Smoothed Complexity & Knapsack

Constantinos Vrohidis

February 2021

Smoothed Analysis

algorithmic paradigm introduced to explain the apparent discrepancy between

Worst-case performance ✗

Practical performance ✓

Smoothed Analysis

Algorithmic paradigm introduced to explain the apparent discrepancy between

Worst-case performance ✗

Practical performance ✓

Archetypical example of this phenomenon:

Simplex algorithm for linear programming

[Teng, Spielman '01]

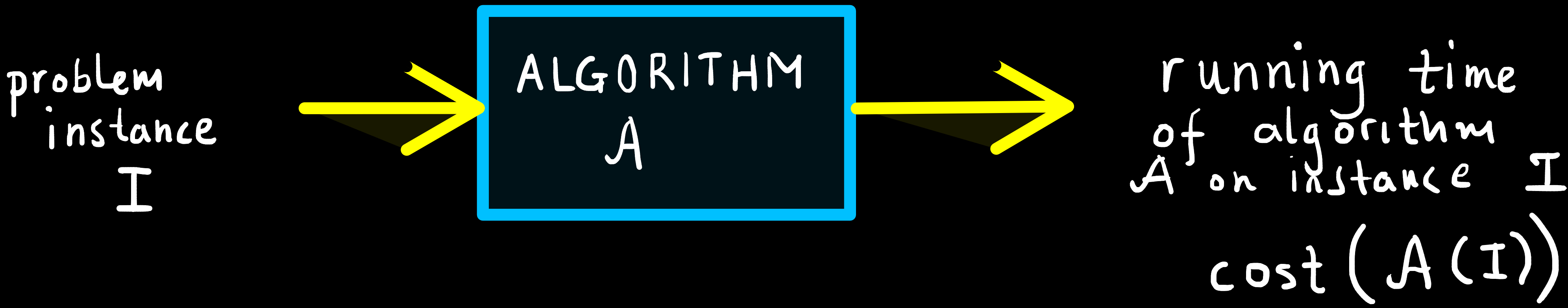
Gödel Prize '08

Shang-Hua Teng

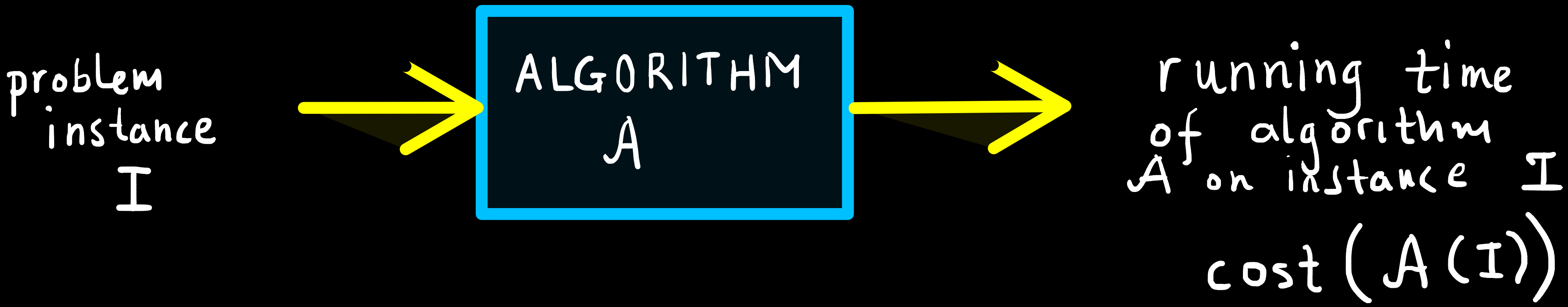
Daniel Spielman

"Smoothed analysis of algorithms:
Why the simplex algorithm usually
takes polynomial time"

Worst-case Analysis



Worst-case Analysis



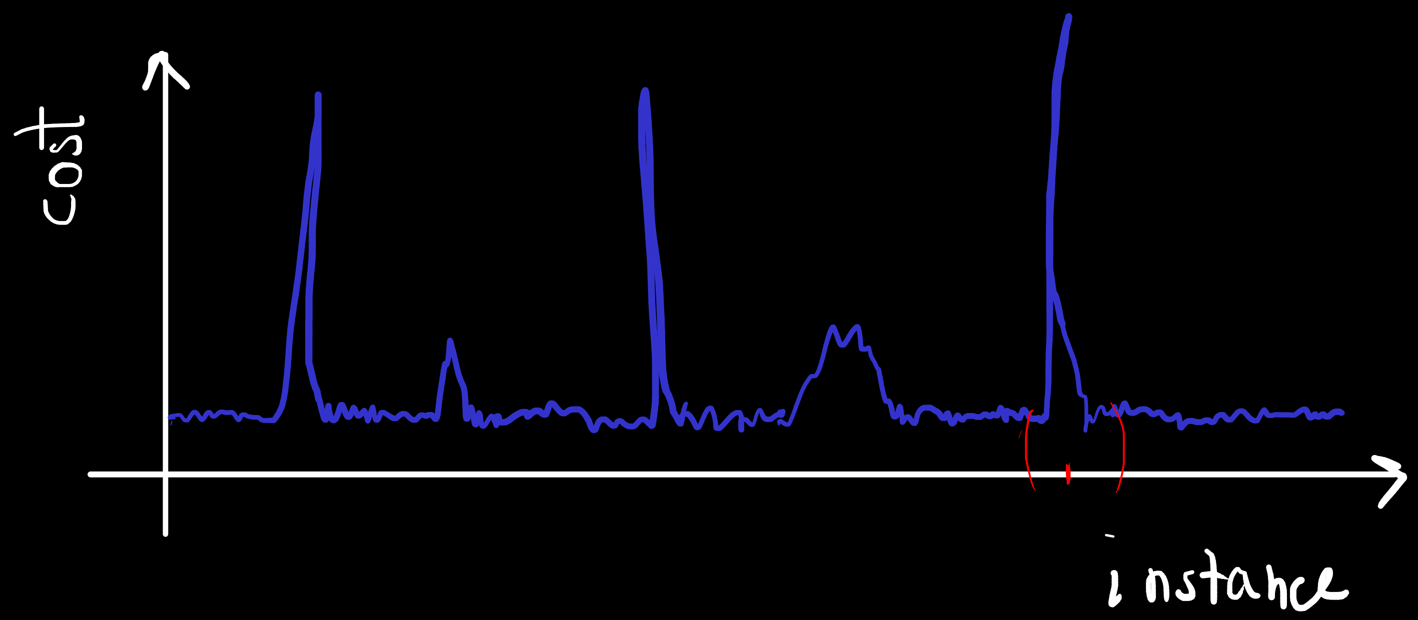
Suppose I_n is the set of all instances of "size" n

