Approximation Algorithms

Lecture 12: SteinerForest via Primal–Dual

Part I: SteinerForest

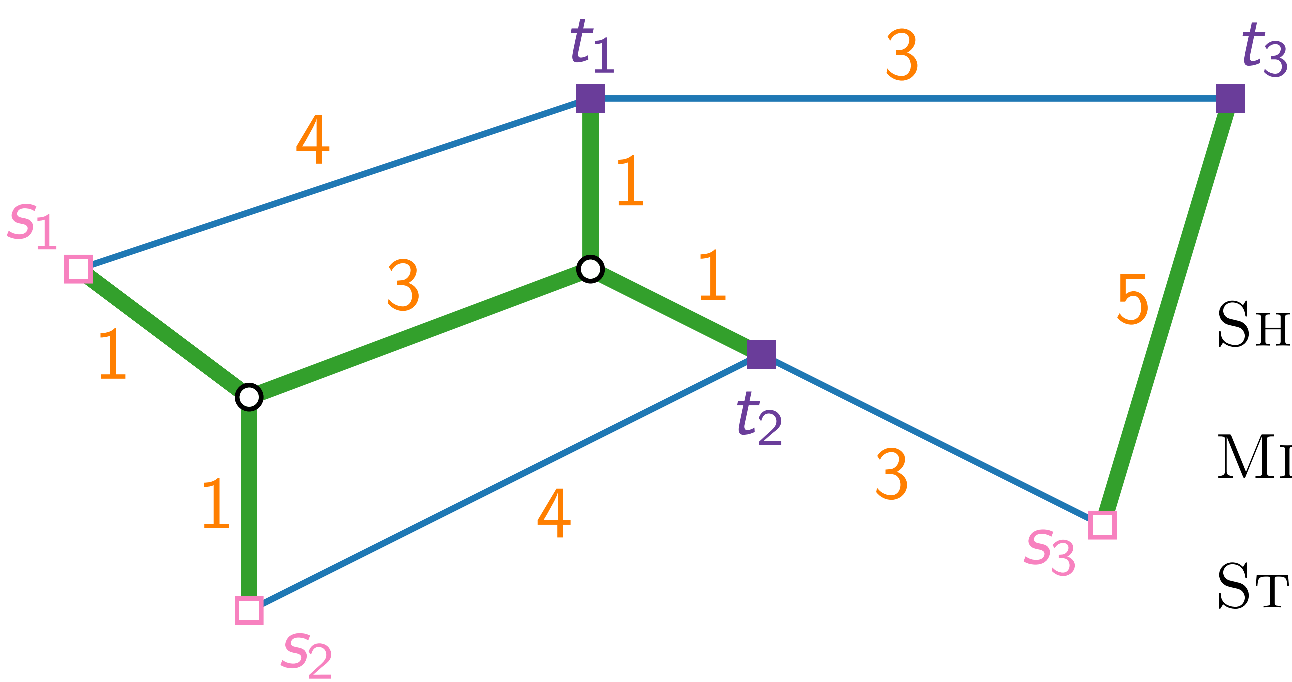
Alexander Wolff

Winter term 2025

SteinerForest

Given: A graph G with edge costs $c \colon E(G) \to \mathbb{N}$ 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 $(R = \{s, t\})$

