Approximation Algorithms (ADM III) 4- Deterministic Rounding of Linear Programs

Guillaume Sagnol



Outline

1 Minimizing Sum of Completion Times on a Single Machine

- 2 Minimizing Weighted Sum of Completion Times
- 3 Prize-Collecting Steiner Tree Problem
- 4 Uncapacitated Facility Location Problem
- 5 Bin Packing Revisited

Given: jobs with processing time $p_j > 0$, release date $r_j \ge 0$, j = 1, ..., n. Task: Schedule the jobs nonpreemptively on a single machine;

minimize the total completion time $\sum_{i=1}^{n} C_i$.

Given: jobs with processing time $p_j > 0$, release date $r_j \ge 0, j = 1, ..., n$. Task: Schedule the jobs nonpreemptively on a single machine; minimize the *total completion time* $\sum_{j=1}^{n} C_j$. Remarks:

This problem is known to be strongly NP-hard.

Given: jobs with processing time $p_j > 0$, release date $r_j \ge 0, j = 1, ..., n$. Task: Schedule the jobs nonpreemptively on a single machine; minimize the *total completion time* $\sum_{j=1}^{n} C_j$. Remarks:

- This problem is known to be strongly NP-hard.
- The preemptive relaxation, however, can be solved efficiently.

Given: jobs with processing time $p_j > 0$, release date $r_j \ge 0, j = 1, ..., n$. Task: Schedule the jobs nonpreemptively on a single machine; minimize the *total completion time* $\sum_{j=1}^{n} C_j$. Remarks:

- This problem is known to be strongly NP-hard.
- The preemptive relaxation, however, can be solved efficiently.

Shortest Remaining Processing Time (SRPT) Rule

At any point in time, process an available and uncompleted job with shortest remaining processing time.

Given: jobs with processing time $p_j > 0$, release date $r_j \ge 0, j = 1, ..., n$. Task: Schedule the jobs nonpreemptively on a single machine; minimize the *total completion time* $\sum_{j=1}^{n} C_j$. Remarks:

- This problem is known to be strongly NP-hard.
- The preemptive relaxation, however, can be solved efficiently.

Shortest Remaining Processing Time (SRPT) Rule

 At any point in time, process an available and uncompleted job with shortest remaining processing time.

Theorem 4.1

The SRPT Rule finds an optimal preemptive schedule in time $O(n \log n)$.

Given: jobs with processing time $p_j > 0$, release date $r_j \ge 0, j = 1, ..., n$. Task: Schedule the jobs nonpreemptively on a single machine; minimize the *total completion time* $\sum_{j=1}^{n} C_j$. Remarks:

- This problem is known to be strongly NP-hard.
- The preemptive relaxation, however, can be solved efficiently.

Shortest Remaining Processing Time (SRPT) Rule

 At any point in time, process an available and uncompleted job with shortest remaining processing time.

Theorem 4.1

The SRPT Rule finds an optimal preemptive schedule in time $O(n \log n)$.

Proof: Use an exchange argument.

G. Sagnol

4- Deterministic LP Rounding 3 / 25

Idea: Use optimal preemptive solution to get good nonpreemptive solution.

Idea: Use optimal preemptive solution to get good nonpreemptive solution.

Algorithm

1 compute optimal preemptive schedule with job completion times C_i^P ;

Idea: Use optimal preemptive solution to get good nonpreemptive solution.

Algorithm

- **1** compute optimal preemptive schedule with job completion times C_i^P ;
- **2** sort jobs such that $C_1^P < C_2^P < \cdots < C_n^P$;

Idea: Use optimal preemptive solution to get good nonpreemptive solution.

Algorithm

- **1** compute optimal preemptive schedule with job completion times C_i^P ;
- 2 sort jobs such that $C_1^P < C_2^P < \cdots < C_n^P$;
- 3 schedule all jobs nonpreemptively and as early as possible in this order;

Idea: Use optimal preemptive solution to get good nonpreemptive solution.

Algorithm

- **1** compute optimal preemptive schedule with job completion times C_i^P ;
- 2 sort jobs such that $C_1^P < C_2^P < \cdots < C_n^P$;
- 3 schedule all jobs nonpreemptively and as early as possible in this order;

Step 3: set $C_1 := r_1 + p_1$; for j = 2 to n set $C_j := \max\{r_j, C_{j-1}\} + p_j$;

Idea: Use optimal preemptive solution to get good nonpreemptive solution.

Algorithm

- **1** compute optimal preemptive schedule with job completion times C_i^P ;
- **2** sort jobs such that $C_1^P < C_2^P < \cdots < C_n^P$;
- 3 schedule all jobs nonpreemptively and as early as possible in this order;

Step 3: set
$$C_1 := r_1 + p_1$$
; for $j = 2$ to n set $C_j := \max\{r_j, C_{j-1}\} + p_j$;

Lemma 4.2

For each job j = 1, ..., n, it holds that $C_j \leq 2 \cdot C_j^P$.

Idea: Use optimal preemptive solution to get good nonpreemptive solution.

Algorithm

- **1** compute optimal preemptive schedule with job completion times C_i^P ;
- **2** sort jobs such that $C_1^P < C_2^P < \cdots < C_n^P$;
- 3 schedule all jobs nonpreemptively and as early as possible in this order;

Step 3: set
$$C_1 := r_1 + p_1$$
; for $j = 2$ to n set $C_j := \max\{r_j, C_{j-1}\} + p_j$;

Lemma 4.2

For each job
$$j = 1, \ldots, n$$
, it holds that $C_j \leq 2 \cdot C_j^P$.

Theorem 4.3

The algorithm above is a 2-approximation algorithm.

Idea: Use optimal preemptive solution to get good nonpreemptive solution.