worst case analysis is interested in

$$\max_{I \in I_n} \text{cost}(A(I))$$

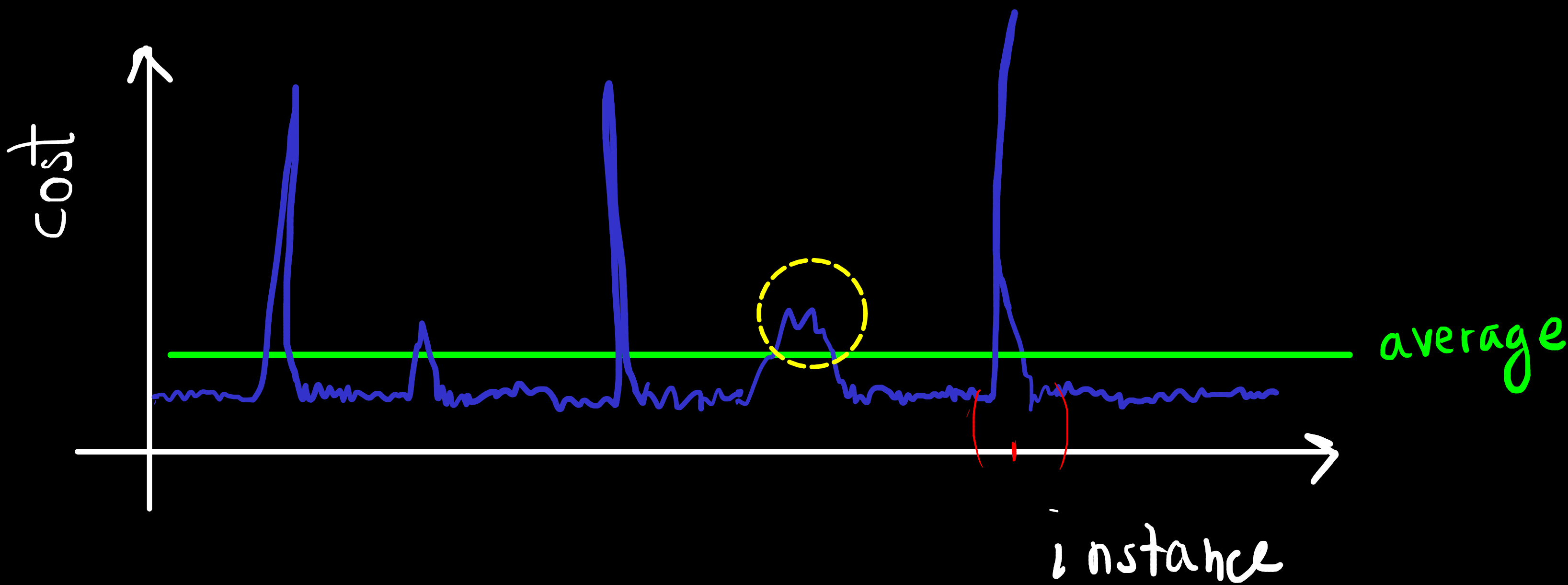
Worst-case Analysis

maximizers can be outliers



~~Worst~~ Average -case Analysis

Alternatives : Average case analysis

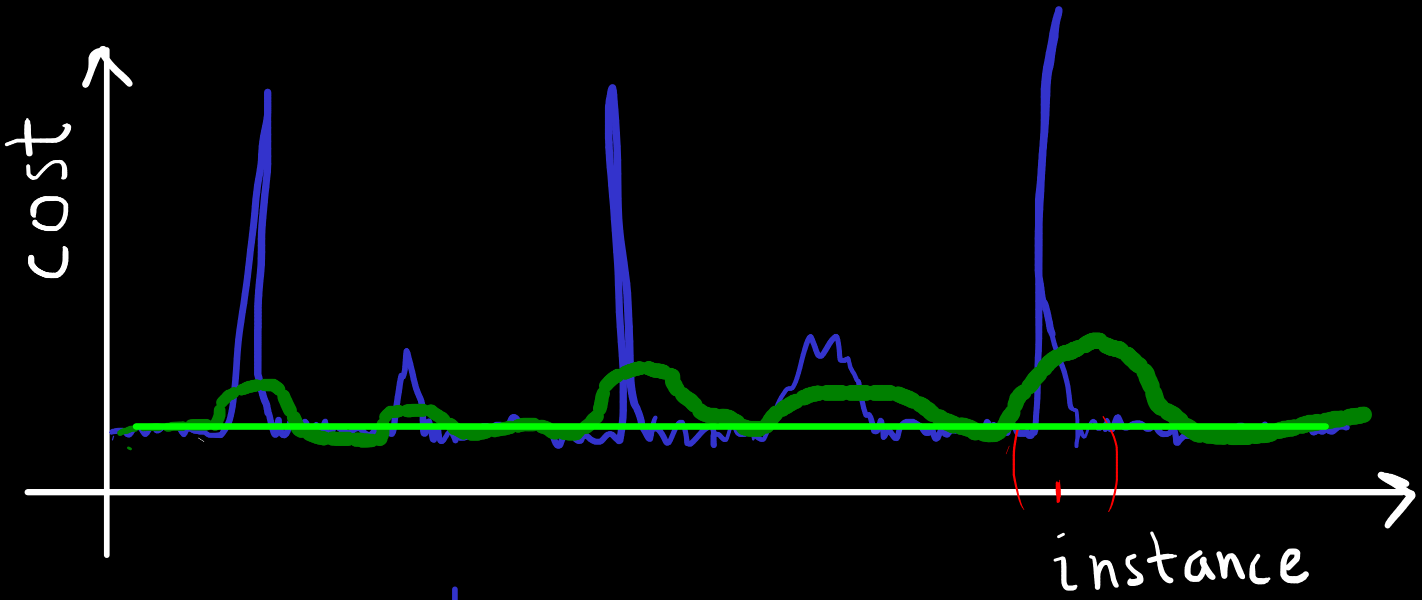


Between

worst-case

Smoothed

average-case

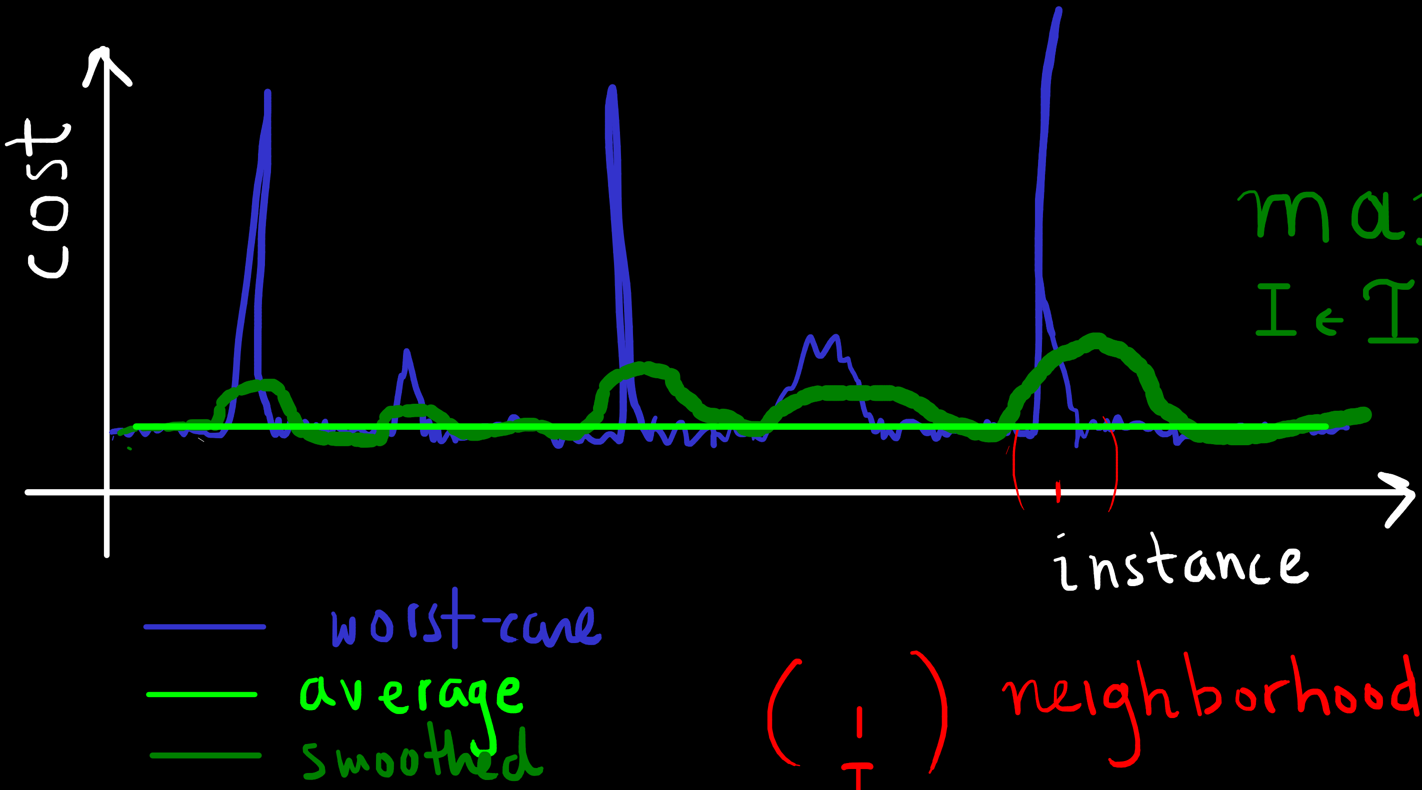


worst-case
average
smoothed

(I) neighborhood of instance I, with some real parameter σ

Between

worst-case — Smoothed — average-case



$$\max_{I \in \mathcal{I}_n} \mathbb{E}_{I' \sim \text{nhood}_\sigma(I)} \left[\text{cost}(A(I')) \right]$$

(I) neighborhood of instance I, with some real parameter σ

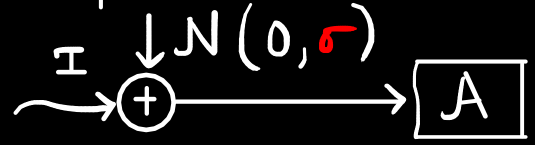
Two models of smoothed analysis

Two models of smoothed analysis

Classical model: instances are perturbed by zero-mean Gaussians.

Two models of smoothed analysis

Classical model: instances are perturbed by zero-mean Gaussians.

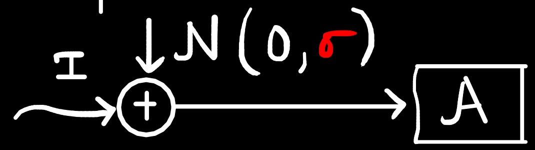


$\sigma \longrightarrow 0$ (worstcase)

$\sigma \longrightarrow \infty$ (average case)

Two models of smoothed analysis

Classical model: instances are perturbed by zero-mean Gaussians.



