Tail Bounds and the Probabilistic Method

- $\boxed{\text{Probabilistic Method}}$ – Given a large collection of objects, show that at least one has a certain property by arguing that a random object from the set has the property
- $\boxed{\text{Threshold Phenomena}}$ – Further, if we scale n (i.e., consider larger collections of objects), then either almost all or almost none of the objects have the property.

<u>Eg</u> – A graph $G(V,E)$ with $|V|=n$ is said to be an n-clique (denoted K_n) if E contains <u>all</u> possible edges (i,j) between nodes in V.

– A 2-coloring of G is a function $f: E \to \{r, b\}$ which associates a color in $\{r, b\}$ to each edge

---

<u>Thm</u> – If $\binom{n}{2} 2^{-\binom{k}{2}+1} < 1$, then $\exists$ a two coloring of K_n s.t. there are no monochromatic K_k subgraph.

<u>Pf</u> – Let $\Omega = \{\text{2-colorings of } K_n\}$, $|\Omega| = 2^{\binom{n}{2}}$

- Let $\mathbb{P}$ = Uniform distr on Ω
  $\equiv$ Color each (i,j) with $\{r, b\}$ u.a.r independently
- Let $\{1, 2, \dots, \binom{n}{k}\}$ be an enumeration of K_k subgraphs of K_n

- Now let $A_i = \mathbb{1}[\text{Subgraph } i \in [\binom{n}{k}] \text{ is monochromatic}]$

$$\Rightarrow \mathbb{P}[A_i] = \underset{\uparrow \text{ choice of color}}{2} \cdot 2^{-\binom{k}{2}} \leftarrow \mathbb{P}[\text{all edges in } i \text{ have chosen color}]$$

- $\mathbb{P}[\text{No monochromatic } K_k] = 1 - \mathbb{P}[\exists i \text{ s.t. } A_i = 1]$

$$= 1 - \mathbb{P}\left[\bigcup_{i=1}^{\binom{n}{k}} A_i\right]$$

- $$\mathbb{P}\left[\bigcup_{i=1}^{\binom{n}{k}} A_i\right] \leq \sum_{i=1}^{\binom{n}{k}} \mathbb{P}[A_i] \quad \text{(Union bound)}$$

$$= \underbrace{\binom{n}{k} 2^{-\binom{k}{2}+1}}_{=1-\delta} < 1 \quad \text{(Assumption)}$$

- $\Rightarrow \mathbb{P}[\text{No monochromatic } K_k] = \delta > 0$

$\therefore \exists$ 2-coloring of K_n s.t. no monochromatic K_k

---

The above result is existential, but can be made constructive via a randomized algorithm. This takes 2 forms -

- Monte Carlo Algo (Deterministic time, randomized correctness) - Color m graphs randomly as above. Then at least 1 satisfies property w.p $\geq 1-(1-\delta)^m \geq 1 - e^{-m\delta}$
- Las Vegas Algo (Random time, deterministic correctness) - Color single graph randomly, check if condition true, else repeat...

The First Moment Methods

These refer to arguments which only need $\mathbb{E}[X]$

- <u>Lemma</u> - For $X \in \mathbb{N}$, $\mathbb{P}[X \geq \mathbb{E}[X]] > 0$, $\mathbb{P}[X \leq \mathbb{E}[X]] > 0$

<u>Pf</u> - Suppose $\mathbb{P}[X \geq \mathbb{E}[X]] = 0 \Rightarrow \mathbb{E}[X] = \sum_{i=1}^{\infty} x \mathbb{P}[X = x]$

$= \sum_{x=1}^{\mathbb{E}[X]-1} x \mathbb{P}[X = x] < \mathbb{E}[X]$ which is a contradiction

- <u>Lemma</u> - If $X \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$, then

$$\mathbb{P}[X \neq 0] \leq \mathbb{E}[X]$$

<u>Pf</u> - $\mathbb{P}[X > 0] = \mathbb{P}[X \geq 1] \leq \mathbb{E}[X]$ by Markov.

<u>Eg</u> - If $m = (1+\varepsilon) n \log n$ balls thrown in n bins uar. Then the prob that there is an empty bis goes to 0.

