Sum of squares.

Quotient Rayleigh: $\max_{x \perp 1} \frac{x^T M x}{x^t x}$.

Alternatively: $\mu = \min_{x \perp 1} \frac{x^T (L/d) x}{x^t x}$.

$L = dI - A$ or $L/d = I - A/d = I - M$.

Also:
$$x^T L x = \sum_i dx_i^2 - \sum_{e=(i,j)} 2x_i x_j$$
$$= \sum_{e=(i,j)} (x_i^2 + x_j^2 - 2x_i x_j)$$
$$= \sum_{e=(i,j)} (x_i - x_j)^2.$$

Warmup exercise: no larger eigenvalue on cycle.

Consider x for path.

Sort by x-value: $x_1, \ldots, x_n$.

Let $x_1^2 + x_n^2 = 1$. Shift so that $\sum_i x = 0$.

Now, since any cut ≥ 1
$x^T L x \geq 1 \times \sum_i (x_i - x_{i+1})^2$.

Cuz: $\geq$ one edge "crossing" over the interval $[x_i, x_{i+1}]$.

Take $a = (x_i - x_{i+1})$ and $b = 1/\sqrt{n}$.

Cauchy-Schwartz $|a||b| \geq a \cdot b$.

Yields: $\sum_i (x_i - x_{i+1})^2 \geq (x_1 - x_n)^2 / n$

Furthermore: $\sum_i x_i^2 \leq n(x_1^2 + x_n^2) \leq n(x_1^2 + x_n^2 - 2x_1 x_n) = n(x_1 - x_n)^2$.

$\rightarrow \mu = \frac{x^T L x}{x^T x} \geq 1/n^2$.

Argument uses cut size to lower bound eigenvalue.
Or use eigenvalue to upper bound on cut size.
Is $\mu \geq 2/n^2$?

Interpretation of quadratic form

Consider $\mu = \min_{x \perp 1} \frac{x^T L x}{d x^T x}$.

$\sum_i x_i = 0$

Claim: $n \sum_i x_i^2 = \frac{1}{n} \sum_{i \leq j} (x_i - x_j)^2$

$\sum_i (2(n-1)x_i^2 - 2x_i \sum_{j \neq i} x_j)$.

$-2x_i \sum_{j \neq i} x_j) = -2x_i((\sum_j x_j) - x_i) = 2x_i^2 + x_i \sum_i x_j = 2x_i^2$

$\sum_i n x_i^2 - (\sum_i x_i)(\sum_j x_j)$.

Claim: $x^T L x = \sum_{e=(i,j)} M_{ij} (x_i - x_j)^2$

$\mu = \min_{x \perp 1} \frac{\sum_e M_{ij} (x_i - x_j)^2}{\frac{d}{n} \sum_{i,j} (x_i - x_j)^2}$

Ratio of average edge "length" to pair length.

Conductance $\phi(G) = n \frac{E(S, \overline{S})}{|S||V - S|}$.

Idea: pick random threshold.
Average number of edges cut compared to pairs is μ?

Random threshold.

$\mu = \min_{x \perp 1} \frac{\sum_e M_{ij} (x_i - x_j)^2}{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2}$

Conductance $\phi(G) = n \frac{E(S, \overline{S})}{|S||V - S|}$.

Idea: pick random threshold.

Average number of edges cut compared to pairs is μ.
True if "length" corresponded to $|x_i - x_j|$.
Off when length is $(x_i - x_j)^2$.

Off by average length of edges.
For path:
edges have $|x_i - x_j|$ being $1/n$ of that for pairs.
For expanders:
edges go a constant fraction of what random pairs do.

Cheeger Hard Part.

Now let's get to the hard part of Cheeger $h(G) \leq \sqrt{2(1 - \lambda_2)}$.

Idea: We have $1 - \lambda_2$ as a continuous relaxation of $\phi(G)$

Take the 2^{nd} eigenvector $x = \text{argmin}_{x \in \mathbb{R}^V - \text{Span}\{\mathbf{1}\}} \frac{\sum_{i,j} M_{ij} (x_i - x_j)^2}{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2}$

Consider x as an embedding of the vertices to the real line.

