The Knapsack Problem

René Beier

rbeier@mpi-sb.mpg.de

Max-Planck-Institut für Informatik
Saarbrücken, Germany

The Knapsack Problem

Given n items with

- weights $w_1, \ldots, w_n$
- profits $p_1, \ldots, p_n$

and a knapsack of capacity b.

Find a subset $K \subseteq \{1, 2, \ldots, n\}$ such that

- the capacity is not exceeded, i.e., satisfy $\sum_{i \in K} w_i \leq b$,
- and the profit is maximized, i.e., maximize $\sum_{i \in K} p_i$.

Is it difficult?

Worst Case: Knapsack is difficult (NP-hard)

Is it difficult?

Worst Case: Knapsack is difficult (NP-hard)

Average Case: Knapsack is easy ...

Is it difficult?

Worst Case: Knapsack is difficult (NP-hard)

Average Case: Knapsack is easy ...

Example: Uniform Distribution, 1 Mio items

Preprocessing:	850 ms
Finding optimal solution:	10 ms

Is it difficult?

Worst Case: Knapsack is difficult (NP-hard)

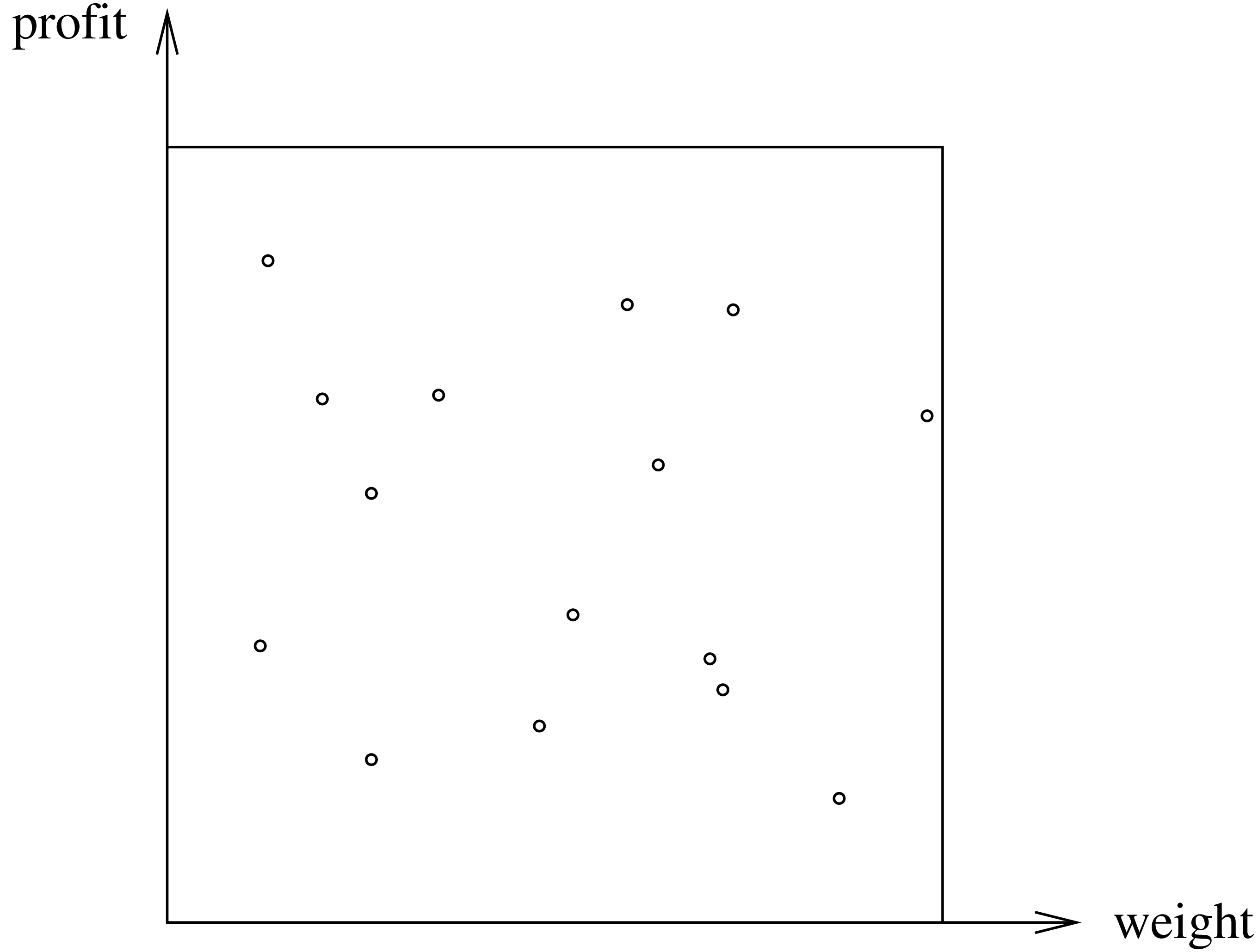
Average Case: Knapsack is easy ...

Example: Uniform Distribution, 1 Mio items

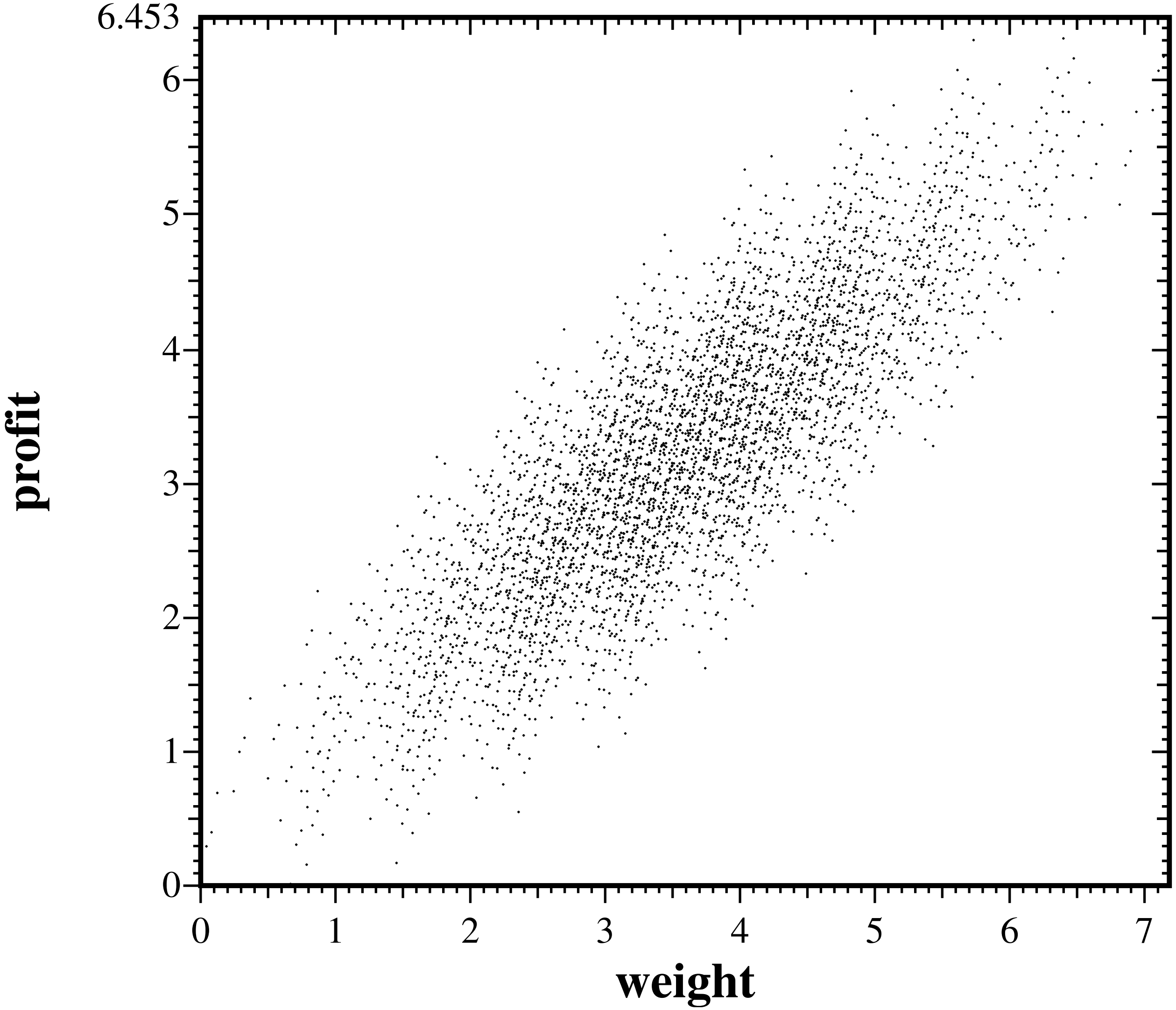
Preprocessing:	850 ms
Finding optimal solution:	10 ms