MinSpannTree $(R = E(G))$

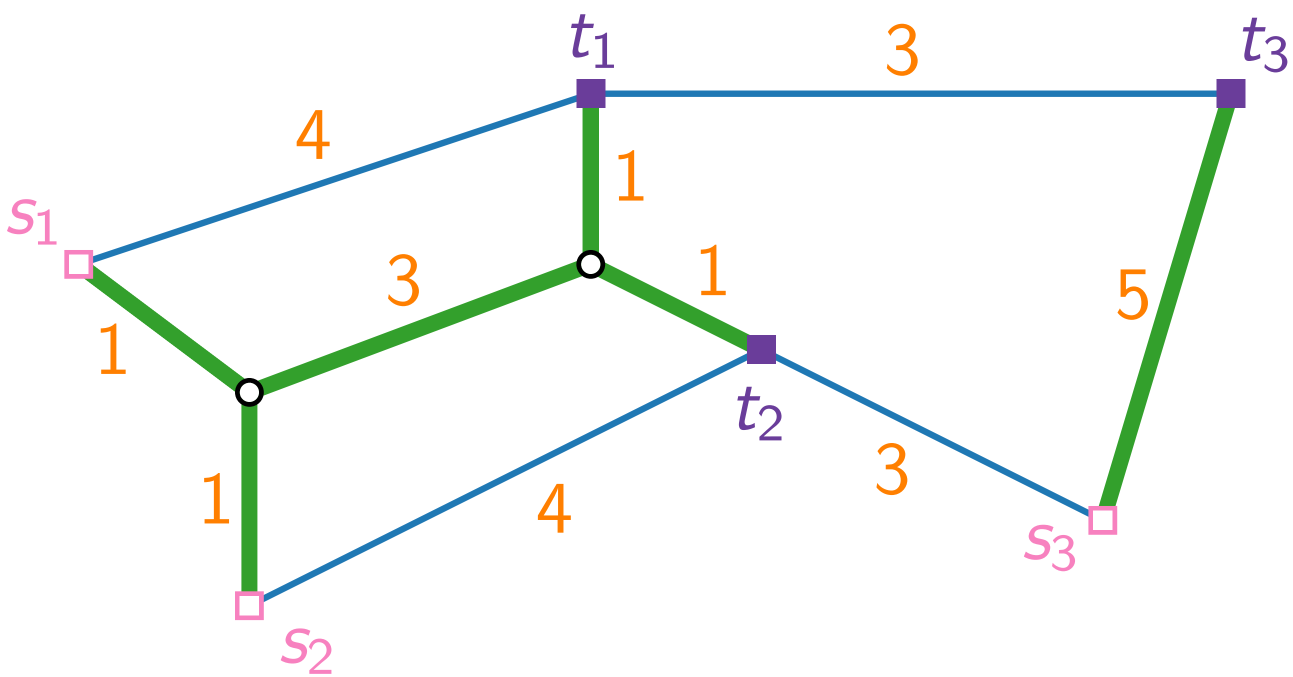
SteinerTree $(R = T \times T)$

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

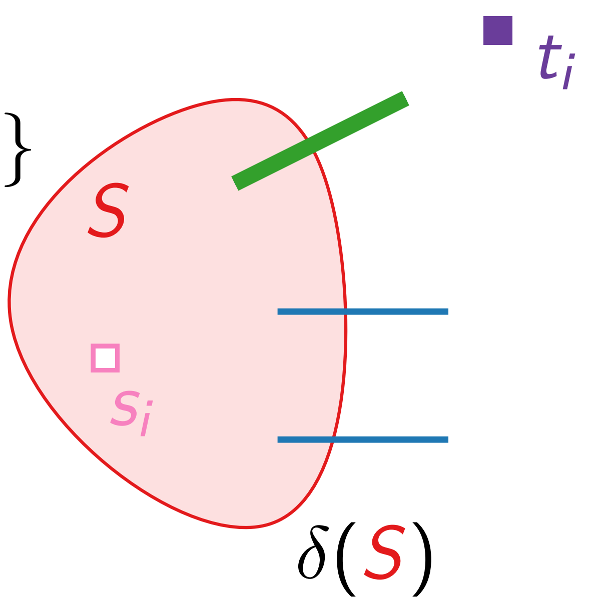
An ILP

$$\begin{aligned} \textbf{minimize} \quad & \sum_{e \in E(G)} c_e x_e \\ \textbf{subject to} \quad & \sum_{e \in \delta(S)} x_e \geq 1 \qquad S \in \mathcal{S}_i,\ i \in \{1, \ldots, k\} \\ & x_e \in \{0, 1\} \qquad e \in E(G) \end{aligned}$$

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

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

$\Rightarrow$ exponentially many constraints!



