

Approximation Algorithms

Lecture 12:

STEINERFOREST via Primal–Dual

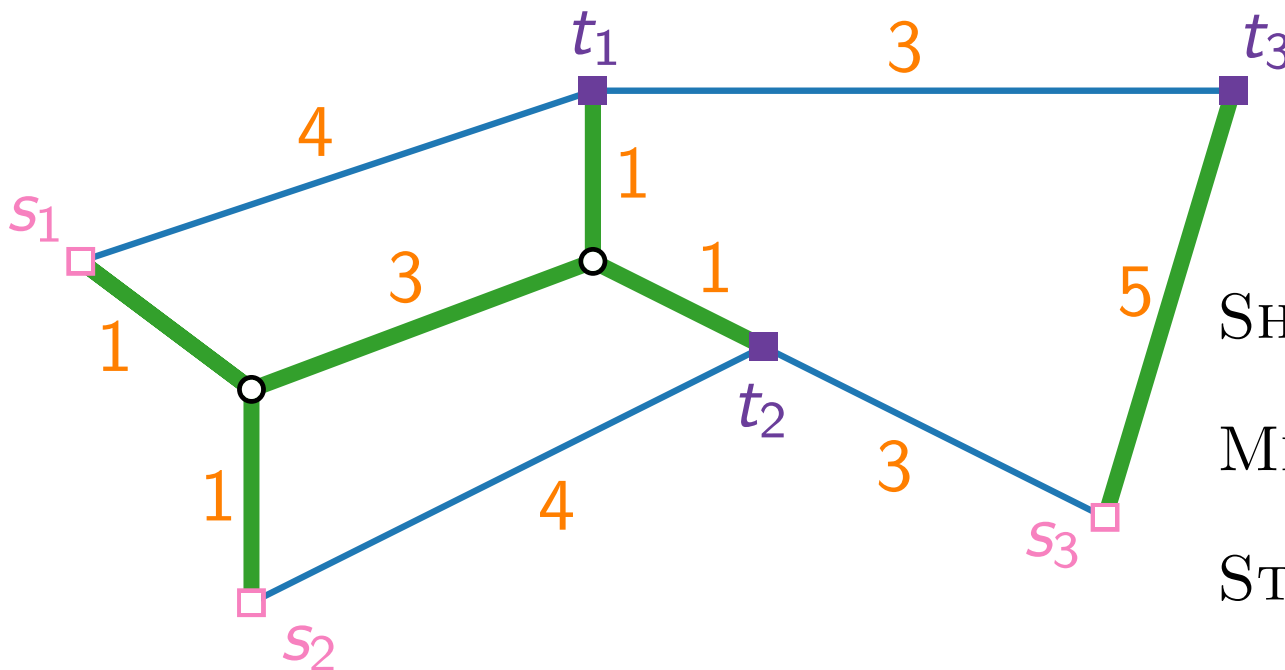
Part I:

STEINERFOREST

STEINERFOREST

Given: A graph G with edge costs $c: E(G) \rightarrow \mathbb{N}$ and a set $R = \{(s_1, t_1), \dots, (s_k, t_k)\}$ of k vertex pairs.

Task: Find an edge set $F \subseteq E(G)$ of minimum total cost $c(F)$ such that the subgraph $(V(G), F)$ connects all vertex pairs $(s_1, t_1), \dots, (s_k, t_k)$.



Special cases?

SHORTESTPATH ($R = \{s, t\}$)

MINSPANNINGTREE ($R = E(G)$)

STEINERTREE ($R = T \times T$)