Key concepts: Core algorithms, Sparse Dynamic Programming

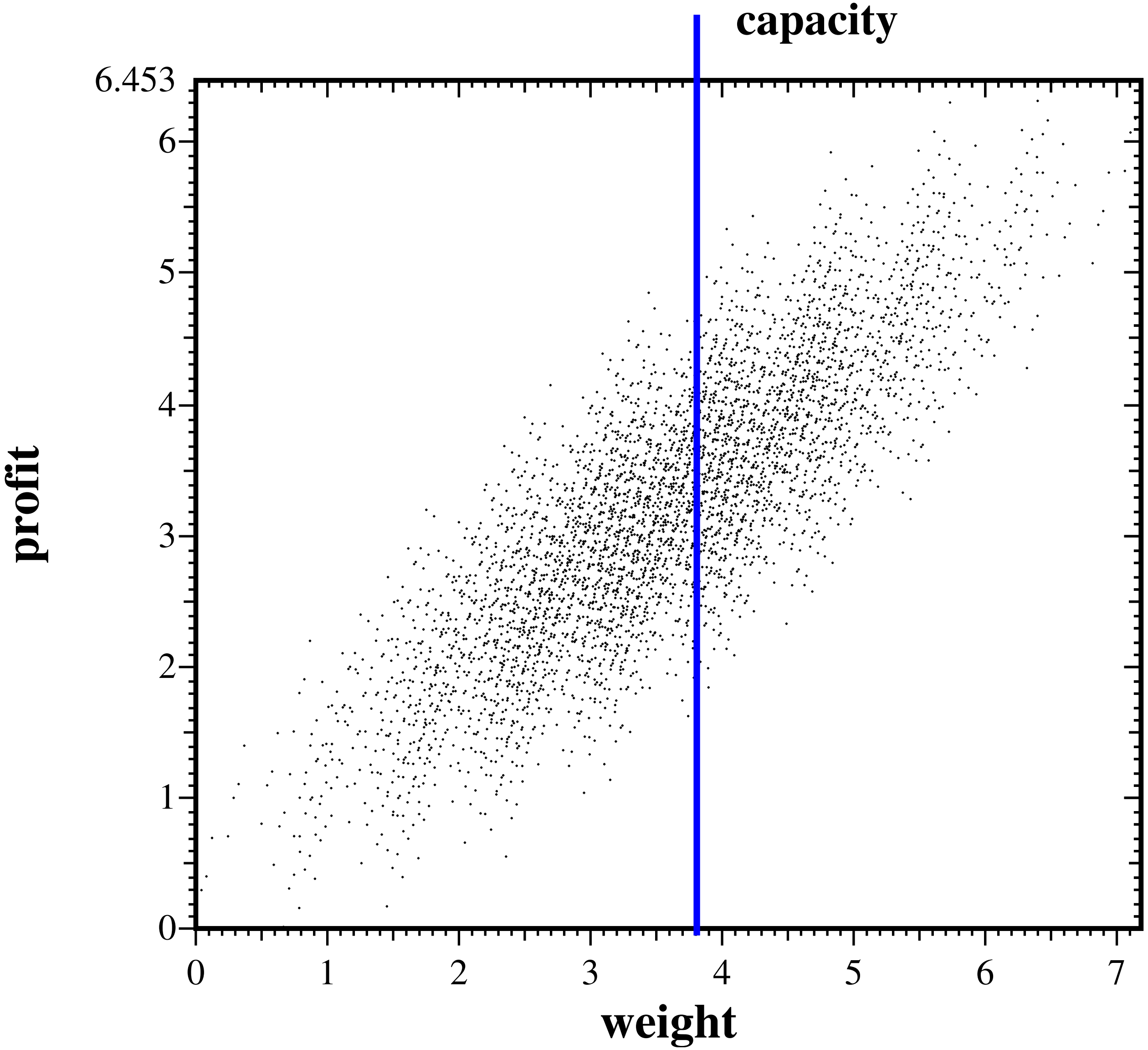
uniform random instance



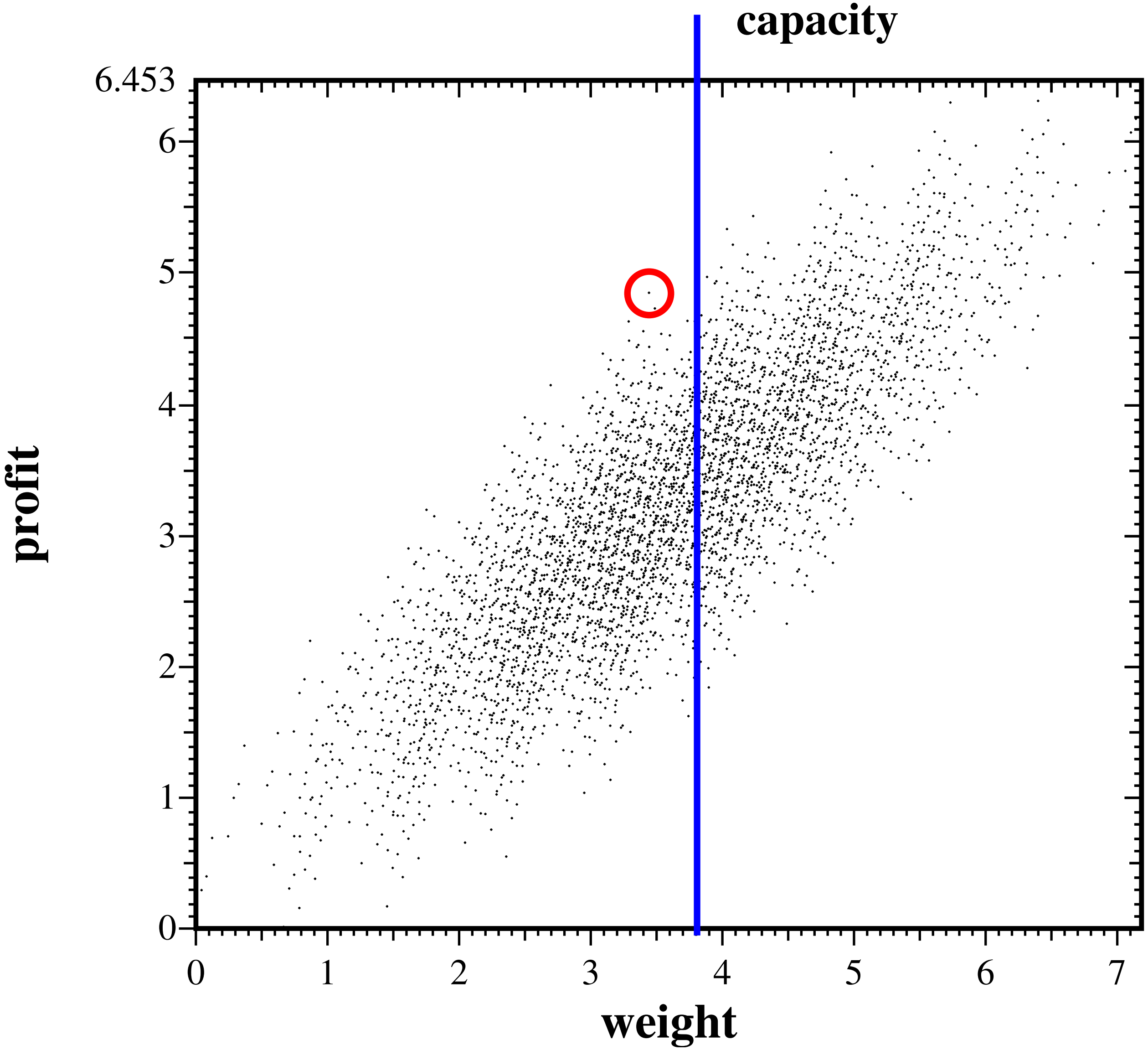
. . . find the best knapsack filling



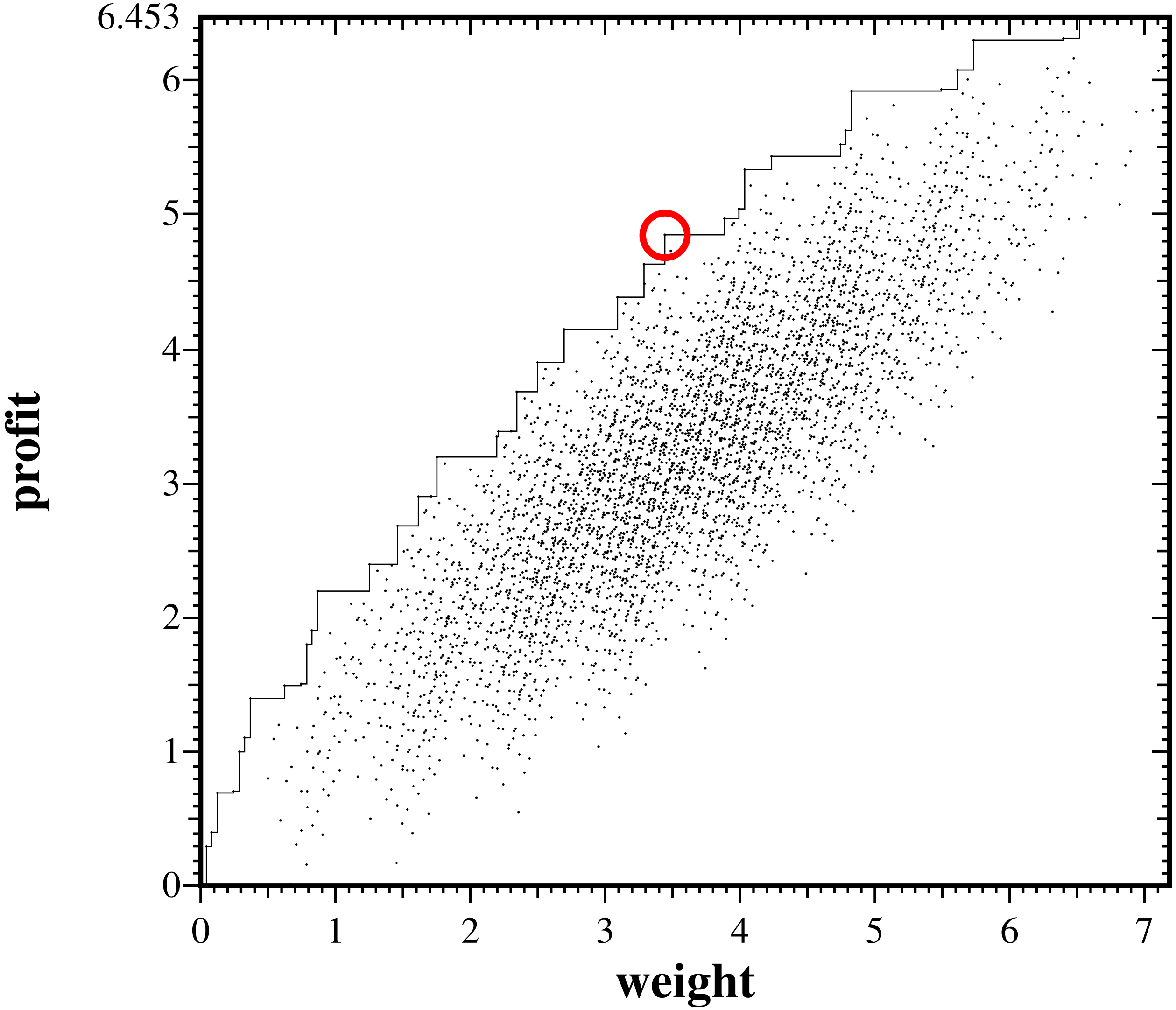
. . . find the best knapsack filling



. . . find the best knapsack filling



. . . find the best knapsack filling



Dominating Sets

A subset $K \subseteq [n]$ is called dominating set if

$$p(K) > p(J), \text{ for all } J \subseteq [n] \text{ with } w(J) \leq w(K).$$

Dominating Sets

A subset $K \subseteq [n]$ is called dominating set if

$$p(K) > p(J), \text{ for all } J \subseteq [n] \text{ with } w(J) \leq w(K).$$

dominating set = Pareto-optimal solution

Dominating Sets

A subset $K \subseteq [n]$ is called dominating set if

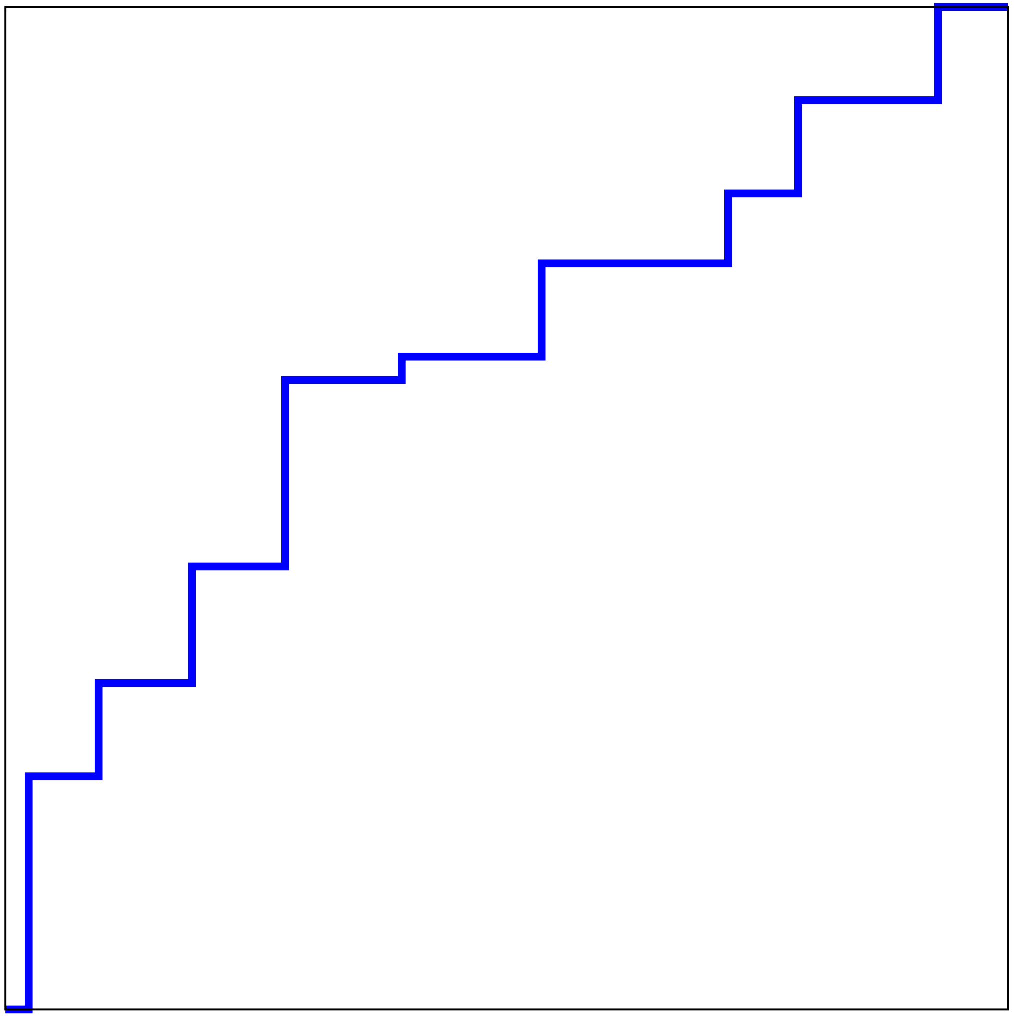
$$p(K) > p(J), \text{ for all } J \subseteq [n] \text{ with } w(J) \leq w(K).$$