Algorithm

- **1** compute optimal preemptive schedule with job completion times C_i^P ;
- **2** sort jobs such that $C_1^P < C_2^P < \cdots < C_n^P$;
- 3 schedule all jobs nonpreemptively and as early as possible in this order;

Step 3: set
$$C_1 := r_1 + p_1$$
; for $j = 2$ to n set $C_j := \max\{r_j, C_{j-1}\} + p_j$;

Lemma 4.2

For each job
$$j = 1, \ldots, n$$
, it holds that $C_j \leq 2 \cdot C_j^P$.

Theorem 4.3

The algorithm above is a 2-approximation algorithm.

Proof:...

G. Sagnol

Outline

1 Minimizing Sum of Completion Times on a Single Machine

2 Minimizing Weighted Sum of Completion Times

- 3 Prize-Collecting Steiner Tree Problem
- 4 Uncapacitated Facility Location Problem
- 5 Bin Packing Revisited

Given: As before, but now all jobs *j* also have a weight $w_j \ge 0$.

Given: As before, but now all jobs *j* also have a weight $w_j \ge 0$. Task: Minimize the total *weighted* completion time $\sum_{j=1}^{n} w_j C_j$.

Given: As before, but now all jobs *j* also have a weight $w_j \ge 0$. Task: Minimize the total *weighted* completion time $\sum_{j=1}^{n} w_j C_j$.

Remarks:

■ Unfortunately, already the weighted preemptive problem is NP-hard.

Given: As before, but now all jobs *j* also have a weight $w_j \ge 0$. Task: Minimize the total *weighted* completion time $\sum_{i=1}^{n} w_j C_j$.

Remarks:

- Unfortunately, already the weighted preemptive problem is NP-hard.
- Thus, instead of preemptive relaxation use LP relaxation:

Given: As before, but now all jobs *j* also have a weight $w_j \ge 0$. Task: Minimize the total *weighted* completion time $\sum_{i=1}^{n} w_j C_j$.

Remarks:

- Unfortunately, already the weighted preemptive problem is NP-hard.
- Thus, instead of preemptive relaxation use LP relaxation:

$$\begin{array}{ll} \min & \sum_{j=1}^{n} w_j C_j \\ \text{s.t.} & C_j \ge r_j + p_j \\ & \sum_{j \in S} p_j C_j \ge \frac{1}{2} p(S)^2 \end{array} \quad \quad \text{for all } jobs \ j = 1, \dots, \\ & \text{for all } S \subseteq \{1, \dots, n\}. \end{array}$$

n.

Given: As before, but now all jobs *j* also have a weight $w_j \ge 0$. Task: Minimize the total *weighted* completion time $\sum_{i=1}^{n} w_j C_j$.

Remarks:

- Unfortunately, already the weighted preemptive problem is NP-hard.
- Thus, instead of preemptive relaxation use LP relaxation:

min
$$\sum_{j=1}^{n} w_j C_j$$

s.t. $C_j \ge r_j + p_j$ for all jobs $j = 1, ..., n$,
 $\sum_{j \in S} p_j C_j \ge \frac{1}{2} p(S)^2$ for all $S \subseteq \{1, ..., n\}$.

Lemma 4.4

The completion times C_j of a feasible schedule satisfy the LP constraints.

G. Sagnol

4- Deterministic LP Rounding 6 / 25

Lemma 4.5

Despite the exponential number of constraints, an optimal solution C^* to the LP relaxation can be computed in polynomial time.

Lemma 4.5

Despite the exponential number of constraints, an optimal solution C^* to the LP relaxation can be computed in polynomial time.

Proof:...

Lemma 4.5

Despite the exponential number of constraints, an optimal solution C^* to the LP relaxation can be computed in polynomial time.

Proof:...

Algorithm

1 compute optimal solution C^* to the LP relaxation;

Lemma 4.5

Despite the exponential number of constraints, an optimal solution C^* to the LP relaxation can be computed in polynomial time.

Proof:...

Algorithm

- **1** compute optimal solution C^* to the LP relaxation;
- 2 sort jobs such that $C_1^* \leq C_2^* \leq \cdots \leq C_n^*$;

Lemma 4.5

Despite the exponential number of constraints, an optimal solution C^* to the LP relaxation can be computed in polynomial time.

Proof:...

Algorithm

- **1** compute optimal solution C^* to the LP relaxation;
- 2 sort jobs such that $C_1^* \leq C_2^* \leq \cdots \leq C_n^*$;
- schedule all jobs nonpreemptively and as early as possible in this order;

Lemma 4.5

Despite the exponential number of constraints, an optimal solution C^* to the LP relaxation can be computed in polynomial time.

Proof:...

Algorithm

- **1** compute optimal solution C^* to the LP relaxation;
- 2 sort jobs such that $C_1^* \leq C_2^* \leq \cdots \leq C_n^*$;
- schedule all jobs nonpreemptively and as early as possible in this order;

Theorem 4.6

The algorithm above is a 3-approximation algorithm.

Lemma 4.5

Despite the exponential number of constraints, an optimal solution C^* to the LP relaxation can be computed in polynomial time.

Proof:...

Algorithm

- **1** compute optimal solution C^* to the LP relaxation;
- 2 sort jobs such that $C_1^* \leq C_2^* \leq \cdots \leq C_n^*$;
- schedule all jobs nonpreemptively and as early as possible in this order;

Theorem 4.6

The algorithm above is a 3-approximation algorithm.

Proof:...

G. Sagnol

Outline

- 1 Minimizing Sum of Completion Times on a Single Machine
- 2 Minimizing Weighted Sum of Completion Times
- 3 Prize-Collecting Steiner Tree Problem
- 4 Uncapacitated Facility Location Problem
- 5 Bin Packing Revisited

Given: Graph G = (V, E), root node $r \in V$, edge costs $c_e \ge 0$, $e \in E$, and penalties $\pi_i \ge 0$, $i \in V$.

Given: Graph G = (V, E), root node $r \in V$, edge costs $c_e \ge 0$, $e \in E$, and penalties $\pi_i \ge 0$, $i \in V$.

