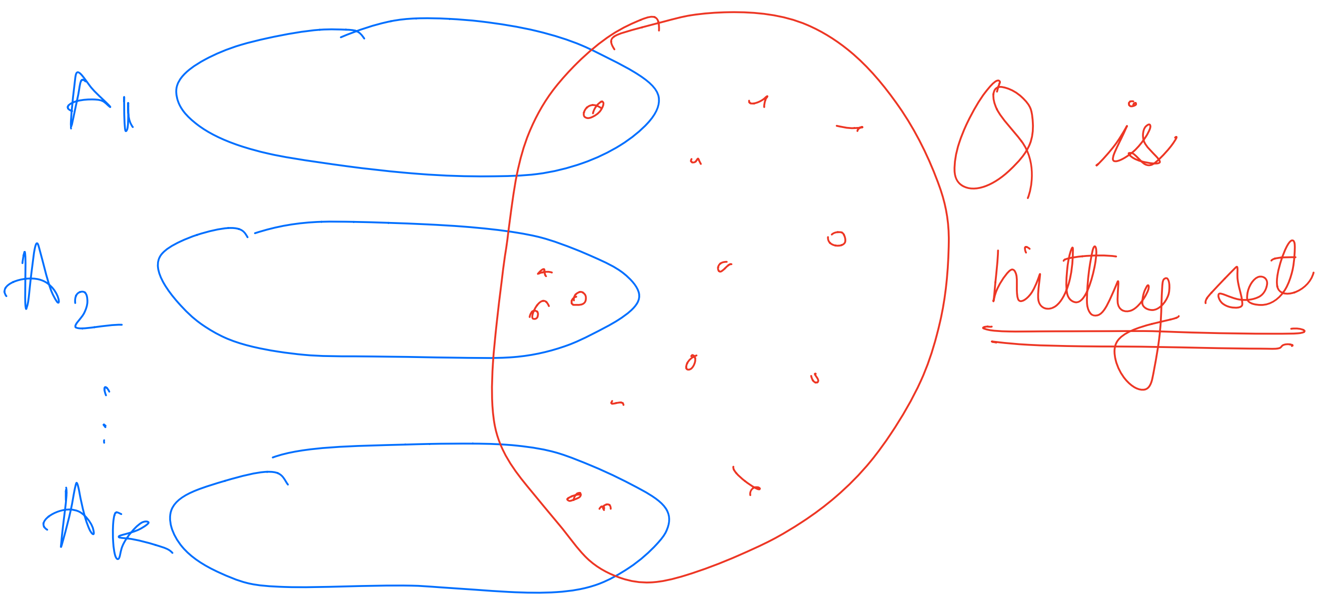
Today's class:

- (+2)-Additive spanners (i.e $\beta = 2$)
- Hitting Sets ← combinatorial object used in our +2 spanner.

Def^n (Hitting set): Let U be universe set, and $\mathcal{F} = \{A_1, A_2, \ldots, A_k\}$ be a family of sets such that $A_1, \ldots, A_k \subseteq U$.

Then a set $Q \subseteq U$ is said be hitting set for $\mathcal{F}$ iff $A_i \cap Q$ is non-empty, for $1 \le i \le k$.