<u>Pf</u> - Let $X = \#$ of empty bins $\Rightarrow \mathbb{P}[X \neq 0] < \mathbb{E}[X]$

Also $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[\mathbb{1}[\text{Bin } i \text{ is empty}]]$

$$= n\left(1 - \tfrac{1}{n}\right)^m = n\left(1 - \tfrac{1}{n}\right)^{(1+\varepsilon) n \log n}$$

$$\leq n e^{-(1+\varepsilon)\log n} = n^{-\varepsilon}$$

$\Rightarrow \mathbb{P}[X \neq 0] < n^{-\varepsilon} \searrow 0$ as $n \to \infty$

Eg - For any undirected graph $G(V,E)$ with $|V|=n$ and $|E|=m$, $\exists$ partition of V into sets A,B s.t the cut $\delta(A,B)=\{(i,j)\in E \mid i\in A, j\in B\}$ has $|\delta(A,B)|\geq m/2$

Pf - For each $i\in V$, put i in A wp $1/2$, else $i\in B$

$\Rightarrow$ For any $(i,j)\in E$, we have $\mathbb{P}[(i,j)\in\delta(A,B)]=1/2$

Let $X_{ij}=\mathbb{1}_{\{(i,j)\in\delta(A,B)\}} \Rightarrow |\delta(A,B)|=\sum_{(i,j)\in E} X_{ij}$

$$\Rightarrow \mathbb{E}\big[|\delta(A,B)|\big]=\sum_{(i,j)\in E}\mathbb{E}[X_{ij}]=\frac{m}{2}$$

$$\Rightarrow \mathbb{P}\big[|\delta(A,B)|\geq m/2\big]>0$$

First moment arguments are sometimes strengthened if you first thin the underlying set by sub-sampling.

Eg - In a given graph $G(V,E)$, a set $B\subseteq V$ is said to be independent if no pair of nodes i,j in B have an edge between them. The size of the largest independent set is denoted as $\alpha(G)$.

- Finding $\alpha(G)$ is computationally hard! In fact, even approximating it is hard...

<u>Thm</u> - For any $G(V,E)$ with $|V| = n$, $|E| = m$, we have $\alpha(G) \geqslant n^2/4m$.

<u>Pf</u> - We first thin the graph by removing each vertex independently with prob $1-p$ (Removing vertex $\Rightarrow$ remove all incident edges)

- Let X = # of vertices which remain after thinning
  $\Rightarrow \mathbb{E}[X] = np$
- Let Y = # of edges which remain after thinning
  $\Rightarrow \mathbb{E}[Y] = \sum_{(i,j)\in E} \mathbb{E}[\mathbb{1}\{i,j \text{ not thinned}\}] = m \cdot p^2$
- Finally, we remove each remaining edge and one of its neighbors arbitrarily. The remaining nodes form an independent set, with expected size at least $\mathbb{E}[X-Y] = p(n-mp)$. Setting $p = \frac{n}{2m}$ (i.e., 1/ 'average degree' of G), we get $\mathbb{E}[X-Y] = \frac{n^2}{4m}$
  $\Rightarrow \alpha(G) \geqslant n^2/4m$

The Second-Moment Method

- Till now we tried showing some rare event happens (ie, $X \neq 0$)

What if instead we want to show a highly likely event always happens?

<u>Lemma</u> - For $X \in \mathbb{N}_0$, $\mathbb{P}[X=0] \leq \operatorname{Var}(X)/\mathbb{E}[X]^2$

<u>Pf</u> - $\mathbb{P}[X=0] \leq \mathbb{P}[|X-\mathbb{E}[X]| \geq \mathbb{E}[X]] \leq \operatorname{Var}(X)/\mathbb{E}[X]^2$ (Chebyshev)

<u>Eg</u> - Suppose we throw $m=(1-\varepsilon)n\log n$ balls in n bins.

- As before, Let $X = \#$ of empty bins

$\Rightarrow \mathbb{E}[X] = n.\left(1-1/n\right)^m \approx n.n^{-1+\varepsilon} \nearrow \infty$

Let $\left(1-\frac{1}{n}\right)^m = \delta$

$\Rightarrow \mathbb{E}[X] = n\delta$

Does this mean $\mathbb{P}[X=0] \searrow 0$?

