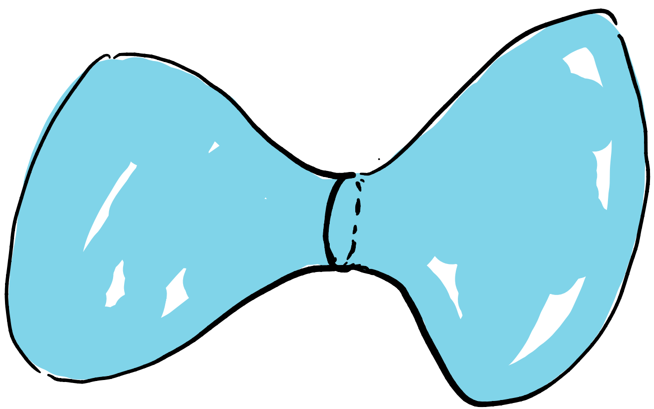
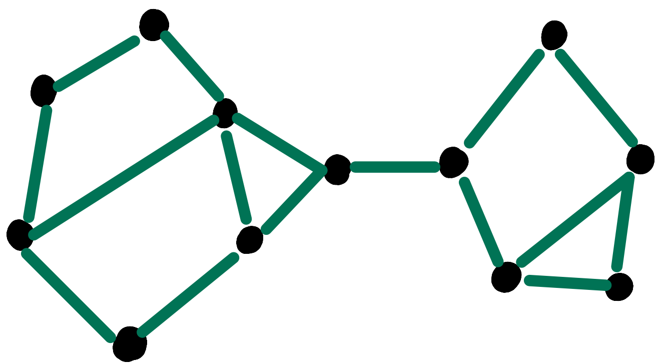
BOTTLENECKS

Bottle



Graph

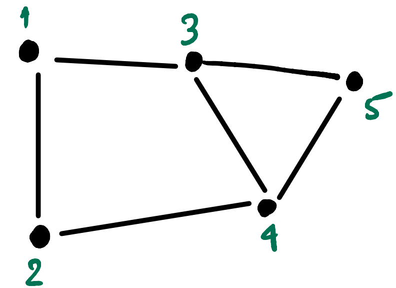


GRAPH IS OFTEN JUST A MATRIX

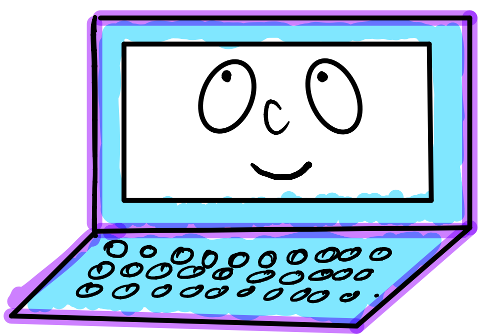
$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

Adjacency matrix

vs.

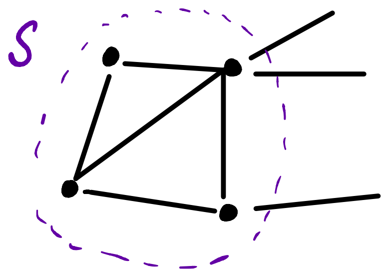


Drawing



HOW TO DEFINE BOTTLENECKS?

- Let $G=(V,E)$ be a d-regular graph. Let $S \subseteq V$.

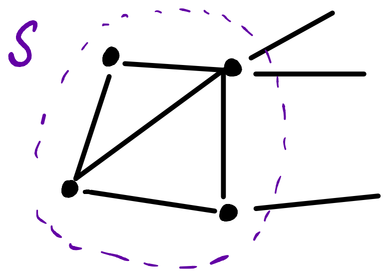


Expansion of S

$$\phi(S) = \frac{E(S,\bar{S})}{d|S|}$$

HOW TO DEFINE BOTTLENECKS?

- Let $G=(V,E)$ be a d-regular graph. Let $S \subseteq V$.



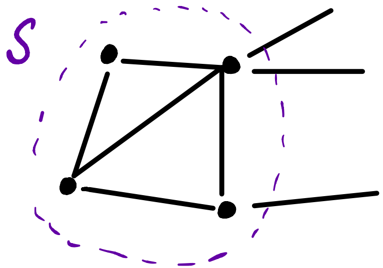
Expansion of S

$$\phi(S) = \frac{E(S,\bar{S})}{d|S|}$$

- Note $d|S|$ is the maximum number of edges that could go out of S.

HOW TO DEFINE BOTTLENECKS?

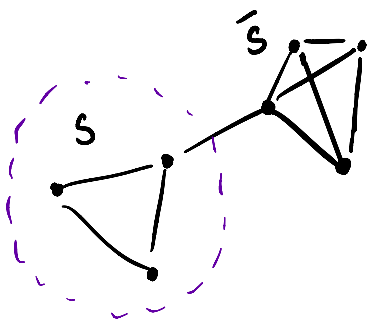
- Let $G=(V,E)$ be a d-regular graph. Let $S \subseteq V$.



Expansion of S

$$\phi(S) = \frac{E(S,\bar{S})}{d|S|}$$

- Note $d|S|$ is the maximum number of edges that could go out of S.

- 

Expansion of a cut $(S, V-S)$

$$\phi(S,\bar{S}) := \max\{\phi(S), \phi(\bar{S})\} = \frac{E(S,\bar{S})}{d \cdot \min\{|S|, |\bar{S}|\}}$$

CHEEGER'S INEQUALITY

- Let $\mathcal{L}$ denote the normalized Laplacian. $\mathcal{L} := I - \frac{1}{d}A$.

CHEEGER'S INEQUALITY

- Let $\mathcal{L}$ denote the normalized Laplacian. $\mathcal{L} := I - \frac{1}{d}A$.
- Let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ denote the eigenvalues of $\mathcal{L}$.

CHEEGER'S INEQUALITY

- Let $\mathcal{L}$ denote the normalized Laplacian. $\mathcal{L} := I - \frac{1}{d}A$.
- Let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ denote the eigenvalues of $\mathcal{L}$.

$$\boxed{\frac{\lambda_2}{2} \leq \phi(G) \leq \sqrt{2\lambda_2}}$$

$$\phi(G) = \min_{S : |S| \leq \frac{|V|}{2}} \phi(S)$$

WHY LAPLACIAN ?

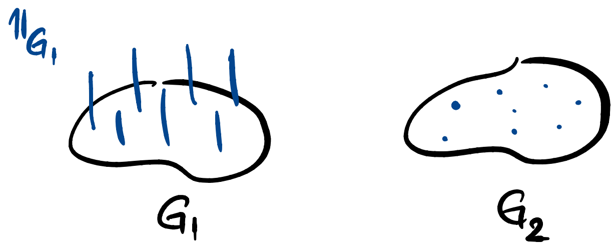
- $\mathcal{L} = I - \frac{1}{d} A.$

Some calculation $\Rightarrow \forall f \in \mathbb{R}^V : \langle \mathcal{L} f, f \rangle = \frac{1}{d} \sum_{(u,v) \in E} (f(u) - f(v))^2.$

WHY LAPLACIAN ?

