Aalto University
School of Science

Department of Computer Science

Combinatorics of Efficient Computations

Approximation Algorithms

Lecture 5: LP-Based Approaches for Set Cover

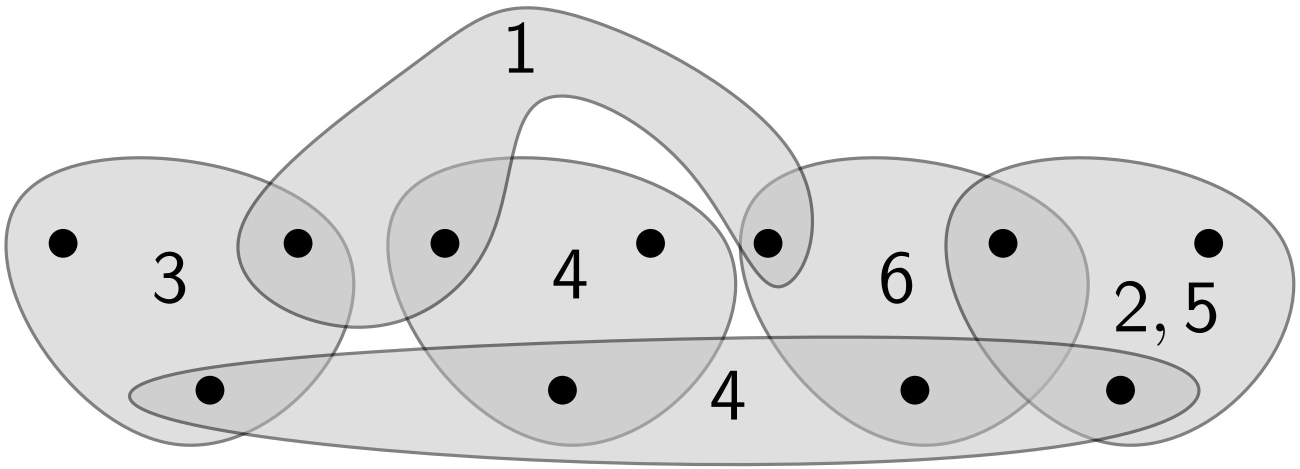
Joachim Spoerhase

2019

SetCover – as an ILP

$$\begin{aligned} \min \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\ & x_S \in \{0, 1\} \qquad S \in \mathcal{S} \end{aligned}$$

- Ground set U
- Set Families $\mathcal{S} \subseteq 2^U$ with $\bigcup \mathcal{S} = U$
- Costs $c \colon \mathcal{S} \to \mathbb{Q}^+$



Find a cover $\mathcal{S}' \subseteq \mathcal{S}$ of U with minimum cost.

SetCover – LP-Relaxation

$$
\begin{array}{lll}
\text{min} & \displaystyle\sum_{S \in \mathcal{S}} c_S x_S & \\
\text{s.t.} & \displaystyle\sum_{S \ni e} x_S \geq 1 & e \in U \\
 & \boxed{x_S \geq 0} & S \in \mathcal{S}
\end{array}
$$

LP-Rounding

LP-Rounding: Approach I

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\ & x_S \geq 0 \qquad\qquad S \in \mathcal{S} \end{aligned}$$

LP-Rounding-I$(U, \mathcal{S}, c)$

Find an optimal solution $\mathbf{x}$ for the LP-Relaxation

Round each x_S with $x_S > 0$ to 1

LP-Rounding: Approach I

$$\begin{array}{lll} \text{min} & \displaystyle\sum_{S \in \mathcal{S}} c_S x_S & \\ \text{s.t.} & \displaystyle\sum_{S \ni e} x_S \geq 1 & e \in U \\ & x_S \geq 0 & S \in \mathcal{S} \end{array}$$

LP-Rounding-I$(U, \mathcal{S}, c)$

Find an optimal solution $\mathbf{x}$ for the LP-Relaxation

Round each x_S with $x_S > 0$ to 1

- generates a feasible solution :-)