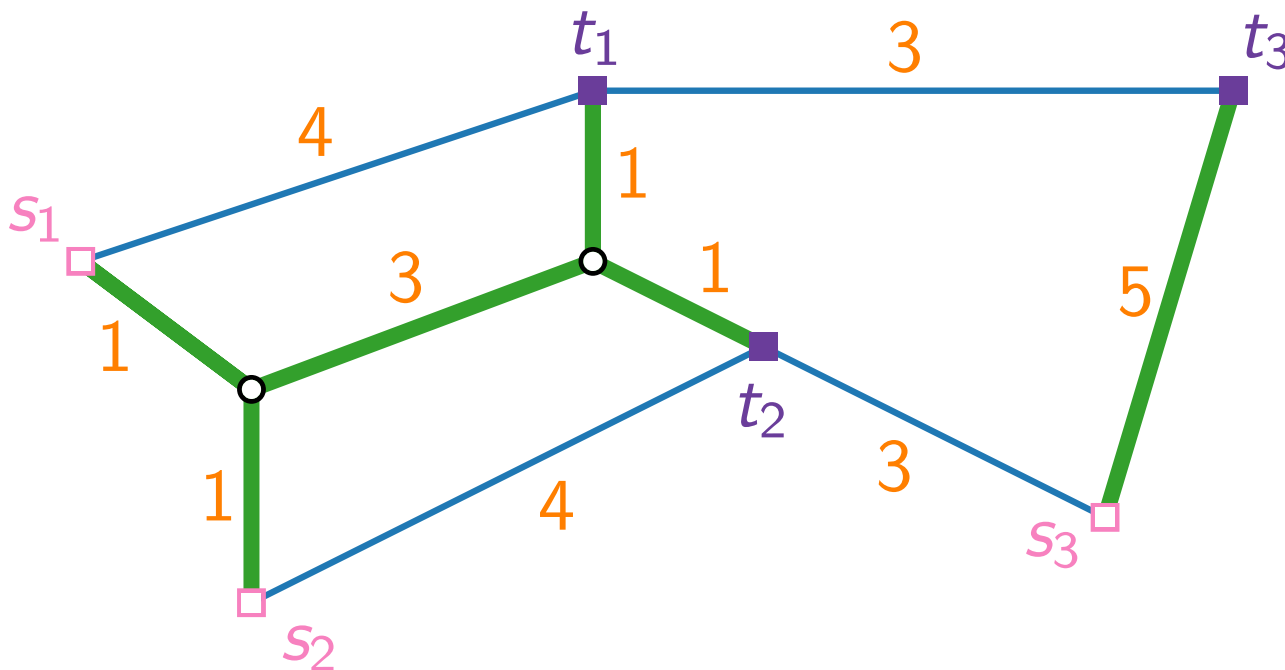
Computational Approaches?

- Merge k shortest s_i-t_i paths
- STEINERTREE on the set of terminals

Homework: Both above approaches perform poorly :-)

Difficulty:

Which terminals belong to the same tree of the forest?



Approximation Algorithms

Lecture 12:

STEINERFOREST via Primal–Dual

Part II:

Primal and Dual LP

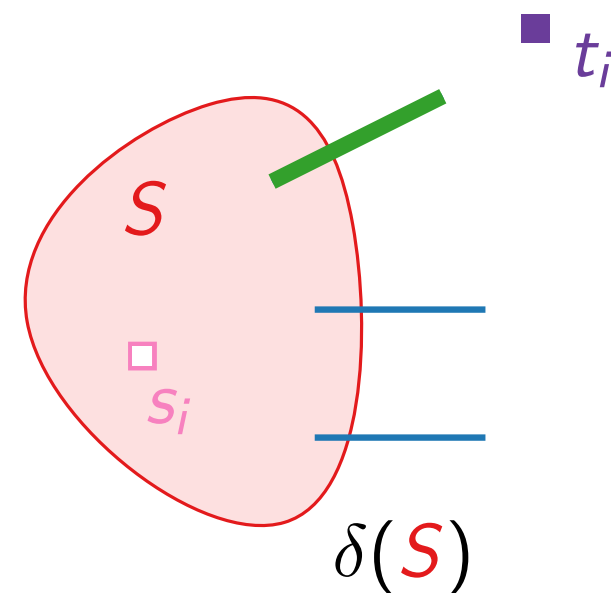
An ILP

$$\begin{array}{ll}
 \text{minimize} & \sum_{e \in E(G)} c_e x_e \\
 \text{subject to} & \sum_{e \in \delta(S)} x_e \geq 1 \quad S \in \mathcal{S}_i, i \in \{1, \dots, k\} \\
 & x_e \in \{0, 1\} \quad e \in E(G)
 \end{array}$$

where $\mathcal{S}_i := \{S \subseteq V : s_i \in S, t_i \notin S\}$

and $\delta(S) := \{(u, v) \in E : u \in S \text{ and } v \notin S\}$

\Rightarrow exponentially many constraints!