- $\mathcal{L} = I - \frac{1}{d} A.$

  Some calculation $\Rightarrow \forall f \in \mathbb{R}^V : \langle \mathcal{L} f, f \rangle = \frac{1}{d} \sum_{(u,v) \in E} (f(u) - f(v))^2$

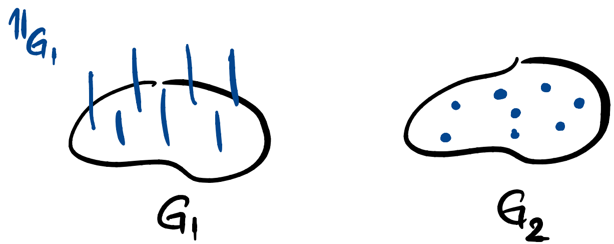
- 

$\langle \mathcal{L} \mathbb{1}_{G_1}, \mathbb{1}_{G_1} \rangle = 0$ and $\langle \mathcal{L} \mathbb{1}_{G_2}, \mathbb{1}_{G_2} \rangle = 0$

WHY LAPLACIAN ?

- $\mathcal{L} = I - \frac{1}{d}A.$

  Some calculation $\Rightarrow$ $\forall f \in \mathbb{R}^V: \langle \mathcal{L}f, f \rangle = \frac{1}{d} \sum_{(u,v)\in E} (f(u)-f(v))^2$

- 

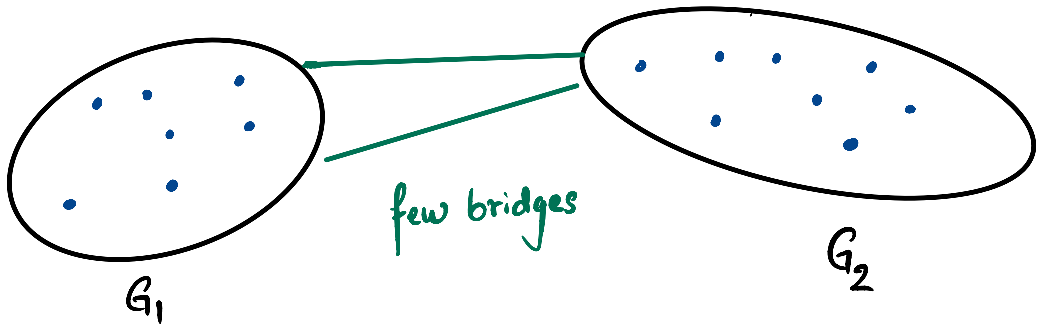
  $\langle \mathcal{L}\mathbb{1}_{G_1}, \mathbb{1}_{G_1} \rangle = 0$ and $\langle \mathcal{L}\mathbb{1}_{G_2}, \mathbb{1}_{G_2} \rangle = 0$

- Multiplicity of 0 is the number of connected components.

  $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0 \Rightarrow$ k components

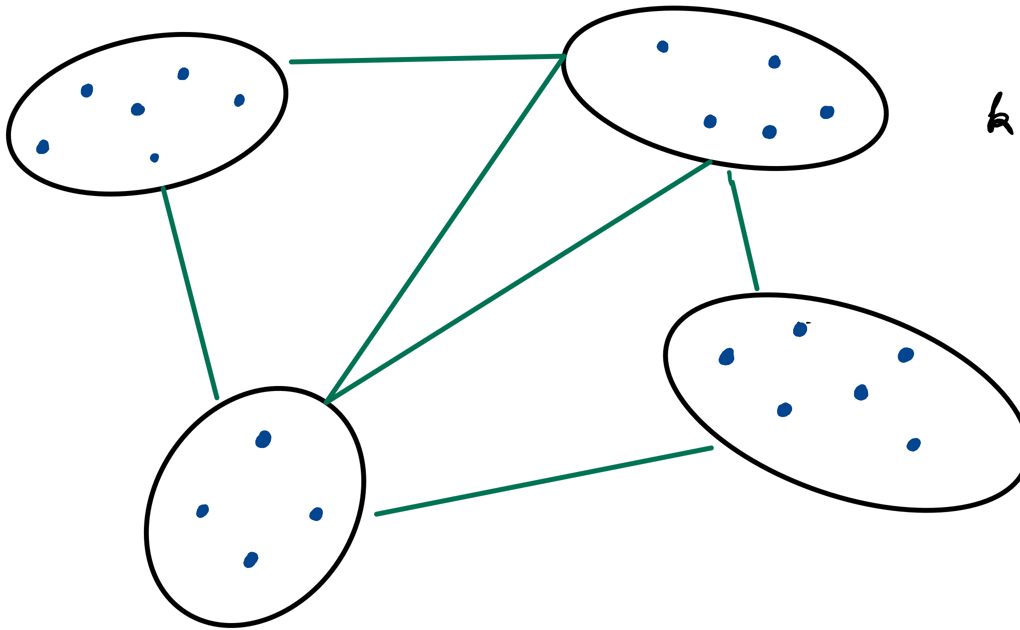
WHAT IF λ_2 IS CLOSE TO ZERO?

Wishful thinking



WHAT IF λ_k IS CLOSE TO ZERO?

Wishful thinking



k components

HIGHER ORDER CHEEGER

<u>Thm</u>. For every graph G, and every $k \in \mathbb{N}$, we have

$$\frac{\lambda_k}{2} \le \rho_G(k) \le O(k^2)\sqrt{\lambda_k}$$

$$\rho_G(k) = \min_{\substack{S_1,\dots,S_k \\ \text{disjoint}}} \max\{\phi_G(S_i) : i = 1, 2, \dots, k\}$$

$$\phi_G(S_i) = \frac{w(E(S,\bar{S}))}{w(S)}$$

WEAKER VERSION

Thm. For any weighted graph $G = (V, E, \omega)$, there exists a partition $V = S_1 \cup S_2 \cup \dots \cup S_k$, such that

$$\phi_G(S_i) \lesssim k^4 \sqrt{\lambda_k}$$

WEAKER VERSION PROOF

Thm. For any weighted graph $G = (V, E, \omega)$, there exists a partition $V = S_1 \cup S_2 \cup \dots \cup S_k$, such that

$$\phi_G(S_i) \lesssim k^4 \sqrt{\lambda_k}$$

Step 1: It suffices to find disjointly supported functions $\psi_1, \dots, \psi_k : V \to \mathbb{R}$ such that $R_G(\psi_i) \lesssim k^8 \lambda_k$