Round x to get a $x \in \{0, 1\}^{|V|}$

Rounding: Take a threshold t,
$$\begin{cases} x_i \geq t & \rightarrow x_i = 1 \\ x_i < t & \rightarrow x_i = 0 \end{cases}$$

What will be a good t?

We don't know. Try all possible thresholds ($n - 1$ possibilities), and hope there is a t leading to a good cut!

Sweep Cut Algorithm

Input: $G = (V, E)$, $x \in \mathbb{R}^V, x \perp \mathbf{1}$

Sort the vertices in non-decreasing order with respect to x
WLOG $V = \{1, \ldots, n\}$ $\quad x_1 \leq x_2 \leq \ldots \leq x_n$

Let $S_i = \{1, \ldots, i\}$ $\quad i = 1, \ldots, n - 1$

Return $S = \text{argmin}_{S_i} h(S_i)$

Main Lemma: $G = (V, E)$, d-regular
$x \in \mathbb{R}^V, x \perp \mathbf{1}, \mu = \frac{\sum_{i,j} M_{ij} (x_i - x_j)^2}{\frac{d}{n} \sum_{i,j} (x_i - x_j)^2}$

If S is the ouput of the sweep cut algorithm, then $h(S) \leq \sqrt{2\mu}$

Note: Applying the Main Lemma with the 2^{nd} eigenvector v_2, we have $\mu = 1 - \lambda_2$, and $h(G) \leq h(S) \leq \sqrt{2(1 - \lambda_2)}$. Done!

Proof of Main Lemma

WLOG $V = \{1, \ldots, n\}$ $\quad x_1 \le x_2 \le \ldots \le x_n$

Want to show

$$\exists i \text{ s.t. } h(S_i) = \frac{\frac{1}{d}|E(S_i, V - S_i)|}{\min(|S_i|, |V - S_i|)} \le \sqrt{2\mu}$$

Probabilistic Argument: Construct a distribution D over $\{S_1, \ldots, S_{n-1}\}$ such that

$$\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S, V - S)|]}{\mathbb{E}_{S\sim D}[\min(|S|, |V - S|)]} \le \sqrt{2\mu}$$

$\to \mathbb{E}_{S\sim D}[\frac{1}{d}|E(S, V - S)| - \sqrt{2\mu}\min(|S|, |V - S|)] \le 0$

$\exists S \quad \frac{1}{d}|E(S, V - S)| - \sqrt{2\mu}\min(|S|, |V - S|) \le 0$

The distribution D

WLOG, shift and scale so that $x_{\lfloor \frac{n}{2} \rfloor} = 0$, and $x_1^2 + x_n^2 = 1$

Take t from the range $[x_1, x_n]$ with density function $f(t) = 2|t|$.

Check: $\int_{x_1}^{x_n} f(t)\mathrm{d}t = \int_{x_1}^{0} -2t\mathrm{d}t + \int_{0}^{x_n} 2t\mathrm{d}t = x_1^2 + x_n^2 = 1$

$S = \{i : x_i \le t\}$

Let D be distribution over $S_1, \ldots, S_{n-1}$ from the above process.

Denominator.

Goal: $\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S, V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|, |V-S|)]} \le \sqrt{2\mu}$

Denominator:

Let T_i = indicator for "i is in the smaller set of $S, V - S$"

Can check

$$\mathbb{E}_{S\sim D}[T_i] = Pr[T_i = 1] = x_i^2$$

Idea: i in smaller set if $\tau \in [0, x_i]$ or $[x_i, 0]$.

$$\begin{aligned}
\mathbb{E}_{S\sim D}[min(|S|, |V - S|)] &= \mathbb{E}_{S\sim D}[\sum_i T_i] \\
&= \sum_i \mathbb{E}_{S\sim D}[T_i] \\
&= \sum_i x_i^2
\end{aligned}$$

Numerator

Goal: $\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S, V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|, |V-S|)]} \le \sqrt{2\mu}$

Numerator:

Let $T_{i,j}$ = indicator for i, j is cut by $(S, V - S)$

$$\begin{cases} x_i, x_j \text{ same sign:} & Pr[T_{i,j} = 1] = |x_i^2 - x_j^2| \\ x_i, x_j \text{ different sign:} & Pr[T_{i,j} = 1] = x_i^2 + x_j^2 \end{cases}$$

A common upper bound: $\mathbb{E}[T_{i,j}] = Pr[T_{i,j} = 1] \le |x_i - x_j|(|x_i| + |x_j|)$

$$\begin{aligned}
\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S, V - S)|] &= \frac{1}{2}\sum_{i,j} M_{ij}\mathbb{E}[T_{i,j}] \\
&\le \frac{1}{2}\sum_{i,j} M_{ij}|x_i - x_j|(|x_i| + |x_j|)
\end{aligned}$$

Cauchy-Schwarz Inequality

$|a \cdot b| \le \|a\|\|b\|$, as $a \cdot b = \|a\|\|b\|cos(a, b)$

Applying with $a, b \in \mathbb{R}^{n^2}$ with $a_{ij} = \sqrt{M_{ij}}|x_i - x_j|, b_{ij} = \sqrt{M_{ij}}(|x_i| + |x_j|)$

Numerator:

$$\begin{aligned}
\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S, V - S)|] &= \frac{1}{2}\sum_{i,j} M_{ij}\mathbb{E}[T_{i,j}] \\
&\le \frac{1}{2}\sum_{i,j} M_{ij}|x_i - x_j|(|x_i| + |x_j|) \\
&= \frac{1}{2} a \cdot b \\
&\le \frac{1}{2}\|a\|\|b\|
\end{aligned}$$

Simplify numerator.

Recall $\mu = \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i - x_j)^2}, a_{ij} = \sqrt{M_{ij}}|x_i - x_j|, b_{ij} = \sqrt{M_{ij}}(|x_i| + |x_j|)$

$$\begin{aligned}
\|a\|^2 &= \sum_{i,j} M_{ij}(x_i - x_j)^2 = \frac{\mu}{n}\sum_{i,j}(x_i - x_j)^2 \\
&= \frac{\mu}{n}\sum_{i,j}(x_i^2 + x_j^2) - \sum_{i,j} 2x_i x_j \\
&= \frac{\mu}{n}\sum_{i,j}(x_i^2 + x_j^2) - 2(\sum_i x_i)^2 \\
&\le \frac{\mu}{n}\sum_{i,j}(x_i^2 + x_j^2) = 2\mu\sum_i x_i^2
\end{aligned}$$

$$\begin{aligned}
\|b\|^2 &= \sum_{i,j} M_{ij}(|x_i| + |x_j|)^2 \\
&\le \sum_{i,j} M_{ij}(2x_i^2 + 2x_j^2) \\
&= 4\sum_i x_i^2
\end{aligned}$$

Put together.

Goal: $\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \leq \sqrt{2\mu}$

Numerator:

$$\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|] =\leq \frac{1}{2}\|a\|\|b\|$$
$$\leq \frac{1}{2}\sqrt{2\mu\sum_i x_i^2}\sqrt{4\sum_i x_i^2} \quad = \sqrt{2\mu}\sum_i x_i^2$$

Recall **Denominator:**

$$\mathbb{E}_{S\sim D}[min(|S|,|V-S|)] = \sum_i x_i^2$$

Combining:

$$\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|] - \sqrt{2\mu}\mathbb{E}_{S\sim D}[min(|S|,|V-S|)] \leq 0$$

$$\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \leq \sqrt{2\mu}$$

Thus $\exists S_i$ such that $h(S_i) \leq \sqrt{2\mu}$, which gives $h(G) \leq \sqrt{2(1-\lambda)}$ □

Kind of a proof.

$G=(V,E)$, $h=h(G)$.

Claim:
From $S \subset V$ of vertices, $|E(N_{1/h}(S))| \geq 2|E(S)|$.

Claim': there are $\Omega(h|S|)$ paths of length $\ell = 1/h$ in $N_{1/h}(S)$.

Cut size is $\geq h(G)|S| \implies$ flow of value $h(G)|S|$.
Max flow-min cut theorem.

From path argument: $\implies \mu \geq \frac{1}{\ell^2} = h(G)^2$.