dominating set = Pareto-optimal solution

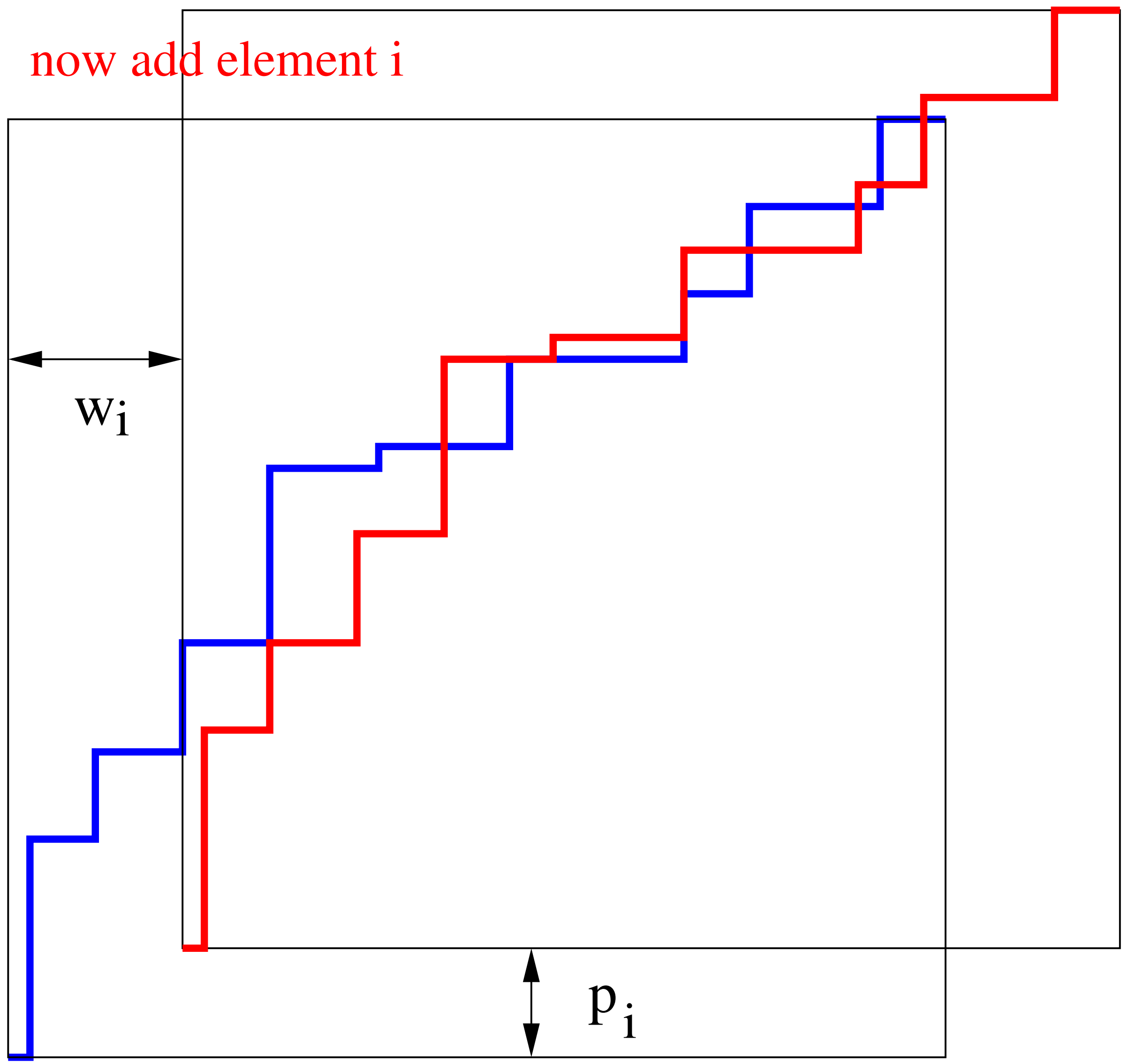
A list with all dominating sets can be calculated efficiently ...

The Nemhauser/Ullmann algorithm (1969)

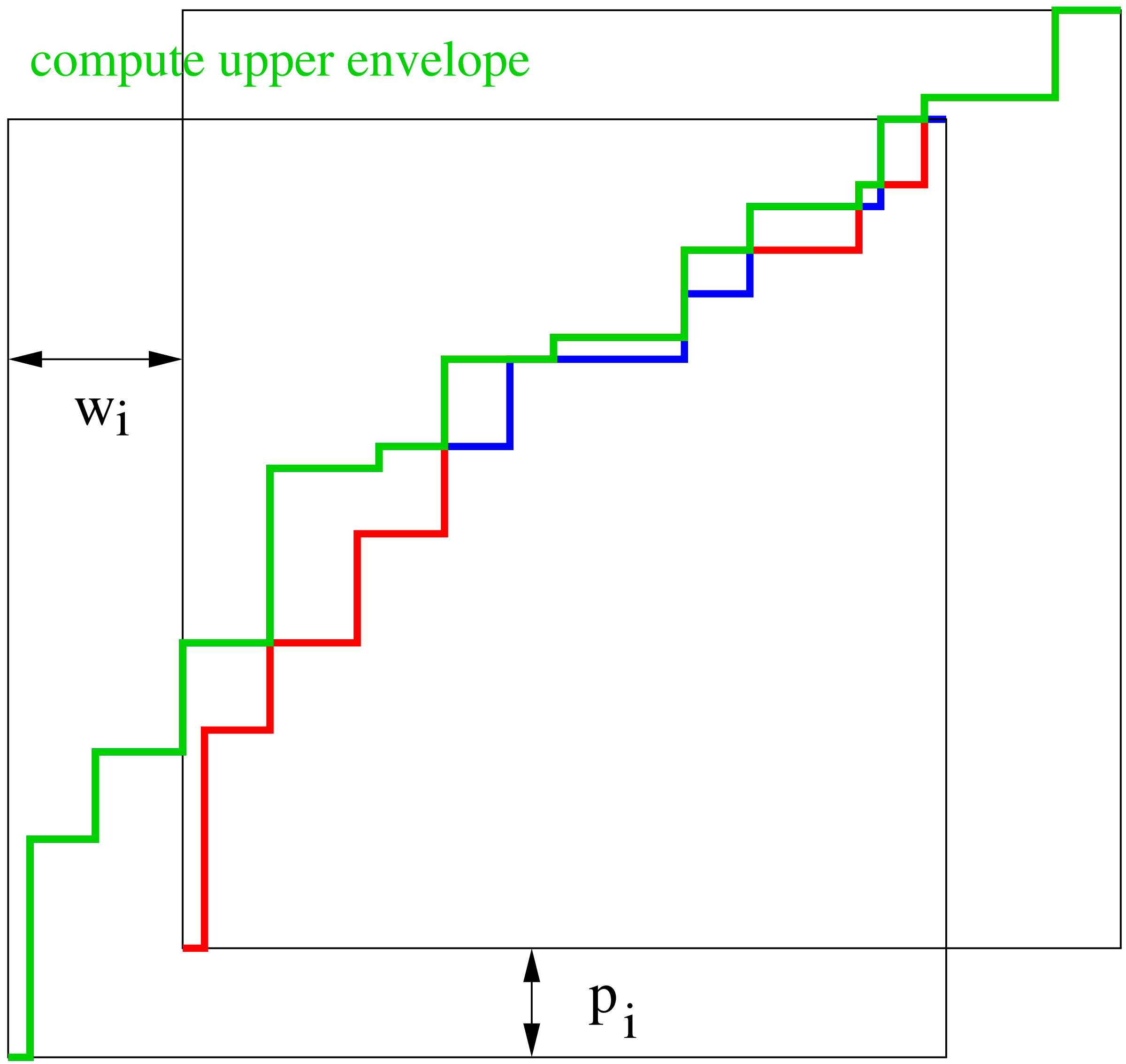
after adding elements 1 to i−1



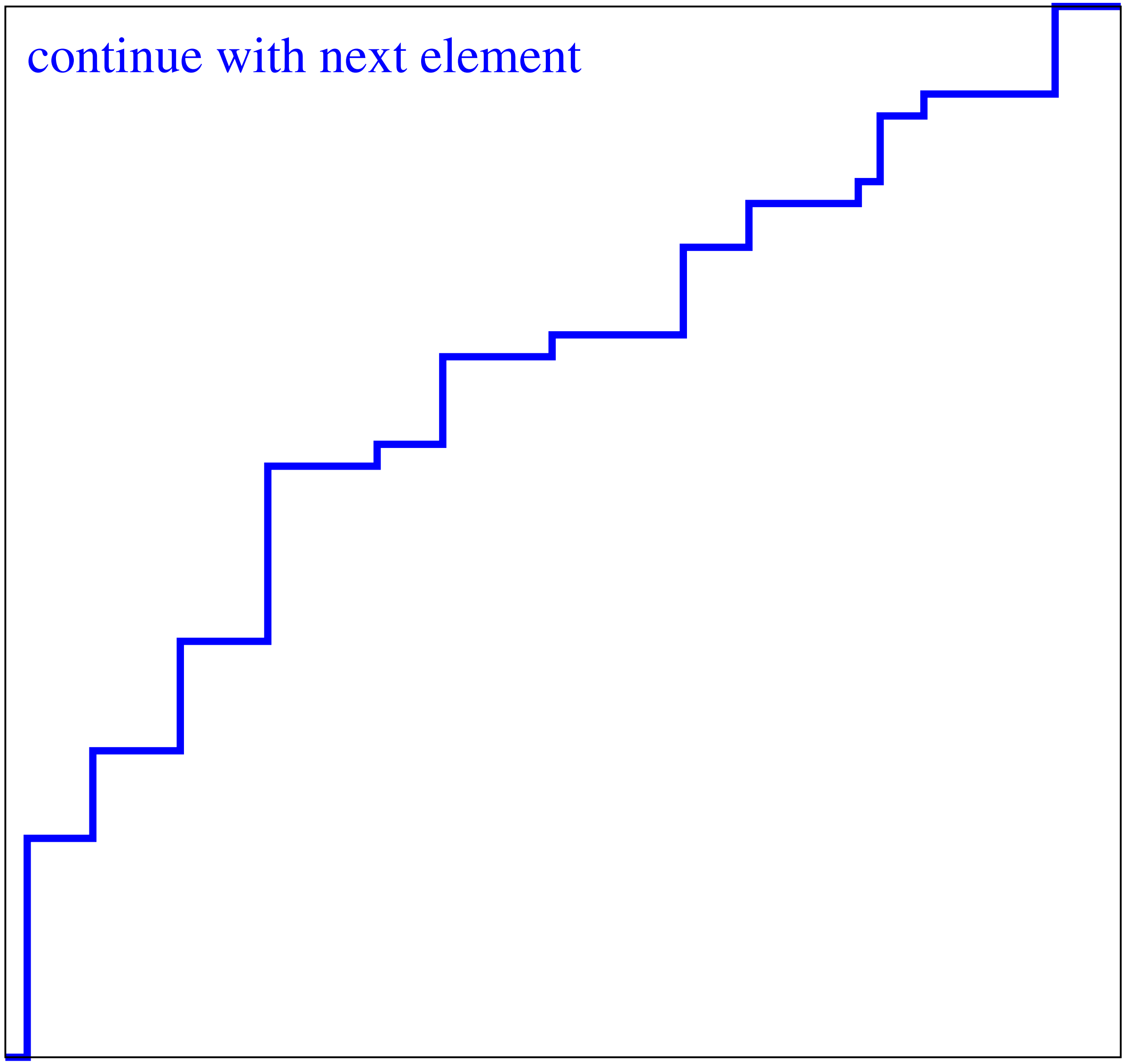
The Nemhauser/Ullmann algorithm (1969)



The Nemhauser/Ullmann algorithm (1969)