Task: Find subtree *T* containing root *r* minimizing $\sum_{e \in E(T)} c_e + \sum_{i \in V \setminus V(T)} \pi_i$.

Given: Graph G = (V, E), root node $r \in V$, edge costs $c_e \ge 0$, $e \in E$, and penalties $\pi_i \ge 0$, $i \in V$.

Task: Find subtree T containing root r minimizing $\sum c_e + \sum$

Example:



 π_i .

 $e \in E(T)$ $i \in V \setminus V(T)$

Given: Graph G = (V, E), root node $r \in V$, edge costs $c_e \ge 0$, $e \in E$, and penalties $\pi_i \ge 0$, $i \in V$.

Task: Find subtree *T* containing root *r* minimizing $\sum_{e \in E(T)} c_e + \sum_{i \in V \setminus V(T)} \pi_i.$

Example:



Remark: The Steiner Tree Problem is a special case with $\pi_i = 0$ for all non-terminals and $\pi_i = \infty$ for terminals *i*.

Given: Graph G = (V, E), root node $r \in V$, edge costs $c_e \ge 0$, $e \in E$, and penalties $\pi_i \ge 0$, $i \in V$.

Task: Find subtree *T* containing root *r* minimizing $\sum_{e \in E(T)} c_e + \sum_{i \in V \setminus V(T)} \pi_i$.

Remark: The Steiner Tree Problem is a special case with $\pi_i = 0$ for all non-terminals and $\pi_i = \infty$ for terminals *i*.

IP formulation:

$$\begin{array}{ll} \min & \sum_{e \in E} c_e \cdot x_e + \sum_{i \in V} \pi_i \cdot (1 - y_i) \\ \text{s.t.} & \sum_{e \in \delta(S)} x_e \geq \max_{i \in S} y_i \\ & y_r = 1, \\ & x_e, y_i \in \{0, 1\} \end{array} \qquad \qquad \text{for all } S \subseteq V \setminus \{r\}, \\ \end{array}$$

G. Sagnol

4- Deterministic LP Rounding 9 / 25
Prize-Collecting Steiner Tree Problem

Given: Graph G = (V, E), root node $r \in V$, edge costs $c_e \ge 0$, $e \in E$, and penalties $\pi_i \ge 0$, $i \in V$.

Task: Find subtree *T* containing root *r* minimizing $\sum_{e \in E(T)} c_e + \sum_{i \in V \setminus V(T)} \pi_i$.

Remark: The Steiner Tree Problem is a special case with $\pi_i = 0$ for all non-terminals and $\pi_i = \infty$ for terminals *i*.

IP formulation:

G. Sagnol

$$\begin{array}{ll} \min & \sum_{e \in E} c_e \cdot x_e + \sum_{i \in V} \pi_i \cdot (1 - y_i) \\ \text{s.t.} & \sum_{e \in \delta(S)} x_e \geq \max_{i \in S} y_i & \text{for all } S \subseteq V \setminus \{r\}, \\ & y_r = 1, \\ & x_e, y_i \in \{0, 1\} & \text{for all } e \in E, i \in V. \end{array}$$

LP relaxation: $x_e \ge 0$ for all $e \in E$ and $y_i \le 1$ for all $i \in V$.

4- Deterministic LP Rounding 9 / 25

Deterministic LP Rounding Algorithm Let $0 \le \alpha < 1$.

- Let $0 \le \alpha < 1$.
 - **1** compute optimal LP solution (x^*, y^*) in polytime with ellipsoid algo;

Let $0 \le \alpha < 1$.

1 compute optimal LP solution (x^*, y^*) in polytime with ellipsoid algo;

```
2 set U := \{i \in V \mid y_i^* \ge \alpha\};
```

Let $0 \le \alpha < 1$.

- **1** compute optimal LP solution (x^*, y^*) in polytime with ellipsoid algo;
- **2** set $U := \{i \in V \mid y_i^* \ge \alpha\};$
- **3** Find Steiner tree *T* on terminals *U* using some primal-dual algorithm.

Let $0 \le \alpha < 1$.

- **1** compute optimal LP solution (x^*, y^*) in polytime with ellipsoid algo;
- **2** set $U := \{i \in V \mid y_i^* \ge \alpha\};$
- Find Steiner tree T on terminals U using some primal-dual algorithm. We will prove the following lemma later, in an exercise:

Lemma 4.7

There is a primal-dual algorithm that returns a Steiner tree T on terminals U with cost at most $\frac{2}{\alpha} \sum_{e \in E} c_e \cdot x_e^*$.

Let $0 \le \alpha < 1$.

- **1** compute optimal LP solution (x^*, y^*) in polytime with ellipsoid algo;
- **2** set $U := \{i \in V \mid y_i^* \ge \alpha\};$
- Find Steiner tree T on terminals U using some primal-dual algorithm. We will prove the following lemma later, in an exercise:

Lemma 4.7

There is a primal-dual algorithm that returns a Steiner tree T on terminals U with cost at most $\frac{2}{\alpha} \sum_{e \in E} c_e \cdot x_e^*$.

Theorem 4.8

For $\alpha={\rm 2/3}$ the cost of the solution returned by the algorithm is

$$c(E(T)) + \pi(V \setminus V(T)) \leq \frac{2}{\alpha} \sum_{e \in E} c_e \cdot x_e^* + \frac{1}{1 - \alpha} \sum_{i \in V} \pi_i \cdot (1 - y_i^*) \leq 3 \cdot \text{OPT}$$

G. Sagnol

4- Deterministic LP Rounding 10 / 25

Outline

- 1 Minimizing Sum of Completion Times on a Single Machine
- 2 Minimizing Weighted Sum of Completion Times
- 3 Prize-Collecting Steiner Tree Problem
- 4 Uncapacitated Facility Location Problem
- 5 Bin Packing Revisited

Given: Set of facilities *F* with opening costs $f_i \ge 0$, $i \in F$; set of clients *D* with connection costs $c_{ij} \ge 0$, $i \in F$, $j \in D$.