Sparse subgraph "H" of G preserving all-pairs distances approximately

$\equiv$

Spanners

β-Additive Spanner

H satisfies that

$\forall x, y$

$$d_H(x,y) \le d_G(x,y) + \beta$$

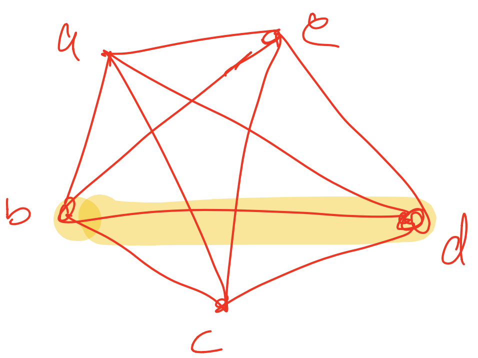
α-Multiplicative Spanner

H satisfies that $\forall x, y$

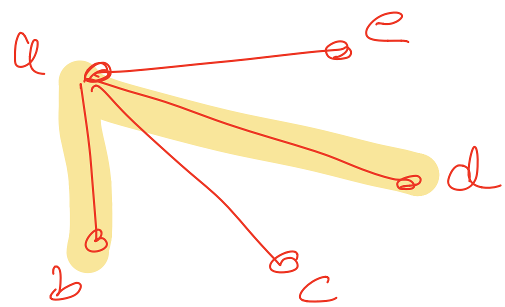
$$d_H(x,y) \le \alpha \cdot d_G(x,y)$$

Eg. $G = K_n$

(complete graph of size n)



H = Star graph



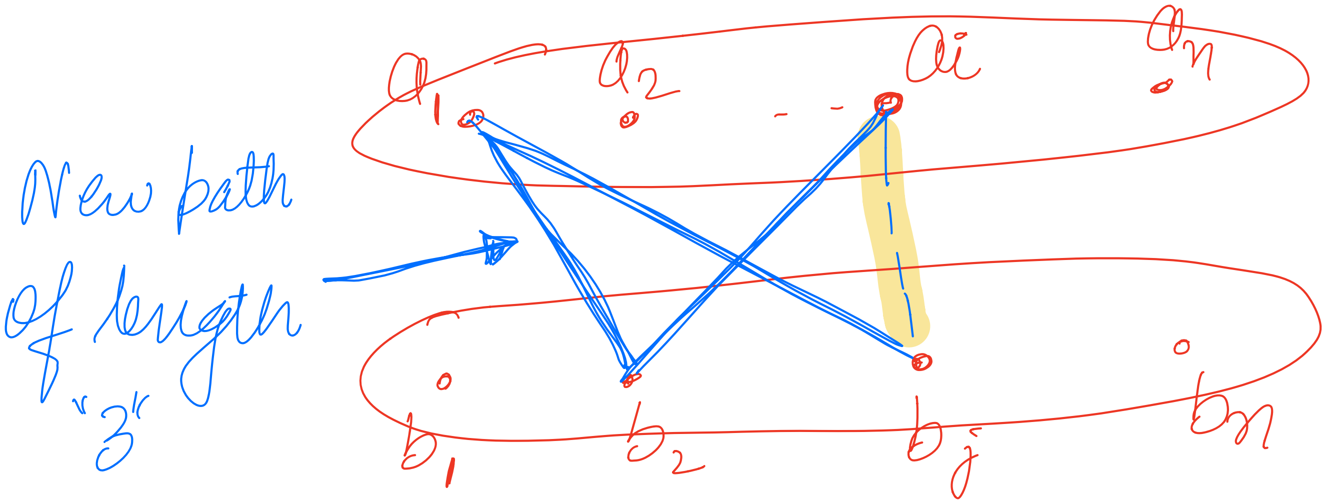
distortion $\le +1$

$\therefore \beta = 1$

Ques: Can we have $\beta \lneq 2$?

Ans: No. The counter eg. is $K_{n,n}$

$K_{n,n}$ — $2n$ vertices, n^2 edges



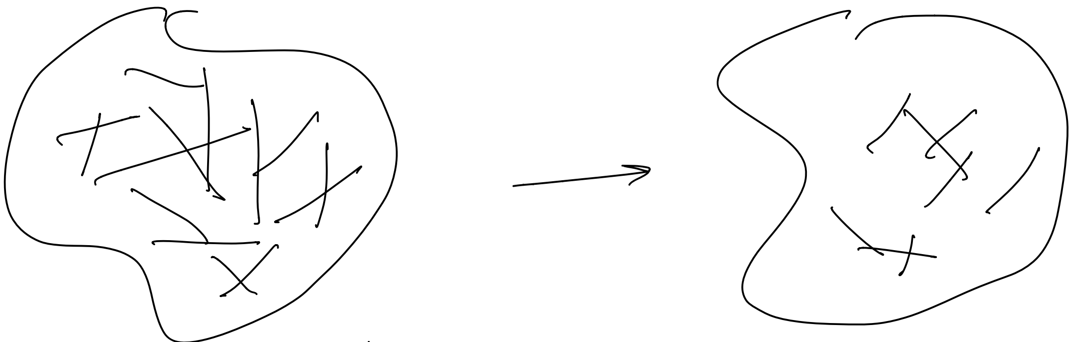
Let $H := G$ minus the (a_i, b_j)

$Dist(a_i, b_j, G) = 1$	$Dist(a_i, b_j, H) = 3$

So stretch is $+2$ even when H has $(n^2 - 1)$ edges.