The Nemhauser/Ullmann algorithm (1969)



running time analysis

- $q_i \ldots$ # dominating sets over items $\{1, \ldots, i\}$.
- running time $T = O\left(\sum_{i=1}^{n} q_{i-1}\right) \quad \Rightarrow \quad \mathbf{E}[T] = O\left(\sum_{i=1}^{n} \mathbf{E}[q_{i-1}]\right)$

running time analysis

- $q_i \ldots$ # dominating sets over items $\{1, \ldots, i\}$.
- running time $T = O\left(\sum_{i=1}^{n} q_{i-1}\right) \quad \Rightarrow \quad \mathbf{E}[T] = O\left(\sum_{i=1}^{n} \mathbf{E}[q_{i-1}]\right)$
- **Lemma:** For uniform random instances it holds

$$\forall i \geq 1, \mathbf{E}[q_i] \leq 16 i^3$$

running time analysis

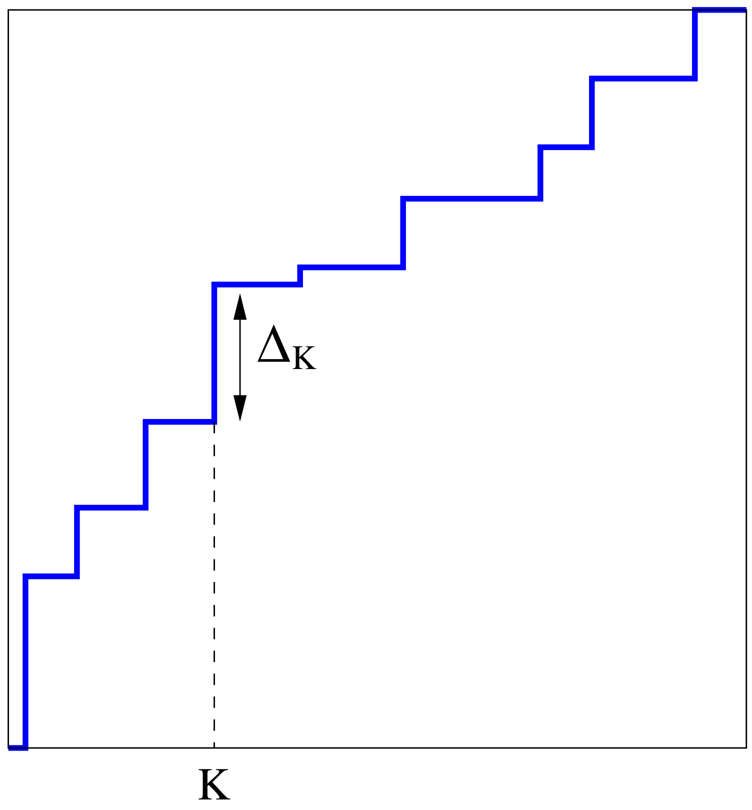
- $q_i \ldots$ # dominating sets over items $\{1, \ldots, i\}$.
- running time $T = O\left(\sum_{i=1}^{n} q_{i-1}\right) \Rightarrow \mathbf{E}[T] = O\left(\sum_{i=1}^{n} \mathbf{E}[q_{i-1}]\right)$
- **Lemma:** For uniform random instances it holds

$$\forall i \geq 1, \mathbf{E}[q_i] \leq 16 i^3$$

- thus $\mathbf{E}[T] = O\left(\sum_{i=1}^{n} i^3\right) = O\left(n^4\right)$

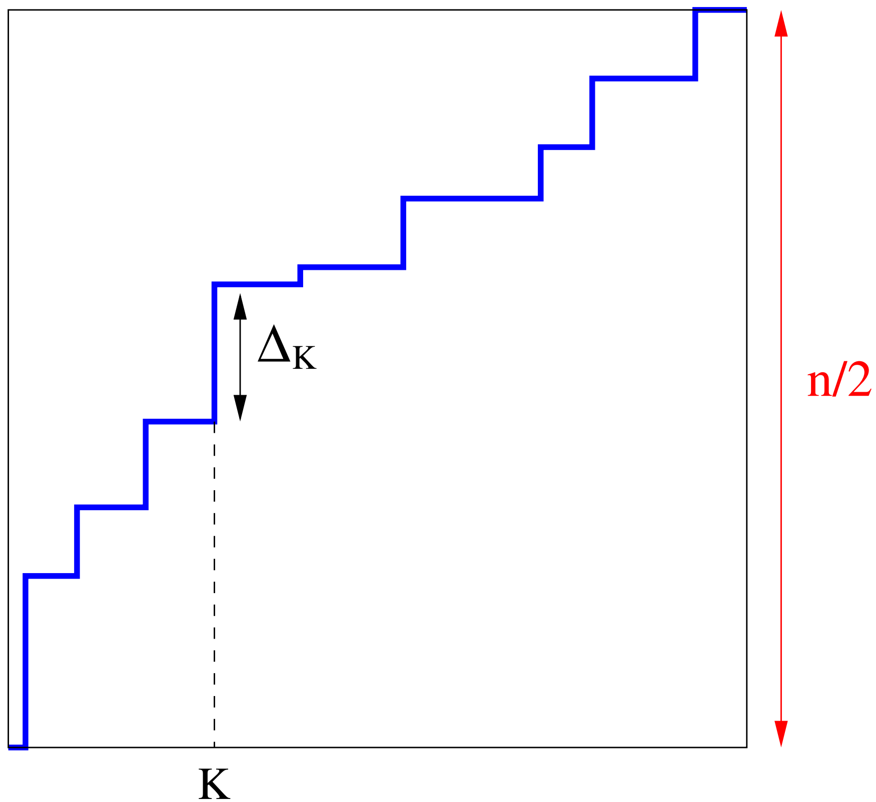
average case analysis - uniform distribution

proof idea: show that the expected step size is large



average case analysis - uniform distribution

proof idea: show that the expected step size is large



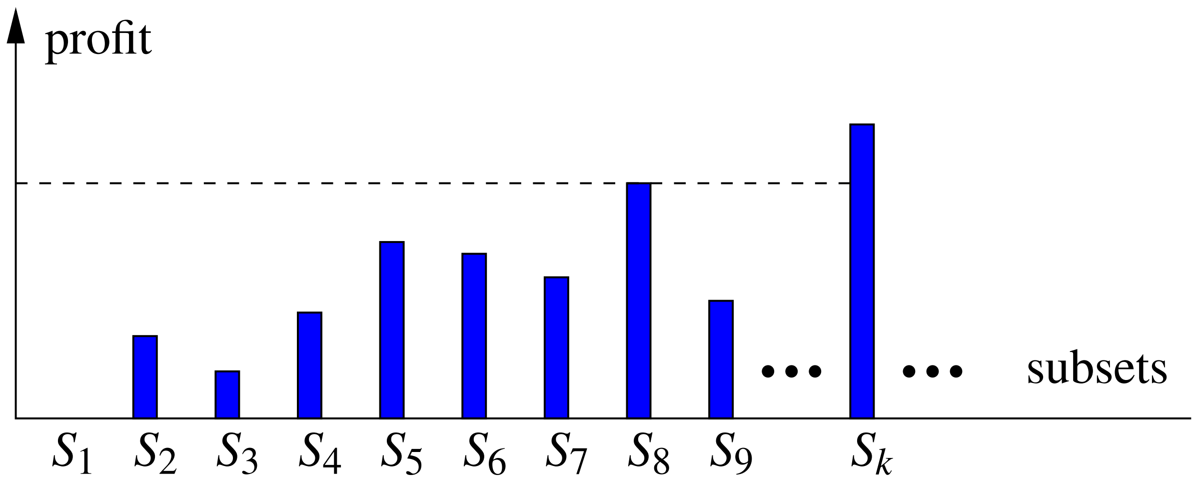
We show $\mathbf{E}\left[\Delta_K \mid \Delta_K > 0\right] \geq \frac{1}{32n^2}$.

average case analysis - uniform distribution

- We show $\mathbf{Pr}\left[\Delta_K \geq \frac{1}{16n^2} \,\middle|\, \Delta_K > 0\right] \geq \frac{1}{2}$.
- This implies $\mathbf{E}\left[\Delta_K \,|\, \Delta_K > 0\right] \geq \frac{1}{32n^2}$.

average case analysis - uniform distribution

- We show $\mathbf{Pr}\left[\Delta_K \geq \frac{1}{16n^2} \,\middle|\, \Delta_K > 0\right] \geq \frac{1}{2}$.
- This implies $\mathbf{E}\left[\Delta_K \,|\, \Delta_K > 0\right] \geq \frac{1}{32n^2}$.
- Let $S_1, \ldots, S_{2^n}$ denote the subsets in non-decreasing order of weight.
- Let K be the k-th subset in this order, i.e., $K = S_k$.



average case analysis - uniform distribution

$$\mathbf{Pr}\left[\Delta_K \geq \frac{1}{16n^2} \,\middle|\, \Delta_K > 0\right]$$

$$= \quad \mathbf{Pr}\left[\forall j < k : p(S_j) \leq p(S_k) - \tfrac{1}{16n^2} \;\middle|\; \forall j < k : p(S_j) \leq p(S_k)\right]$$

$$=$$

$$=$$

average case analysis - uniform distribution

$$\mathbf{Pr}\left[\Delta_K \geq \frac{1}{16n^2} \,\middle|\, \Delta_K > 0\right]$$

$$= \quad \mathbf{Pr}\left[\forall j < k : p(S_j) \leq p(S_k) - \tfrac{1}{16n^2} \,\middle|\, \forall j < k : p(S_j) \leq p(S_k)\right]$$

$$= \quad \mathbf{Pr}\left[\forall j : p(S_j \setminus S_k) \leq P(S_k \setminus S_j) - \tfrac{1}{16n^2} \,\middle|\, \forall j : p(S_j \setminus S_k) \leq P(S_k \setminus S_j)\right]$$

$$=$$

average case analysis - uniform distribution

$$\mathbf{Pr}\left[\Delta_K \geq \frac{1}{16n^2} \,\middle|\, \Delta_K > 0\right]$$

$$= \quad \mathbf{Pr}\left[\forall j < k : p(S_j) \leq p(S_k) - \tfrac{1}{16n^2} \,\middle|\, \forall j < k : p(S_j) \leq p(S_k)\right]$$

$$= \quad \mathbf{Pr}\left[\forall j : p(S_j \setminus S_k) \leq P(S_k \setminus S_j) - \tfrac{1}{16n^2} \,\middle|\, \forall j : p(S_j \setminus S_k) \leq P(S_k \setminus S_j)\right]$$

$$= \quad \mathbf{Pr}\left[\forall j : p(S_j \setminus S_k) \leq \quad U_j \quad - \tfrac{1}{16n^2} \,\middle|\, \forall j < k : p(S_j \setminus S_k) \leq \quad U_j \quad \right]$$

average case analysis - uniform distribution

$$
\begin{aligned}
&\mathbf{Pr}\left[\Delta_K \geq \frac{1}{16n^2} \,\middle|\, \Delta_K > 0\right] \\
&= \mathbf{Pr}\left[\forall j < k : p(S_j) \leq p(S_k) - \tfrac{1}{16n^2} \;\middle|\; \forall j < k : p(S_j) \leq p(S_k)\right] \\
&= \mathbf{Pr}\left[\forall j : p(S_j \setminus S_k) \leq P(S_k \setminus S_j) - \tfrac{1}{16n^2} \;\middle|\; \forall j : p(S_j \setminus S_k) \leq P(S_k \setminus S_j)\right] \\
&= \mathbf{Pr}\Big[\underbrace{\forall j : p(S_j \setminus S_k) \leq \quad U_j \quad - \tfrac{1}{16n^2}}_{\text{polytope } \mathcal{A}} \;\Big|\; \underbrace{\forall j < k : p(S_j \setminus S_k) \leq \quad U_j}_{\text{polytope } \mathcal{B}} \quad\Big]
\end{aligned}
$$

average case analysis - uniform distribution

- Let $\ell = |S_k|$. W.l.o.g. $S_k = \{n-\ell+1, \ldots, n\}$.
- We obtain the $(n-\ell)$-dimensional polytopes:

$$\mathcal{A} = \left\{ (p_1 \times \ldots \times p_{n-\ell}) \in [0,1]^{n-\ell} \;\middle|\; \forall j < k : \sum_{i \in S_j \setminus S_k} p_i \leq U_j - \frac{1}{16n^2} \right\}$$

$$\mathcal{B} = \left\{ (p_1 \times \ldots \times p_{n-\ell}) \in [0,1]^{n-\ell} \;\middle|\; \forall j < k : \sum_{i \in S_j \setminus S_k} p_i \leq U_j \right\}$$

average case analysis - uniform distribution

- Let $\ell = |S_k|$. W.l.o.g. $S_k = \{n-\ell+1, \ldots, n\}$.
- We obtain the $(n-\ell)$-dimensional polytopes:

$$\mathcal{A} = \left\{ (p_1 \times \ldots \times p_{n-\ell}) \in [0,1]^{n-\ell} \;\middle|\; \forall j < k : \sum_{i \in S_j \setminus S_k} p_i \leq U_j - \frac{1}{16n^2} \right\}$$

$$\mathcal{B} = \left\{ (p_1 \times \ldots \times p_{n-\ell}) \in [0,1]^{n-\ell} \;\middle|\; \forall j < k : \sum_{i \in S_j \setminus S_k} p_i \leq U_j \right\}$$

- $\mathbf{Pr}\left[\exists j < k : U_j \leq \frac{1}{4n}\right] \leq \mathbf{Pr}\left[\exists i \in K : p_i \leq \frac{1}{4n}\right] \leq \frac{1}{4}$. Fix $U_j \geq \frac{1}{4n}$.

average case analysis - uniform distribution

- Let $\ell = |S_k|$. W.l.o.g. $S_k = \{n-\ell+1, \ldots, n\}$.
- We obtain the $(n-\ell)$-dimensional polytopes:

$$\mathcal{A} = \left\{ (p_1 \times \ldots \times p_{n-\ell}) \in [0,1]^{n-\ell} \;\middle|\; \forall j < k : \sum_{i \in S_j \setminus S_k} p_i \leq U_j - \frac{1}{16n^2} \right\}$$

$$\mathcal{B} = \left\{ (p_1 \times \ldots \times p_{n-\ell}) \in [0,1]^{n-\ell} \;\middle|\; \forall j < k : \sum_{i \in S_j \setminus S_k} p_i \leq U_j \right\}$$