Given: Set of facilities *F* with opening costs $f_i \ge 0$, $i \in F$; set of clients *D* with connection costs $c_{ij} \ge 0$, $i \in F$, $j \in D$.

Task: Choose $F' \subseteq F$ and assign each client to nearest facility in F'.

Given: Set of facilities F with opening costs $f_i \ge 0$, $i \in F$; set of clients D with connection costs $c_{ij} \ge 0$, $i \in F$, $j \in D$. Task: Choose $F' \subseteq F$ and assign each client to nearest facility in F'. Objective: Minimize $\sum_{i \in F'} f_i + \sum_{i \in D} \min_{i \in F'} c_{ij}$.

Given: Set of facilities F with opening costs $f_i \ge 0, i \in F$; set of clients D with connection costs $c_{ij} \ge 0, i \in F, j \in D$. Task: Choose $F' \subseteq F$ and assign each client to nearest facility in F'. Objective: Minimize $\sum_{i \in F'} f_i + \sum_{j \in D} \min_{i \in F'} c_{ij}$.

Remarks:

• This is a generalization of the Set Cover Problem.

Given: Set of facilities F with opening costs $f_i \ge 0, i \in F$; set of clients D with connection costs $c_{ij} \ge 0, i \in F, j \in D$. Task: Choose $F' \subseteq F$ and assign each client to nearest facility in F'. Objective: Minimize $\sum_{i \in F'} f_i + \sum_{j \in D} \min_{i \in F'} c_{ij}$.

Remarks:

- This is a generalization of the Set Cover Problem.
- In the following, we consider the special case with metric costs *c*_{ij}.

Given: Set of facilities F with opening costs $f_i \ge 0, i \in F$; set of clients D with connection costs $c_{ij} \ge 0, i \in F, j \in D$. Task: Choose $F' \subseteq F$ and assign each client to nearest facility in F'. Objective: Minimize $\sum_{i \in F'} f_i + \sum_{j \in D} \min_{i \in F'} c_{ij}$.

Remarks:

- This is a generalization of the Set Cover Problem.
- In the following, we consider the special case with metric costs *c*_{ij}.

IP formulation:

$$\begin{array}{ll} \min_{x_{ij},y_i \in \{0,1\}} & \sum_{i \in F} f_i \cdot y_i + \sum_{i \in F, j \in D} c_{ij} \cdot x_{ij} \\ \text{s.t.} & \sum_{i \in F} x_{ij} = 1 & \text{for all } j \in D, \\ & y_i - x_{ij} \ge 0 & \text{for all } i \in F, j \in D. \end{array}$$

G. Sagnol

4- Deterministic LP Rounding 12 / 25

min
$$\sum_{i \in F} f_i \cdot y_i + \sum_{i \in F, j \in D} c_{ij} \cdot x_{ij}$$

s.t.
$$\sum_{i \in F} x_{ij} = 1$$
$$y_i - x_{ij} \ge 0$$
$$x_{ij}, y_i \ge 0$$

for all $j \in D$,

for all $i \in F$, $j \in D$, for all $i \in F$, $j \in D$.

$$\begin{array}{ll} \min & \sum_{i \in F} f_i \cdot y_i + \sum_{i \in F, j \in D} c_{ij} \cdot x_{ij} \\ \text{s.t.} & \sum_{i \in F} x_{ij} = 1 \\ & y_i - x_{ij} \geq 0 \\ & x_{ij}, y_i \geq 0 \end{array}$$
$$\begin{array}{ll} \text{p:} & \max_{v_j, w_{ij} \geq 0} & \sum_{j \in D} v_j \\ & \text{s.t.} & \sum_{j \in D} w_{ij} \leq f_i \\ & v_j - w_{ij} \leq c_{ij} \end{array}$$

for all $j \in D$,

for all $i \in F$, $j \in D$, for all $i \in F$, $j \in D$.

dual LP

for all $i \in F$, for all $i \in F$, $j \in D$.

$$\begin{array}{ll} \min & \sum_{i \in F} f_i \cdot y_i + \sum_{i \in F, j \in D} c_{ij} \cdot x_{ij} \\ \text{s.t.} & \sum_{i \in F} x_{ij} = 1 & \text{for all } j \in D, \\ & y_i - x_{ij} \ge 0 & \text{for all } i \in F, j \in D, \\ & x_{ij}, y_i \ge 0 & \text{for all } i \in F, j \in D. \end{array}$$

$$\begin{array}{ll} \text{LP:} & \max_{v_j, w_{ij} \ge 0} & \sum_{j \in D} v_j \\ & \text{s.t.} & \sum_{j \in D} w_{ij} \le f_i & \text{for all } i \in F, \\ & v_j - w_{ij} \le c_{ij} & \text{for all } i \in F, j \in D. \end{array}$$

Interpretation of the dual LP:

dual

• v_j is the total amount that client *j* wants to pay for being served.

$$\begin{array}{ll} \min & \sum_{i \in F} f_i \cdot y_i + \sum_{i \in F, j \in D} c_{ij} \cdot x_{ij} \\ \text{s.t.} & \sum_{i \in F} x_{ij} = 1 & \text{for all } j \in D, \\ & y_i - x_{ij} \ge 0 & \text{for all } i \in F, j \in D, \\ & x_{ij}, y_i \ge 0 & \text{for all } i \in F, j \in D. \end{array}$$

$$\begin{array}{ll} \text{LP:} & \max_{v_j, w_{ij} \ge 0} & \sum_{j \in D} v_j \\ & \text{s.t.} & \sum_{j \in D} w_{ij} \le f_i & \text{for all } i \in F, \\ & v_j - w_{ij} \le c_{ij} & \text{for all } i \in F, j \in D. \end{array}$$

Interpretation of the dual LP:

dual

- *v_j* is the total amount that client *j* wants to pay for being served.
- client j might contribute w_{ij} to facility i for being connected to i.
 G. Sagnol
 4- Deterministic LP Rounding
 13 / 25

Let (x^*, y^*) and (v^*, w^*) be optimal solutions to the primal and dual LP, respectively.