- Let $X = X_1 + X_2 + \dots + X_n$, where $X_i = \mathbb{1}\{\text{Bin } i \text{ is empty}\}$

$$\operatorname{Var}(X) = \sum_{i=1}^{n} \operatorname{Var}(X_i) + \sum_{i\neq j} \operatorname{Cov}(X_i, X_j)$$

where $\operatorname{Cov}(X_i, X_j) = \mathbb{E}[(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)]$

$X_i \sim \operatorname{Ber}(\delta)$

- $\operatorname{Var}(X_i) = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = \delta(1-\delta)$

$\operatorname{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - (\mathbb{E}[X_i])^2 = \left(1-2/n\right)^m - \delta^2$

To use the 2nd moment method we now want to upper bound $\operatorname{Var}(X) = n\operatorname{Var}(X_i) + \binom{n}{2}\operatorname{Cov}(X_i, X_j)$

- First for the covariance terms

$$\mathrm{Cov}(X_i, X_j) = \left(1-\frac{2}{n}\right)^m - \left(\left(1-\frac{1}{n}\right)^2\right)^m = \left(1-\frac{2}{n}\right)^m - \left(1-\frac{2}{n}+\frac{1}{n^2}\right)^m$$

$\Rightarrow \mathrm{Cov}(X_i, X_j) < 0 \quad \forall\, n, m \geq 0$!

$\Rightarrow \mathrm{Var}(X) \leq n\,\mathrm{Var}(X_i) = n\delta(1-\delta)$

- Using the lemma, we have

$$\mathbb{P}[X=0] \leq \frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2} = \frac{1-\delta}{n\delta} \leq \frac{1}{n\delta} = \frac{1}{\mathbb{E}[X]}!$$

- To complete the proof, we need to check that $\mathbb{E}[X]$ does indeed $\uparrow \infty$ with n, when $m = (1-\varepsilon)\, n \log n$.

This is true (try to show it....)

What is more important is to observe what we did at a higher level -

- We want to go from $\mathbb{E}[X] \uparrow \infty$ to $\mathbb{P}[X=0] \searrow 0$
- To do so, we show $\mathrm{Var}(X) \leq \mathbb{E}[X]$ (or $\mathrm{Var}(X) = \theta(\mathbb{E}[X]^2)$)
- Finally using $\mathbb{P}[X=0] \leq \frac{\mathrm{Var}(X)}{\mathbb{E}[X]^2}$, we are done!

To summarize - If $X_n \in \mathbb{N}_0$, then we have as $n \uparrow \infty$

i) If $\mathbb{E}[X_n] \searrow 0 \Rightarrow \mathbb{P}[X_n \neq 0] \searrow 0$

ii) If $\mathbb{E}[X_n] \uparrow \infty$ & $\mathbb{E}[X_n^2] = \theta(\mathbb{E}[X_n]^2) \Rightarrow \mathbb{P}[X_n = 0] \searrow 0$

Threshold Phenomena in large random systems

The moment methods are useful for studying threshold phenomena in large systems – settings where as we scale a system, then a certain property is always true or never true depending on some underlying parameter.

Eg- If we throw m balls in n bins, then as $n \nearrow \infty$

i) If $m = (1-\varepsilon) n \log n \Rightarrow \mathbb{P}[\exists \text{ empty bin}] \nearrow 1$

ii) If $m = (1+\varepsilon) n \log n \Rightarrow \mathbb{P}[\exists \text{ empty bin}] \searrow 0$

In computer science, this is known as the coupon collector problem, and in fact, we know sharper bounds than this (see assignment) – however the 1st/2nd moment methods give us these bounds in an easy way!

Eg- A $G(n,p)$ random graph is a random graph $\mathcal{Y} = G(V,E)$ where $|V| = n$ and $E_{ij} \sim Ber(p)\ \forall i \neq j$ (i.e., each edge is present w.p. p iid)

Let $C_4^{n,p}$ = # of K_4 in a given $G(n,p)$ graph. Then we have

i) $\mathbb{P}[C_4^{n,p} = 0] \searrow 0$ if $p = \omega(n^{-2/3})$

ii) $\mathbb{P}[C_4^{n,p} \neq 0] \searrow 0$ if $p = o(n^{-2/3})$

Thus $p = \Theta(n^{-2/3})$ is the threshold for existence of 4-cliques in a $G(n,p)$ graph.