This class:

"+2"-additive spanner with $O(n^{1.5} \log n)$ edges.



$n^{1.5}$ edges

G can be very dense & have n^2 edges

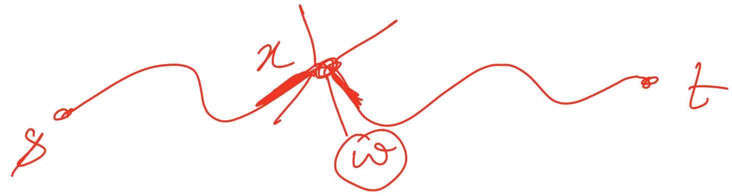
Subgraph H is relatively very sparse and yet distortion = +2

Notation

$\deg(v, G) \equiv$ degree of v in G

$BFS(v, G) \equiv$ BFS-tree of v in G

Theorem 1:



For a pair (s,t) with P as shortest path if subgraph H contains each edge incident to $x \in P$, then

$$d_H(s,t) = d_G(s,t).$$

Theorem 2:

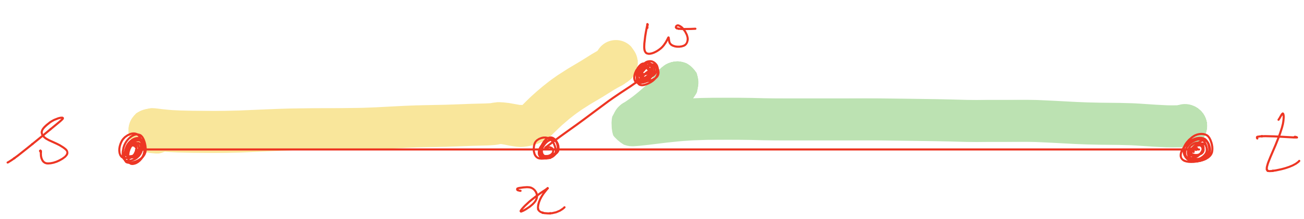
For a pair (s,t) with P as shortest path if subgraph H contains $BFS(w,G)$ for some $w \in \bigcup_{x \in P} N[x]$, then

edges of tree

$$d_H(s,t) \leq d_G(s,t) + 2.$$

Proof of Thm 2

Let 'x' be a vertex on s–t shortest path in G.
Further let $w \in N(x)$ (i.e. let w be a neighbor of x) and suppose H contain $BFS(w, G)$.



Then,

(i) $d_G(s,w) + d_G(w,t) \leq d_G(s,t) + 2$

<u>Proof of (i)</u>: By Δ-inequality,

$$d_G(s,w) \leq d_G(s,x) + 1 \quad \text{—— (1)}$$

$$d_G(w,t) \leq d_G(x,t) + 1 \quad \text{—— (2)}$$

On adding (1), (2), we get

$$d_G(s,w) + d_G(w,t) \leq d_G(s,t) + 2$$

(ii) $d_H(s,t) \leq d_G(s,t) + 2$.