$\sigma \longrightarrow 0$ (worst case)

$\sigma \longrightarrow \infty$ (average case)

One-step model: Instance is drawn directly by sampling from some continuous probability distribution with p.d.f bounded by φ.

Lemma: Let X be a continuous real random variable with pdf $f_X : \mathbb{R} \to [0, \varphi]$. Then, for any $\alpha \in \mathbb{R}$ and any $\varepsilon > 0$,

$$\Pr\Big[X \in [\alpha, \alpha + \varepsilon] \Big] \leq \varepsilon \cdot \varphi.$$

Lemma: Let X be a continuous real random variable with pdf $f_X : \mathbb{R} \to [0, \varphi]$. Then, for any $\alpha \in \mathbb{R}$ and any $\epsilon > 0$,

$$\Pr\left[X \in [\alpha, \alpha+\varepsilon] \right] \leq \varepsilon \cdot \varphi.$$

proof:

$$\Pr\left[X \in [\alpha, \alpha+\varepsilon] \right] = \int_{\alpha}^{\alpha+\varepsilon} f_X(t)\,dt$$

$$\leq \int_{\alpha}^{\alpha+\varepsilon} \varphi\,dt = \varphi(\alpha+\varepsilon-\alpha)$$

$$= \varepsilon \cdot \varphi.$$

Lemma: Let X be a continuous real random variable with pdf $f_X : \mathbb{R} \to [0, \varphi]$. Then, for any $\alpha \in \mathbb{R}$ and any $\varepsilon > 0$,

$$\Pr\left[X \in [\alpha, \alpha + \varepsilon] \right] \leq \varepsilon \cdot \varphi.$$

proof:

$$\Pr\left[X \in [\alpha, \alpha+\varepsilon] \right] = \int_{\alpha}^{\alpha+\varepsilon} f_X(t)\, dt$$

$$\leq \int_{\alpha}^{\alpha+\varepsilon} \varphi \, dt = \varphi(\alpha + \varepsilon - \alpha)$$

$$= \varepsilon \cdot \varphi.$$

e.g., if one considers uniform distributions, small φ implies large support.

Knapsack

We are given n items $1, 2, \dots, n$ with non-negative weights $w_1, w_2, \dots, w_n$ and profits $p_1, \dots, p_n$ and a capacity t.

Knapsack

We are given n items $1, 2, \dots, n$ with non-negative weights $w_1, w_2, \dots, w_n$ and profits $p_1, \dots, p_n$ and a capacity t.

Find a subset $I \subseteq [n]$ such that

$$I \in \operatorname*{argmax}_{S \subseteq [n]} \sum_{j \in S} p_j$$

$$\text{subject to} \quad \sum_{j \in S} w_j \leq t$$

Knapsack

We are given n items $1, 2, \ldots, n$

with non-negative weights $w_1, w_2, \ldots, w_n$

and profits $p_1, \ldots, p_n$ and a capacity t.

Find a subset $I \subseteq [n]$ such that

$$I \in \operatorname*{argmax}_{S \subseteq [n]} \sum_{j \in S} p_j$$

$$\text{subject to} \quad \sum_{j \in S} w_j \leq t$$

For convenience, we view subsets of $[n]$ as vectors in

$\{0,1\}^n$.

Knapsack

NP-hard (Karp, '72)

We are given n items $1, 2, \ldots, n$
with non-negative weights $w_1, w_2, \ldots, w_n$
and profits $p_1, \ldots, p_n$ and a capacity t.

Find a subset $I \subseteq [n]$ such that

$$I \in \operatorname*{argmax}_{S \subseteq [n]} \sum_{j \in S} p_j$$

$$\text{subject to} \quad \sum_{j \in S} w_j \leq t$$

For convenience, we view subsets of $[n]$ as vectors in $\{0,1\}^n$.

$$w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n_{\geq 0}$$

$$p = (p_1, \ldots, p_n) \in \mathbb{R}^n_{\geq 0}$$

$$t \in \mathbb{R}_{\geq 0}$$

$$\boxed{\begin{array}{ll} \text{maximize} & p^T x \\ \text{subject to} & w^T x \leq t \\ & x \in \{0,1\}^n \end{array}}$$

$$w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n_{\geq 0}$$

$$p = (p_1, \ldots, p_n) \in \mathbb{R}^n_{\geq 0}$$

$$t \in \mathbb{R}_{\geq 0}$$

$$\boxed{\begin{array}{ll} \text{maximize} & p^T x \\ \text{subject to} & w^T x \leq t \\ & x \in \{0,1\}^n \end{array}}$$

Naive approach: enumerate all feasible solutions and select one with maximum profit.

$w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n_{\geq 0}$

$p = (p_1, \ldots, p_n) \in \mathbb{R}^n_{\geq 0}$

$t \in \mathbb{R}_{\geq 0}$

$$\boxed{\begin{array}{ll} \text{maximize} & p^T x \\ \text{subject to} & w^T x \leq t \\ & x \in \{0,1\}^n \end{array}}$$

Naive approach: enumerate all feasible solutions and select one with maximum profit.

exponentially many feasible solutions

$w = (w_1, w_2, \dots, w_n) \in \mathbb{R}^n_{\geq 0}$

$p = (p_1, \dots, p_n) \in \mathbb{R}^n_{\geq 0}$

$t \in \mathbb{R}_{\geq 0}$

$$\begin{array}{ll} \text{maximize} & p^T x \\ \text{subject to} & w^T x \leq t \\ & x \in \{0,1\}^n \end{array}$$

Naive approach: enumerate all feasible solutions and select one with maximum profit.

exponentially many feasible solutions

Is there a way around this?

__Definition:__ Let $x, y \in \{0,1\}^n$ be two solutions. Then, y __dominates__ x if

$$p^T y \geq p^T x$$
$$w^T y \leq w^T x$$

and at least one of the inequalities is __strict.__

Pareto Optimality

<u>Definition:</u> A solution $x \in \{0,1\}^n$ that is not dominated by any other solution is called <u>Pareto-optimal</u>.

Pareto Optimality

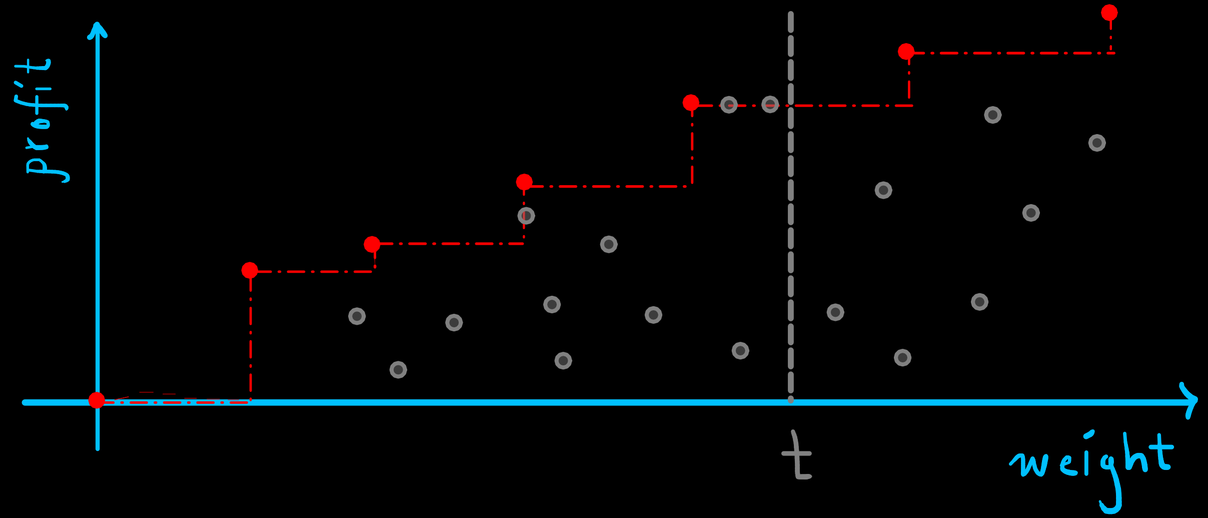
Definition: A solution $x \in \{0,1\}^n$ that is not dominated by any other solution is called Pareto-optimal.

The Pareto set is the set of all Pareto optimal solutions.

Pareto Optimality

Definition: A solution $x \in \{0,1\}^n$ that is not dominated by any other solution is called Pareto-optimal.