DISJOINT SETS TO DISJOINT FUNCTIONS

<u>Thm</u> For any $\psi : V \to \mathcal{R}$, there exists a subset $S \subseteq \mathrm{supp}(\psi)$ with

$$\phi_G(S) \leq \sqrt{2 R_G(\psi)}$$

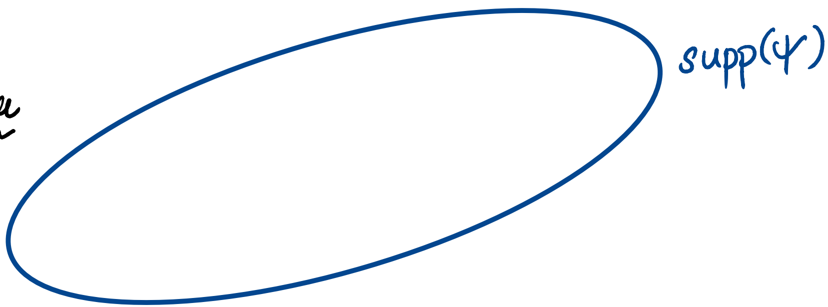
DISJOINT SETS TO DISJOINT FUNCTIONS

Thm For any $\psi : V \to \mathbb{R}$, there exists a subset $S \subseteq \mathrm{supp}(\psi)$ with

$$\phi_G(S) \le \sqrt{2 R_G(\psi)}$$

$$R_G(\psi) = \frac{\sum_{u \sim v} w(u,v) \|\psi(u) - \psi(v)\|^2}{\sum_u w(u) \|\psi(u)\|^2}$$

Idea



DISJOINT SETS TO DISJOINT FUNCTIONS

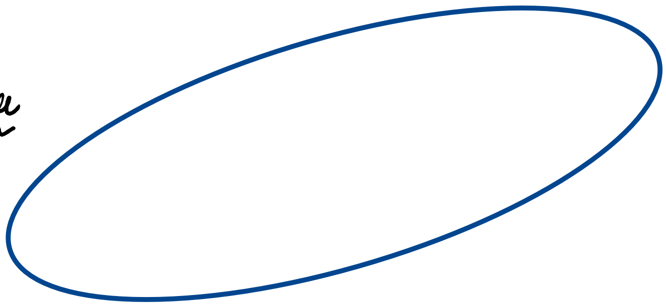
Thm For any $\psi : V \to \mathbb{R}$, there exists a subset $S \subseteq supp(\psi)$ with

$$\phi_G(S) \leq \sqrt{2 R_G(\psi)}$$

$$R_G(\psi) = \frac{\sum_{u \sim v} w(u,v) \|\psi(u) - \psi(v)\|^2}{\sum_u w(u) \|\psi(u)\|^2}$$

$= \dfrac{\text{Variation of } \psi \text{ across edges}}{\text{Total variation}}$

Idea



$supp(\psi)$

DISJOINT SETS TO DISJOINT FUNCTIONS

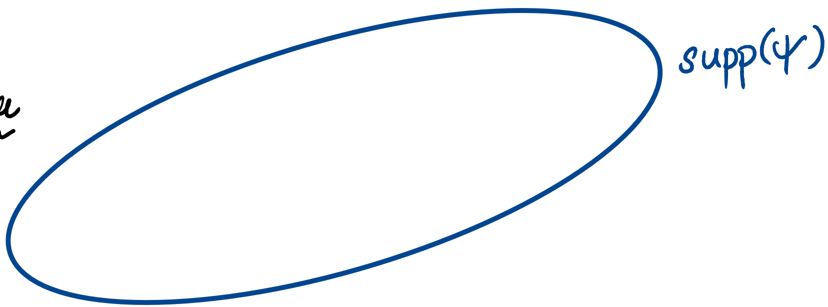
Thm For any $\psi : V \to \mathbb{R}$, there exists a subset $S \subseteq \mathrm{supp}(\psi)$ with

$$\phi_G(S) \leq \sqrt{2 R_G(\psi)}$$

$$R_G(\psi) = \frac{\sum_{u \sim v} w(u,v) \|\psi(u) - \psi(v)\|^2}{\sum_u w(u) \|\psi(u)\|^2}$$

$$= \frac{\text{Variation of } \psi \text{ across edges}}{\text{Total variation}}$$

Idea



If $R_G(\psi)$ is small, maybe we can find a set $S \subseteq \mathrm{supp}(\psi)$ such that $\phi_G(S)$ is small.

DISJOINT SETS TO DISJOINT FUNCTIONS

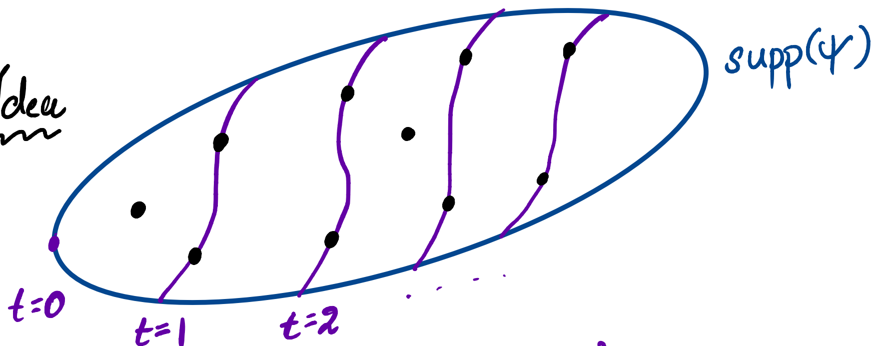
<u>Thm</u> For any $\psi : V \to \mathbb{R}$, there exists a subset $S \subseteq \text{supp}(\psi)$ with

$$\phi_G(S) \leq \sqrt{2 R_G(\psi)}$$

$$R_G(\psi) = \frac{\sum_{u \sim v} w(u,v) \|\psi(u) - \psi(v)\|^2}{\sum_u w(u) \|\psi(u)\|^2}$$

$$= \frac{\text{Variation of } \psi \text{ across edges}}{\text{Total variation}}$$

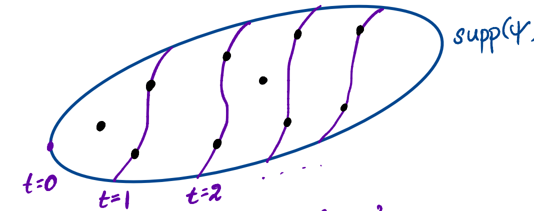
Idea



$$S_t = \{ u \in V : \|\psi(u)\|^2 > t \}$$

If $R_G(\psi)$ is small, maybe we can find a set $S \subseteq \text{supp}(\psi)$ such that $\phi_G(S)$ is small.

$$\int_0^\infty w(S_t) = \sum_{u \in V} w(u) \|\psi(u)\|^2$$



$$S_t = \{ u \in V : \|\psi(u)\|^2 > t \}$$

$$\int_0^\infty w(E(S_t, \bar{S}_t))\, dt = \sum_{u \sim v} w(u,v) \left| \|\psi(u)\|^2 - \|\psi(v)\|^2 \right|$$

$$\leq \sum_{u \sim v} w(u,v) \|\psi(u) - \psi(v)\| \, \|\psi(u) + \psi(v)\|$$

$$\leq \sqrt{\sum_{u \sim v} w(u,v) \|\psi(u) - \psi(v)\|^2} \sqrt{\sum_{u \sim v} w(u,v) \|\psi(u) + \psi(v)\|^2}$$

$$\leq \sqrt{\sum_{u \sim v} w(u,v) \|\psi(u) - \psi(v)\|^2} \sqrt{2 \sum_{u \in V} w(u) \|\psi(u)\|^2}$$

So,

$$\frac{\int_0^\infty w(E(S_t, \bar{S}_t))\, dt}{\int_0^\infty w(S_t)\, dt} \leq \sqrt{2 R_G(\psi)}$$

$\Rightarrow \exists\, t \in [0, \infty]$ for which $\phi(S_t) \leq \sqrt{2 R_G(\psi)}$

GRAPH THEORY TO LINEAR ALGEBRA

Take orthonormal eigenfunctions of $\mathcal{L}$ — $f_1, f_2, \ldots, f_k : V \to \mathbb{R}$ (f_i has eigenvalue λ_i).