<u>Proof of (ii)</u>:

$$d_H(s,t) \leq d_H(s,w) + d_H(w,t) \quad \text{(By } \Delta\text{-inequality)}$$

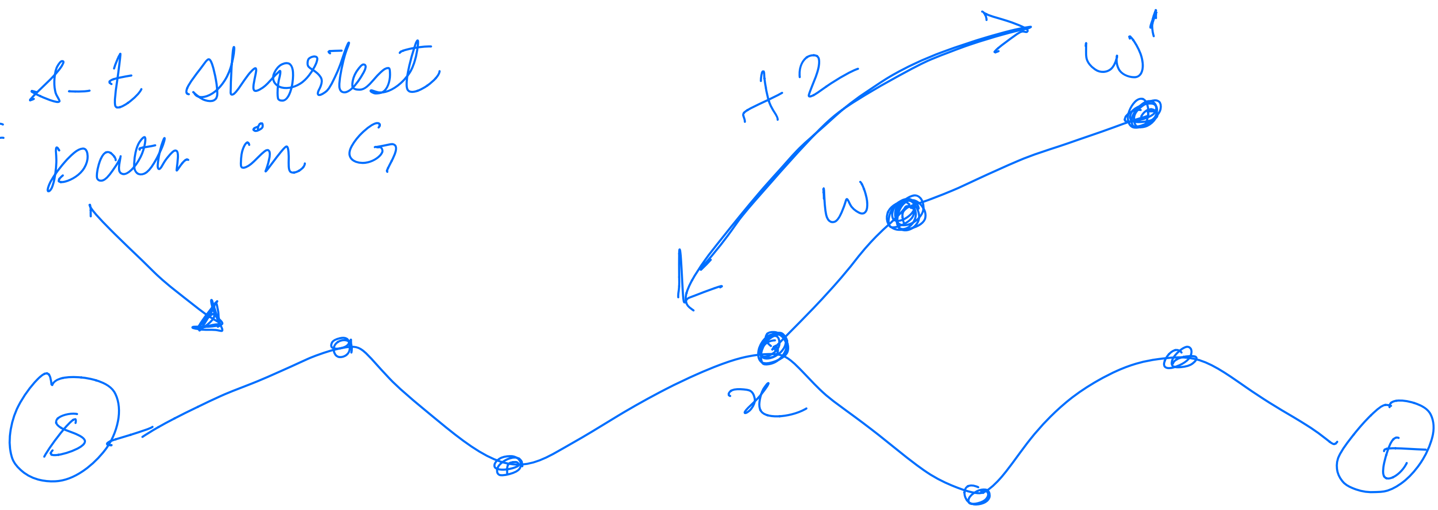
$$= d_G(s,w) + d_G(w,t) \quad \left(\begin{array}{c}H \text{ contains} \\ BFS(w,G)\end{array}\right)$$

$$\leq d_G(s,t) + 2$$

—□

Exercise:

$P \equiv$ s-t shortest path in G



If $BFS(w', G)$ lies in H, and $w' \in N(w)$ for some $w \in \bigcup_{x \in P} N(x)$,