Run argument over sets of size 2^i and one gets the upper bound.

Why no $\log n$ factor?
The mass splits, and **every level** has $\mu_i \geq h(G)^2$.

Cheeger proof magically does this!

Metric spaces.

A metric space X, $d(i,j)$ where
$d(i,j) \leq d(i,k)+d(k,j)$, $d(i,j)=d(j,i)$, $d(i,i)=0$,
and $d(i,j) \geq 0$.

Which are metric spaces?

(A) X from R^d and $d(\cdot,\cdot)$ is Euclidean distance.

(B) X from R^d and $d(\cdot,\cdot)$ is squared Euclidean distance.

(C) X- vertices in graph, $d(i,j)$ is shortest path distances in graph.

(D) X is a set of vectors and $d(u,v)$ is $u\cdot v$.

(A) Obeys triangle inequality. (B) $a^2+b^2 \leq (a+b)^2 = a^2+2ab+b^2$
(C) Shortest! (D) $1\cdot -1 < 0$.

Input to TSP, facility location, some layout problems, ..., metric labelling.

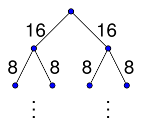
Hard problems. Easier to solve on trees.
Dynamic programming on trees. Linear solving on trees.

Approximate metric on trees?

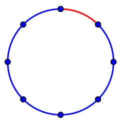
Approximate metric using a tree.

Tree metric:
X is nodes of tree with edge weights
$d_T(i,j)$ shortest path metric on tree.

Hierarchically well separated tree metric:
Tree weights are geometrically decreasing.



Probabilistic Tree embedding.
Map X into tree.
(i) No distance shrinks. (dominating)
(ii) Every distance stretches $\leq \alpha$
in expectation.



Distance 1 goes to $n-1$!
Bummer.

Map metric onto tree?

Fix it up chappie!

For cycle, remove a random edge get a tree.

Stretch of edge: $\frac{n-1}{n} \times 1 + \frac{1}{n} \times (n-1) \approx 2$
General metrics?

Probabilistic Tree embedding.

Probabilistic Tree embedding.
Map X into tree.
(i) No distance shrinks (dominating).
(ii) Every distance stretches $\leq \alpha$ in expecation.

Today: the tree will be Hierarchically well-separated (HST).
Elements of X are leaves of tree.

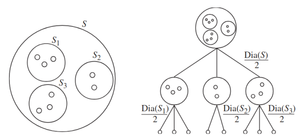
Useful: **use** spanning tree for graphical metrics.

The Idea:
HST $\equiv$ recursive decomposition of metric space.

Decompose space by diameter $\approx \Delta$ balls.
Recurse on each ball for $\Delta/2$.

Use randomness in
selection of ball centers.
the $\approx$ diameter of the balls.

Idea of decomposition.



Algorithm

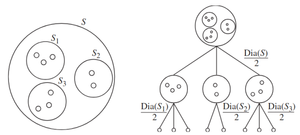
Algorithm: (X,d), $\text{diam}(X) \le D$, $|X| = n$, $d(i,j) \ge 1$
1. π – random permutation of X.
2. Choose β in $[\frac{3}{8}, \frac{1}{2}]$ uniformly at random.

```
def subtree(S,Δ):
  T = []
  if Δ < 1 return [S]
  foreach i in π:
    if i ∈ S
      B = ball(i, βΔ) ; S = S/B
      T.append(B)
  return map (λ x: subtree(x,Δ/2), T);
```

3. subtree(X, D)

Tree has internal node for each level of call.
Tree edges have weight Δ to children.

Analysis: no distance shrinks.



Claim 1: $d_T(x,y) \ge d(x,y)$.

When $\Delta \le d(x,y)$, x and y in diff. balls, $\implies$ cut at $\Delta \ge d(x,y)/2$.

$\rightarrow d_T(x,y) \ge \Delta + \Delta \ge d(x,y)$

Analysis: idea

Claim: $E[d_T(x,y)] = O(\log n) d(x,y)$.

Cut at level $\Delta \rightarrow d_T(x,y) \le 4\Delta$. (Level of subtree call.)