LP-Relaxation and Dual LP

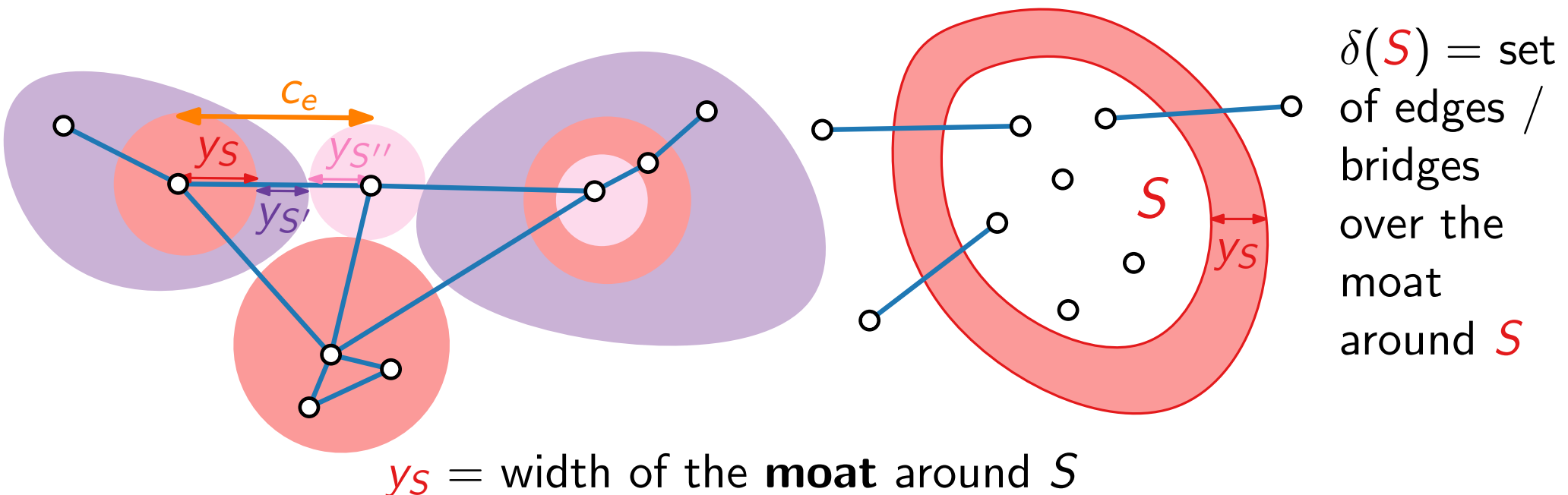
$$\begin{array}{ll}
 \text{minimize} & \sum_{e \in E(G)} c_e x_e \\
 \text{subject to} & \sum_{e \in \delta(S)} x_e \geq 1 \quad S \in \mathcal{S}_i, i \in \{1, \dots, k\} \quad (y_S) \\
 & x_e \geq 0 \quad e \in E(G)
 \end{array}$$

$$\begin{array}{ll}
 \text{maximize} & \sum_{\substack{S \in \mathcal{S}_i \\ i \in \{1, \dots, k\}}} y_S \\
 \text{subject to} & \sum_{S: e \in \delta(S)} y_S \leq c_e \quad e \in E(G) \\
 & y_S \geq 0 \quad S \in \mathcal{S}_i, i \in \{1, \dots, k\}
 \end{array}$$

Intuition for the Dual

$$\begin{array}{ll}
 \text{maximize} & \sum_{\substack{S \in \mathcal{S}_i \\ i \in \{1, \dots, k\}}} y_S \\
 \text{subject to} & \sum_{S: e \in \delta(S)} y_S \leq c_e \quad e \in E(G) \\
 & y_S \geq 0 \quad S \in \mathcal{S}_i, i \in \{1, \dots, k\}
 \end{array}$$

The graph is a network of **bridges**, spanning the **moats**.