GRAPH THEORY TO LINEAR ALGEBRA

Take orthonormal eigenfunctions of $\mathcal{L}$ — $f_1, f_2, \ldots, f_k : V \to \mathbb{R}$ (f_i has eigenvalue λ_i).

Spectral embedding — $F : V \to \mathbb{R}^k$

$$F(v) = (f_1(v), f_2(v), \ldots, f_k(v))$$

GRAPH THEORY TO LINEAR ALGEBRA

Take orthonormal eigenfunctions of $\mathcal{L}$ — $f_1, f_2, \ldots, f_k : V \to \mathbb{R}$ (f_i has eigenvalue λ_i).

Spectral embedding — $F : V \to \mathbb{R}^k$

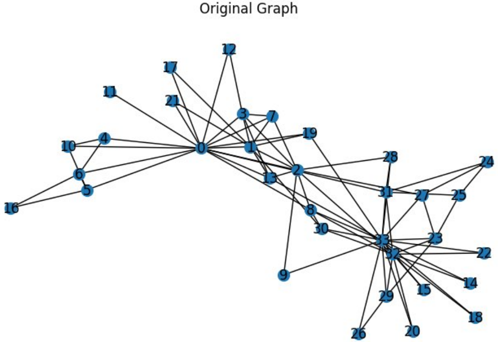
$$F(v) = (f_1(v), f_2(v), \ldots, f_k(v))$$

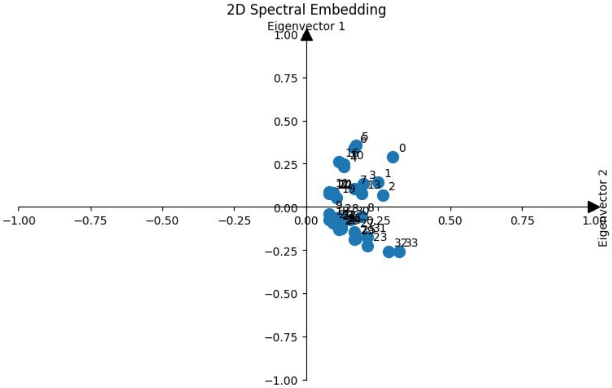
Observe

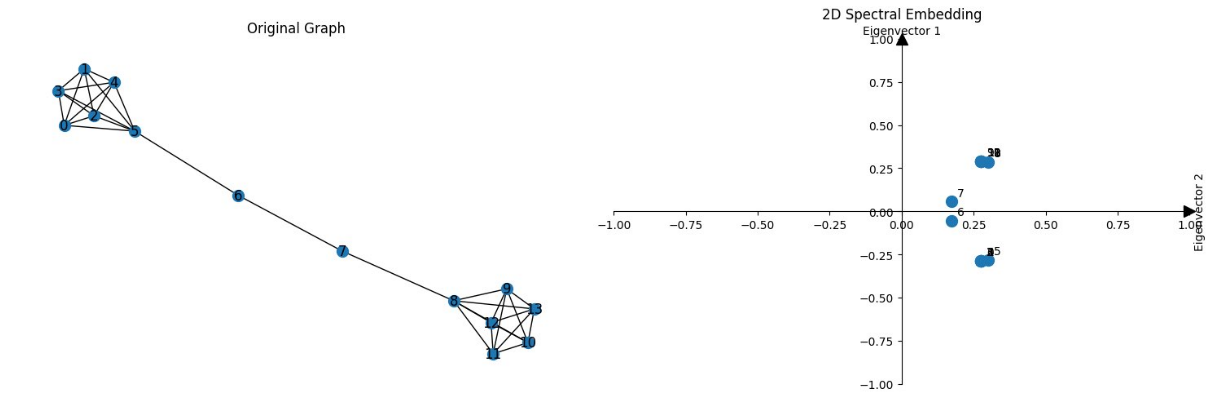
$$R_G(F) = \frac{\sum_{u \sim v} w(u,v) \, \|F(u) - F(v)\|^2}{\sum_{u \in V} w(u) \, \|F(u)\|^2}$$

$$= \frac{\sum_{i=1}^{k} \sum_{u \sim v} w(u,v) \, |f_i(u) - f_i(v)|^2}{\sum_{i=1}^{k} \sum_{u \in V} w(u) \, f_i(u)^2}$$

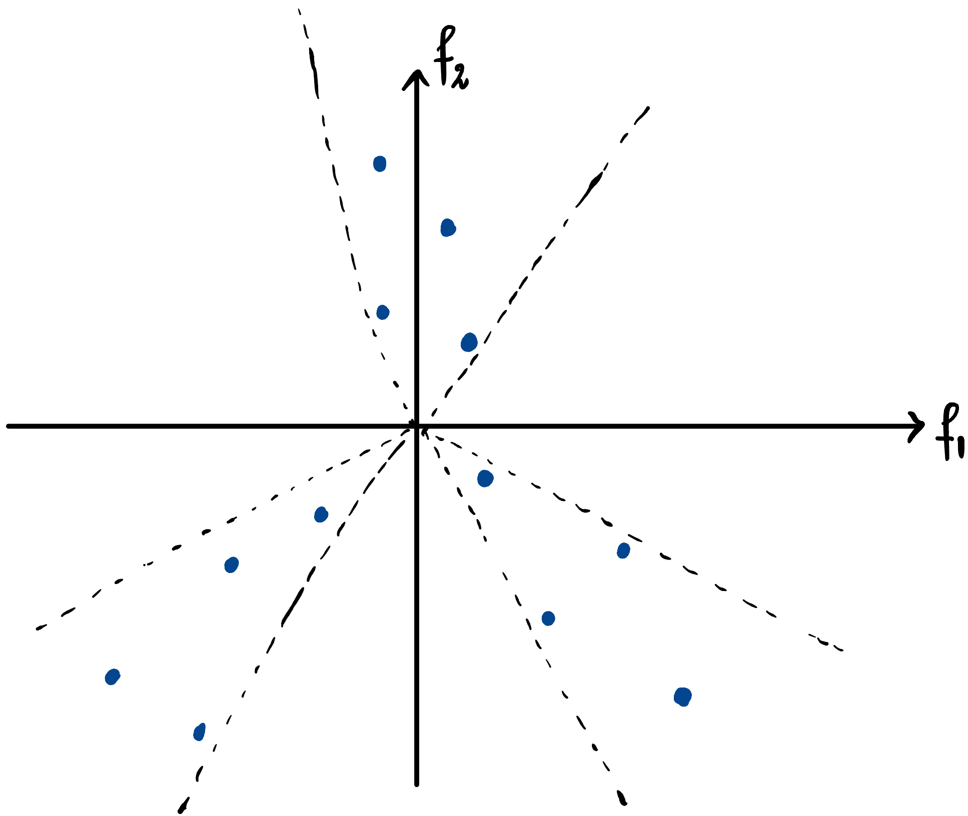
$$= \frac{\lambda_1 + \cdots + \lambda_k}{k} \leq \lambda_k$$