LP-Rounding: Approach I

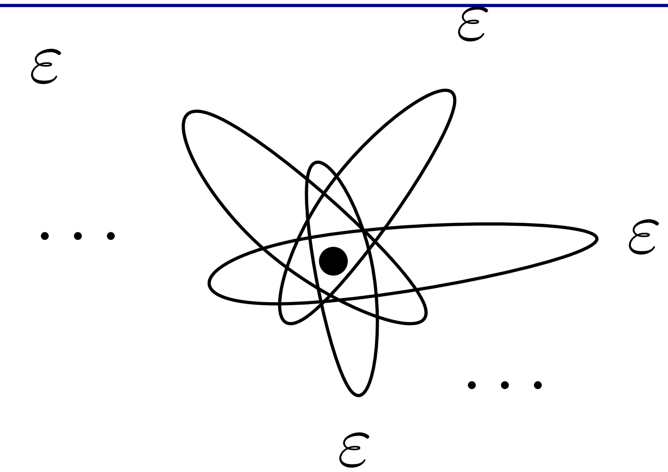
$$\begin{aligned} \min \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 && e \in U \\ & x_S \geq 0 && S \in \mathcal{S} \end{aligned}$$

LP-Rounding-I$(U, \mathcal{S}, c)$

Find an optimal solution $\mathbf{x}$ for the LP-Relaxation

Round each x_S with $x_S > 0$ to 1

- generates a feasible solution :-)
- scaling factor is arbitrarily large :-(



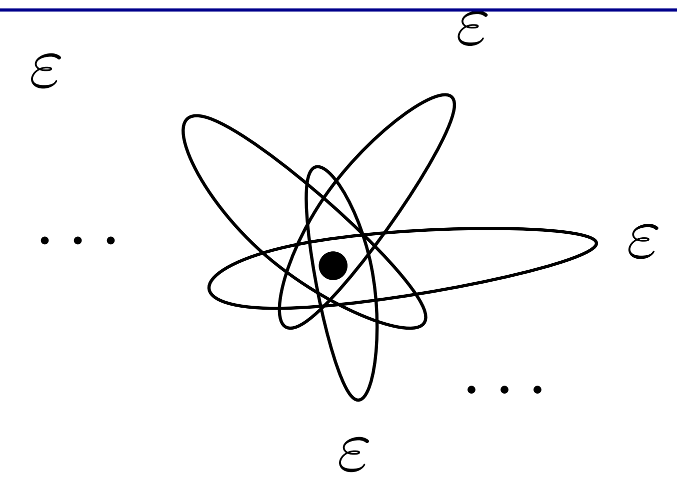
LP-Rounding: Approach I

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\ & x_S \geq 0 \qquad\qquad S \in \mathcal{S} \end{aligned}$$

LP-Rounding-I$(U, \mathcal{S}, c)$
  Find an optimal solution $\mathbf{x}$ for the LP-Relaxation
  Round each x_S with $x_S > 0$ to 1

- generates a feasible solution :-)
- scaling factor is arbitrarily large :-(

better rounding policy?



LP-Rounding: Approach II

$$\begin{aligned}
\text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\
\text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\
& x_S \geq 0 \qquad\qquad S \in \mathcal{S}
\end{aligned}$$

Let f be the frequency of (number of sets containing) the most frequent element.

LP-Rounding-II$(U, \mathcal{S}, c)$
Find an optimal solution $\mathbf{x}$ for the LP-Relaxation
Round each x_S with $x_S \geq 1/f$ to 1; others to 0

LP-Rounding: Approach II

$$\begin{array}{lll} \text{min} & \sum_{S \in \mathcal{S}} c_S x_S & \\ \text{s.t.} & \sum_{S \ni e} x_S \geq 1 & e \in U \\ & x_S \geq 0 & S \in \mathcal{S} \end{array}$$

Let f be the frequency of (number of sets containing) the most frequent element.

LP-Rounding-II$(U, \mathcal{S}, c)$
  Find an optimal solution $\mathbf{x}$ for the LP-Relaxation
  Round each x_S with $x_S \geq 1/f$ to 1; others to 0

Thm. LP-Rounding-II is a factor-f approximation algorithm for SETCOVER.

LP-Rounding: Approach II

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \quad & e \in U \\ & x_S \geq 0 \quad & S \in \mathcal{S} \end{aligned}$$

Let f be the frequency of (number of sets containing) the most frequent element.

LP-Rounding-II$(U, \mathcal{S}, c)$
Find an optimal solution $\mathbf{x}$ for the LP-Relaxation
Round each x_S with $x_S \geq 1/f$ to 1; others to 0

Thm. LP-Rounding-II is a factor-f approximation algorithm for SetCover.