Approximation Algorithms

Lecture 12:

STEINERFOREST via Primal–Dual

Part III:

A First Primal–Dual Approach

Complementary Slackness (Reminder)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \geq b \\ & x \geq 0 \end{array}$$

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y \leq c \\ & y \geq 0 \end{array}$$

Theorem. Let $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_m)$ be valid solutions for the primal and dual program (resp.). Then x and y are optimal if and only if the following conditions are met:

Primal CS:

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

Dual CS:

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

A First Primal–Dual Approach

Complementary slackness: $x_e > 0 \Rightarrow \sum_{S: e \in \delta(S)} y_S = c_e.$

\Rightarrow Pick “critical” edges (and only these)!

Idea: Iteratively build a feasible integral primal solution.

How to find a violated primal constraint? $(\sum_{e \in \delta(S)} x_e < 1)$

- Consider related connected component C !

How do we iteratively improve the dual solution?

- Increase y_C (until some edge in $\delta(C)$ becomes critical)!

A First Primal–Dual Approach

PrimalDualSteinerForestNaive(graph G , costs c , pairs R)

$y \leftarrow 0, F \leftarrow \emptyset$

while $\exists (s, t) \in R$ not connected in $(V(G), F)$ **do**

$C \leftarrow$ component in $(V(G), F)$ with $|C \cap \{s, t\}| = 1$

 Increase y_C

 until $\sum_{S: e' \in \delta(S)} y_S = c_{e'}$ for some $e' \in \delta(C)$.

$F \leftarrow F \cup \{e'\}$

return F

Running time??

Trick: Handle all y_S with $y_S = 0$ implicitly.

Analysis

The cost of the solution F can be written as

$$\sum_{e \in F} c_e \stackrel{\text{CS}}{=} \sum_{e \in F} \sum_{S: e \in \delta(S)} y_S = \sum_S |\delta(S) \cap F| \cdot y_S.$$

Compare to the value of the dual objective function $\sum_S y_S$.

There are examples with $|\delta(S) \cap F| = k$ for each $y_S > 0$:-
(Homework!)

But: Average degree of “active components” is less than 2.

\Rightarrow Increase y_C for all active components C simultaneously!

Approximation Algorithms

Lecture 12:

STEINERFOREST via Primal–Dual

Part IV:

Primal–Dual with Synchronized Increases

Primal–Dual with Synchronized Increases

PrimalDualSteinerForest(graph G , edge costs c , pairs R)

$y \leftarrow 0, F \leftarrow \emptyset, \ell \leftarrow 0$

while $\exists (s, t) \in R$ not connected in $(V(G), F)$ **do**

$\ell \leftarrow \ell + 1$

$\mathcal{C} \leftarrow \{\text{component } C \text{ in } (V(G), F) \text{ with } |C \cap \{s_i, t_i\}| = 1 \text{ for some } i\}$

 Increase y_C for all $C \in \mathcal{C}$ simultaneously

 until $\sum_{S: e_\ell \in \delta(S)} y_S = c_{e_\ell}$ for some $e_\ell \in \delta(C), C \in \mathcal{C}$.