Let (x^*, y^*) and (v^*, w^*) be optimal solutions to the primal and dual LP, respectively.

Notation:

Facility *i* neighbors client *j* if $x_{ij}^* > 0$; $N(j) := \{i \in F \mid x_{ij}^* > 0\}$.

• $N^2(j) := \{\ell \in D \mid \text{client } \ell \text{ neighbors some facility } i \in N(j)\}.$

Let (x^*, y^*) and (v^*, w^*) be optimal solutions to the primal and dual LP, respectively.

Notation:

Facility *i* neighbors client *j* if $x_{ij}^* > 0$; $N(j) := \{i \in F \mid x_{ij}^* > 0\}$.

• $N^2(j) := \{\ell \in D \mid \text{client } \ell \text{ neighbors some facility } i \in N(j)\}.$

Lemma 4.9

If clients j_1, \ldots, j_k have disjoint neighborhoods $N(j_1), \ldots, N(j_k)$, then opening cheapest facility in each neighborhood costs $\leq \sum_{i \in F} f_i \cdot y_i^* \leq \text{OPT}$.

Let (x^*, y^*) and (v^*, w^*) be optimal solutions to the primal and dual LP, respectively.

Notation:

Facility *i* neighbors client *j* if $x_{ij}^* > 0$; $N(j) := \{i \in F \mid x_{ij}^* > 0\}$.

• $N^2(j) := \{\ell \in D \mid \text{client } \ell \text{ neighbors some facility } i \in N(j)\}.$

Lemma 4.9

If clients j_1, \ldots, j_k have disjoint neighborhoods $N(j_1), \ldots, N(j_k)$, then opening cheapest facility in each neighborhood costs $\leq \sum_{i \in F} f_i \cdot y_i^* \leq \text{OPT}$.

Lemma 4.10

For each client j, $v_i^* \ge c_{ij}$ for all $i \in N(j)$.

Let (x^*, y^*) and (v^*, w^*) be optimal solutions to the primal and dual LP, respectively.

Notation:

Facility *i* neighbors client *j* if $x_{ij}^* > 0$; $N(j) := \{i \in F \mid x_{ij}^* > 0\}$.

• $N^2(j) := \{\ell \in D \mid \text{client } \ell \text{ neighbors some facility } i \in N(j)\}.$

Lemma 4.9

If clients j_1, \ldots, j_k have disjoint neighborhoods $N(j_1), \ldots, N(j_k)$, then opening cheapest facility in each neighborhood costs $\leq \sum_{i \in F} f_i \cdot y_i^* \leq \text{OPT}$.

Lemma 4.10

For each client j, $v_i^* \ge c_{ij}$ for all $i \in N(j)$.

Proofs: ...

G. Sagnol

- **1** compute optimal LP solutions (x^*, y^*) and (v^*, w^*) ;
- 2 while $D \neq \emptyset$
- 3 choose $j := \operatorname{argmin}_{j' \in D} v_{j'}^*$ and $i := \operatorname{argmin}_{i' \in N(j)} f_{i'};$
- 4 assign all unassigned clients in $N^2(j)$ to facility *i*;
- $5 \qquad \text{set } D := D \setminus N^2(j);$

- **1** compute optimal LP solutions (x^*, y^*) and (v^*, w^*) ;
- 2 while $D \neq \emptyset$
- 3 choose $j := \operatorname{argmin}_{j' \in D} v_{j'}^*$ and $i := \operatorname{argmin}_{i' \in N(j)} f_{i'};$
- 4 assign all unassigned clients in $N^2(j)$ to facility *i*;
- $5 \qquad \text{set } D := D \setminus N^2(j);$

Theorem 4.11

The algorithm above is a 4-approximation algorithm.

- **1** compute optimal LP solutions (x^*, y^*) and (v^*, w^*) ;
- 2 while $D \neq \emptyset$
- 3 choose $j := \operatorname{argmin}_{j' \in D} v_{j'}^*$ and $i := \operatorname{argmin}_{i' \in N(j)} f_{i'};$
- 4 assign all unassigned clients in $N^2(j)$ to facility *i*;
- $5 \qquad \text{set } D := D \setminus N^2(j);$

Theorem 4.11

The algorithm above is a 4-approximation algorithm.

Proof:...

- **1** compute optimal LP solutions (x^*, y^*) and (v^*, w^*) ;
- 2 while $D \neq \emptyset$
- 3 choose $j := \operatorname{argmin}_{j' \in D} v_{j'}^*$ and $i := \operatorname{argmin}_{i' \in N(j)} f_{i'};$
- **4** assign all unassigned clients in $N^2(j)$ to facility *i*;
- $5 \qquad \text{set } D := D \setminus N^2(j);$

Theorem 4.11

The algorithm above is a 4-approximation algorithm.

Proof:...

We finally mention the following non-approximability result without proof.

Theorem 4.12

There is no α -approximation algorithm for the metric uncapacitated facility location problem with $\alpha < 1.463$ unless each problem in *NP* has an $O(n^{O(\log \log n)})$ time algorithm.

Outline

- **1** Minimizing Sum of Completion Times on a Single Machine
- 2 Minimizing Weighted Sum of Completion Times
- 3 Prize-Collecting Steiner Tree Problem
- 4 Uncapacitated Facility Location Problem
- 5 Bin Packing Revisited

In the previous chapter we showed how to find a solution to instance *I* with at most $(1 + \varepsilon)$ OPT (I) + 1 bins in polynomial time.

In the previous chapter we showed how to find a solution to instance *I* with at most $(1 + \varepsilon)$ OPT (I) + 1 bins in polynomial time.

