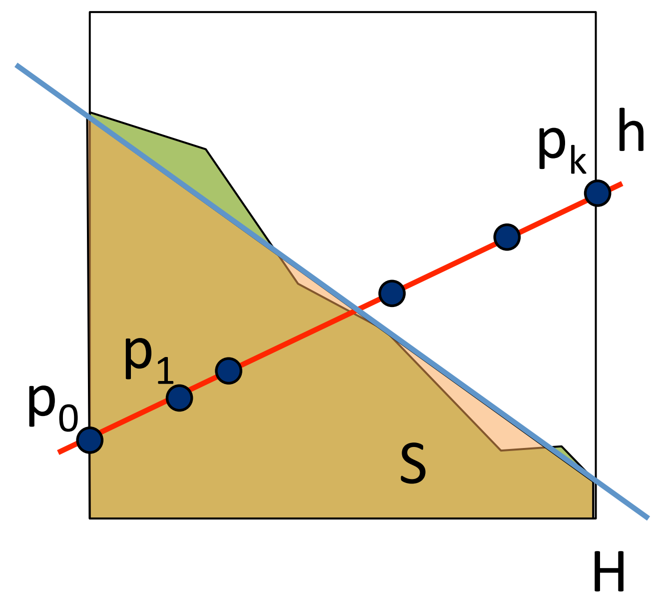
Taming ε-efficient maps (cont.)

Lemma: [Lee,S]

If $|C^*(S) - C(S)| = O(\varepsilon^2)$,

then there exists half-plane H, such that

$$|S \triangle (H \cap [0,1]^2)| = O(\varepsilon)$$



Lemma: [Lee,S]

If $|C^*(S) - C(S)| = O(\varepsilon^2)$, and $|S|<1/16$,

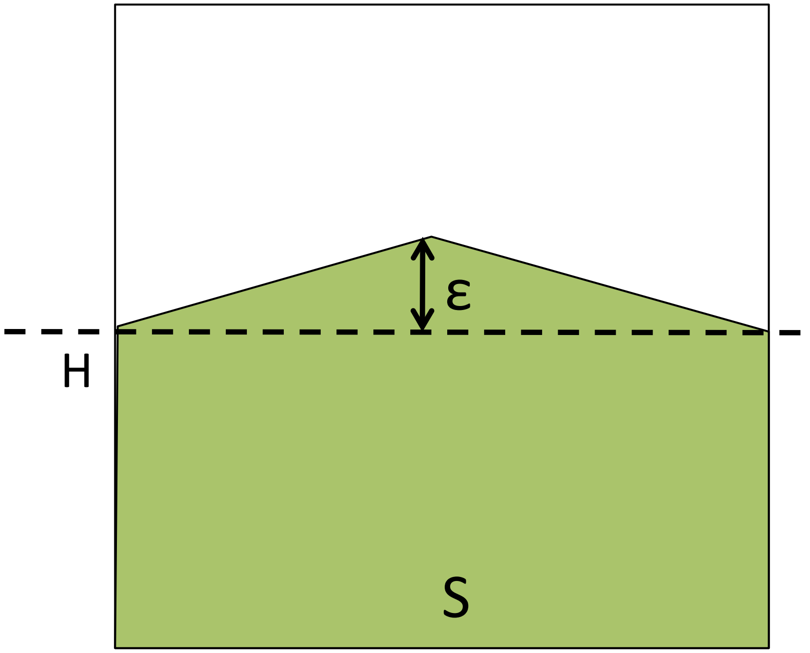
then there exists half-plane H, such that

$$\left|(S \triangle H) \cap \left[\frac{1}{3}, \frac{3}{4}\right]^2\right| = O(\varepsilon^2)$$

Tightness of the analysis

$$|(S \triangle H) \cap [0,1]^2| = O(\varepsilon)$$

$$|C(S) - C^*(S)| = O(\varepsilon^2)$$



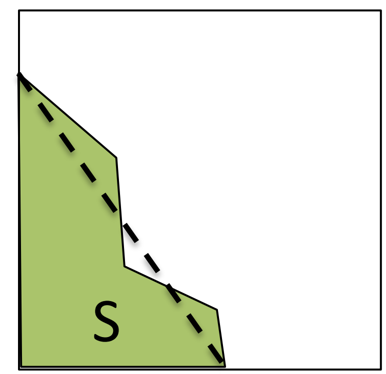
Obtaining the distortion bound

Consider two parallel “sheets”

Since the boundaries are identified,

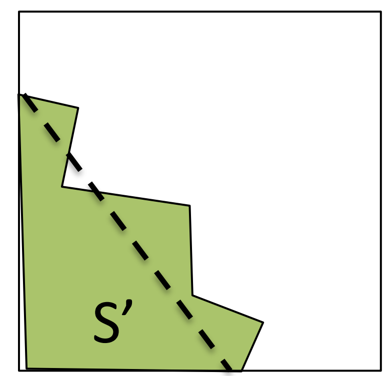
both S and S’ are close to the

same half-plane.



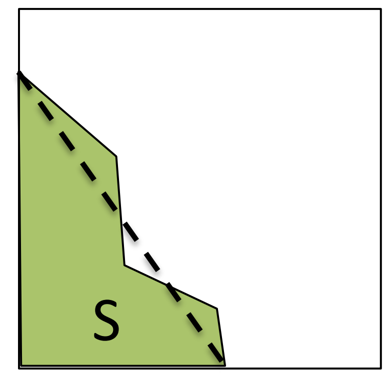
Thus, S and S’ are close to

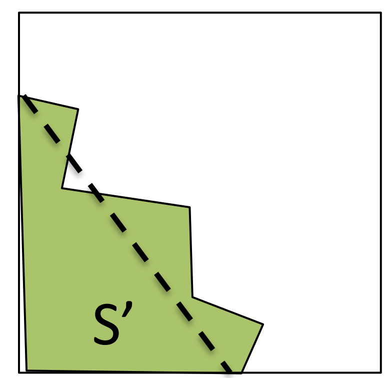
each other.



Obtaining the distortion bound (cont.)

If S and S' are close to each other, then the distance between most antipodals that are close to the center of $[0,1]^2$, is too small!





Further directions

- NEG vs L_1?
- Can these techniques be used to obtain computational hardness?
- Gupta-Newman-Rabinovich-Sinclair conjecture: Minor-free graphs into L_1? The diamondfold contains arbitrarily large clique minors.