$F \leftarrow F \cup \{e_\ell\}$

$F' \leftarrow F$

// Pruning

for $j \leftarrow \ell$ **downto** 1 **do**

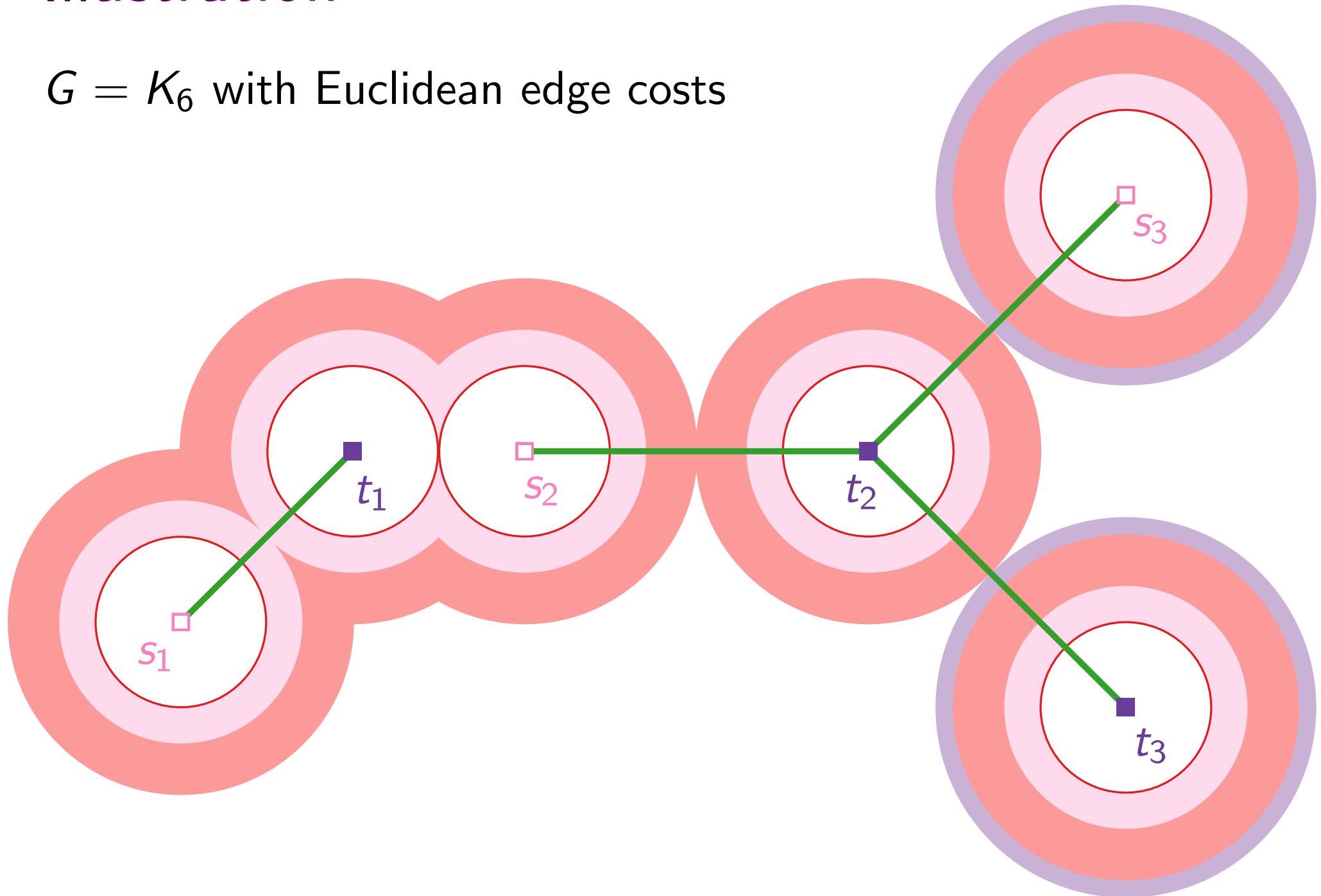
if $F' \setminus \{e_j\}$ is feasible solution **then**

$F' \leftarrow F' \setminus \{e_j\}$

return F'

Illustration

$G = K_6$ with Euclidean edge costs



Approximation Algorithms

Lecture 12:

STEINERFOREST via Primal–Dual

Part V:

Structure Lemma

Structure Lemma

Lemma. In any iteration of the algorithm, it holds that

$$\sum_{C \in \mathcal{C}} |\delta(C) \cap F'| \leq 2|\mathcal{C}|.$$

Proof. First the intuition...

Every connected component C of F is a forest in F' .