Proof. (discuss).

LP-Rounding: Approach II

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \quad & e \in U \\ & x_S \geq 0 \quad & S \in \mathcal{S} \end{aligned}$$

Let f be the frequency of (number of sets containing) the most frequent element.

LP-Rounding-II$(U, \mathcal{S}, c)$
Find an optimal solution $\mathbf{x}$ for the LP-Relaxation
Round each x_S with $x_S \geq 1/f$ to 1; others to 0

Thm. LP-Rounding-II is a factor-f approximation algorithm for SETCOVER.

Proof. (discuss). □

even better rounding?

Randomized LP-Rounding

$$
\begin{aligned}
\text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\
\text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\
& x_S \geq 0 \qquad\quad S \in \mathcal{S}
\end{aligned}
$$

Randomized LP-Rounding

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\ & x_S \geq 0 \qquad\qquad S \in \mathcal{S} \end{aligned}$$

$x_S \in [0, 1]$ in opt. solution

Randomized LP-Rounding

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\ & x_S \geq 0 \qquad\quad S \in \mathcal{S} \end{aligned}$$

interpret x_S as a probability

$x_S \in [0, 1]$ in opt. solution

Randomized LP-Rounding

$$\min \quad \sum_{S \in \mathcal{S}} c_S x_S$$

$$\text{s.t.} \quad \sum_{S \ni e} x_S \geq 1 \qquad e \in U$$

$$x_S \geq 0 \qquad S \in \mathcal{S}$$

interpret x_S as a probability

$x_S \in [0, 1]$ in opt. solution

Random-LP-Rounding$(U, \mathcal{S}, c)$

Find an optimal solution $\mathbf{x}$ for the LP-Relaxation

Round each x_S to 1 with probability x_S (o.w. set x_S to 0)

Randomized LP-Rounding

$$\begin{aligned} \min \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\ & x_S \geq 0 \qquad\qquad S \in \mathcal{S} \end{aligned}$$

interpret x_S as a probability

$x_S \in [0, 1]$ in opt. solution

Random-LP-Rounding$(U, \mathcal{S}, c)$
Find an optimal solution $\mathbf{x}$ for the LP-Relaxation
Round each x_S to 1 with probability x_S (o.w. set x_S to 0)

Lemma.

Let OPT_f be the cost of an optimal solution for the above LP-Relaxation and $\mathcal{S}' \subseteq \mathcal{S}$ be a corresponding solution generated by Random-LP-Rounding. Then $E[c(\mathcal{S}')] = \text{OPT}_\text{f}$.

Randomized LP-Rounding

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\ & x_S \geq 0 \qquad\quad S \in \mathcal{S} \end{aligned}$$

interpret x_S as a probability

$x_S \in [0, 1]$ in opt. solution

Random-LP-Rounding$(U, \mathcal{S}, c)$

Find an optimal solution $\mathbf{x}$ for the LP-Relaxation

Round each x_S to 1 with probability x_S (o.w. set x_S to 0)

$\mathcal{S}'$ is not necessarily feasible!

Lemma.

Let OPT_f be the cost of an optimal solution for the above LP-Relaxation and $\mathcal{S}' \subseteq \mathcal{S}$ be a corresponding solution generated by Random-LP-Rounding. Then $E[c(\mathcal{S}')] = \text{OPT}_\text{f}$.

Randomized LP-Rounding

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 & e \in U \\ & x_S \geq 0 & S \in \mathcal{S} \end{aligned}$$

Random-LP-Rounding($U, \mathcal{S}, c$)
Find an optimal solution $\mathbf{x}$ for the LP-Relaxation
Round each x_S to 1 with probability x_S (o.w. set x_S to 0)

Lemma.

Let $\mathcal{S}' \subseteq \mathcal{S}$ be a result of Random-LP-Rounding. For each $u \in U$, u is covered by $\mathcal{S}'$ with probability $1 - 1/e \approx 0.632\ldots$.