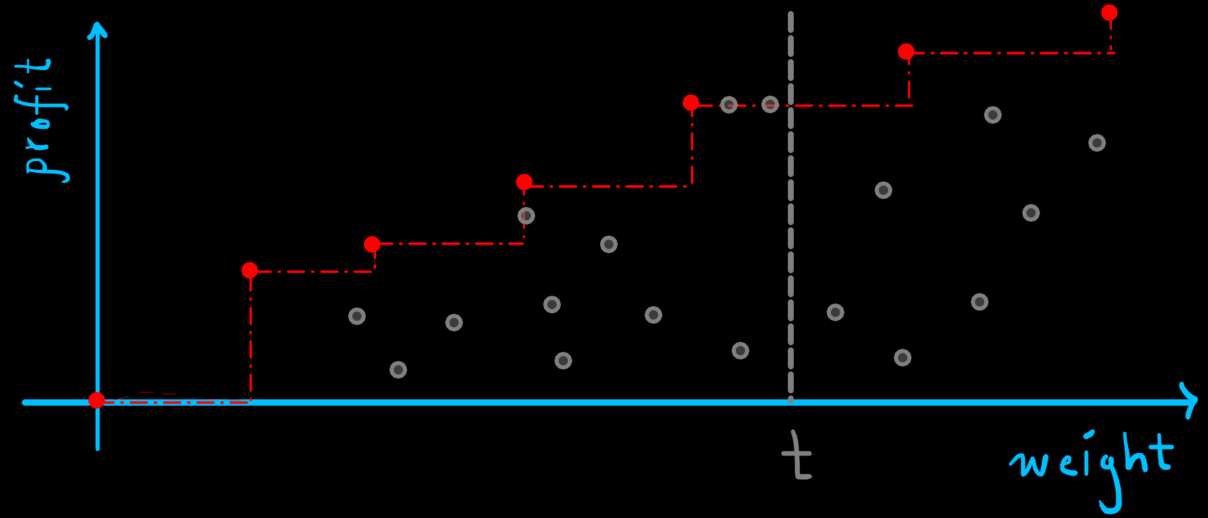
The Pareto set is the set of all Pareto optimal solutions.

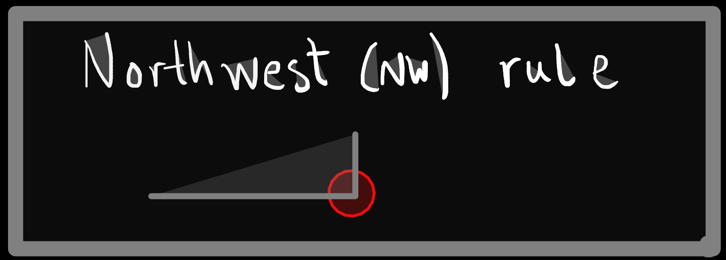


Pareto Optimality

Definition: A solution $x \in \{0,1\}^n$ that is not dominated by any other solution is called Pareto-optimal.

The Pareto set is the set of all Pareto optimal solutions.

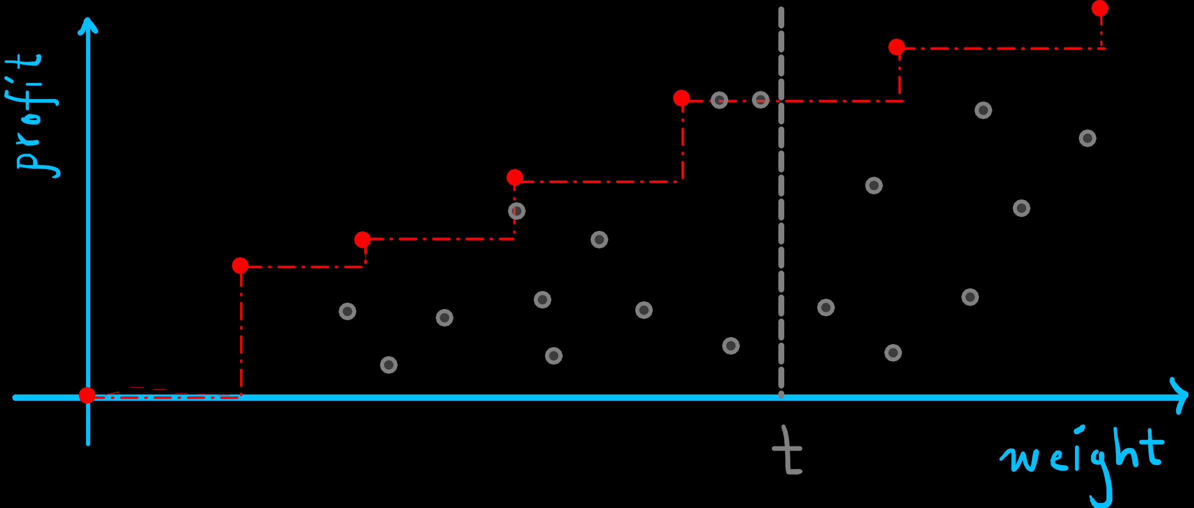


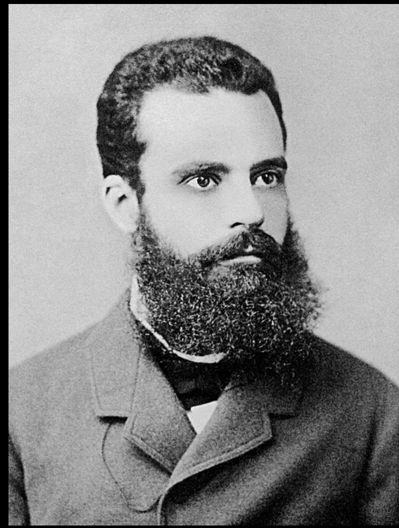


Pareto Optimality

Definition: A solution $x \in \{0,1\}^n$ that is not dominated by any other solution is called **Pareto-optimal**.

The **Pareto set** is the set of all Pareto optimal solutions.





Vilfredo Pareto

Northwest (NW) rule

Lemma: There always exists an optimal solution that is also Pareto-optimal.

Lemma: There always exists an optimal solution that is also Pareto-optimal.

proof: Say x is an optimal solution.

If x is not pareto optimal, there exists another solution y that dominates x, that is,

$$p^T y \geq p^T x \quad \text{and} \quad w^T y \leq w^T x$$

and at least one is strict.

Lemma: There always exists an optimal solution that is also Pareto-optimal.

proof: Say x is an optimal solution.

If x is not pareto optimal, there exists another solution y that dominates x, that is,

$$p^T y \geq p^T x \quad \text{and} \quad w^T y \leq w^T x$$

and at least one is strict. y is feasible and optimal. So, $w^T y < w^T x$. If y is Pareto optimal, we are done. Otherwise, repeating finitely many times yields a Pareto optimal solution.

Lemma: There always exists an optimal solution that is also Pareto-optimal.

proof: Say x is an optimal solution.

If x is not pareto optimal, there exists another solution y that dominates x, that is,

not strict

$$p^T y \geq p^T x \quad \text{and} \quad w^T y \leq w^T x$$

and at least one is strict. y is feasible and optimal. So, $w^T y < w^T x$. If y is Pareto optimal, we are done. Otherwise, repeating finitely many times yields a Pareto optimal solution.

Lemma: There always exists an optimal solution that is also Pareto-optimal.

proof: Say x is an optimal solution.

If x is not pareto optimal, there exists another solution y that dominates x, that is,

strict

not strict $\quad p^T y \geq p^T x \quad$ and $\quad w^T y \leq w^T x$

and at least one is strict. y is feasible and optimal. So, $w^T y < w^T x$. If y is Pareto optimal, we are done. Otherwise, repeating finitely many times yields a Pareto optimal solution.

An immediate corollary

If $\mathcal{P} \subseteq \{0,1\}^n$ is the Pareto set an element of the set $\{x \in \mathcal{P} : x^T w \leq t\}$ is an optimal solution of the knapsack problem. Moreover, this solution can be found in time linear in the size of $\mathcal{P}$.

An immediate corollary

If $\mathcal{P} = \{0,1\}^n$ is the Pareto set an element of the set $\{x \in \mathcal{P} : x^T w \leq t\}$ is an optimal solution of the knapsack problem.
Moreover, this solution can be found in time linear in the size of $\mathcal{P}$.

But how do we compute $\mathcal{P}$?

The Nemhauser-Ullman Algorithm

The Nemhauser-Ullman Algorithm

(Dynamic programming): For each $i \in [n]$ compute the Pareto set P_i of the modified instance of the knapsack problem that contains only the first i items.

The Nemhauser - Ullman Algorithm

(Dynamic programming): For each $i \in [n]$ compute the Pareto set $\mathcal{P}_i$ of the modified instance of the knapsack problem that contains only the first i items.

Then, $\mathcal{P} = \mathcal{P}_n$.

The Nemhauser-Ullman Algorithm

(Dynamic programming): For each $i \in [n]$ compute the Pareto set $\mathcal{P}_i$ of the modified instance of the knapsack problem that contains only the first i items.

Then, $\mathcal{P} = \mathcal{P}_n$.

$$S_i := \{x \in \{0,1\}^n : x_{i+1} = \dots = x_n = 0\}$$

The Nemhauser-Ullman Algorithm

(Dynamic programming): For each $i \in [n]$ compute the Pareto set P_i of the modified instance of the knapsack problem that contains only the first i items.