Goal: Use at most $OPT(I) + O(\log^2 OPT(I))$ bins! (Karmarkar & Karp, 1982)

In the previous chapter we showed how to find a solution to instance *I* with at most $(1 + \varepsilon)$ OPT (I) + 1 bins in polynomial time.

```
Goal: Use at most OPT (I) + O(\log^2 OPT(I)) bins! (Karmarkar & Karp, 1982)
```

Ingredients:

- Replace dynamic program with integer program plus LP rounding.
- Improved grouping scheme.
- Recursive application of two previous ingredients.

In the previous chapter we showed how to find a solution to instance *I* with at most $(1 + \varepsilon)$ OPT (I) + 1 bins in polynomial time.

```
Goal: Use at most OPT(I) + O(\log^2 OPT(I)) bins!
(Karmarkar & Karp, 1982)
```

Ingredients:

- Replace dynamic program with integer program plus LP rounding.
- Improved grouping scheme.
- Recursive application of two previous ingredients.

Notice:

By Lemma 3.14 we can assume that all items have size at least 1/SIZE(I).

G. Sagnol

4- Deterministic LP Rounding 17 / 25

- let $s_1 > s_2 > \cdots > s_m$ denote the different item sizes;
- for i = 1, ..., m, let b_i denote the number of items of size s_i ;

- let $s_1 > s_2 > \cdots > s_m$ denote the different item sizes;
- for i = 1, ..., m, let b_i denote the number of items of size s_i ;
- an *m*-tuple $(t_1, \ldots, t_m) \in \mathbb{Z}_{\geq 0}^m$ is a configuration if $\sum_{i=1}^m t_i \cdot s_i \leq 1$;

- let $s_1 > s_2 > \cdots > s_m$ denote the different item sizes;
- for i = 1, ..., m, let b_i denote the number of items of size s_i ;
- an *m*-tuple $(t_1, \ldots, t_m) \in \mathbb{Z}_{\geq 0}^m$ is a configuration if $\sum_{i=1}^m t_i \cdot s_i \leq 1$;
- let T₁,..., T_N be a complete enumeration of all configurations and denote by t_{ij} the multiplicity of item *i* in configuration T_j;

- let $s_1 > s_2 > \cdots > s_m$ denote the different item sizes;
- for i = 1, ..., m, let b_i denote the number of items of size s_i ;
- an *m*-tuple $(t_1, \ldots, t_m) \in \mathbb{Z}_{\geq 0}^m$ is a configuration if $\sum_{i=1}^m t_i \cdot s_i \leq 1$;
- let T₁,..., T_N be a complete enumeration of all configurations and denote by t_{ij} the multiplicity of item i in configuration T_j;
- for j = 1,..., N, the integer variable x_j denotes the number of bins that shall be packed according to configuration T_j:
Configuration Integer Program for Bin Packing

- let $s_1 > s_2 > \cdots > s_m$ denote the different item sizes;
- for i = 1, ..., m, let b_i denote the number of items of size s_i ;
- an *m*-tuple $(t_1, \ldots, t_m) \in \mathbb{Z}_{\geq 0}^m$ is a configuration if $\sum_{i=1}^m t_i \cdot s_i \leq 1$;
- let T₁,..., T_N be a complete enumeration of all configurations and denote by t_{ij} the multiplicity of item i in configuration T_j;
- for j = 1,..., N, the integer variable x_j denotes the number of bins that shall be packed according to configuration T_j:

$$\begin{array}{ll} \min & \sum_{j=1}^N x_j \\ \text{s.t.} & \sum_{j=1}^N t_{ij} \cdot x_j \geq b_i \\ & x_j \in \mathbb{Z}_{\geq 0} \end{array}$$

for all $i = 1, \ldots, m$,

for all
$$j = 1, ..., N$$
.

G. Sagnol

4- Deterministic LP Rounding 18 / 25

Configuration LP and its Dual

...

Primal:

min
$$\sum_{j=1}^{N} x_j$$

s.t. $\sum_{j=1}^{N} t_{ij} \cdot x_j \ge b_i$ for all $i = 1, \dots, m$,
 $x_j \ge 0$ for all $j = 1, \dots, N$.

Configuration LP and its Dual

min $\sum_{j=1}^{n} x_j$ Primal: s.t. $\sum_{j=1}^{N} t_{ij} \cdot x_j \ge b_i$ for all $i = 1, \ldots, m$, $x_i > 0$ for all $i = 1, \ldots, N$. $\max \sum_{i=1}^{m} b_i \cdot y_i$ Dual: s.t. $\sum_{ij}^{m} t_{ij} \cdot y_i \leq 1$ for all j = 1, ..., N, $v_i > 0$ for all $i = 1, \ldots, m$.

Configuration LP and its Dual

Primal:min
$$\sum_{j=1}^{N} x_j$$
s.t. $\sum_{j=1}^{N} t_{ij} \cdot x_j \ge b_i$ for all $i = 1, \dots, m$,
 $x_j \ge 0$ Dual:max $\sum_{i=1}^{m} b_i \cdot y_i$ s.t. $\sum_{i=1}^{m} t_{ij} \cdot y_i \le 1$ for all $j = 1, \dots, N$,
 $y_i \ge 0$

Notice: SIZE(I) \leq OPT $_{LP}(I) \leq$ OPT (I)

G. Sagnol

4- Deterministic LP Rounding 19 / 25

■ Configuration LP suffers from exponentially many variables.

- Configuration LP suffers from exponentially many variables.
- Dual separation problem is Knapsack Problem and thus NP-hard.

- Configuration LP suffers from exponentially many variables.
- Dual separation problem is Knapsack Problem and thus NP-hard.
- Remember: optimization and separation are equally difficult.

- Configuration LP suffers from exponentially many variables.
- Dual separation problem is Knapsack Problem and thus NP-hard.
- Remember: optimization and separation are equally difficult.
- Therefore, it is NP-hard to solve the Configuration LP to optimality.

- Configuration LP suffers from exponentially many variables.
- Dual separation problem is Knapsack Problem and thus NP-hard.
- Remember: optimization and separation are equally difficult.
- Therefore, it is NP-hard to solve the Configuration LP to optimality.