Randomized LP-Rounding

Algorithm $\textsc{SetCover}$-Random-Rounding$(U, \mathcal{S}, c)$

$\mathcal{S}' \leftarrow \emptyset$

for t iterations **do**

$\mathcal{S}' \leftarrow \mathcal{S}' \cup$ Random-LP-Rounding$(U, \mathcal{S}, c)$

return $\mathcal{S}'$

Randomized LP-Rounding

Algorithm $\textsc{SetCover}$-Random-Rounding$(U, \mathcal{S}, c)$

$\mathcal{S}' \leftarrow \emptyset$

for t iterations **do**

$\quad \mathcal{S}' \leftarrow \mathcal{S}' \cup$ Random-LP-Rounding$(U, \mathcal{S}, c)$

return $\mathcal{S}'$

Lemma.

For $t := 2 + \ln n$, the above algorithm produces an infeasible solution with probability $\leq \frac{1}{4}$.

Randomized LP-Rounding

Algorithm $\textsc{SetCover}$-Random-Rounding$(U, \mathcal{S}, c)$
$\mathcal{S}' \leftarrow \emptyset$
for t iterations **do**
  $\mathcal{S}' \leftarrow \mathcal{S}' \cup$ Random-LP-Rounding$(U, \mathcal{S}, c)$
return $\mathcal{S}'$

Lemma.

For $t := 2 + \ln n$, the above algorithm produces an infeasible solution with probability $\leq \frac{1}{4}$.

Lemma.

With probability $\leq \frac{1}{4}$, the above algorithm produces an "expensive" solution (i.e., with cost $\geq \text{OPT}_f \cdot 4t$).

Randomized LP-Rounding

Corollary.

With probability $\geq \frac{1}{2}$, our algorithm provides a feasible solution whose cost is less than $4t \cdot \mathsf{OPT} = O(\log n) \cdot \mathsf{OPT}$.

Primal-Dual Schema

Recall Complementary Slackness

$$\begin{array}{ll} \min & \sum_{j=1}^{n} c_j x_j \\ \text{s.t} & \sum_{j=1}^{n} a_{ij} x_j \geq b_i \\ & x_j \geq 0 \end{array} \qquad \begin{array}{ll} \max. & \sum_{i=1}^{m} b_i y_i \\ \text{s.t.} & \sum_{i=1}^{m} a_{ij} y_i \leq c_j \\ & y_i \geq 0 \end{array}$$

Thm.

Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_m)$ be feasible solutions for the primal and dual LPs respectively. The solutions $\mathbf{x}$ and $\mathbf{y}$ are optimal if and only if the following conditions are met:

Primal CS:

For each $j = 1, \ldots, n$: either $x_j = 0$ or $\sum_{i=1}^{m} a_{ij} y_i = c_j$

Dual CS:

For each $i = 1, \ldots, m$: either $y_i = 0$ or $\sum_{j=1}^{n} a_{ij} x_j = b_i$

Relaxing Complementary Slackness

min	$\sum_{j=1}^{n} c_j x_j$	max.	$\sum_{i=1}^{m} b_i y_i$
s.t	$\sum_{j=1}^{n} a_{ij} x_j \geq b_i$	s.t.	$\sum_{i=1}^{m} a_{ij} y_i \leq c_j$
	$x_j \geq 0$		$y_i \geq 0$

Primal CS:

For each $j = 1, \ldots, n$: either $x_j = 0$ or $\sum_{i=1}^{m} a_{ij} y_i = c_j$

Dual CS:

For each $i = 1, \ldots, m$: either $y_i = 0$ or $\sum_{j=1}^{n} a_{ij} x_j = b_i$

$$\Leftrightarrow \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i$$

Relaxing Complementary Slackness

$$\begin{array}{lll} \text{min} & \sum_{j=1}^{n} c_j x_j \\ \text{s.t} & \sum_{j=1}^{n} a_{ij} x_j \geq b_i \\ & x_j \geq 0 \end{array} \qquad \begin{array}{lll} \text{max.} & \sum_{i=1}^{m} b_i y_i \\ \text{s.t.} & \sum_{i=1}^{m} a_{ij} y_i \leq c_j \\ & y_i \geq 0 \end{array}$$

Primal CS: relaxed primal CS

For each $j = 1, \ldots, n$: either $x_j = 0$ or $\sum_{i=1}^{m} a_{ij} y_i = c_j$

$$c_j/\alpha \leq \sum_{i=1}^{m} a_{ij} y_i \leq c_j$$

Dual CS:

For each $i = 1, \ldots, m$: either $y_i = 0$ or $\sum_{j=1}^{n} a_{ij} x_j = b_i$

$$\Leftrightarrow \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i$$

Relaxing Complementary Slackness

min	$\sum_{j=1}^{n} c_j x_j$	max.	$\sum_{i=1}^{m} b_i y_i$
s.t	$\sum_{j=1}^{n} a_{ij} x_j \geq b_i$	s.t.	$\sum_{i=1}^{m} a_{ij} y_i \leq c_j$
	$x_j \geq 0$		$y_i \geq 0$

~~**Primal CS**~~: relaxed primal CS

For each $j = 1, \ldots, n$: either $x_j = 0$ or ~~$\sum_{i=1}^{m} a_{ij} y_i = c_j$~~

$$c_j / \alpha \leq \sum_{i=1}^{m} a_{ij} y_i \leq c_j$$

~~**Dual CS**~~: relaxed dual CS

For each $i = 1, \ldots, m$: either $y_i = 0$ or ~~$\sum_{j=1}^{n} a_{ij} x_j = b_i$~~

$$b_i \leq \sum_{j=1}^{n} a_{ij} x_j \leq \beta \cdot b_i$$

$$\Leftrightarrow \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i$$

Relaxing Complementary Slackness

min	$\sum_{j=1}^{n} c_j x_j$	max.	$\sum_{i=1}^{m} b_i y_i$
s.t	$\sum_{j=1}^{n} a_{ij} x_j \geq b_i$	s.t.	$\sum_{i=1}^{m} a_{ij} y_i \leq c_j$
	$x_j \geq 0$		$y_i \geq 0$

~~**Primal CS**:~~ relaxed primal CS

For each $j = 1, \ldots, n$: either $x_j = 0$ or ~~$\sum_{i=1}^{m} a_{ij} y_i = c_j$~~

$$c_j/\alpha \leq \sum_{i=1}^{m} a_{ij} y_i \leq c_j$$

~~**Dual CS**:~~ relaxed dual CS

For each $i = 1, \ldots, m$: either $y_i = 0$ or ~~$\sum_{j=1}^{n} a_{ij} x_j = b_i$~~

$$b_i \leq \sum_{j=1}^{n} a_{ij} x_j \leq \beta \cdot b_i$$

~~$\Leftrightarrow \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i$~~ $\quad \Rightarrow \sum_{j=1}^{n} c_j x_j \leq \alpha\beta \sum_{i=1}^{m} b_i y_i \leq \alpha\beta \cdot \mathrm{OPT}_{\mathrm{LP}}$

Primal-Dual Schema

- Start with a feasible dual and infeasible primal solution (often trivial).

Primal-Dual Schema

- Start with a feasible dual and infeasible primal solution (often trivial).
- "Improve" the feasibility of the primal solution.

Primal-Dual Schema

- Start with a feasible dual and infeasible primal solution (often trivial).
- “Improve” the feasibility of the primal solution.
- simultaneously improve the obj. value of the dual solution.

Primal-Dual Schema

- Start with a feasible dual and infeasible primal solution (often trivial).
- "Improve" the feasibility of the primal solution.
- simultaneously improve the obj. value of the dual solution.
- do so until the relaxed CS conditions are met.

Primal-Dual Schema

- Start with a feasible dual and infeasible primal solution (often trivial).
- "Improve" the feasibility of the primal solution.
- simultaneously improve the obj. value of the dual solution.
- do so until the relaxed CS conditions are met.
- maintain that the primal solution is integer valued.

Primal-Dual Schema

- Start with a feasible dual and infeasible primal solution (often trivial).
- “Improve” the feasibility of the primal solution.
- simultaneously improve the obj. value of the dual solution.
- do so until the relaxed CS conditions are met.
- maintain that the primal solution is integer valued.
- The feasibility of the primal solution and relaxed CS condition provide an approximation ratio.