Pf – Let $i \in \{1, 2, \ldots, \binom{n}{4}\}$ be an enumeration of all potential 4-cliques, $X_i = \mathbb{1}\{C_i \text{ has a } K_4\}$ ← set of nodes in i^{th} K_4

– $\mathbb{E}[X_i] = p^6$ ($\because$ a K_4 has 6 edges)

– $\mathbb{E}[C_4^{n,p}] = \mathbb{E}\left[\sum_{i=1}^{\binom{n}{4}} X_i\right] = \binom{n}{4} p^6 \approx n^4 p^6$

– $\text{Var}(C_4^{n,p}) = \binom{n}{4} \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$

$\text{Var}(X_i) = \underbrace{\mathbb{E}[X_i^2]}_{= \mathbb{E}[X_i]} - (\mathbb{E}[X_i])^2 \leq \binom{n}{4} p^6 = \Theta(n^4 p^6)$

– $\text{Cov}(X_i, X_j) = 0$ if $|C_i \cap C_j| = 0$ or 1 (i.e., don't share an edge)

This is because such X_i, X_j are $\perp\!\!\!\perp$!

– For $|C_i \cap C_j| = 2$, $\text{Cov}(X_i, X_j) \leq \mathbb{E}[X_i X_j] \leq p^{11}$

This follows as C_i and C_j have 1 edge in common

$\Rightarrow \sum_{|C_i \cap C_j| = 2} \text{Cov}(X_i, X_j) \leq \binom{n}{6} \cdot \left(\frac{6!}{2!\,2!\,2!}\right) \cdot p^{11} = \Theta(n^6 p^{11})$

– For $|C_i \cap C_j| = 3$, $\text{Cov}(X_i, X_j) \leq p^9$ and we have

$$\sum_{|C_i \cap C_j| = 3} \text{Cov}(X_i, X_j) = \Theta(n^5 p^9)$$

$\Rightarrow \text{Var}(C_4^{n,p}) = \Theta\left(n^4 p^6 + n^5 p^9 + n^6 p^{11}\right) = \Theta(\mathbb{E}[C_4^{n,p}])$

Now we use the moment bounds to finish the proof!

The Lovasz Local Lemma

- Our arguments till now depended on showing $\mathbb{E}[X] > 0$, where $X = \mathbb{1}\{E\}$ for some event $E = \bigcap_{i=1}^{n} E_i$. There are 2 'direct' ways for this - i) Via the union bound on 'bad events' $\bar{E}_i$

$$\mathbb{E}[X] = 1 - \mathbb{P}\left[\bigcup_{i=1}^{n} \bar{E}_i\right] \geq 1 - \sum_{i=1}^{n} \mathbb{P}[\bar{E}_i]$$

ii) If X_i are independent $\Rightarrow \mathbb{E}[X] = \prod_{i=1}^{n} \mathbb{E}[X_i]$, where $X_i = \mathbb{1}\{E_i\}$, $\mathbb{E}[X_i] = \mathbb{P}[E_i]$

- The former works when bad events are small (but arbitrarily dependent); the latter only needs $\mathbb{P}[E_i] > 0$ (so $\mathbb{P}[\bar{E}_i] < 1$), but needs them to be independent. The LLL let's us combine these!

- For a set of events $E_1, \dots, E_n$, their dependency graph $G(V,E)$ has $V = \{1, 2, \dots, n\}$ and $\forall i \in V$, event E_i is mutually independent of $\{E_j \mid (i,j) \notin E\}$ (i.e., all non-neighbors)

Lemma (Lovasz Local Lemma) - Let $E_1, E_2, \dots, E_n$ be events with dependency graph $G(V,E)$. Suppose $\exists\, x_i \in (0,1)\ \forall i \in [n]$ s.t

$$\mathbb{P}[\bar{E}_i] \leq x_i \prod_{(i,j) \in E} (1 - x_j)$$

Then $$\mathbb{P}\left[\bigcap_{i=1}^{n} E_i\right] \geq \prod_{i=1}^{n} (1 - x_i)$$