- $\mathbf{Pr}\left[\exists j < k : U_j \leq \frac{1}{4n}\right] \leq \mathbf{Pr}\left[\exists i \in K : p_i \leq \frac{1}{4n}\right] \leq \frac{1}{4}$. Fix $U_j \geq \frac{1}{4n}$.
- We will argue that $\mathbf{Pr}\left[\mathcal{A} | \mathcal{B}\right] = \frac{vol(\mathcal{A})}{vol(\mathcal{B})} \geq \frac{3}{4}$.

average case analysis - uniform distribution

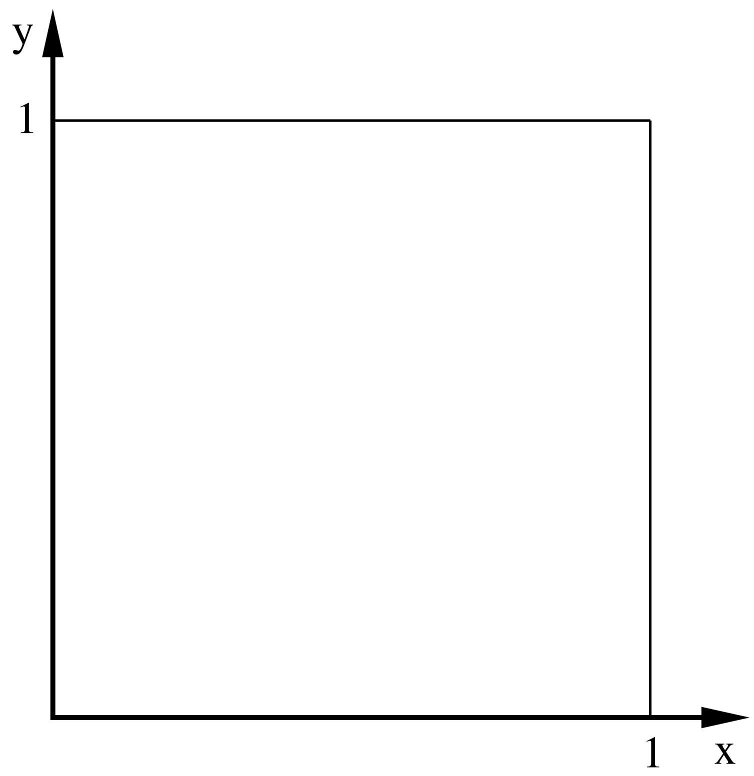
- Let $\ell = |S_k|$. W.l.o.g. $S_k = \{n-\ell+1,\ldots,n\}$.
- We obtain the $(n-\ell)$-dimensional polytopes:

$$\mathcal{A} = \left\{ (p_1 \times \ldots \times p_{n-\ell}) \in [0,1]^{n-\ell} \;\middle|\; \forall j < k : \sum_{i \in S_j \setminus S_k} p_i \leq U_j - \frac{1}{16n^2} \right\}$$

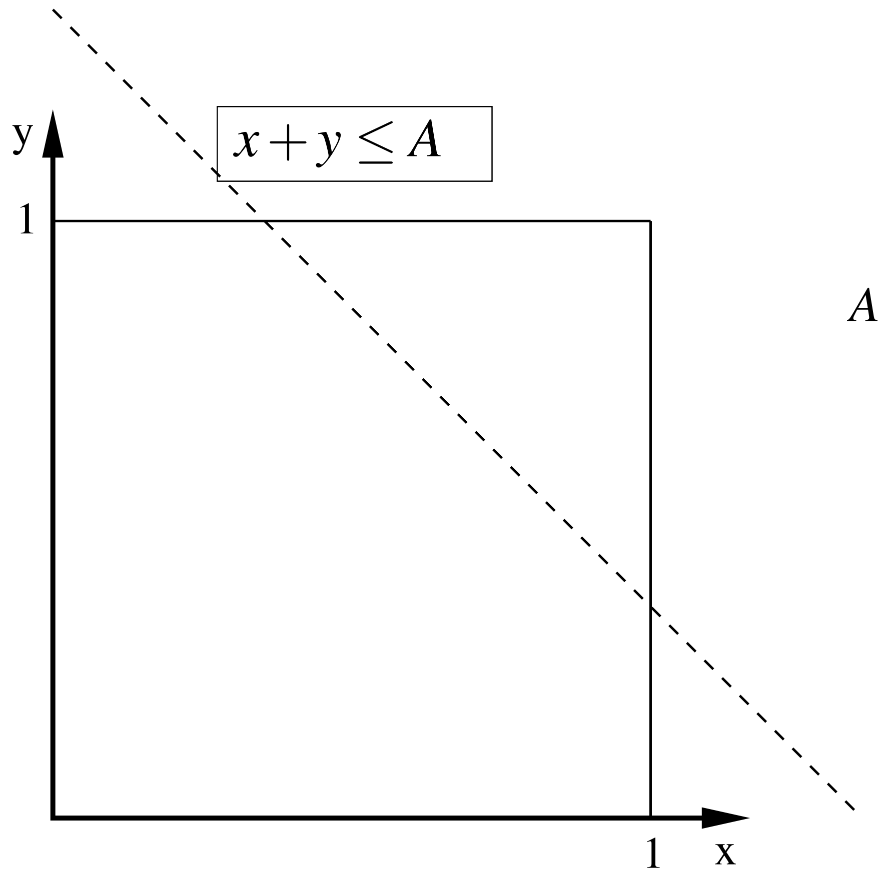
$$\mathcal{B} = \left\{ (p_1 \times \ldots \times p_{n-\ell}) \in [0,1]^{n-\ell} \;\middle|\; \forall j < k : \sum_{i \in S_j \setminus S_k} p_i \leq U_j \right\}$$

- $\mathbf{Pr}\left[\exists j < k : U_j \leq \frac{1}{4n}\right] \leq \mathbf{Pr}\left[\exists i \in K : p_i \leq \frac{1}{4n}\right] \leq \frac{1}{4}$. Fix $U_j \geq \frac{1}{4n}$.
- We will argue that $\mathbf{Pr}\left[\mathcal{A}|\mathcal{B}\right] = \frac{vol(\mathcal{A})}{vol(\mathcal{B})} \geq \frac{3}{4}$.
- This implies $\mathbf{Pr}\left[\Delta_K \geq \frac{1}{16n^2} | \Delta_K > 0\right] \geq \frac{1}{2}$.

the polytopes

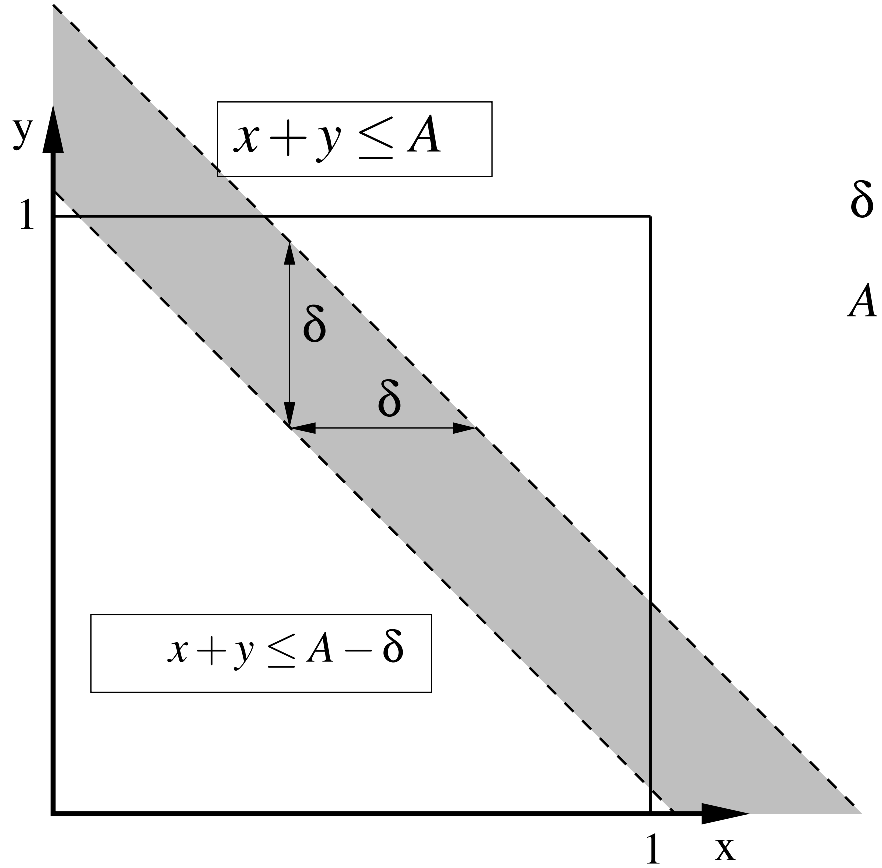


the polytopes



$A \geq \frac{1}{4n}$

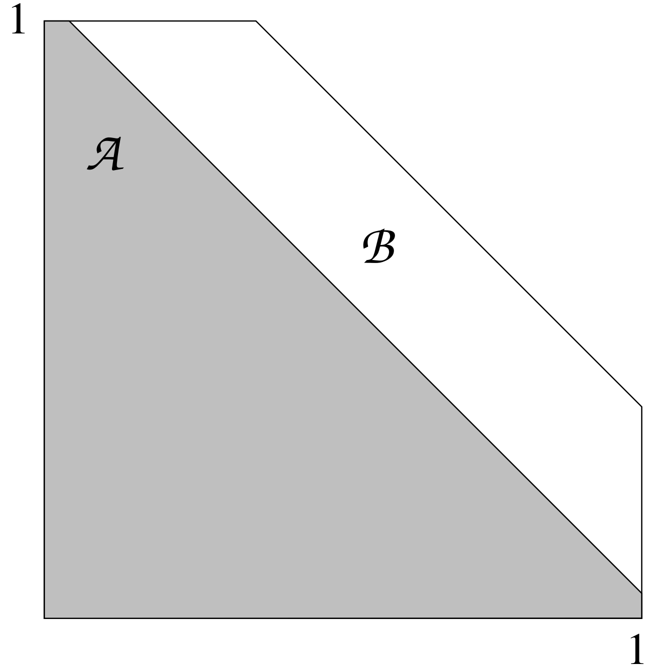
the polytopes



$$\delta = \frac{1}{16n^2}$$

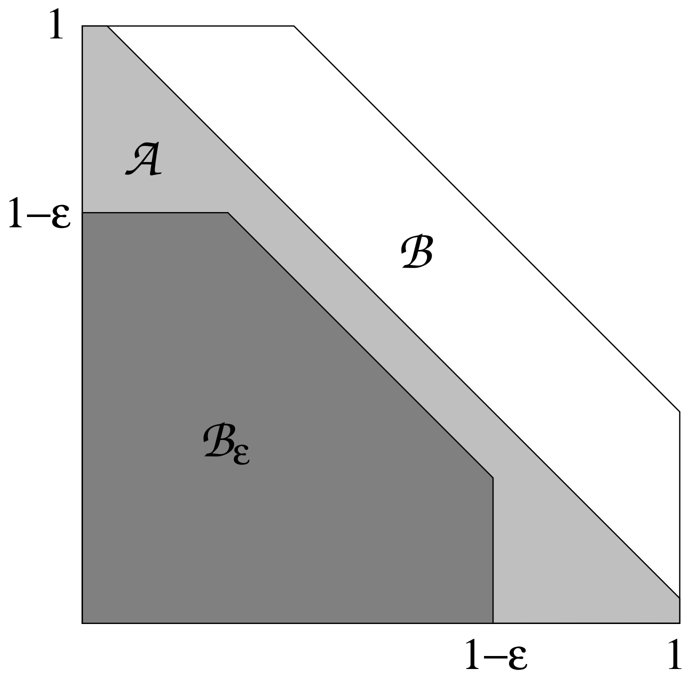
$$A \geq \frac{1}{4n}$$

the polytopes



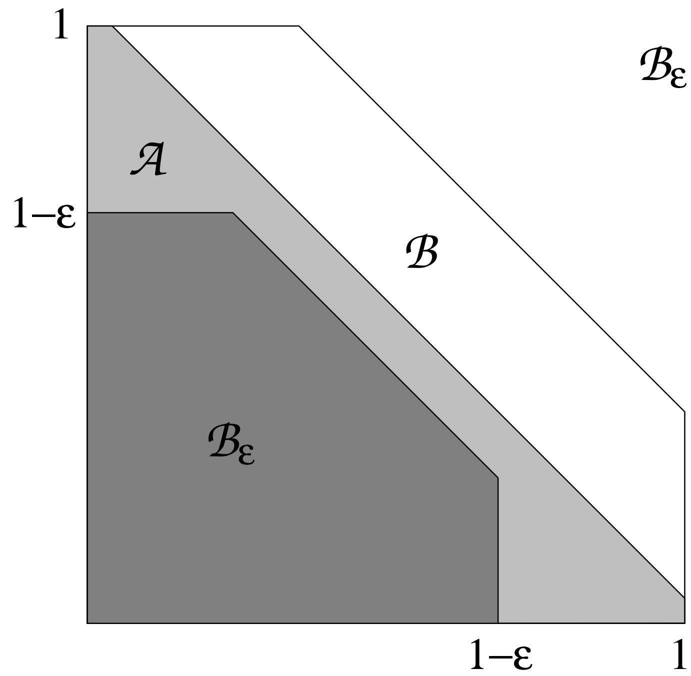
the polytopes

$\mathcal{B}_\varepsilon := \mathcal{B}$ shrinked by factor $1-\varepsilon$ in each dimension.