Original Graph
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33


2D Spectral Embedding
Eigenvector 1
Eigenvector 2
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00


Original Graph
0
1
2
3
4
5
6
7
8
9
10
11
12
13
2D Spectral Embedding
Eigenvector 1
Eigenvector 2
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
7
6

FIND k REGIONS WITH LARGE CONCENTRATION



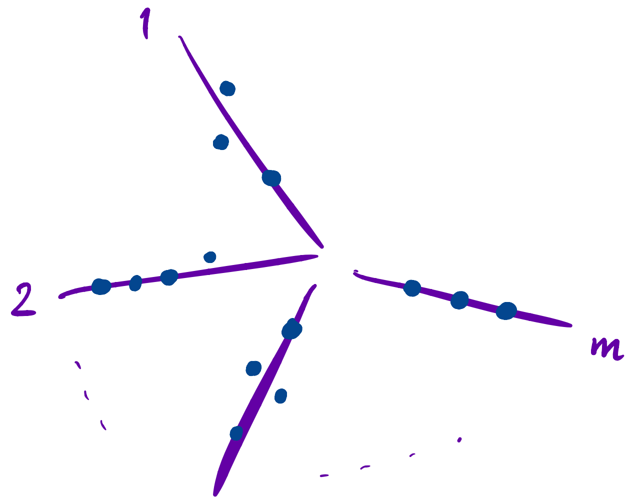
ISOTROPY PROPERTY

We can show that $\sum_{v \in V} \langle x, F(v) \rangle^2 = 1$, for any $x \in S^{k-1}$. Also, $\sum_{v \in V} \|F(v)\|^2 = k$

ISOTROPY PROPERTY

We can show that $\sum_{v \in V} \langle x, F(v) \rangle^2 = 1$, for any $x \in S^{k-1}$. Also, $\sum_{v \in V} \|F(v)\|^2 = k$

This implies

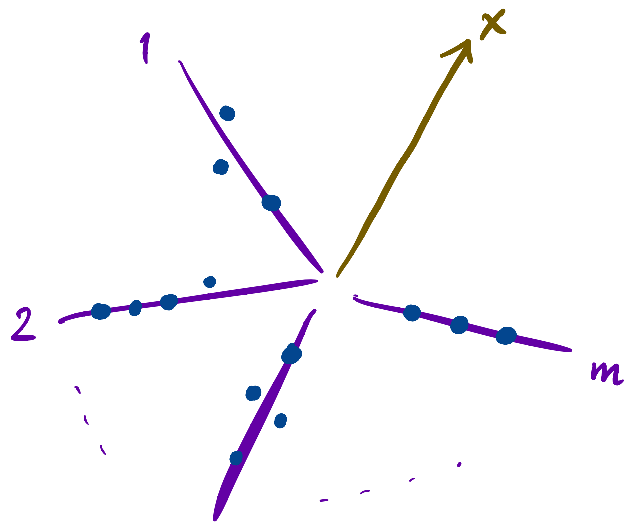


$m < k$

ISOTROPY PROPERTY

We can show that $\sum_{v \in V} \langle x, F(v) \rangle^2 = 1$, for any $x \in S^{k-1}$. Also, $\sum_{v \in V} \|F(v)\|^2 = k$

This implies



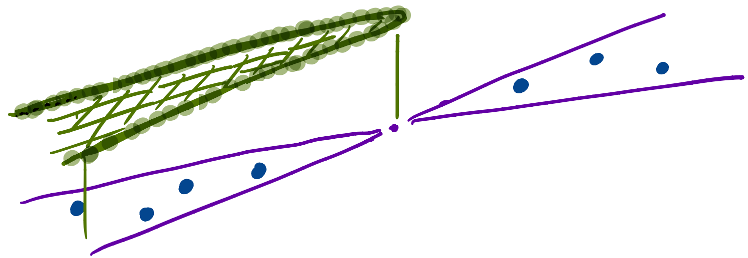
$m < k$

$$\sum_{x \in V} \langle x, F(v) \rangle^2 \approx 0 .$$

F cannot concentrate along fewer than k lines.

FIRST APPROACH TO GET ψ_i

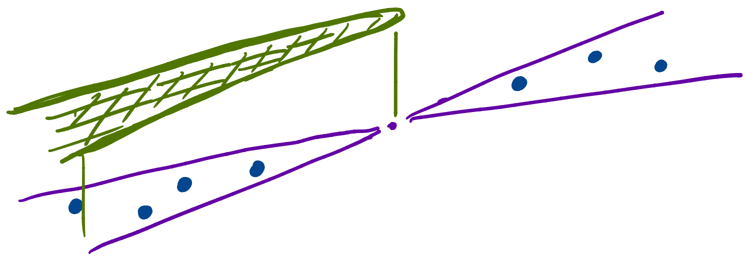
Find k-directions: $x_1, \ldots, x_k$



$$\psi_i(v) = \begin{cases} F(v) & \text{if } F(v) \text{ has large projection on } x_i \\ 0 & \text{otherwise} \end{cases}$$

FIRST APPROACH TO GET ψ_i

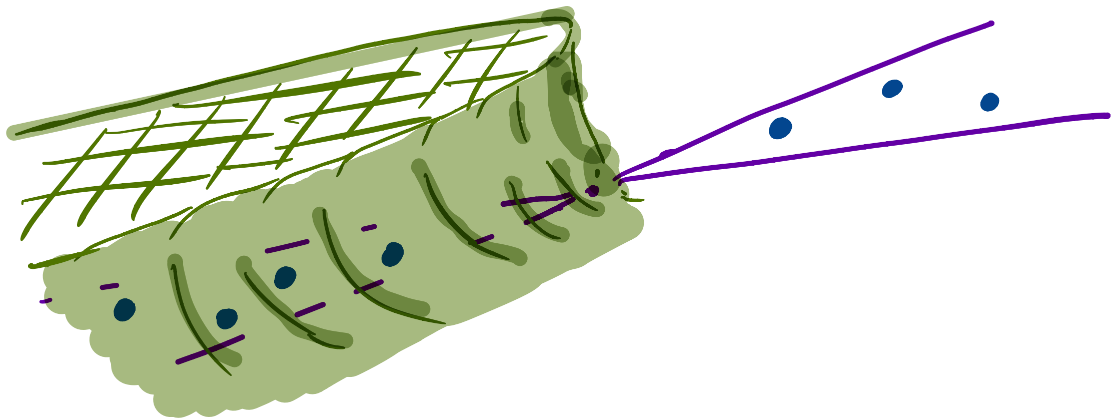
Find k-directions: $x_1, \dots, x_k$



$$\psi_i(v) = \begin{cases} F(v) & \text{if } F(v) \text{ has large projection on } x_i \\ 0 & \text{otherwise} \end{cases}$$