LP-Relaxation and Dual LP

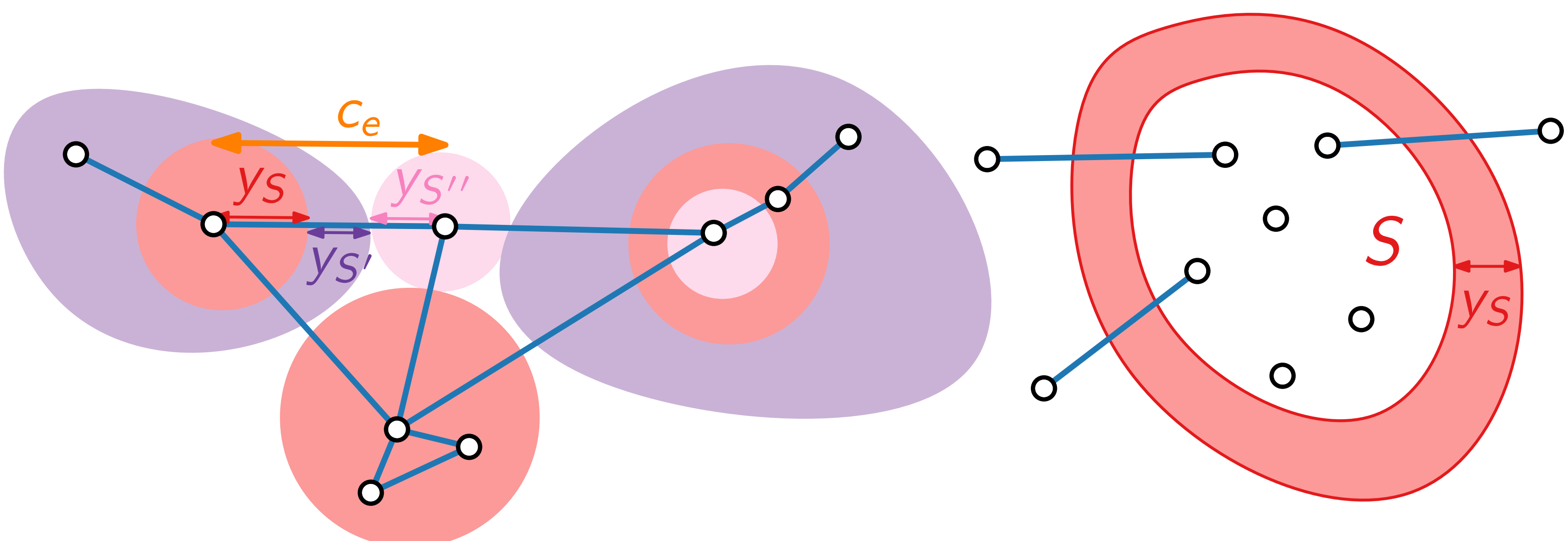
$$\begin{array}{lll}\textbf{minimize} & \displaystyle\sum_{e \in E(G)} c_e x_e & \\ \textbf{subject to} & \displaystyle\sum_{e \in \delta(S)} x_e \geq 1 & S \in \mathcal{S}_i,\ i \in \{1, \ldots, k\} \quad (y_S) \\ & x_e \geq 0 & e \in E(G)\end{array}$$

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

Intuition for the Dual

$$
\begin{aligned}
\textbf{maximize} \quad & \sum_{\substack{S \in \mathcal{S}_i \\ i \in \{1,\ldots,k\}}} y_S \\
\textbf{subject to} \quad & \sum_{S:\ e \in \delta(S)} y_S \le c_e \qquad e \in E(G) \\
& y_S \ge 0 \qquad S \in \mathcal{S}_i,\ i \in \{1, \ldots, k\}
\end{aligned}
$$

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



$\delta(S)$ = set of edges / bridges over the moat around S

y_S = width of the **moat** around S

Approximation Algorithms

Lecture 12:
SteinerForest via Primal–Dual

Part III:
A First Primal–Dual Approach

Complementary Slackness (Reminder)

minimize	$c^\intercal x$		
subject to	Ax	$\geq$	b
	x	$\geq$	0

maximize	$b^\intercal y$		
subject to	$A^\intercal y$	$\leq$	c
	y	$\geq$	0

Theorem. Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_m)$ 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 $j = 1, \ldots, n$: 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 $\sum_{j=1}^{n} a_{ij} x_j = b_i$