the polytopes

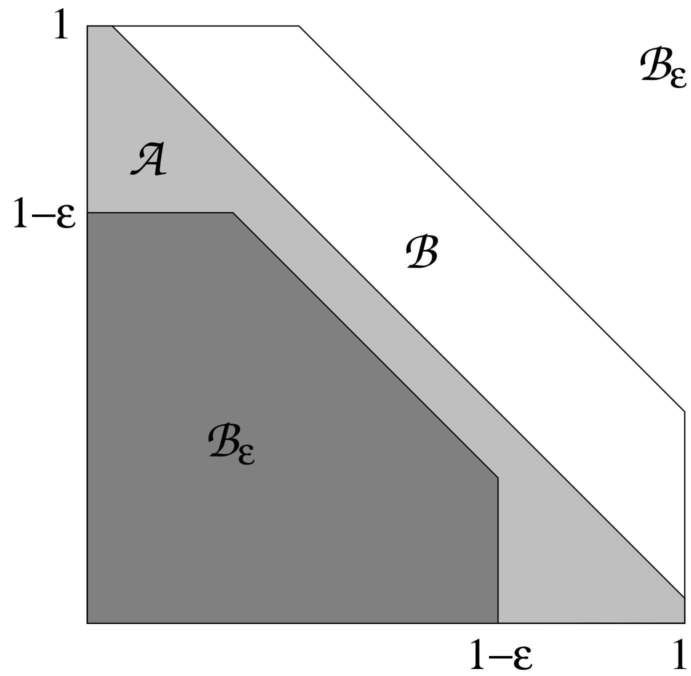
$\mathcal{B}_\varepsilon := \mathcal{B}$ shrinked by factor $1-\varepsilon$ in each dimension.



$\mathcal{B}_\varepsilon \subseteq \mathcal{A}$ for $\varepsilon = \frac{1}{4n}$.

the polytopes

$\mathcal{B}_\varepsilon := \mathcal{B}$ shrinked by factor $1-\varepsilon$ in each dimension.



$\mathcal{B}_\varepsilon \subseteq \mathcal{A}$ for $\varepsilon = \frac{1}{4n}$.

$$vol(\mathcal{A}) \geq vol(\mathcal{B}_\varepsilon) = (1-\varepsilon)^{n-k}\, vol(\mathcal{B}) \geq (1-\varepsilon n)\, vol(\mathcal{B}) \geq \frac{3}{4}\, vol(\mathcal{B})$$

Complexity measures

- **Worst case:** Adversary has full control (e.g. $p_i := w_i$)

Complexity measures

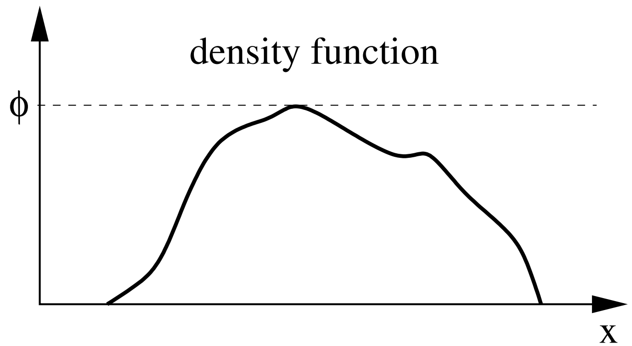
- **Worst case:** Adversary has full control (e.g. $p_i := w_i$)
- **Average case:** Adversary has no control (e.g. p_i uniformly random)

Complexity measures

- **Worst case:** Adversary has full control (e.g. $p_i := w_i$)
- **Average case:** Adversary has no control (e.g. p_i uniformly random)
- **Smoothed complexity:** Adversary has limited control
  1. adversary fixes some instance X with norm at most 1 ($w_i \in [0,1]$)
  2. Instance X is randomly perturbed with parameter δ.
     ($p_i := p_i + r[0,\delta]$)
  3. Smoothed complexity is expected running time w.r.t. random perturbation. (function of n and δ)

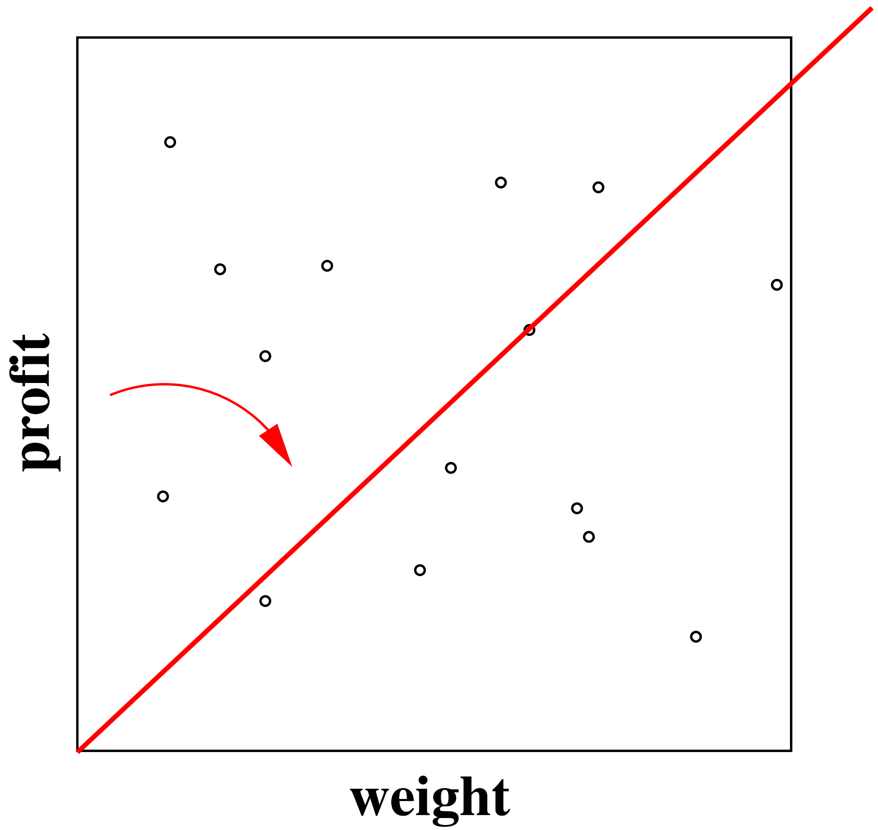
New Result – General Distributions

- Suppose weights are chosen by an adversary.
- Suppose profits are chosen independently according to arbitrary, possibly different probability distributions with finite mean.
- Let μ denote the maximum expected profit.
- Let ϕ denote the supremum of the profit density functions.
- Then the Nemhauser/Ullmann algorithm has expected running time $O(\mu\phi n^5)$.



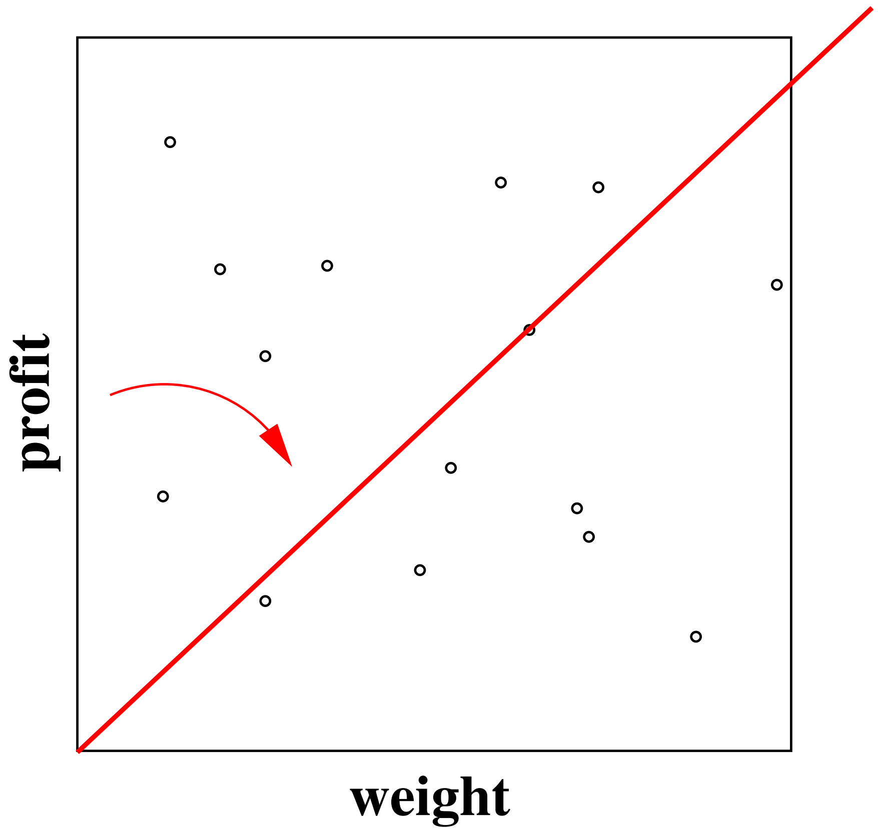
The Core Concept

optimal fractional solution: Dantzig's relaxation (1957)

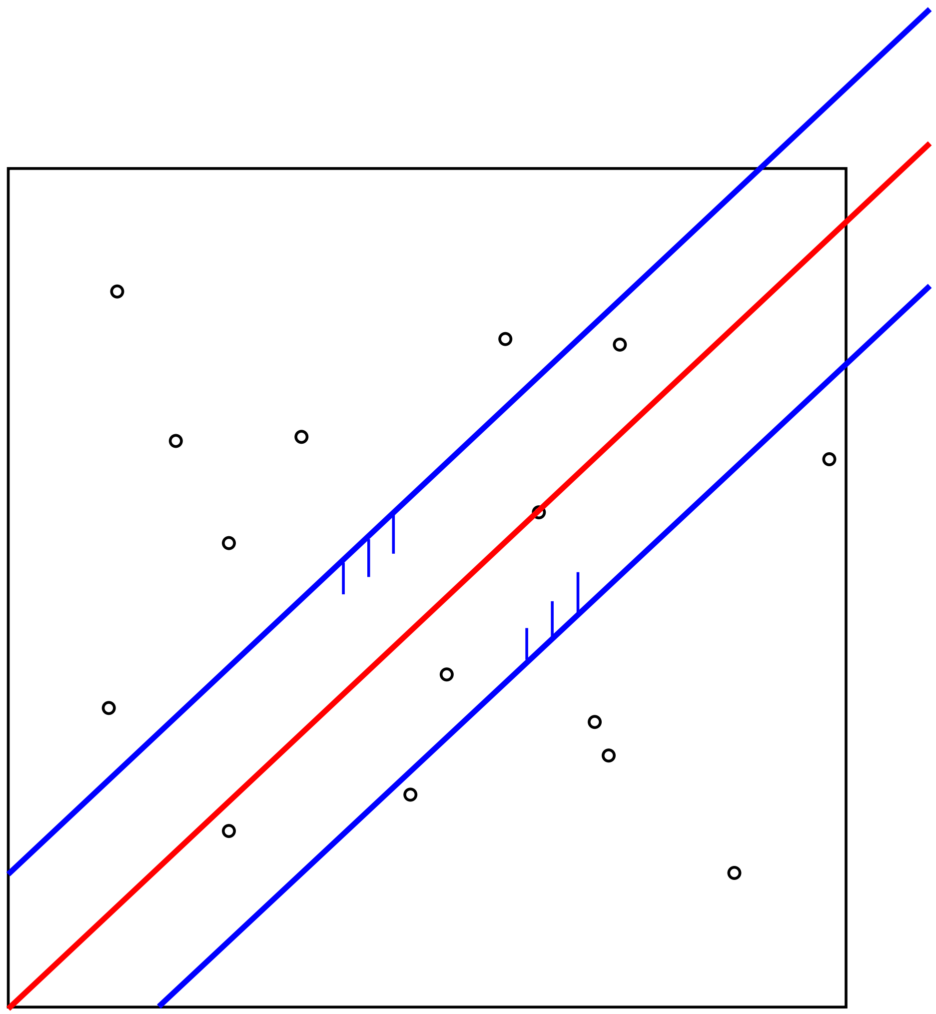


optimal fractional solution: Dantzig's relaxation (1957)