Relaxed CS for SetCover

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S & \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 & e \in U \\ & x_S \geq 0 & S \in \mathcal{S} \end{aligned}$$

LP

$$\begin{aligned} \text{max} \quad & \sum_{e \in U} y_e & \\ \text{s.t.} \quad & \sum_{e \in S} y_e \leq c_S & S \in \mathcal{S} \\ & y_e \geq 0 & e \in U \end{aligned}$$

DLP

Relaxed CS for SetCover

$$\begin{array}{llll} \text{min} & \sum_{S \in \mathcal{S}} c_S x_S & & \\ \text{s.t.} & \sum_{S \ni e} x_S \geq 1 & e \in U \\ & x_S \geq 0 & S \in \mathcal{S} \end{array} \qquad \begin{array}{llll} \text{max} & \sum_{e \in U} y_e & & \\ \text{s.t.} & \sum_{e \in S} y_e \leq c_S & S \in \mathcal{S} \\ & y_e \geq 0 & e \in U \end{array}$$

LP DLP

critical set

(non-relaxed) primal CS: $x_S \neq 0 \Rightarrow \sum_{e \in S} y_e = c_S$

Relaxed CS for SetCover

$$\begin{array}{llll} \text{min} & \sum_{S \in \mathcal{S}} c_S x_S & & \\ \text{s.t.} & \sum_{S \ni e} x_S \geq 1 & e \in U \\ & x_S \geq 0 & S \in \mathcal{S} \end{array}$$

LP

$$\begin{array}{llll} \text{max} & \sum_{e \in U} y_e & & \\ \text{s.t.} & \sum_{e \in S} y_e \leq c_S & S \in \mathcal{S} \\ & y_e \geq 0 & e \in U \end{array}$$

DLP

critical set

(non-relaxed) primal CS: $x_S \neq 0 \Rightarrow \sum_{e \in S} y_e = c_S$

select only critical sets

Relaxed CS for SetCover

$$\begin{aligned} \text{min} \quad & \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \quad & \sum_{S \ni e} x_S \geq 1 \qquad e \in U \\ & x_S \geq 0 \qquad S \in \mathcal{S} \end{aligned}$$

LP

$$\begin{aligned} \text{max} \quad & \sum_{e \in U} y_e \\ \text{s.t.} \quad & \sum_{e \in S} y_e \leq c_S \qquad S \in \mathcal{S} \\ & y_e \geq 0 \qquad e \in U \end{aligned}$$

DLP

critical set

(non-relaxed) primal CS: $x_S \neq 0 \Rightarrow \sum_{e \in S} y_e = c_S$

select only critical sets

Relaxed dual CS: $y_e \neq 0 \Rightarrow \sum_{S \ni e} x_S \leq f$

Relaxed CS for SetCover

$$\begin{array}{lll} \text{min} & \sum_{S \in \mathcal{S}} c_S x_S & \\ \text{s.t.} & \sum_{S \ni e} x_S \geq 1 & e \in U \\ & x_S \geq 0 & S \in \mathcal{S} \end{array}$$

LP

$$\begin{array}{lll} \text{max} & \sum_{e \in U} y_e & \\ \text{s.t.} & \sum_{e \in S} y_e \leq c_S & S \in \mathcal{S} \\ & y_e \geq 0 & e \in U \end{array}$$

DLP

critical set

(non-relaxed) primal CS: $x_S \neq 0 \Rightarrow \sum_{e \in S} y_e = c_S$

select only critical sets

trivial for binary $\mathbf{x}$

Relaxed dual CS: $y_e \neq 0 \Rightarrow \sum_{S \ni e} x_S \leq f$

Primal-Dual-Schema for SetCover

PrimalDualSetCover$(U, \mathcal{S}, c)$

$\mathbf{x} \leftarrow 0, \mathbf{y} \leftarrow 0$

repeat

- select an uncovered element e
- increase y_e until a set S is critical $(\sum_{e' \in S} y_{e'} = c_S)$
- select all critical sets and update $\mathbf{x}$
- mark all elements in these sets as **covered**

until all elements are covered

return $\mathbf{x}$

Primal-Dual-Schema for SetCover

PrimalDualSetCover($U, \mathcal{S}, c$)

$\mathbf{x} \leftarrow 0, \mathbf{y} \leftarrow 0$

repeat

select an uncovered element e

increase y_e until a set S is critical ($\sum_{e' \in S} y_{e'} = c_S$)

select all critical sets and update $\mathbf{x}$

mark all elements in these sets as **covered**

until all elements are covered

return $\mathbf{x}$

Thm. The above is a factor-f approximation algorithm for SetCover. The factor is also tight.