Pr[cut at levelΔ]? Recall: cut at $\beta\Delta$, with $\beta \in [3/8, 1/2]$.

Would like $\le \frac{d(x,y)}{\Delta}$.

$\rightarrow$ expected length is $\sum_{\Delta=D/2^i} (4\Delta) \frac{d(x,y)}{\Delta} = 4 \log D \cdot d(x,y)$.

Why $\propto \frac{d(x,y)}{\Delta}$?
smaller the edge the less likely to be on edge of ball.
larger the delta, more room inside ball.
random diameter β shifts edge across $\Delta/8$.

$\rightarrow Pr[x,y \text{ cut by ball} | x \text{ in ball}] \approx \frac{d(x,y)}{\Delta/8}$

The problem?
Could be cut be many different balls.
For each probability is good, but could be hit by many.
random permutation to deal with this

Analysis: (x,y)

Would like $Pr[x,y \text{ cut by ball} | x \text{ in ball}] \le \frac{8d(x,y)}{\Delta}$
(Only consider cut by x, factor 2 loss.)

At level Δ

At some point x is in some Δ level ball.
Renumber points in order of distance from x.

If $d(x,y) \ge \Delta/8$, $\frac{8d(x,y)}{\Delta} \ge 1$, so claim holds trivially.

Point j cuts (x,y) only if $d(j,x) \in [\frac{3\Delta}{8}, \frac{\Delta}{2}]$.
Call this set X_Δ.

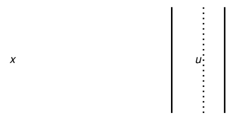
$j \in X_\Delta$ cuts (x,y) if..
$d(j,x) \le \beta\Delta$ and $\beta\Delta \le d(j,y) \le d(j,x) + d(x,y)$
$\rightarrow \beta\Delta \in [d[j,x], d(j,x) + d(x,y)]$.
occurs with prob. $\le \frac{d(x,y)}{\Delta/8} = \frac{8d(x,y)}{\Delta}$.

And j must be before any $i < j$ in $\pi \rightarrow$ prob is $\frac{1}{j}$

$\rightarrow Pr[j \text{ cuts } (x,y)] \le \left(\frac{1}{j}\right) \frac{8d(x,y)}{\Delta}$

$d_T(x,y)$ at level Δ is 4Δ. $\rightarrow E[d_T(x,y)] = \sum_{\Delta=\frac{D}{2^i}} \sum_{j \in X_\Delta} \left(\frac{1}{j}\right) 32d(x,y)$

A picture.



u captures x if first out of j nodes in ball.

The pipes are distinct!

$E(d_T(x,y)] = \sum_{\Delta=D/2^i} \sum_{j \in X_\Delta} \left(\frac{1}{j}\right) 32d(x,y)$

Recall X_Δ has points with $d(x,j) \in [3\Delta/8, \Delta/2]$

"Listen Stash, the pipes are distinct!!"

Uh.. well X_Δ is distinct from $X_{\Delta/2}$.

$E(d_T(x,y)] = \sum_{\Delta=\frac{D}{2^i}} \sum_{j \in X_\Delta} \left(\frac{1}{j}\right) 32d(x,y)$
$\le \sum_j \left(\frac{1}{j}\right) 32d(x,y)$
$\le (32 \ln n)(d(x,y))$.

Claim: $E[d_T(x,y)] = O(\log n) d(x,y)$

Expected stretch is $O(\log n)$.

We gave an algorithm that produces a distribution of trees.

The expected stretch of any pair is $O(\log n)$.

Metric Labelling

Input: graph $G = (V, E)$ with edge weights, $w(\cdot)$, metric labels (X, d), and costs for mapping vertices to labels $c : V \times X$.

Find an labeling of vertices, $\ell : V \rightarrow X$ that minimizes

$\sum_{e=(u,v)} c(e) d(l(u), l(v)) + \sum_{v} c(v, l(v))$

Idea: find HST for metric (X, d).

Solve the problem on a hierarchically well separated tree metric.

Kleinberg-Tardos: constant factor on uniform metric.

Hierarchically well separated tree, “geometric”, constant factor.

$\rightarrow O(\log n)$ approximation.