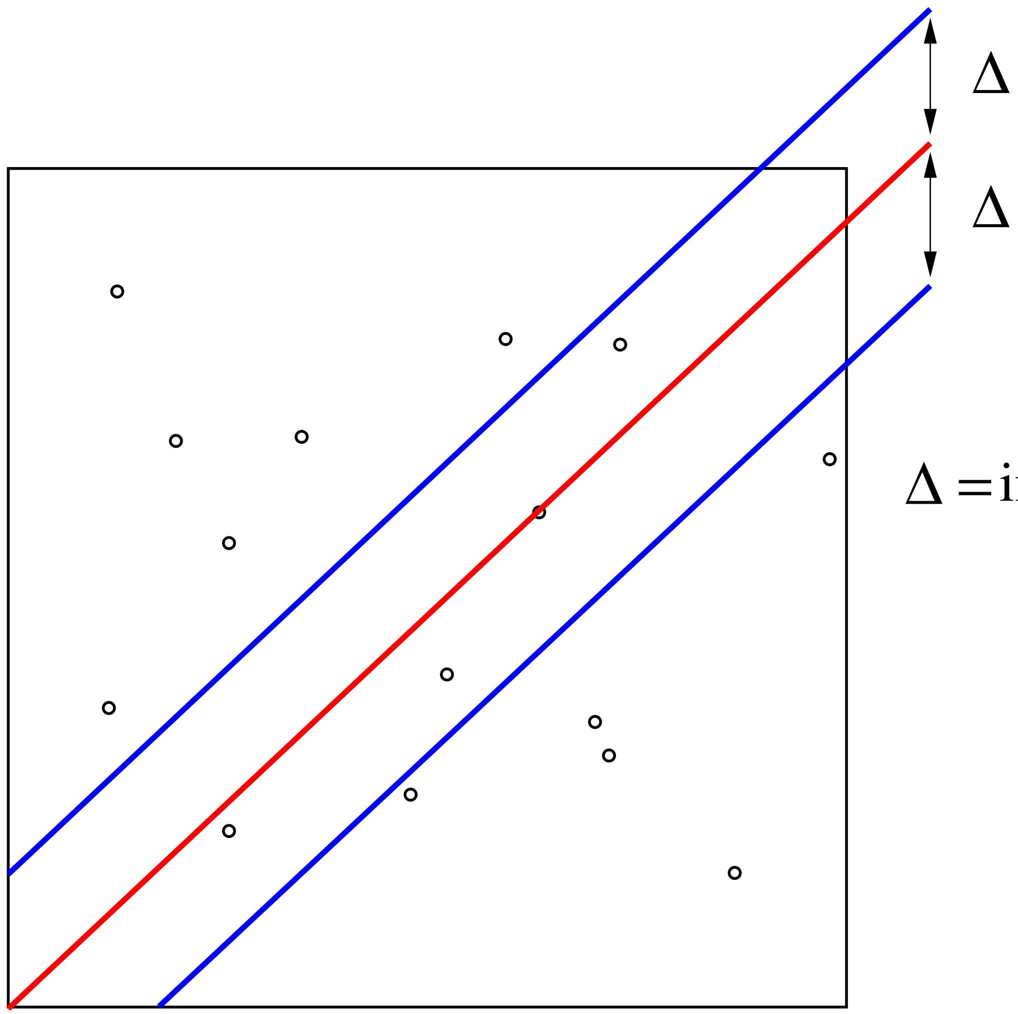
- Linear time (Balas and Zemel, 1980)



core algorithms (Balas and Zemel, 1980)

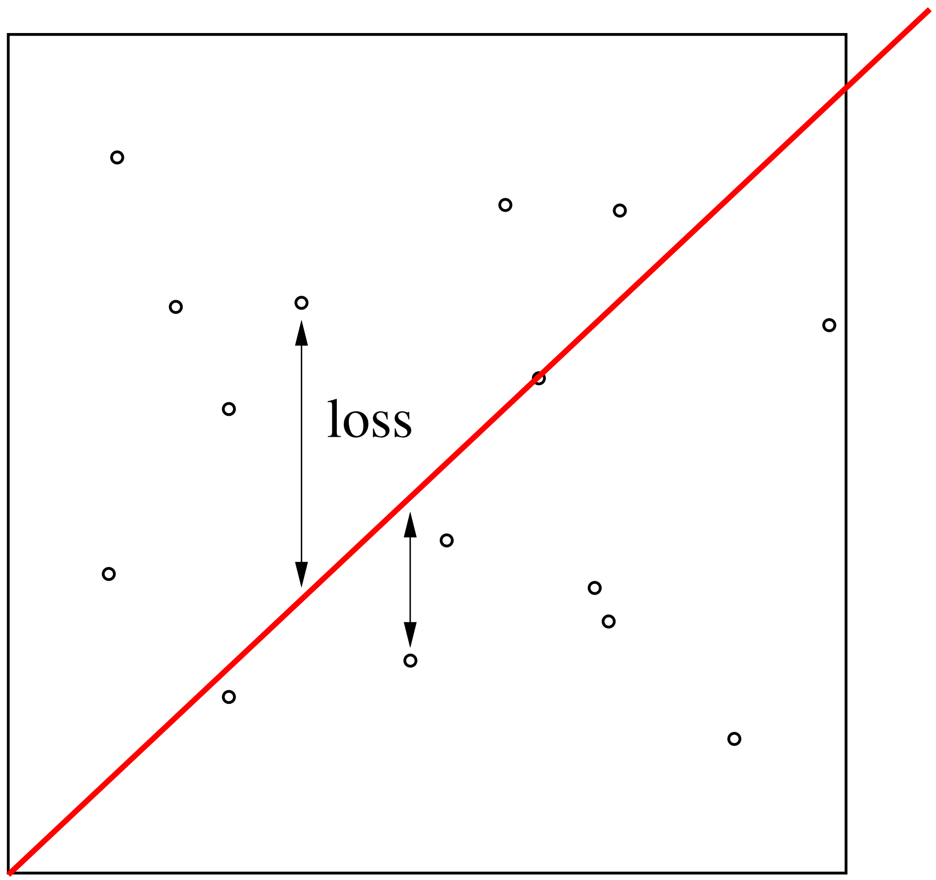


core algorithms – integrality gap

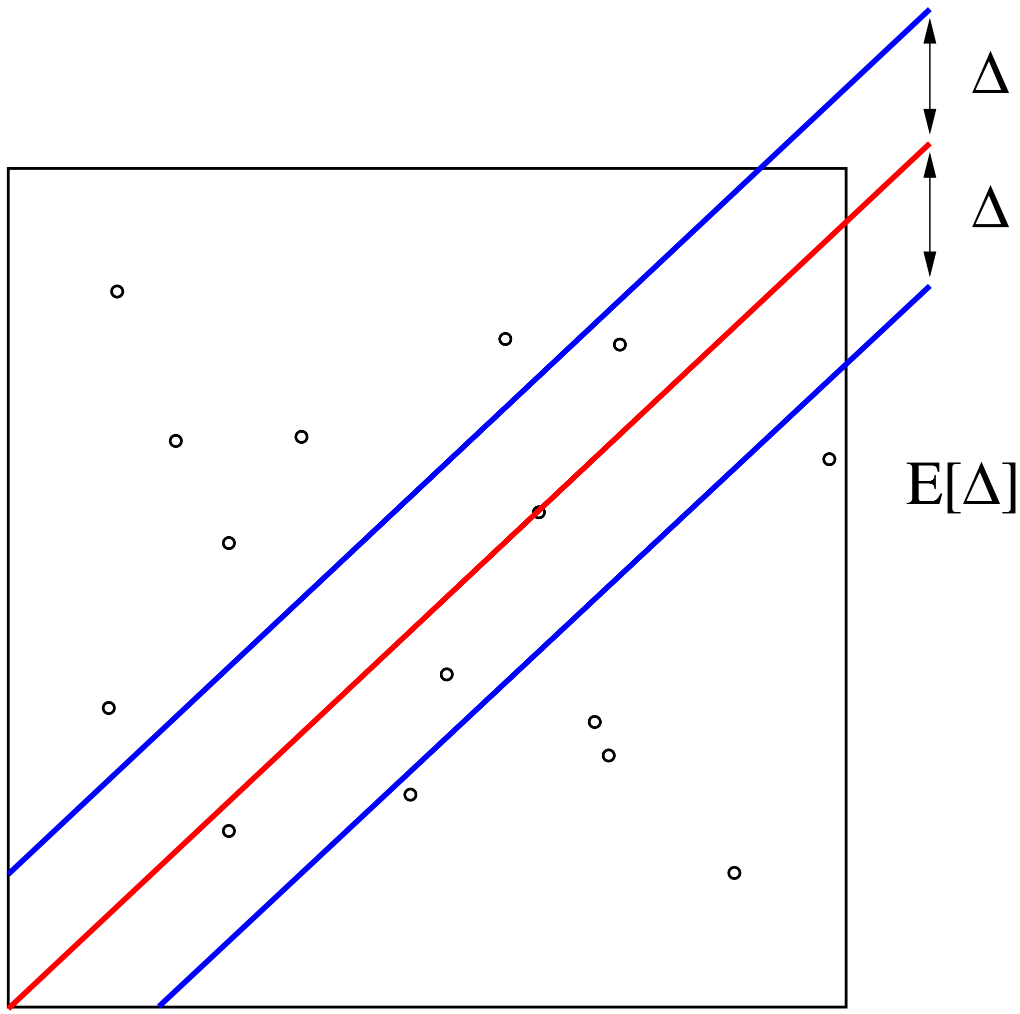


Δ = integrality gap

core algorithms – loss of an element

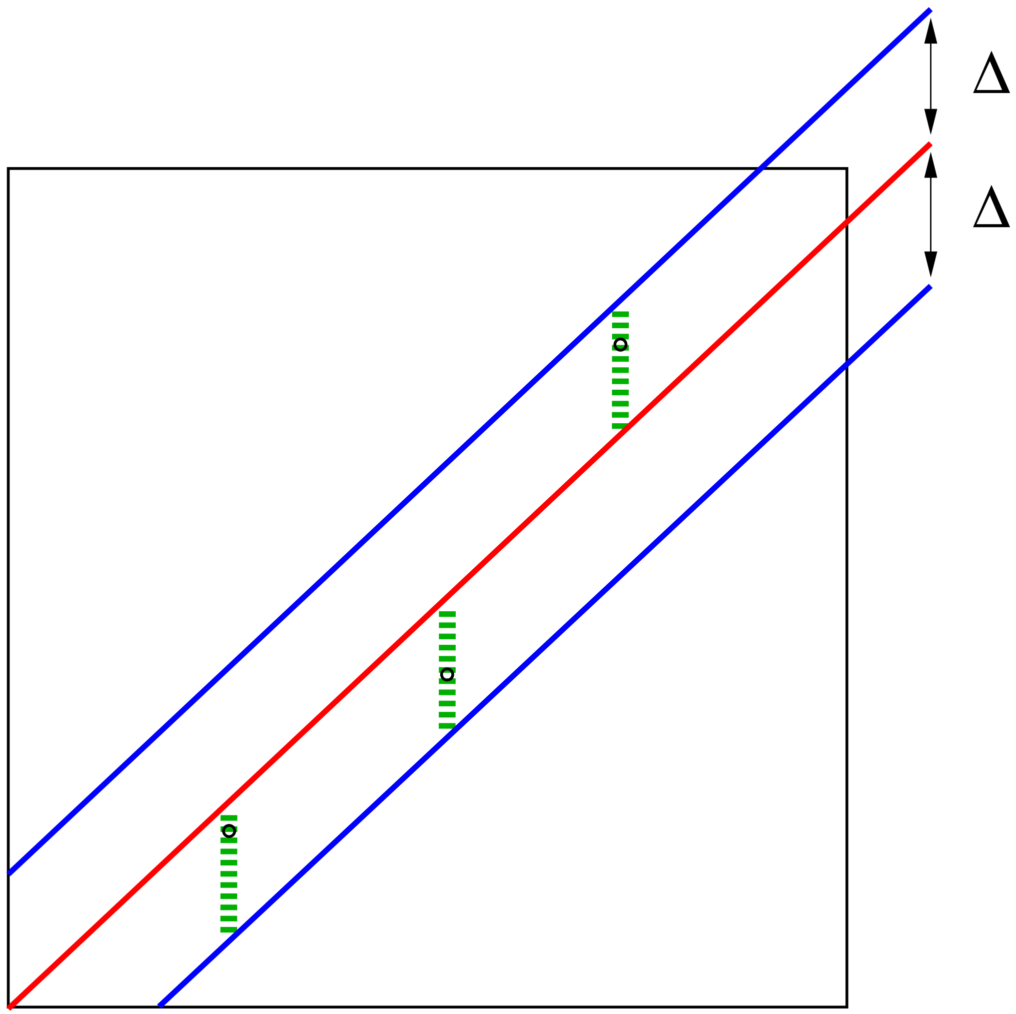


Lueker's result (1982)