Then, $\mathcal{P} = \mathcal{P}_n$.

$$S_i := \{ x \in \{0,1\}^n : x_{i+1} = \dots = x_n = 0 \}$$

For $x \in \{0,1\}^n$, let x^{+i} with $x_j^{+i} = \begin{cases} x_j & , j \neq i \\ 1 & , j = i \end{cases}$.

The Nemhauser-Ullman Algorithm

(Dynamic programming): For each $i \in [n]$ compute the Pareto set $\mathcal{P}_i$ of the modified instance of the knapsack problem that contains only the first i items.

Then, $\mathcal{P} = \mathcal{P}_n$.

$$S_i := \{ x \in \{0,1\}^n : x_{i+1} = \dots = x_n = 0 \}$$

For $x \in \{0,1\}^n$, let x^{+i} with $x_j^{+i} = \begin{cases} x_j & , j \neq i \\ 1 & , j = i \end{cases}$.

For $S \subseteq \{0,1\}^n$, let $S^{+i} \subseteq \{0,1\}^n$

$$S^{+i} = \{ y \in \{0,1\}^n : \exists x \in \mathbb{R}^n , y = x^{+i} \}$$

Lemma: For every $i \in [n]$, the set $\mathcal{P}_i$ is contained in $\mathcal{P}_{i-1} \cup \mathcal{P}_{i-1}^{+i}$.

Lemma: For every $i \in [n]$, the set $\mathcal{P}_i$ is contained in $\mathcal{P}_{i-1} \cup \mathcal{P}_{i-1}^{+i}$.

proof: Let $x \in \mathcal{P}_i$. Either $\underbrace{x_i = 0}_{(i)}$ or $\underbrace{x_i = 1}_{(ii)}$.

Lemma: For every $i \in [n]$, the set $\mathcal{P}_i$ is contained in $\mathcal{P}_{i-1} \cup \mathcal{P}_{i-1}^{+i}$.

proof: Let $x \in \mathcal{P}_i$. Either $\underbrace{x_i = 0}_{(i)}$ or $\underbrace{x_i = 1}_{(ii)}$.

(i) If $x_i = 0$, then $x \in \mathcal{P}_{i-1}$. Indeed, say $x \notin \mathcal{P}_{i-1}$, then there exists some $y \in \mathcal{P}_{i-1} \subseteq S_{i-1} \subseteq S_i$ that dominates x. Since $y \in S_i$, x cannot be Pareto-optimal among the solutions in S_i, i.e., $x \notin \mathcal{P}_i$; a contradiction.

Lemma: For every $i \in [n]$, the set $\mathcal{P}_i$ is contained in $\mathcal{P}_{i-1} \cup \mathcal{P}_{i-1}^{+i}$.

proof: Let $x \in \mathcal{P}_i$. Either $\underbrace{x_i = 0}_{(i)}$ or $\underbrace{x_i = 1}_{(ii)}$.

(ii) If $x_i = 1$, then $x \in \mathcal{P}_{i-1}^{+i}$. $x \in S_i$ and $x_i = 1$. Therefore, $x = y^{+i}$ for some $y \in S_{i-1}$. It suffices to show that $y \in \mathcal{P}_{i-1}$. Suppose, for the sake of contradiction, that y is dominated by some $z \in \mathcal{P}_{i-1}$. Then, $z^{+i} \in S_i$ dominates $x = y^{+i}$ and therefore $x \notin \mathcal{P}_i$; a contradiction.

The Nemhauser-Ullman Algorithm

$P_0 = \{0\}^n$

for $i = 1, 2, \dots, n$ do

  $Q_i = P_{i-1} \cup P_{i-1}^{+i}$

  $P_i = \{ x \in Q_i : \nexists\, y \in Q_i \text{ s.t. } y \text{ dominates } x \}$

return $x^* \in \operatorname*{argmax}_{x \in P_n} \{ p^T x : w^T x \leq t \}$

The Nemhauser-Ullman Algorithm

Lemma: The Nemhauser-Ullman algorithm can be implemented to have a running time of

$$\Theta\left(\sum_{i=0}^{n-1} |\mathcal{P}_i| \right)$$

The Nemhauser-Ullman Algorithm

Lemma: The Nemhauser-Ullman algorithm can be implemented to have a running time of

$$O\left(\sum_{i=0}^{n-1} |\mathcal{P}_i| \right)$$

Proof sketch: Store $val(\mathcal{P}_i) = \{ (p^T x, w^T x) : x \in \mathcal{P}_i \}$ and pointers to the elements of $val(\mathcal{P}_{i-1})$. This allows reconstruction of the solution in time $O(n + |\mathcal{P}|)$. Computing $val(Q_i)$ takes $O(|\mathcal{P}_{i-1}|)$ time (line 3). If $val(\mathcal{P}_{i-1})$ are maintained sorted, sorted $val(Q_i)$ can be computed in $\Theta(|\mathcal{P}_{i-1}|)$ and $val(\mathcal{P}_i)$ in $\Theta(|Q_i|)$.

Smoothed Upper Bound on the number of Pareto optimal solutions

[Beier, Röglin, Vöcking '07]

Smoothed Upper Bound on the number of Pareto optimal solutions [Beier, Röglin, Vöcking '07]

Model:
- Arbitrary fixed profits $w_1, w_2, \ldots, w_n$.
- Each weight w_i is drawn independently from a continuous distribution $f_{w_i}: [0,1] \longrightarrow [0,\varphi]$.

Smoothed Upper Bound on the number of Pareto optimal solutions

[Beier, Röglin, Vöcking '07]

Model:

~ Arbitrary fixed profits $w_1, w_2, \ldots, w_n$.

~ Each weight w_i is drawn independently from a continuous distribution $f_{w_i} : [0,1] \to [0, \varphi]$.

Theorem: Assuming the above smoothness model,

$$\mathbb{E}\left[\, |P| \,\right] \le n^2 \varphi + 1 .$$

Smoothed Upper Bound on the number of Pareto optimal solutions

[Beier, Röglin, Vöcking '07]

Model:

- ~ Arbitrary fixed profits $w_1, w_2, \ldots, w_n$.
- ~ Each weight w_i is drawn independently from a continuous distribution $f_{w_i} : [0,1] \longrightarrow [0, \varphi]$.

Theorem: Assuming the above smoothness model,

$$\mathbb{E}\left[\,|P|\,\right] \leq n^2 \varphi + 1 .$$

Corollary: The Nemhauser – Ullman algorithm has smoothed complexity

$$O(n^3 \varphi)$$

Proving the Theorem: Bound expected size of P_i, $\forall i \in [n]$.

Proving the Theorem : Bound expected size of $\mathcal{P}_i$, $\forall i \in [n]$.

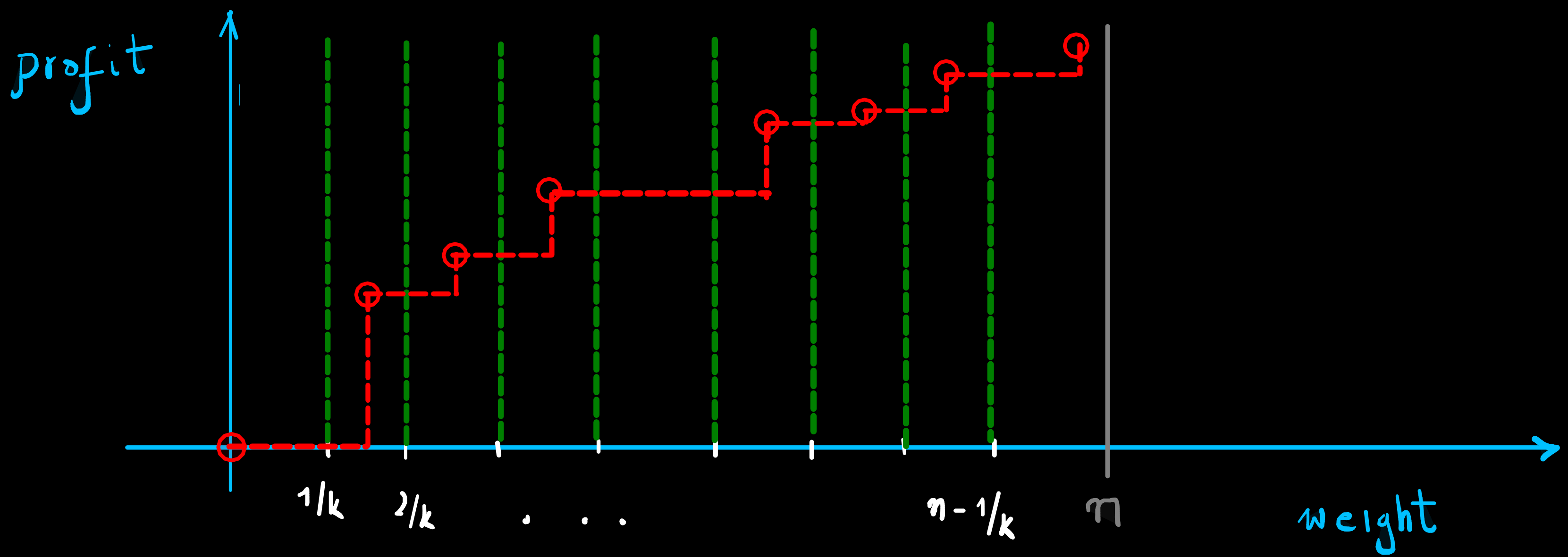
Pick $i \in [n]$, set $P = \mathcal{P}_i$.