then prove that

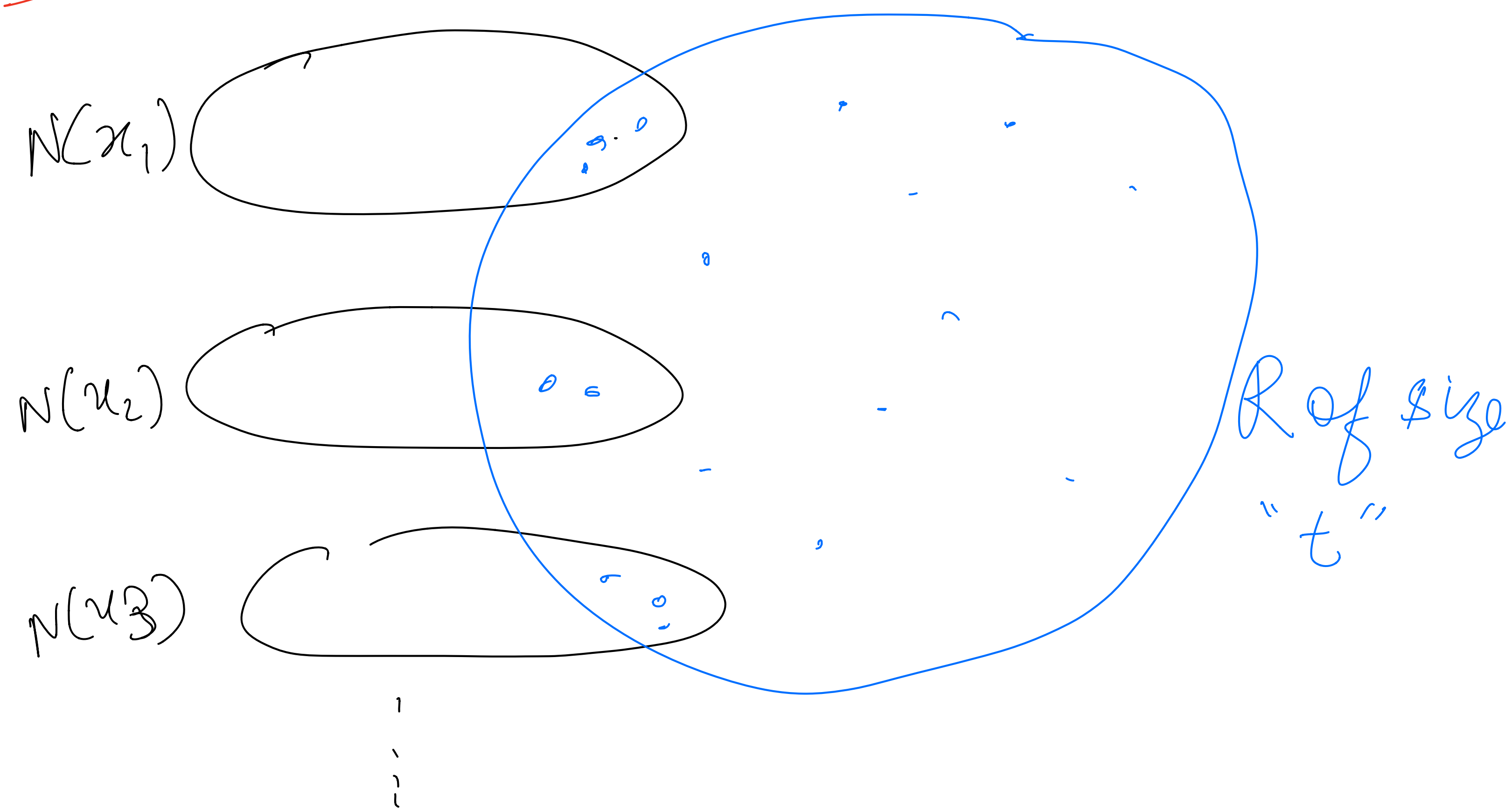
$$d_H(s,t) \le 4 + d_G(s,t)$$

Theorem 3 : $t \leftarrow$ integer

Let $R \subseteq V$ be a uniformly random set of size "t". Then R is a hitting-set for

$$\mathcal{F} = \{ N(x) \mid x \in V(G) \text{ has degree} \geq \frac{n \log n}{t} \}$$

What it means? $x_1, x_2, x_3 \dots \leftarrow$ vertices of deg $\geq \frac{n \log n}{t}$



We next provide algorithm based on Theorems 1, 2, 3.

Algorithm to compute "+2" additive spanner.

1. Initialize $H = (V, E_H = \emptyset)$
2. $R \leftarrow$ uniform random subset of V of size $\sqrt{n}$.
3. For each $w \in R$:

   Add edges of $BFS(w, G)$ to E_H
4. For each $v \in V(G)$ with $deg(v, G) \leq 4\sqrt{n}\log n$:

   Add to E_H all edges incident to v in G.
5. Return H

Time complexity $= O((m+n)|R|) = O\left(m\sqrt{n} + n\sqrt{n}\right)$

Size of H (Edges count) $\leq n|R| + n(4\sqrt{n}\log n)$

$= O(n\sqrt{n}\log n)$

Correctness: Consider a pair $(s,t) \in V \times V$, and let P be s-t shortest path in G