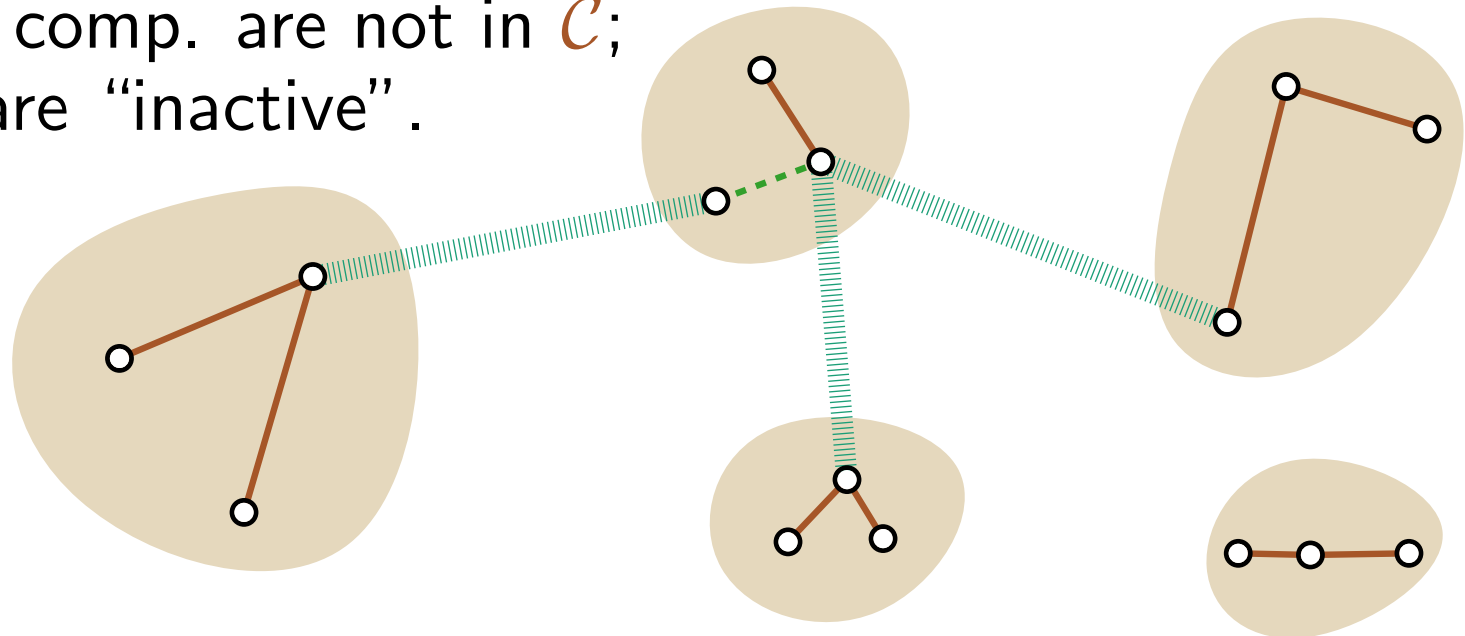
\Rightarrow average degree ≤ 2

Difficulty: Some comp. are not in \mathcal{C} ;
they are “inactive”.

||||| $\delta(C) \cap F'$

— $F' \cap C$

..... $F - F'$



Proof of the Structure Lemma

Lemma. In any iteration of the algorithm, it holds that

$$\sum_{C \in \mathcal{C}} |\delta(C) \cap F'| \leq 2|\mathcal{C}|.$$

Proof.

For $i \in \{1, \dots, \ell\}$, consider the i -th iteration (when e_i was added to F).

Let $F_i = \{e_1, \dots, e_i\}$, $G_i = (V, F_i)$, and $G_i^* = (V, F_i \cup F')$.

Contract every component C of G_i in G_i^* to a single vertex $\rightsquigarrow G'_i$.

(Ignore components C with $\delta(C) \cap F' = \emptyset$.)

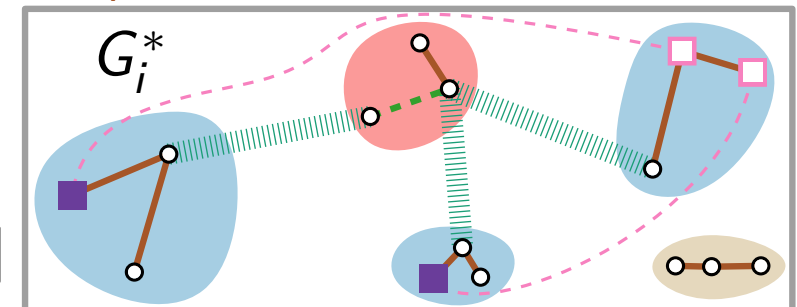
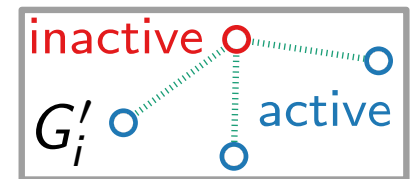
Claim. G'_i is a forest.