Since $w_i \leq 1$, $\forall i \in [n]$, $w^T x \leq n$, $\forall x \in P$.

Proving the Theorem : Bound expected size of $\mathcal{P}_i$, $\forall i \in [n]$.

Pick $i \in [n]$, set $P = \mathcal{P}_i$.

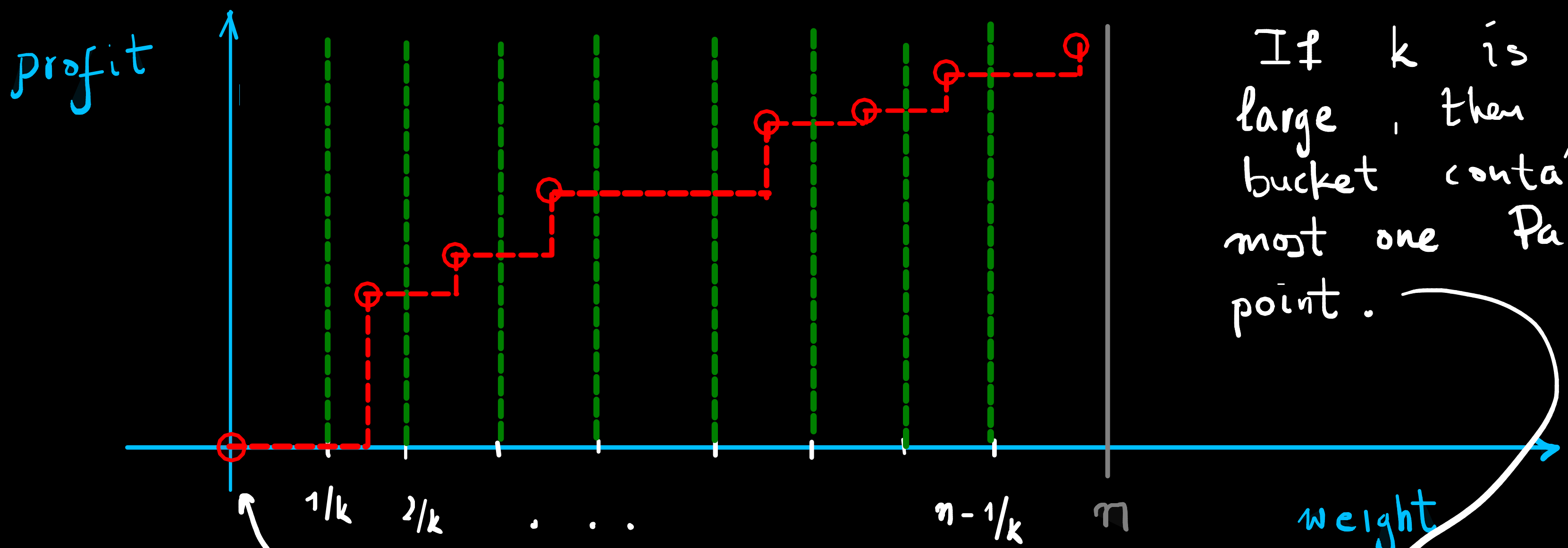
Since $w_i \leq 1$, $\forall i \in [n]$, $w^T x \leq n$, $\forall x \in P$.



Proving the Theorem: Bound expected size of $\mathcal{P}_i$, $\forall i \in [n]$.

Pick $i \in [n]$, set $P = \mathcal{P}_i$.

Since $w_i \le 1$, $\forall i \in [n]$, $w^T x \le n$, $\forall x \in P$.



If k is sufficiently large, then each bucket contains at most one Pareto-optimal point.

$$|P| = 1 + \sum_{\ell=0}^{nk} \mathbb{1}\left\{ \exists x \in P \cap \left(\frac{\ell}{k}, \frac{\ell+1}{k} \right) \right\}, \text{ for } k \text{ large}$$

$$\mathbb{E}[|P|] = 1 + \sum_{\ell=0}^{nk} \Pr\left[\exists p \in P \cap \left(\frac{\ell}{k}, \frac{\ell+1}{k} \right] \right] \quad (k \to \infty)$$

Let's make this more formal

For $k \in \mathbb{N}$, denote by $\mathcal{F}_k$, the event that there exist two solutions $x, y \in \{0,1\}^n$ s.t. $|w^T x - w^T y| \leq 1/k$.

$\boxed{\text{Lemma: For } k \in \mathbb{N}, \quad \Pr[\mathcal{F}_k] \leq \frac{2^{2n+1}\varphi}{k}.}$

$k \to \infty \Rightarrow$ one Pareto-optimal per bin

Let's make this more formal

For $k \in \mathbb{N}$, denote by $\mathcal{F}_k$, the event that there exist two solutions $x, y \in \{0,1\}^n$ s.t. $|w^T x - w^T y| \leq 1/k$.

Lemma: For $k \in \mathbb{N}$, $\Pr[\mathcal{F}_k] \leq \frac{2^{2n+1}\varphi}{k}$.

$k \to \infty \Rightarrow$ one Pareto-optimal per bin

proof: Let $x, y \in \{0,1\}^n$ with $x \neq y$. For some $i \in [n]$, $x_i \neq y_i$.

Wlog, $x_i = 0$ and $y_i = 1$. Say, all weights but w_i have been drawn. Then, $w^T x - w^T y = \kappa - w_i$, where κ depends on x, y and all other weights but w_i. Now, $\Pr[|w^T x - w^T y| \leq 1/k] \leq \Pr[|\kappa - w_i| \leq 1/k]$

↳ principle of deferred decisions

$= \Pr[w_i \in [\kappa - 1/k, \kappa + 1/k]] \leq \frac{2}{k} \cdot \varphi$.

$f_{w_i} : [0,1] \to [0, \varphi]$

Finally, take the union bound over all pairs x, y.

Now, let's bound

$$\Pr\left[\exists x \in P \cap \left(\frac{\ell}{k}, \frac{\ell+1}{k}\right]\right], \quad \text{for some } \ell \in 0 \cup [nk-1]$$

Now, let's bound

$$\Pr\left[\exists x \in P \cap \left(\frac{\ell}{k}, \frac{\ell+1}{k} \right] \right] \text{, for some } \ell \in 0 \cup [nk-1]$$

Let $x^* := \operatorname{argmax}_{x \in P} \left\{ p^T x : w^T x \leq \frac{\ell}{k} \right\}$ (unique with probability 1)

Now, let's bound

$$\Pr\left[\exists x \in P \cap \left(\frac{\ell}{k}, \frac{\ell+1}{k}\right]\right], \quad \text{for some } \ell \in 0 \cup [nk-1]$$

Let $x^* := \underset{x \in P}{\operatorname{argmax}} \left\{ p^T x : w^T x \leq \frac{\ell}{k} \right\}$ (unique with probability 1)

$$\hat{x} := \begin{cases} \underset{x \in P}{\operatorname{argmin}} \left\{ p^T x : w^T x > \ell/k \right\}, & \text{if } \left\{ x \in P : w^T x > \ell/k \right\} \neq \emptyset \\ \perp, & \text{otherwise} \end{cases}$$