Theorem 4.13

An LP solution of value at most $OPT_{LP}(I) + 1$ can be computed in polynomial time.

- Configuration LP suffers from exponentially many variables.
- Dual separation problem is Knapsack Problem and thus NP-hard.
- Remember: optimization and separation are equally difficult.
- Therefore, it is NP-hard to solve the Configuration LP to optimality.

Theorem 4.13

An LP solution of value at most $OPT_{LP}(I) + 1$ can be computed in polynomial time.

Proof:...

G. Sagnol

Main idea: Use FPTAS for Knapsack Problem as approximate separation routine within ellipsoid method.

Main idea: Use FPTAS for Knapsack Problem as approximate separation routine within ellipsoid method. This yields optimal solution y^* to

perturbed dual:
$$\max \sum_{i=1}^{m} b_i \cdot y_i$$
s.t. $\sum_{i=1}^{m} t_{ij} \cdot y_i \le \delta_j$ for all $j = 1, \dots, N$, $y_i \ge 0$ for all $i = 1, \dots, m$,

with $\delta_j \in \{1, 1 + \varepsilon\}$ and $|\{j \mid \delta_j = 1\}|$ polynomially bounded.

Main idea: Use FPTAS for Knapsack Problem as approximate separation routine within ellipsoid method. This yields optimal solution y^* to

perturbed dual:
$$\max \sum_{i=1}^{m} b_i \cdot y_i$$
s.t. $\sum_{i=1}^{m} t_{ij} \cdot y_i \le \delta_j$ for all $j = 1, \dots, N$, $y_i \ge 0$ for all $i = 1, \dots, m$,

with $\delta_j \in \{1, 1 + \varepsilon\}$ and $|\{j \mid \delta_j = 1\}|$ polynomially bounded.

Since $y^*/(1 + \varepsilon)$ is feasible dual solution, $\sum_{i=1}^m b_i \cdot y_i^* \leq (1 + \varepsilon) \mathsf{OPT}_{LP}$.

Main idea: Use FPTAS for Knapsack Problem as approximate separation routine within ellipsoid method. This yields optimal solution y^* to

perturbed dual: max
$$\sum_{i=1}^{m} b_i \cdot y_i$$

s.t. $\sum_{i=1}^{m} t_{ij} \cdot y_i \le \delta_j$ for all $j = 1, \dots, N$,
 $y_i \ge 0$ for all $i = 1, \dots, m$,

with $\delta_j \in \{1, 1 + \varepsilon\}$ and $|\{j \mid \delta_j = 1\}|$ polynomially bounded. Since $y^*/(1 + \varepsilon)$ is feasible dual solution, $\sum_{i=1}^m b_i \cdot y_i^* \leq (1 + \varepsilon) \mathsf{OPT}_{LP}$. Moreover, for $J := \{j \mid \delta_j = 1\}$, vector y^* is optimal solution to

reduced dual: max
$$\sum_{i=1}^{m} b_i \cdot y_i$$

s.t. $\sum_{i=1}^{m} t_{ij} \cdot y_i \le 1$ for all j with $\delta_j = 1$,
 $y_i \ge 0$ for all $i = 1, ..., m$.

Consider the corresponding

reduced primal:

$$\begin{array}{ll} \min & \sum_{j \in J} x_j \\ \text{s.t.} & \sum_{j \in J} t_{ij} \cdot x_j \geq b_i & \text{ for all } i = 1, \dots, m, \\ & x_j \geq 0 & \text{ for all } j \in J. \end{array}$$

Consider the corresponding

reduced primal:min
$$\sum_{j \in J} x_j$$
s.t. $\sum_{j \in J} t_{ij} \cdot x_j \ge b_i$ for all $i = 1, \dots, m$, $x_j \ge 0$ for all $j \in J$.

It has polynomial size and optimal solution value at most $(1 + \varepsilon)OPT_{LP}$.

Consider the corresponding

reduced primal:min
$$\sum_{j \in J} x_j$$
s.t. $\sum_{j \in J} t_{ij} \cdot x_j \ge b_i$ for all $i = 1, \dots, m$, $x_j \ge 0$ for all $j \in J$.

It has polynomial size and optimal solution value at most $(1 + \varepsilon)OPT_{LP}$.

Choose $\varepsilon := 1/n$ such that $(1 + \varepsilon)OPT_{LP} \le OPT_{LP} + \varepsilon n \le OPT_{LP} + 1$.

G. Sagnol

Consider the corresponding

reduced primal: min
$$\sum_{j \in J} x_j$$

s.t. $\sum_{j \in J} t_{ij} \cdot x_j \ge b_i$ for all $i = 1, \dots, m$,
 $x_j \ge 0$ for all $j \in J$.

It has polynomial size and optimal solution value at most $(1 + \varepsilon)OPT_{LP}$.

Choose $\varepsilon := 1/n$ such that $(1 + \varepsilon)OPT_{LP} \le OPT_{LP} + \varepsilon n \le OPT_{LP} + 1$.

Reduced primal and its optimal solution \bar{y} can be computed in polynomial time (FPTAS for Knapsack!).

G. Sagnol

Consider the corresponding

reduced primal: min
$$\sum_{j \in J} x_j$$

s.t. $\sum_{j \in J} t_{ij} \cdot x_j \ge b_i$ for all $i = 1, \dots, m$,
 $x_j \ge 0$ for all $j \in J$.

It has polynomial size and optimal solution value at most $(1 + \varepsilon)OPT_{LP}$.

Choose $\varepsilon := 1/n$ such that $(1 + \varepsilon)OPT_{LP} \le OPT_{LP} + \varepsilon n \le OPT_{LP} + 1$.

Reduced primal and its optimal solution \bar{y} can be computed in polynomial time (FPTAS for Knapsack!).

 \bar{y} is feasible solution to original primal LP of value at most OPT_{LP} + 1.