Corollary - If $\mathbb{P}[\bar{E}_i] \leq p$, $\deg(i) \leq d$ and $ep(d+1) < 1 \Rightarrow \mathbb{P}\left[\bigcap_{i=1}^{n} E_i\right] > 0$

Easier condition
$6pd < 1$
$\Rightarrow ep(d+1) < 1$

Pf - Choose $x_i = \frac{1}{d+1} \Rightarrow x_i \prod_{(i,j) \in E} (1 - x_j) = \left(\frac{1}{d+1}\right)\left(1 - \frac{1}{d+1}\right)^d \geq ep\, e^{-1}$

Now we can use the LLL

Pf - The idea is to write $\mathbb{P}\left[\bigcap_{i=1}^{n} E_i\right] = \mathbb{P}[E_1]\,\mathbb{P}[E_2 | E_1]\,\mathbb{P}[E_3 | E_1, E_2] \ldots \mathbb{P}[E_n | E_1, \ldots, E_{n-1}]$

Now if $\mathbb{P}[\bar{E}_i | E_1, E_2, \ldots, E_{i-1}] \leq x_i$ for all i, then we are done!

- To show the above, we show that $\forall i$, and all $S \subseteq [n] \setminus i$, we have $\boxed{\mathbb{P}\left[\bar{E}_i \mid \bigcap_{j \in S} E_j\right] \leq x_i}$. We do so by $\boxed{\text{induction on } |S|}$.

- Fix any i. For $S = \phi$, we have $\mathbb{P}[\bar{E}_i] \leq x_i \prod_{(i,j) \in E} (1 - x_j) \leq x_i$

- Now suppose it's true for all $S' \subseteq [n] \setminus i$, $|S'| < s$

Consider $S \subseteq [n] \setminus i$ s.t $|S| = s$. We need to show that

(i) $\mathbb{P}[E_S] > 0$ (so that $\mathbb{P}[\bar{E}_i | E_S]$ is defined)

(ii) $\mathbb{P}[\bar{E}_i | E_S] \leq x_i$

- $$\mathbb{P}[E_S] = \mathbb{P}\left[\bigcap_{k=1}^{s} E_{(k)}\right] = \prod_{k=1}^{s} \mathbb{P}\left[E_k \mid \bigcap_{j=1}^{k-1} E_j\right]$$

$\nwarrow$ some enumeration of S

$$\geq \prod_{k=1}^{s} (1 - x_k) > 0 \quad \text{(by induction hypothesis)}$$

- Next let $S = S_c \cup S_d$, where $S_c = S \cap N(i)$

($N(i) \equiv$ Neighbors of i in G) $\quad S_d = S \cap ([n] \setminus N(i))$

($c \equiv$ Connected, $d \equiv$ disconnected)

- If $S = S_d \Rightarrow \mathbb{P}[\bar{E}_i | E_{S_d}] = \mathbb{P}[\bar{E}_i] \leq x_i$

- Otherwise
$$\mathbb{P}[\bar{E}_i | E_S] = \frac{\mathbb{P}[\bar{E}_i \cap E_{S_c} \cap E_{S_d}]}{\mathbb{P}[E_{S_c} \cap E_{S_d}]} = \frac{\mathbb{P}[\bar{E}_i \cap E_{S_c} | E_{S_d}]\,\cancel{\mathbb{P}[E_{S_d}]}}{\mathbb{P}[E_{S_c} | E_{S_d}]\,\cancel{\mathbb{P}[E_{S_d}]}}$$

Now we have $\mathbb{P}[\bar{E}_i \cap E_{S_c} | E_{S_d}] \leq \mathbb{P}[\bar{E}_i | E_{S_d}] = \mathbb{P}[\bar{E}_i] \leq x_i \prod_{(i,j) \in E} (1 - x_j)$