Cutoff is too sharp! $\sum_{\{u,v\}\in E} \|\psi_i(u) - \psi_i(v)\|^2 \gg \sum_{\{u,v\}\in E} \|F(u) - F(v)\|^2$

SMOOTH IT OUT



New distance:

$$d_F(u,v) = \left\| \frac{F(u)}{\|F(u)\|} - \frac{F(v)}{\|F(v)\|} \right\|$$

Radial distance

GOAL

Find k-regions $S_1, S_2, \ldots, S_k$ such that

GOAL

Find k-regions $S_1, S_2, \ldots, S_k$ such that

- each region contains a large fraction of the ℓ^2 mass of F.

GOAL

Find k-regions $S_1, S_2, \ldots, S_k$ such that

- each region contains a large fraction of the ℓ^2 mass of F.
- far enough apart to allow ψ_i to smooth out.

RANDOM PARTITIONS

Let (X,d) be a finite metric space. Let $B(x,R) = \{y \in X : d(x,y) \leq R\}$ denote the closed **ball** of radius R around x.

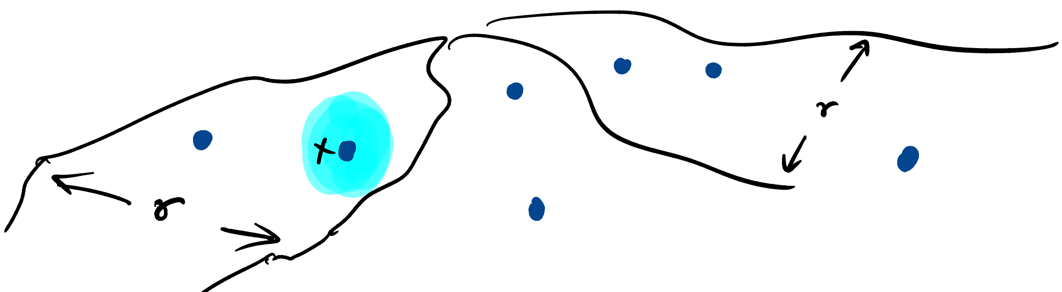
RANDOM PARTITIONS

- Let (X,d) be a finite metric space. Let $B(x,R) = \{y \in X : d(x,y) \leq R\}$ denote the closed ball of radius R around x.
- An (r,ε)-padded decomposition of a metric space (X,d) is a distribution μ over $\mathcal{P}$ (collection of partitions of X) satisfying

  1) Bounded diameter: $\mathrm{diam}(C) \leq r \;\; \forall$ Cluster C in every partition P in support of μ.

  2) Padding: $\Pr_{\mu}[\pi_P(x) \geq \varepsilon r] \geq \frac{1}{2} \;\; \forall\, x \in X$

  where $\pi_P(x) = \sup\{t : \exists C \in P \text{ with } B(x,t) \subset C\}$

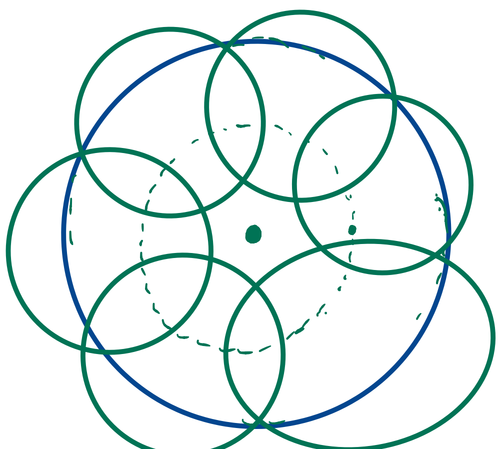


<u>Thm.</u> [Gupta, Krauthgamer, Lee] Let (X, d) be a finite metric space. Then for every $r > 0$, there exists an (r, ε)-padded probabilistic decomposition of X with $\frac{1}{\varepsilon} \leq 64 \dim(X)$. (Assume $X \subseteq \mathbb{R}^k$)

Thm. [Gupta, Krauthgamer, Lee] Let (X,d) be a finite metric space. Then for every $r>0$, there exists an (r,ε)-padded probabilistic decomposition of X with $\frac{1}{\varepsilon} \le 64\dim(X)$. (Assume $X \subseteq \mathbb{R}^K$)

Doubling dimension

- doubling constant is the smallest value λ such that every ball in X can be covered by λ balls of half the radius
- doubling dimension $\dim(X) = \log_2 \lambda$.

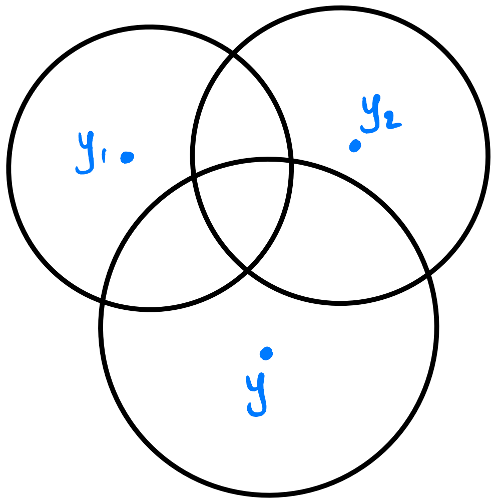


$\mathbb{R}^2$ has doubling constant 7

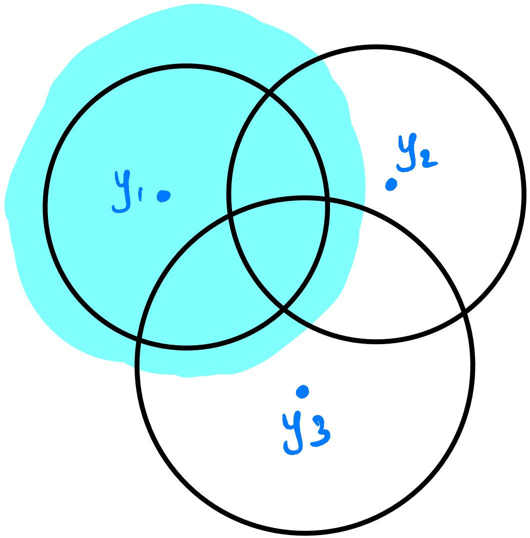
$\Rightarrow$ doubing dimension $\log_2 7$