G. Sagnol

4- Deterministic LP Rounding 22 / 25

Grouping

- consider items in order of non-increasing size;
- open a group and start putting items in current group, one at a time;
- close current group if its total size is at least 2 and start new group;

Grouping

- consider items in order of non-increasing size;
- open a group and start putting items in current group, one at a time;
- close current group if its total size is at least 2 and start new group;

Let r := number of groups; let G_i denote *i*th group; $n_i := |G_i|$.

Grouping

- consider items in order of non-increasing size;
- open a group and start putting items in current group, one at a time;
- close current group if its total size is at least 2 and start new group;

Let r := number of groups; let G_i denote *i*th group; $n_i := |G_i|$.

Notice that $r \leq \lceil SIZE(I)/2 \rceil$ and $n_i \geq n_{i-1}$, for i = 2, ..., r - 1.

Grouping

- consider items in order of non-increasing size;
- open a group and start putting items in current group, one at a time;
- close current group if its total size is at least 2 and start new group;

Let r := number of groups; let G_i denote *i*th group; $n_i := |G_i|$.

Notice that $r \leq \lceil SIZE(I)/2 \rceil$ and $n_i \geq n_{i-1}$, for i = 2, ..., r - 1.

Rounding: Construct new instance I' as follows:

- discard items in G_1 and G_r ;
- for i = 2, ..., r 1 discard the $n_i n_{i-1}$ smallest items in G_i ;
- for i = 2, ..., r 1 round sizes of remaining items in G_i to largest one.

Grouping

- consider items in order of non-increasing size;
- open a group and start putting items in current group, one at a time;
- close current group if its total size is at least 2 and start new group;

Let r := number of groups; let G_i denote *i*th group; $n_i := |G_i|$.

Notice that $r \leq \lceil SIZE(I)/2 \rceil$ and $n_i \geq n_{i-1}$, for i = 2, ..., r - 1.

Rounding: Construct new instance I' as follows:

- discard items in G_1 and G_r ;
- for i = 2, ..., r 1 discard the $n_i n_{i-1}$ smallest items in G_i ;
- for i = 2, ..., r 1 round sizes of remaining items in G_i to largest one.

Lemma 4.14

There are at most SIZE(I)/2 distinct item sizes in I'; the total size of all discarded items is $O(\log SIZE(I))$.

G. Sagnol

4- Deterministic LP Rounding 23 / 25

BinPack(1)

if SIZE(*I*) < 10 then pack remaining items using First-Fit and stop;

BinPack(1)

- if SIZE(1) < 10 then pack remaining items using First-Fit and stop;
- **2** apply harmonic grouping scheme to create instance *I*';
- **3** pack discarded items in $O(\log SIZE(I))$ bins using First-Fit;

BinPack(1)

- if SIZE(1) < 10 then pack remaining items using First-Fit and stop;
- **2** apply harmonic grouping scheme to create instance *I*';
- **3** pack discarded items in $O(\log SIZE(I))$ bins using First-Fit;
- compute near-optimal solution x to Configuration LP for instance l';
- **5** for j = 1, ..., N pack $\lfloor x_j \rfloor$ bins in configuration T_j ;
- **6** call the packed items instance I_1 and the remaining items I_2 ;

BinPack(1)

- if SIZE(1) < 10 then pack remaining items using First-Fit and stop;
- **2** apply harmonic grouping scheme to create instance I';
- **3** pack discarded items in $O(\log SIZE(I))$ bins using First-Fit;
- compute near-optimal solution x to Configuration LP for instance l';
- **5** for j = 1, ..., N pack $\lfloor x_j \rfloor$ bins in configuration T_j ;
- **6** call the packed items instance I_1 and the remaining items I_2 ;
- pack *I*₂ recursively via BinPack(*I*₂);

G. Sagnol

Lemma 4.15

$\operatorname{OPT}_{LP}(I_1) + \operatorname{OPT}_{LP}(I_2) \leq \operatorname{OPT}_{LP}(I') \leq \operatorname{OPT}_{LP}(I).$

Lemma 4.15

$\operatorname{OPT}_{LP}(I_1) + \operatorname{OPT}_{LP}(I_2) \leq \operatorname{OPT}_{LP}(I') \leq \operatorname{OPT}_{LP}(I).$

Proof:...

Lemma 4.15

$$\operatorname{OPT}_{LP}(I_1) + \operatorname{OPT}_{LP}(I_2) \leq \operatorname{OPT}_{LP}(I') \leq \operatorname{OPT}_{LP}(I).$$

Proof:...

Theorem 4.16 (Karmarkar & Karp, 1982)

Algorithm BinPack runs in polynomial time and finds a solution using at most $OPT(I) + O(\log^2 OPT(I))$ bins.

Lemma 4.15

$$\operatorname{OPT}_{LP}(I_1) + \operatorname{OPT}_{LP}(I_2) \leq \operatorname{OPT}_{LP}(I') \leq \operatorname{OPT}_{LP}(I).$$

Proof:...

Theorem 4.16 (Karmarkar & Karp, 1982)

Algorithm BinPack runs in polynomial time and finds a solution using at most $OPT(I) + O(\log^2 OPT(I))$ bins.

Proof:...

Lemma 4.15

$$\operatorname{OPT}_{LP}(I_1) + \operatorname{OPT}_{LP}(I_2) \leq \operatorname{OPT}_{LP}(I') \leq \operatorname{OPT}_{LP}(I).$$

Proof:...

Theorem 4.16 (Karmarkar & Karp, 1982)

Algorithm BinPack runs in polynomial time and finds a solution using at most $OPT(I) + O(log^2 OPT(I))$ bins.

Proof:...

Theorem 4.17 (Hoberg & Rothvoß, 2015)

A solution using at most $OPT(I) + O(\log OPT(I))$ bins can be found in polynomial time.

G. Sagnol

4- Deterministic LP Rounding 25 / 25