$\uparrow$

$\because E_i \perp\!\!\!\perp E_j \;\forall j \in S_d$

Moreover $\mathbb{P}[E_{S_c} | E_{S_d}] = \mathbb{P}\left[\bigcap_{j \in S_c} E_j \,\middle|\, \bigcap_{k \in S_d} E_k\right]$ (let $S_c = \{j_1, j_2, \dots, j_r\}$)

$$= \prod_{\ell=1}^{r}\left(1 - \mathbb{P}\left[E_{j_\ell} \,\middle|\, \underbrace{\bigcap_{\ell' < \ell} E_{j_{\ell'}}, \bigcap_{k \in S_d} E_k}_{\subseteq S \setminus j_\ell}\right]\right)$$

$$\geq \prod_{\ell=1}^{r}(1 - x_{j_\ell}) \geq \prod_{(i,j) \in E}(1 - x_j)$$

$\leftarrow$ induction hypothesis

$\Rightarrow \mathbb{P}[\bar{E}_i | E_S] \leq x_i \quad \forall i, \forall S$ with $|S| = s$

By induction, we complete the proof ▨

Eg - (Satisficbility) A k-SAT formula is a function of binary variables $x_1, x_2, \dots, x_n \in \{0,1\}^n$ of the form

$$f(x_1, \dots, x_n) = \bigwedge_{c=1}^{m}(x_{1c} \vee x_{2c} \vee \dots \vee x_{kc})$$

where $\wedge \equiv$ AND (or conjunction), $\vee \equiv$ OR (or disjunction), $x_{ic} \in \{x_1, x_2, \dots, x_n, \bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\} \equiv$ literal, and $(x_{1c} \vee \dots \vee x_{kc}) \equiv$ clause.

(In words - f is a conjunction of m disjunctive clauses, each with n literals)

Thm - A k-SAT formula where no variable appears in more than $2^k/6k$ clauses is satisfiable (i.e., $\exists x \in \{0,1\}^n$ s.t. $f(x) = 1$)

Pf - Set $X_i \sim Ber(1/2)$. Want to show $\mathbb{P}[f(X_1, \dots, X_n) = 1] > 0$

- Let E_i = Event clause i not satisfied. $\mathbb{P}[E_i] \leq 2^{-k}$. Want $\mathbb{P}[\bigcap_{i=1}^{m} \bar{E}_i] > 0$
- $E_i \perp\!\!\!\perp E_j$ if E_i and E_j share no common literals
  $\Rightarrow d(E_i) \leq k \cdot$ (max # of clauses with same literal) $\leq 2^k/6$
  $\Rightarrow 6 \cdot \mathbb{P}[E_i] \cdot d[E_i] < 1 \Rightarrow$ By LLL, $\mathbb{P}[\bigcap_{i=1}^{n} \bar{E}_i] > 0$ ▨

- ## A segue - 2-SAT

- A 2-SAT formula on $\{x_1, \bar{x}_1, x_2, \ldots, x_n, \bar{x}_n\}$ is of the form
$$f(\underline{x}) = (x_1 \vee x_2) \wedge (x_3 \vee \bar{x}_1) \wedge \ldots$$

Claim - A satisfying assignment for a 2-SAT formula can be found in poly-time
- There are multiple ways to do this; we will now see a simple randomized algorithm, which then takes us into Markov Chains

---

(Papadimitriou's WALK-SAT algo) - Given CNF formula f and any starting assignment $X(0)$

For $i \in \{0, 1, 2, \ldots, cn^2\}$
- If $X(i)$ is feasible, STOP and return $x(i)$
- Else pick any unsatisfied clause, choose one of its literals l_i u.a.r
- Set $x(i+1) = x(i)$ with bit l_i flipped

Claim - If f is satisfiable, then WALK-SAT finds a satisfying assignment with probability $\geq 1/2$

Pf - Let S be any satisfying assignment, and let N_i = # of variables in $X(i)$ which agree with S. If $X(i)$ is not feasible $\Rightarrow$
$$X(i+1) = \begin{cases} X(i) + 1 & \text{wp} \geq 1/2 \\ X(i) - 1 & \text{wp} \leq 1/2 \end{cases}, \text{if } X(i) \in \{1, 2, \ldots, n-1\}$$

- Now let $h_j = \mathbb{E}[\min\{T \mid N_T = n, N_0 = j\}]$

$$\Rightarrow \quad h_j \leq \frac{1}{2}(h_{j-1} + 1) + \frac{1}{2}(h_{j+1} + 1), \quad h_n = 0$$

Solve to get $h_j \leq n^2 - j^2 \leq n^2$

$\Rightarrow \mathbb{E}[\text{Time to return correct solution}] \leq n^2$