$\mathbb{R}^K$ has doubling dimension $\Theta(K)$

Proof. Take a r-net N. $\operatorname{diam}(\text{Ball}) \le r$. Also ensure $d(y_i, y_j) \ge r$

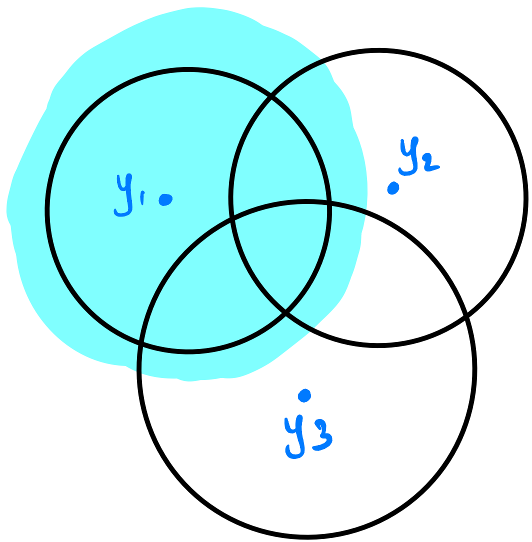


Proof. Take a r-net N. $\text{diam}(\text{Ball}) \leq r$. Also ensure $d(y_i, y_j) \geq r$



- Choose $R \sim U[r, 2r]$
  Create a ball $B(y, R)$ around all $y \in N$.

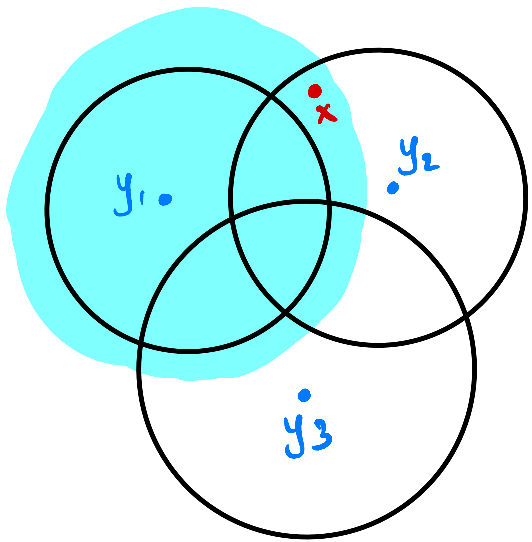
Proof: Take a r-net N. $\text{diam}(\text{Ball}) \le r$. Also ensure $d(y_i, y_j) \ge r$



- Choose $R \sim U[r, 2r]$
  Create a ball $B(y, R)$ around all $y \in N$.

- generate a random permutation σ
  of the net points

Proof: Take a r-net N. $\mathrm{diam}(\text{Ball}) \le r$. Also ensure $d(y_i, y_j) \ge r$



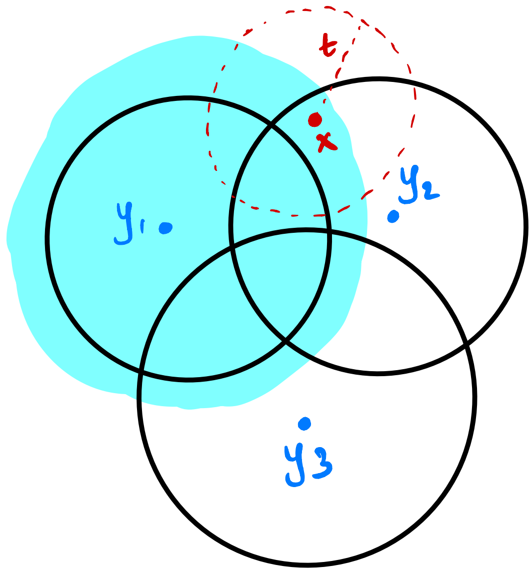
- Choose $R \sim U[r, 2r]$
  Create a ball $B(y, R)$ around all $y \in N$.

- generate a random permutation σ of the net points

- Define clusters:
  $\forall y \in N$, $C_y = \{ x \in X : x \in B(y,R)$ and $\sigma(y) < \sigma(z)$ for all other z where $x \in B(z,R) \}$

Example. $x \in C_{y_1}$ if $\sigma(y_1) < \sigma(y_2)$

Proof: Take a r-net N. $\text{diam}(\text{Ball}) \le r$. Also ensure $d(y_i, y_j) \ge r$



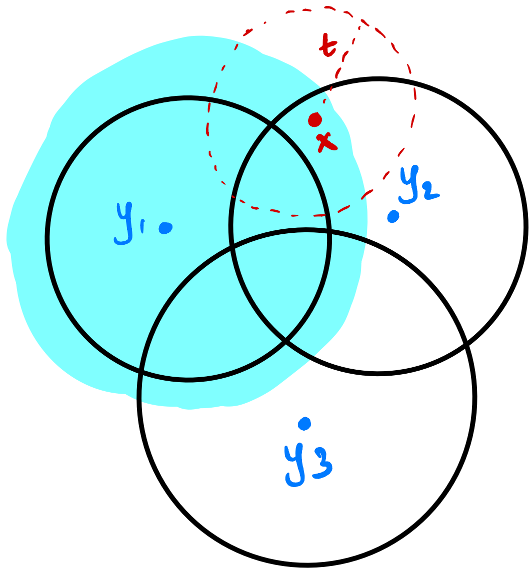
- Choose $R \sim U[r, 2r]$
  Create a ball $B(y, R)$ around all $y \in N$.

- generate a random permutation σ of the net points

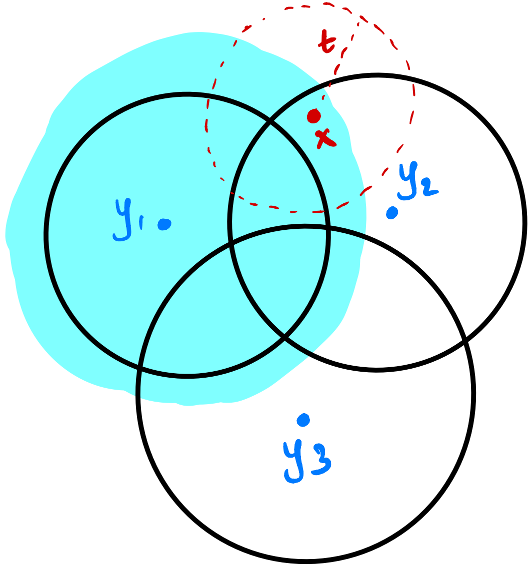
- Define clusters:
  $\forall y \in N$, $C_y = \{x \in X : x \in B(y, R)$ and $\sigma(y) < \sigma(z)$ for all other z where $x \in B(z, R)\}$

- For $x \in X$, take $B(x, t)$ with $t \le \varepsilon r$
  $W := B(x, 2r+t) \cap N$
  These are the only net points whose ball might cut $B(x, t)$

Proof. Take a r-net N. $\text{diam}(\text{Ball}) \le r$. Also ensure $d(y_i, y_j) \ge r$