CASE 1: All vertices in P have degree $\leq 4\sqrt{n}\log n$

$\Rightarrow$ Edges incident to all vertices in P lie in H

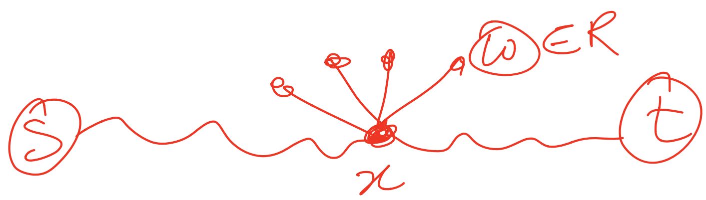
$\Rightarrow$ By Theorem 1,

$$d_H(s,t) = d_G(s,t)$$

—— x ——

CASE 2: At least one vertex in P has degree more than $4\sqrt{n}\log n$, let this vertex be x.

By theorem 3, $R \cap N(x)$ contains a vertex (say w).



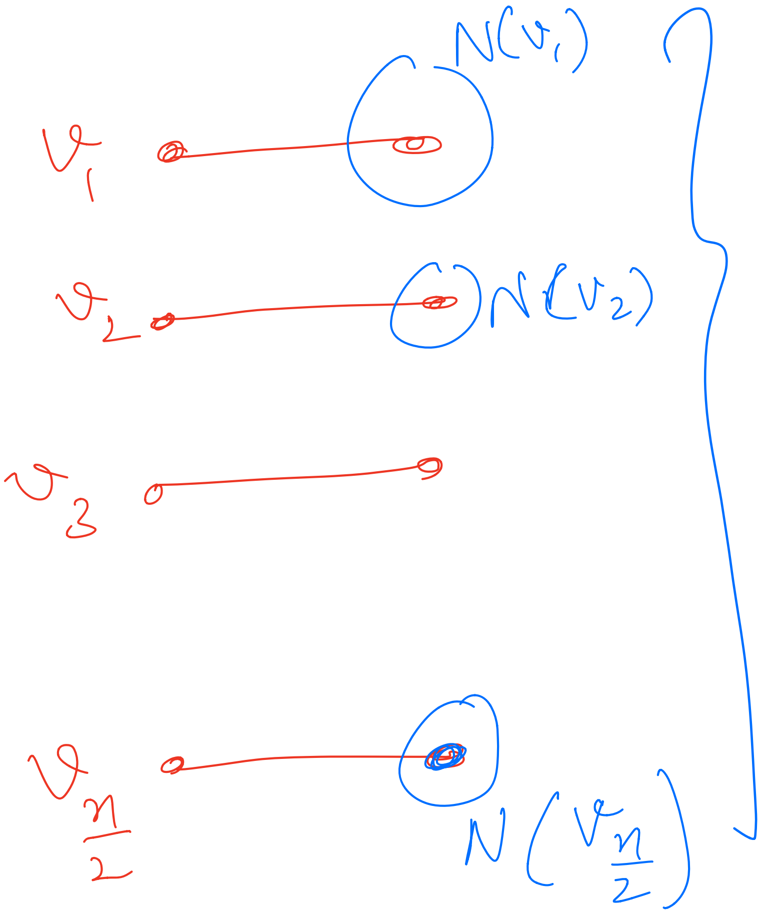
As BFS(w, G) lie in H, by Theorem 2,

$$d_H(s,t) \leq d_G(s,t) + 2.$$

Question: Why can't we use hitting set idea to take care of vertices of small degree?

Ans: Not possible bcoz then size of hitting set will be large.

For example:



Here all neighborhoods are disjoint

So hitting set will have $\frac{n}{2}$ vertices

We will now prove Theorem 3.

Theorem 3 (Hitting Set Property):

Let $R \subseteq V$ be a uniformly random set of size "t". Then R is a hitting-set for $F = \{ N(x) \mid x \in V(G) \text{ has degree} \geq \frac{n \log n}{t} \}$

Proof: Let $R = \{ r_1, r_2, \ldots, r_t \}$ be uniform random subset of V chosen WITH REPLCEMENT.