$$E[\Delta] = \Theta\left(\frac{\log^2 n}{n}\right)$$

implications for core algorithms



implications for core algorithms

- Apply the Nemhauser/Ullmann algorithm to core elements
- Profits follow a uniform distribution with density $1/\Delta$
- $k \Rightarrow$ number of core elements
- expected running time of Nemhauser/Ullmann is $O(k^5/\Delta)$

$$\Delta = \Theta\left(\frac{\log^2 n}{n}\right); \qquad k \sim 2\Delta n = \Theta(\log^2 n)$$

implications for core algorithms

- Apply the Nemhauser/Ullmann algorithm to core elements
- Profits follow a uniform distribution with density $1/\Delta$
- $k \Rightarrow$ number of core elements
- expected running time of Nemhauser/Ullmann is $O(k^5/\Delta)$

$$\Delta = \Theta\left(\frac{\log^2 n}{n}\right); \qquad k \sim 2\Delta n = \Theta(\log^2 n)$$

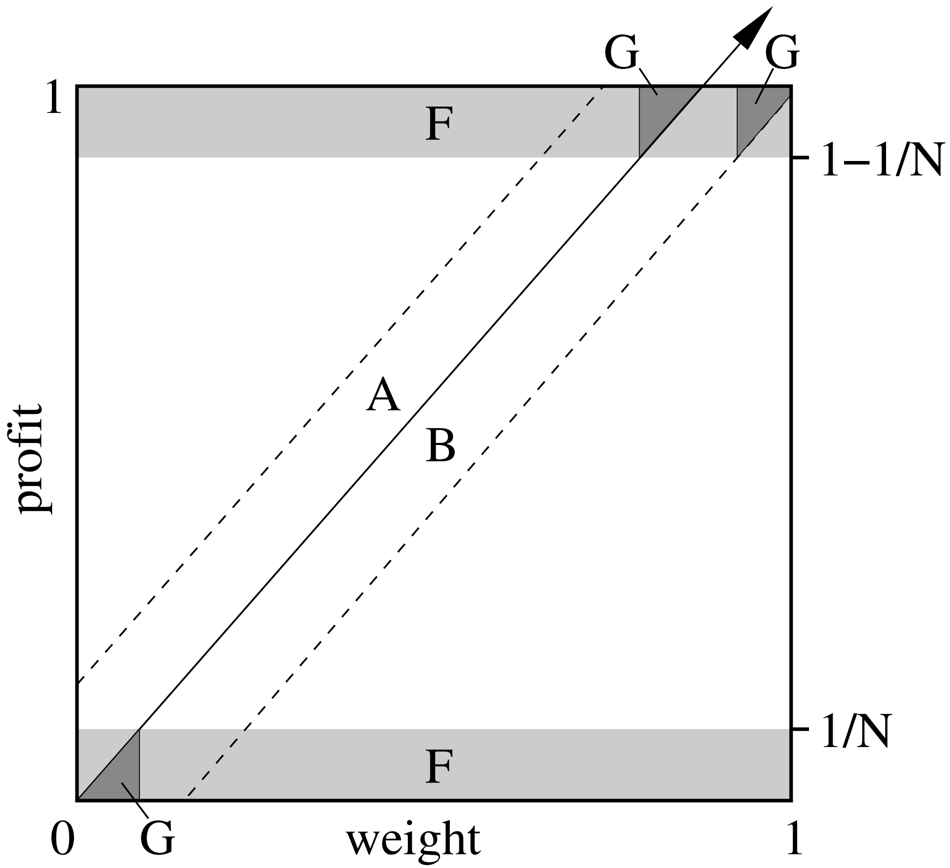
$$\Rightarrow O(n\,\text{polylog}(n))$$

some problematic details ...

- not all core items have sufficient randomness

some problematic details ...

- not all core items have sufficient randomness



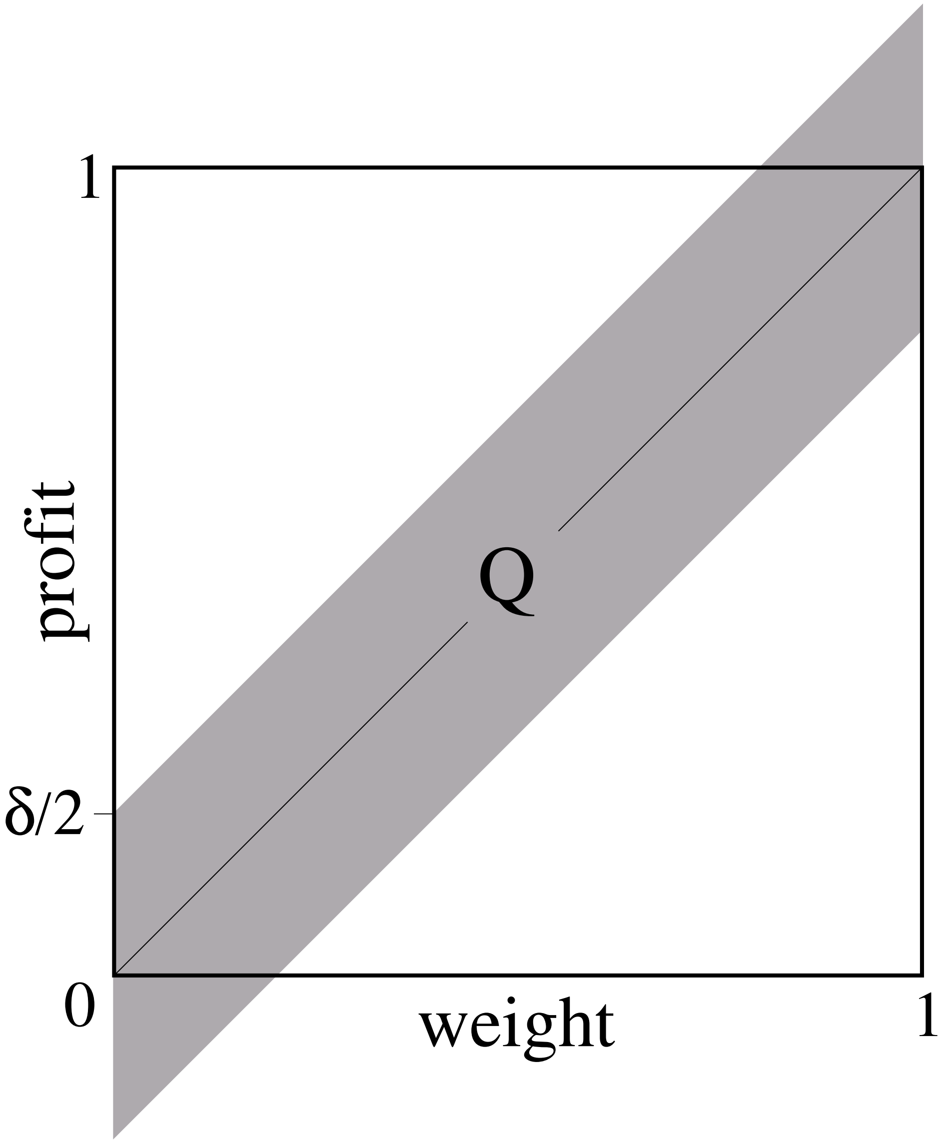
some problematic details ...

- not all core items have sufficient randomness
- number of core items

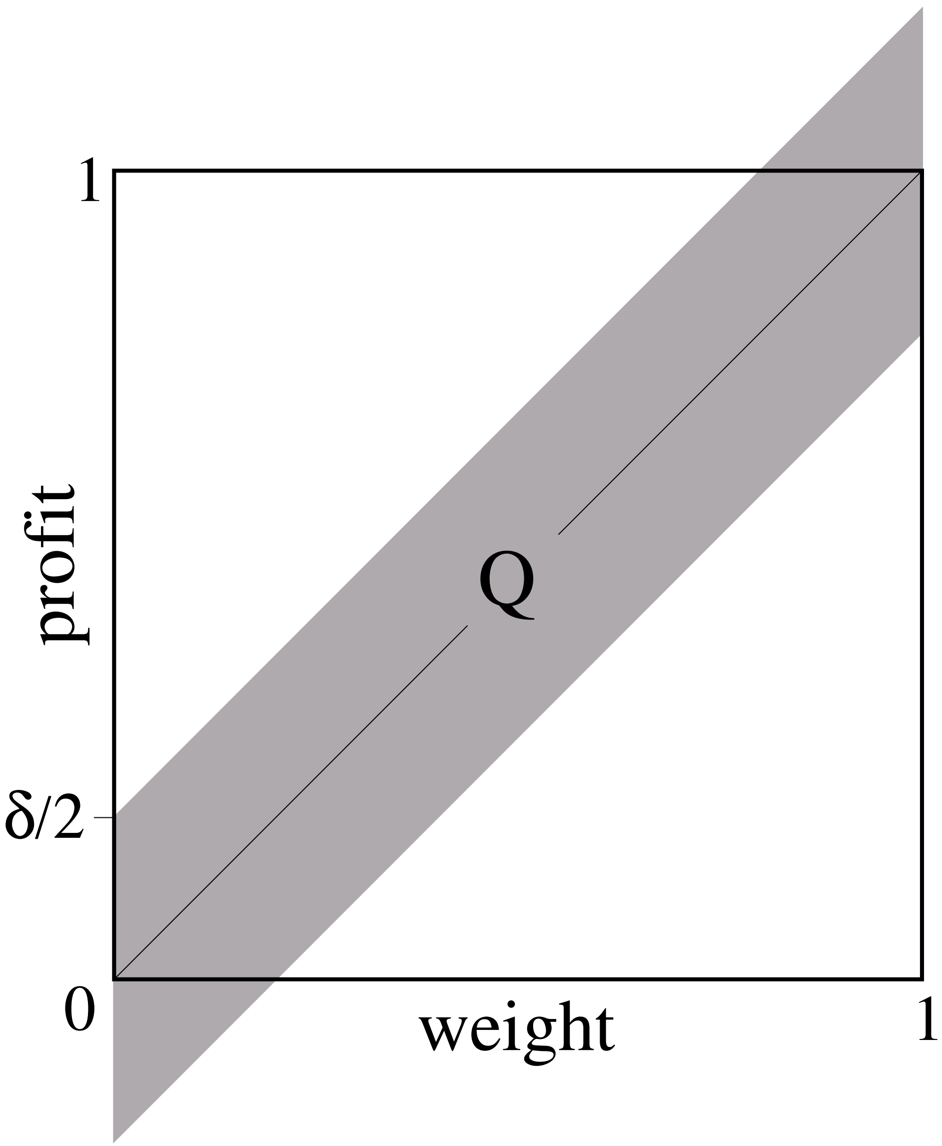
some problematic details ...

- not all core items have sufficient randomness
- number of core items
- integrality gap is not known
  - Choose large core region s.t. $\mathrm{P}\,[\textit{failure}] \leq n^{-5}$
  - in case of *failure*:
    - compute second list of dominating sets for remaining items
    - Combine two lists in linear time (Horowitz and Sahni ’74)

δ-correlated instances



δ-correlated instances



$\mathbf{E}[\Delta] = O(\frac{\delta}{n} \log^2 \frac{n}{\delta})$

running time: $O(\frac{n}{\delta}$ polylog $\frac{n}{\delta})$