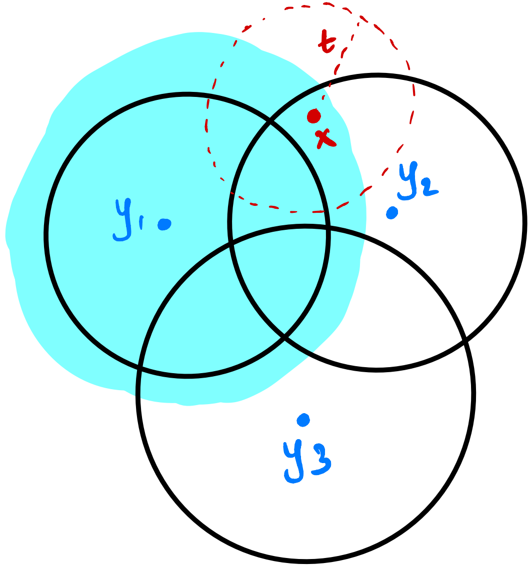
- Choose $R \sim U[r, 2r]$
  Create a ball $B(y, R)$ around all $y \in N$.
- generate a random permutation σ of the net points
- Define clusters:
  $\forall y \in N$, $C_y = \{ x \in X : x \in B(y, R)$ and $\sigma(y) < \sigma(z)$ for all other z where $x \in B(z, R) \}$
- For $x \in X$, take $B(x, t)$ with $t \le \varepsilon r$
  $W = B(x, 2r+t) \cap N$
  These are the only net points whose ball might cut $B(x, t)$
- $|W| \le 2^k$? skip.



- Order points in $W = \{w_1, \dots, w_m\}$ by distance from x.

  Let $I_k = [\, d(x, w_k) - t,\ d(x, w_k) + t \,]$

  Let E_k event that w_k's cluster cuts $B(x, t)$.



- Order points in $W = \{w_1, \dots, w_m\}$ by distance from x.

  Let $I_k = [\, d(x, w_k) - t,\ d(x, w_k) + t\,]$

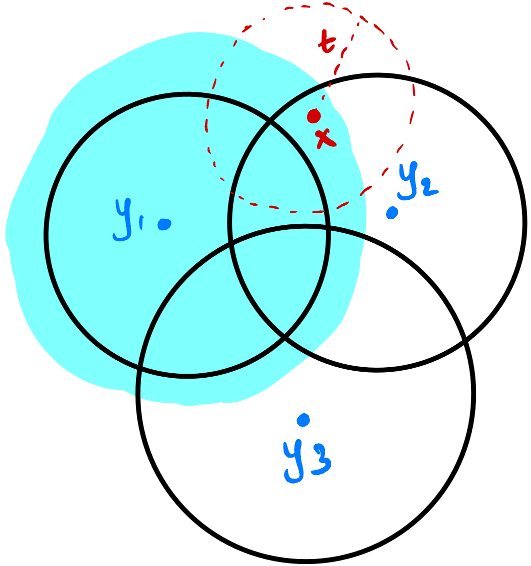
  Let E_k event that w_k's cluster cuts $B(x, t)$.

- $\Pr[E_k] = \Pr\big[\, R \in [d(x, w_k) - t, d(x, w_k) + t]$ and w_k has highest priority i.e $\sigma(w_k)$ is smallest $\big]$

$$\leq \Pr[R \in I_k] \cdot \Pr[\sigma(w_k) \text{ is smallest}]$$

$$= \frac{|I_k|}{|[r, 2r]|} \cdot \frac{1}{k}$$

$$= \frac{2t}{r} \cdot \frac{1}{k}$$



- Order points in $W = \{w_1, \dots, w_m\}$ by distance from x.

  Let $I_k = [\, d(x,w_k) - t,\ d(x,w_k) + t \,]$

  Let E_k event that w_k's cluster cuts $B(x,t)$.

- $$\Pr[E_k] = \Pr\Big[R \in [d(x,w_k)-t,\ d(x,w_k)+t] \text{ and } w_k \text{ has highest priority i.e } \sigma(w_k) \text{ is smallest} \Big]$$

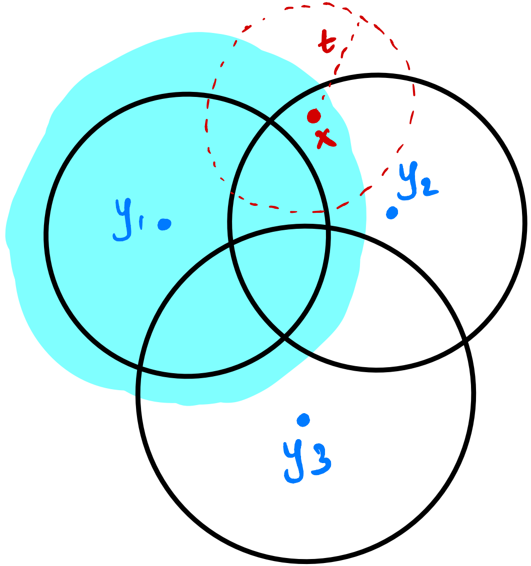
  $$\leq \Pr[R \in I_k] \cdot \Pr[\sigma(w_k) \text{ is smallest}]$$

  $$= \frac{|I_k|}{|[r, 2r]|} \cdot \frac{1}{k}$$

  $$= \frac{2t}{r} \cdot \frac{1}{k}$$

- Union Bound $\displaystyle\sum_{k=1}^{m} \Pr[E_k] \leq \sum_{k=1}^{m} \frac{2t}{r} \frac{1}{k}$

  $$\leq \frac{2t}{r}\left(1 + \ln m\right)$$



- Order points in $W = \{w_1, \dots, w_m\}$ by distance from x.

  Let $I_K = [\, d(x, w_K) - t,\ d(x, w_K) + t \,]$

  Let E_K event that w_K's cluster cuts $B(x, t)$.

- $\Pr[E_K] = \Pr\left[R \in [d(x, w_K) - t, d(x, w_K) + t] \text{ and } w_K \text{ has highest priority i.e. } \sigma(w_K) \text{ is smallest} \right]$

$$\leq \Pr[R \in I_K] \cdot \Pr[\sigma(w_K) \text{ is smallest}]$$

$$= \frac{|I_K|}{|[r, 2r]|} \cdot \frac{1}{K}$$

$$= \frac{2t}{r} \cdot \frac{1}{K}$$

- Union Bound $\sum_{K=1}^{m} \Pr[E_K] \leq \sum_{K=1}^{m} \frac{2t}{r} \frac{1}{K}$

$$\leq \frac{2t}{r} (1 + \ln m)$$

So, $\Pr[B(x,t) \text{ is cut}] \leq \frac{2t}{r} (1 + K \ln 2)$

Set $t = \frac{r}{8K}$

$$\Pr[B(x,t) \text{ is cut}] \leq \frac{2}{8K} (1 + K \ln 2)$$

$$= \frac{1}{4K} + \frac{\ln 2}{4} \leq \frac{1}{2}$$