V_{high} = Set of all vertices of deg $\geq \frac{n \log n}{t}$

vertices of high degree

For $x \in V_{high}$, let

E_x be event that $R \cap N(x)$ = empty set

$$\text{Prob}(r_i \in N(x)) = \frac{\deg(x)}{n} \geq \frac{4\log n}{t}$$

$$\text{Prob}(r_i \notin N(x)) \leq \left(1 - \frac{4\log n}{t}\right)$$

$$\text{Prob}\begin{pmatrix} r_1 \notin N(w) \\ \vdots \\ r_t \notin N(w) \end{pmatrix} \leq \left(1 - \frac{4\log n}{t}\right)^t$$

$$= \left(\left(1 - \frac{4\log n}{t}\right)^{\frac{t}{4\log n}}\right)^{4\log n}$$

$$\leq \left(\frac{1}{2}\right)^{4\log n} = \frac{1}{n^4}$$

$$\boxed{\begin{array}{c}\text{H.W.} \\ \left(1-\frac{1}{n}\right)^n \leq \frac{1}{2}\end{array}}$$

So, $\text{Prob}(E_x) \leq \frac{1}{n^4}$

By union bound, we have,

$$\text{Prob.}\left(\begin{array}{c} E_x \text{ holds} \\ \text{for at least} \\ \text{one } x \in V_{high} \end{array}\right) \leq \frac{|V_{high}|}{n^4} \leq \frac{1}{n^3}$$

So with probability $\left(1-\frac{1}{n^3}\right)$

"R" will have non-empty intersection with each $N(x)$, for $x \in V_{high}$. $\square$

Remark: It is also possible to have a deterministic construction for hitting set R.

$\boxed{\text{Application of Hitting Set}}$

Exercise: Given n pairs: (s_i, t_i), $1 \le i \le n$.

Goal: Find a subgraph H of $O(n\sqrt{n}\log n)$ size s.t. $d_H(s_i, t_i) = d_G(s_i, t_i)$.

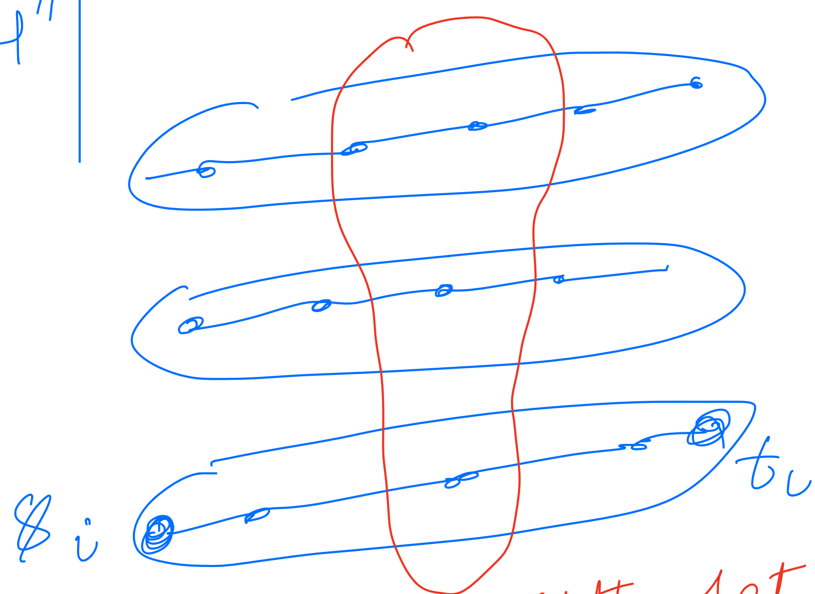
Sol^n

CASE 1: Pairs (s_i, t_i) with $d_G(s_i, t_i) \le \sqrt{n}\log n$

→ All edges on s_i-t_i shortestpath for such pairs to "H"

CASE 2: Pairs (s_i, t_i) with $d_G(s_i, t_i) > \sqrt{n}\log n$

→ collect all such pairs



$R \equiv$ Hitting set of size $\sqrt{n}$

We add BFS(w) to H

for each $w \in R$

Homework: Prove the correctness, time complexity of above sol^n.

Homework: How will you modify above sol^n to make it work for weighted graphs?

Homework: If G is directed, how will you find H?