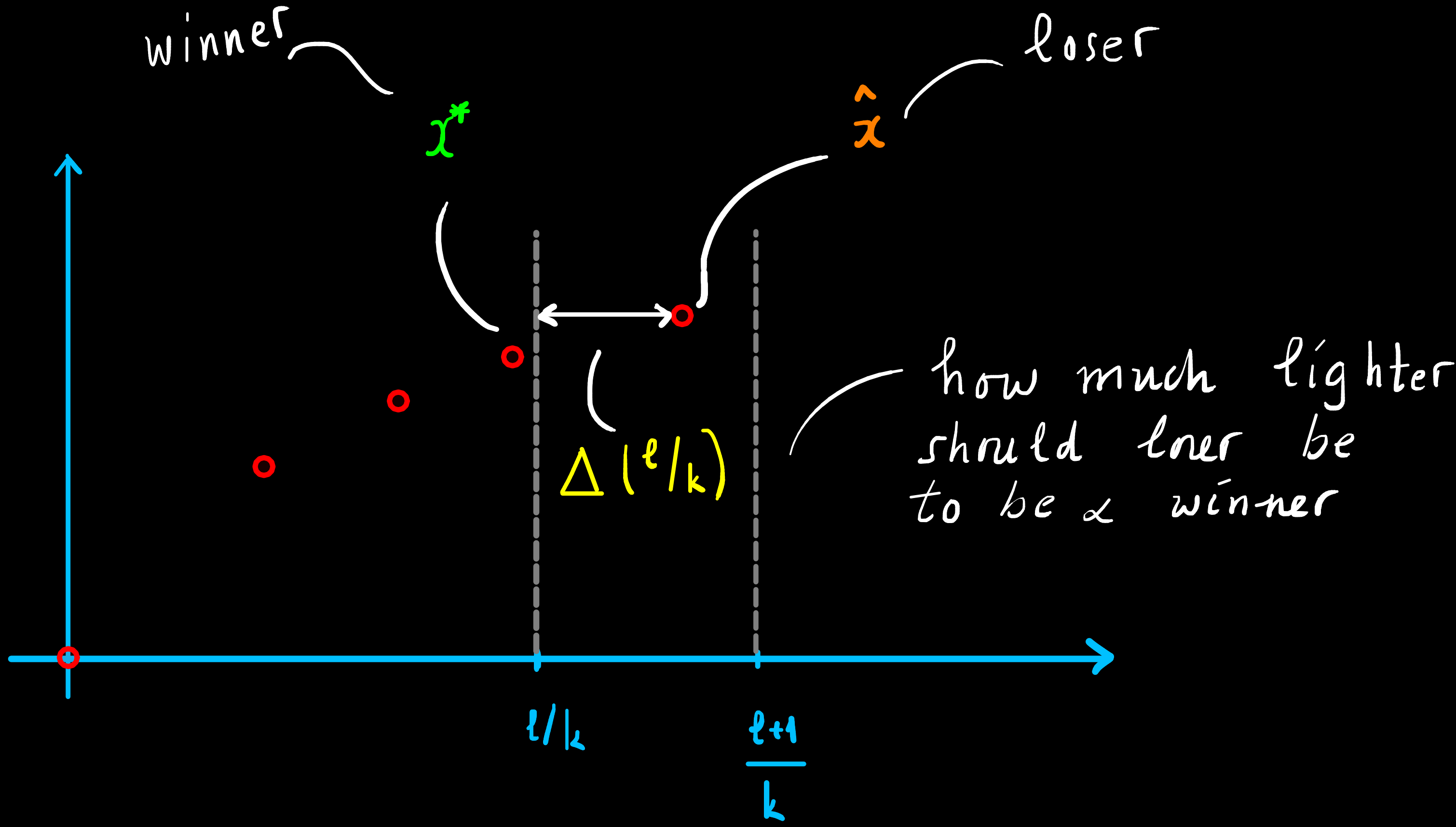
Now, let's bound

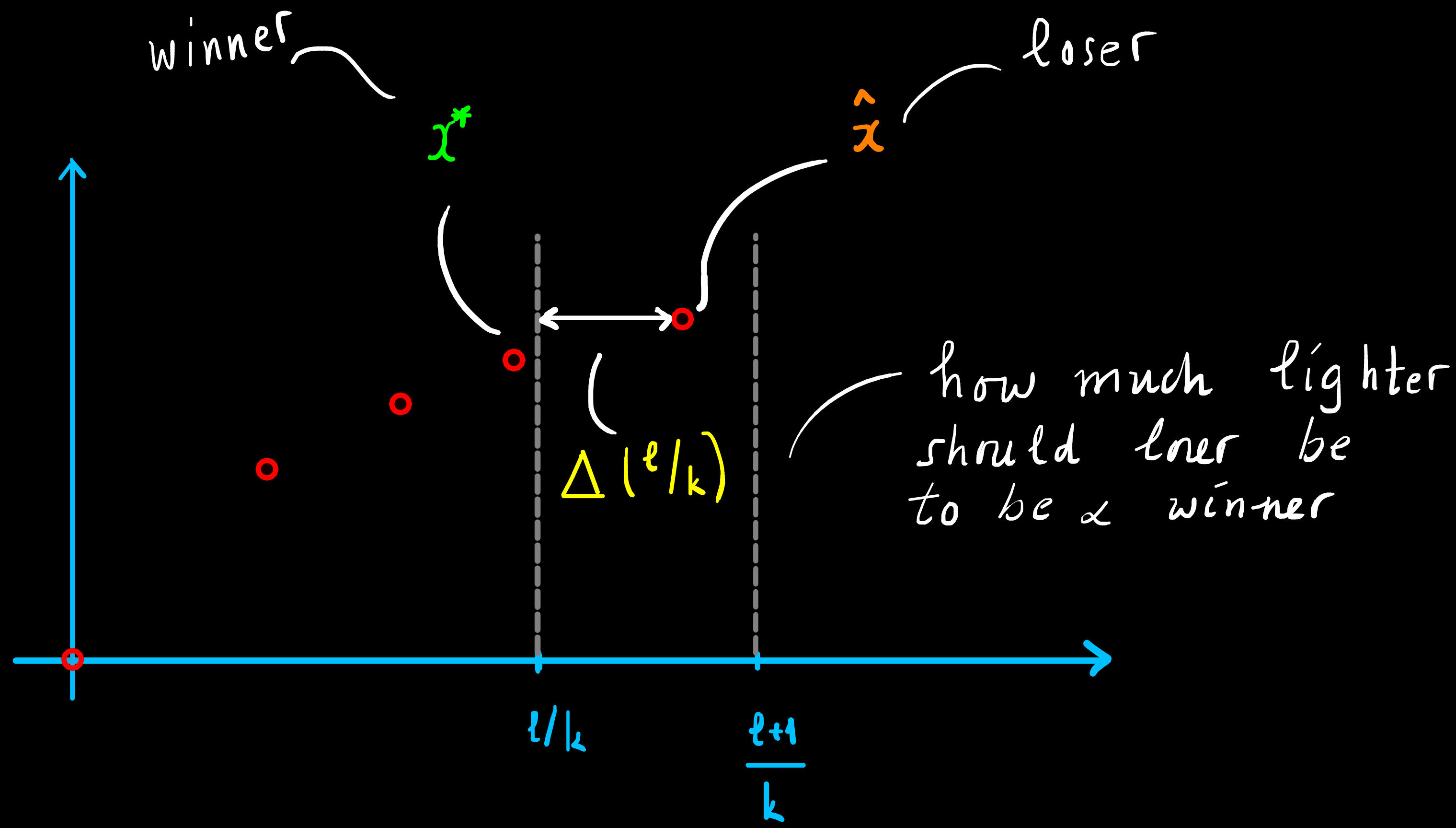
$$\Pr\left[\exists x \in P \cap \left(\frac{\ell}{k}, \frac{\ell+1}{k}\right]\right], \quad \text{for some } \ell \in 0 \cup [nk-1]$$

Let $x^* := \underset{x \in P}{\operatorname{argmax}} \left\{ p^T x : w^T x \leq \frac{\ell}{k} \right\}$ (unique with probability 1)