$$\begin{aligned} \text{Note: } \sum_{C \text{ comp.}} |\delta(C) \cap F'| &= \sum_{v \in V(G'_i)} \deg_{G'_i}(v) \\ &= 2|E(G'_i)| < 2|V(G'_i)| \end{aligned}$$

Claim. Inactive vertices have degree ≥ 2 .

$$\begin{aligned} \Rightarrow \sum_{v \text{ active}} \deg_{G'_i}(v) &\leq \\ 2 \cdot |V(G'_i)| - 2 \cdot \#(\text{inactive}) &= 2|\mathcal{C}|. \end{aligned}$$

□



Approximation Algorithms

Lecture 12:

STEINERFOREST via Primal–Dual

Part VI:
Analysis

Analysis

Theorem. The Primal–Dual algorithm with synchronized increases yields a 2-approximation for STEINERFOREST.

Proof.

As mentioned before,

$$\sum_{e \in F'} c_e \stackrel{\text{CS}}{=} \sum_{e \in F'} \sum_{S: e \in \delta(S)} y_S = \sum_S |\delta(S) \cap F'| \cdot y_S.$$

We prove by induction over the number of iterations of the algorithm that

$$\sum_S |\delta(S) \cap F'| \cdot y_S \leq 2 \sum_S y_S. \quad (*)$$

From that, the claim of the theorem follows.

Analysis

Theorem. The Primal–Dual algorithm with synchronized increases yields a 2-approximation for STEINERFOREST.

Proof.
$$\sum_S |\delta(S) \cap F'| \cdot y_S \leq 2 \sum_S y_S. \quad (*)$$

Base case trivial since we start with $y_S = 0$ for every S .

Assume that $(*)$ holds at the start of the current iteration.

In the current iteration, we increase y_C for every $C \in \mathcal{C}$ by the same amount, say $\varepsilon \geq 0$.

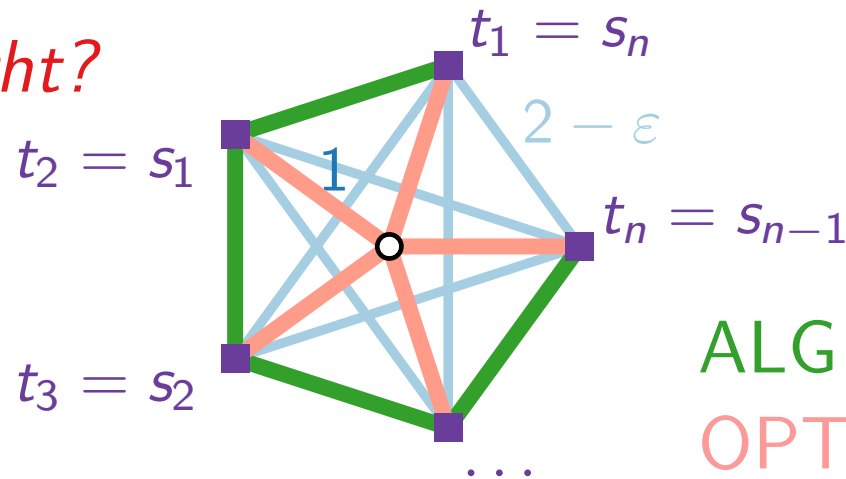
This increases the left side of $(*)$ by $\varepsilon \cdot \sum_{C \in \mathcal{C}} |\delta(C) \cap F'|$ and the right side by $\varepsilon \cdot 2|\mathcal{C}|$.

Structure lemma $\Rightarrow (*)$ also holds after the current iteration. □

Summary

Theorem. The Primal–Dual algorithm with synchronized increases yields a **2**-approximation for STEINERFOREST.

Is our analysis tight?



$$\text{ALG} = (2 - \varepsilon)(n - 1)$$

$$\text{OPT} = n$$

Can we do better?

No better approximation factor is known. :-)

The integrality gap is $2 - 1/n$.

STEINERFOREST (as STEINERTREE) cannot be approximated within factor $\frac{96}{95} \approx 1.0105$ (unless $P = NP$). [Chlebík, Chlebíková '08]