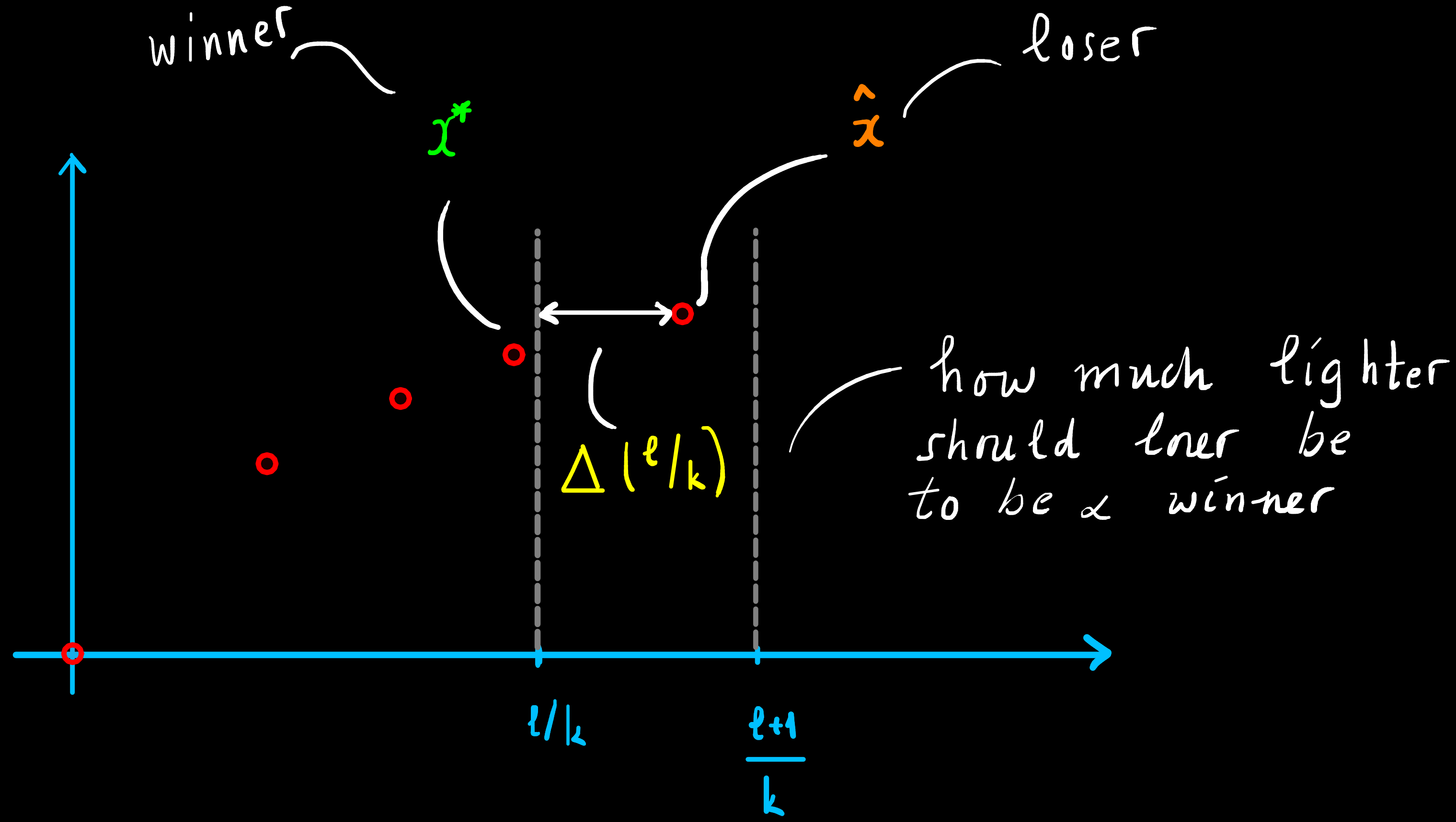
$$\hat{x} := \begin{cases} \underset{x \in P}{\operatorname{argmin}} \left\{ p^T x : w^T x > \ell/k \right\}, & \text{if } \left\{ x \in P : w^T x > \ell/k \right\} \neq \emptyset \\ \perp, & \text{otherwise} \end{cases}$$

$$\Delta\left(\frac{\ell}{k}\right) := \begin{cases} w^T \hat{x} - \ell/k, & \text{if } \hat{x} \neq \perp \\ \infty, & \text{otherwise} \end{cases}$$


winner
x^*
$\hat{x}$
loser
$\Delta(\ell/k)$
how much lighter should loser be to be a winner
ℓ/k
$\frac{\ell+1}{k}$



Observation: $P \cap \left(\frac{\ell}{k}, \frac{\ell+1}{k}\right] \neq \emptyset \iff \Delta(\ell/k) \leq 1/k$



Observation: $P \cap \left(\frac{\ell}{k}, \frac{\ell+1}{k}\right] \neq \emptyset \iff \Delta(\ell/k) \leq 1/k$

Thus, it suffices to bound $\Pr[\Delta(\ell/k) \in (0, 1/k]]$.

Claim: $\Pr\left[\Delta(\ell/k) \in (0, 1/k] \right] \leq \frac{n}{k} \cdot \varphi.$

$$\boxed{\underline{\text{Claim}}: \ \Pr\left[\Delta(\ell/k) \in (0, 1/k]\right] \leq \frac{n}{k} \cdot \varphi.}$$

Assuming the claim:

$$\mathbb{E}[|P|] \leq 1 + \sum_{\ell=0}^{nk} \Pr\left[P \cap \left(\tfrac{\ell}{k}, \tfrac{\ell+1}{k}\right] \neq \emptyset\right]$$

$$= 1 + \sum_{\ell=0}^{nk} \Pr\left[\Delta(\ell/k) \in (0, 1/k]\right]$$

$$\leq 1 + \sum_{\ell=0}^{nk} \frac{n}{k} \varphi$$

Main Theorem ✓

$$= 1 + n\cancel{k} \cdot \frac{n}{\cancel{k}} \varphi = 1 + n^2 \varphi.$$

Now, let's prove the Claim!

First, let's introduce some more definitions.

First, let's introduce some more definitions.

For every $i \in [n]$,

$$x^{*,i} := \underset{\substack{x \in P \\ x_i = 0}}{\operatorname{argmax}} \left\{ p^T x \;:\; w^T x \leq 1/k \right\}$$

$$\hat{x}^i := \begin{cases} \underset{\substack{x \in P \\ x_i = 1}}{\operatorname{argmin}} \left\{ p^T x : w^T x > \ell/k \right\}, & \text{if } \left\{ x \in P : x_i = 1,\ w^T x > \ell/k \right\} \neq \emptyset \\ \perp, & \text{otherwise} \end{cases}$$

$$\Delta^i(\ell/k) := \begin{cases} w^T \hat{x}^i - \ell/k, & \text{if } \hat{x}^i \neq \perp \\ \infty, & \text{otherwise} \end{cases}$$

If $\hat{x}$ exists, then $\hat{x}$ contains at least one item that x^* does not contain. (why?)

If $\hat{x}$ exists, then $\hat{x}$ contains at least one item that x^* does not contain. (why?)

Therefore, if i is the index of such an item

$$\Delta(\ell/k) = \Delta^i(\ell/k).$$

If $\hat{x}$ exists, then $\hat{x}$ contains at least one item that x^* does not contain. (why?)

Therefore, if i is the index of such an item

$$\Delta(\ell/k) = \Delta^i(\ell/k).$$

<u>Claim′</u>: $\Pr\left[\Delta^i(\ell/k) \in (0, 1/k]\right] \leq \frac{\varphi}{k}.$

If $\hat{x}$ exists, then $\hat{x}$ contains at least one item that x^* does not contain. (why?)

Therefore, if i is the index of such an item

$$\Delta(\ell/k) = \Delta^i(\ell/k).$$

$$\boxed{\underline{\text{Claim}'}: \quad \Pr[\Delta^i(\ell/k) \in (0, 1/k]] \le \frac{\varphi}{k}.}$$

+ union bound

If $\hat{x}$ exists, then $\hat{x}$ contains at least one item that x^* does not contain. (why?)

Therefore, if i is the index of such an item

$$\Delta(\ell/k) = \Delta^i(\ell/k).$$

$$\boxed{\underline{\text{Claim}'}: \quad \Pr[\Delta^i(\ell/k) \in (0, 1/k]] \leq \frac{\varphi}{k}.}$$

+ union bound

<u>proof</u>: Fix $i \in [n]$, and suppose all weights but w^i are drawn.

Since $x^{*,i}$ does not contain item i it can be determined.

Now, $\hat{x}^i$ does not depend on the value of w_i. Thus,

$$\Pr[w^T \hat{x}^i \in (\ell/k, \ell/k + 1/k]] = \Pr[w' + w_i \in (\ell/k, \ell/k + 1/k]] =$$

$$= \Pr[w_i \in (\underline{\ell/k - w'}, \underline{\ell/k - w'} + 1/k]] \leq \varphi/k.$$

✓

Matching Lower Bound

[Beier, Vöcking '04]

Theorem: Consider the knapsack instance with n items, $p_i = 2^i$ for $i \in [n]$, $w_i \overset{iid}{\sim} \text{Uniform}([0,1])$. Then,

$$\mathbb{E}[|\mathcal{P}|] = \Theta(n^2).$$

Generalization

[Beier, Röglin, Vöcking '07]

Consider the following general "combinatorial optimization" problem

$$\boxed{\begin{array}{lll} \text{maximize} & p^T x & \\ \text{subject to} & Ax \le b & \qquad (\Pi) \\ & x \in S \subseteq \{0,1\}^n & \end{array}}$$

Generalization [Beier, Röglin, Vöcking '07]

Consider the following general "combinatorial optimization" problem

$$\begin{aligned} &\text{maximize} && p^T x \\ &\text{subject to} && Ax \le b && (\Pi) \\ & && x \in S \subseteq \{0,1\}^n \end{aligned}$$

<u>Theorem:</u> Problem Π has polynomial smoothed complexity if solving Π on unitarily encoded instances can be done in polynomial time.

~ References ~

[Teng, Spielman '01]

[Beier, Vöcking '04]

[Beier, Röglin, Vöcking '07]

Heiko Röglin, Probabilistic Analysis of Algorithms, Lecture notes

Bodo Manthey, Algorithms Beyond the Worst Case, Lecture notes