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

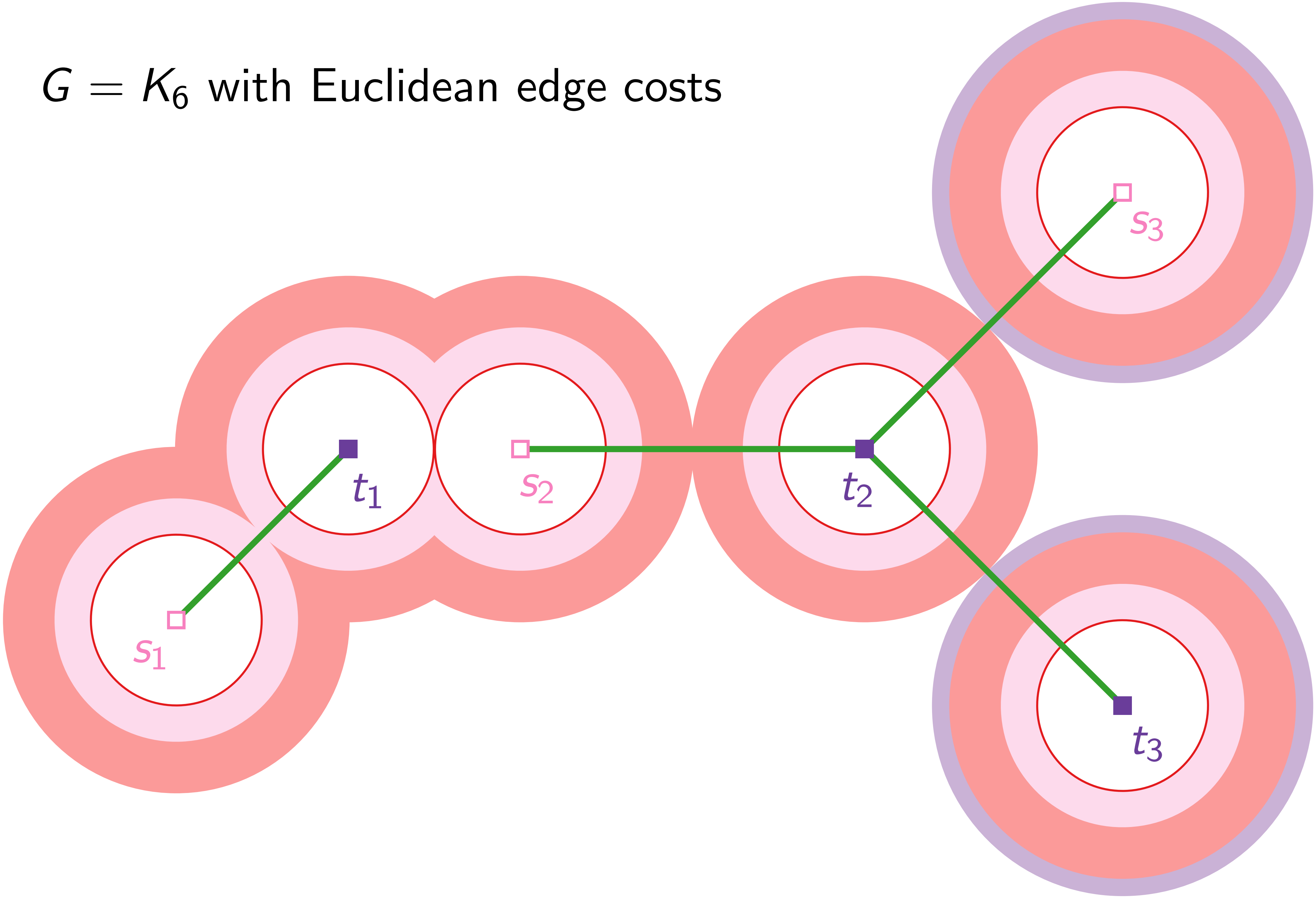
Primal–Dual with Synchronized Increases

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**
  $\ell \leftarrow \ell + 1$
  $\mathcal{C} \leftarrow \{$component C in $(V(G), F)$ with $|C \cap \{s_i, t_i\}| = 1$ for some $i\}$
  Increase y_C for all $C \in \mathcal{C}$ simultaneously
    until $\sum_{S\colon e_\ell \in \delta(S)} y_S = c_{e_\ell}$ 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

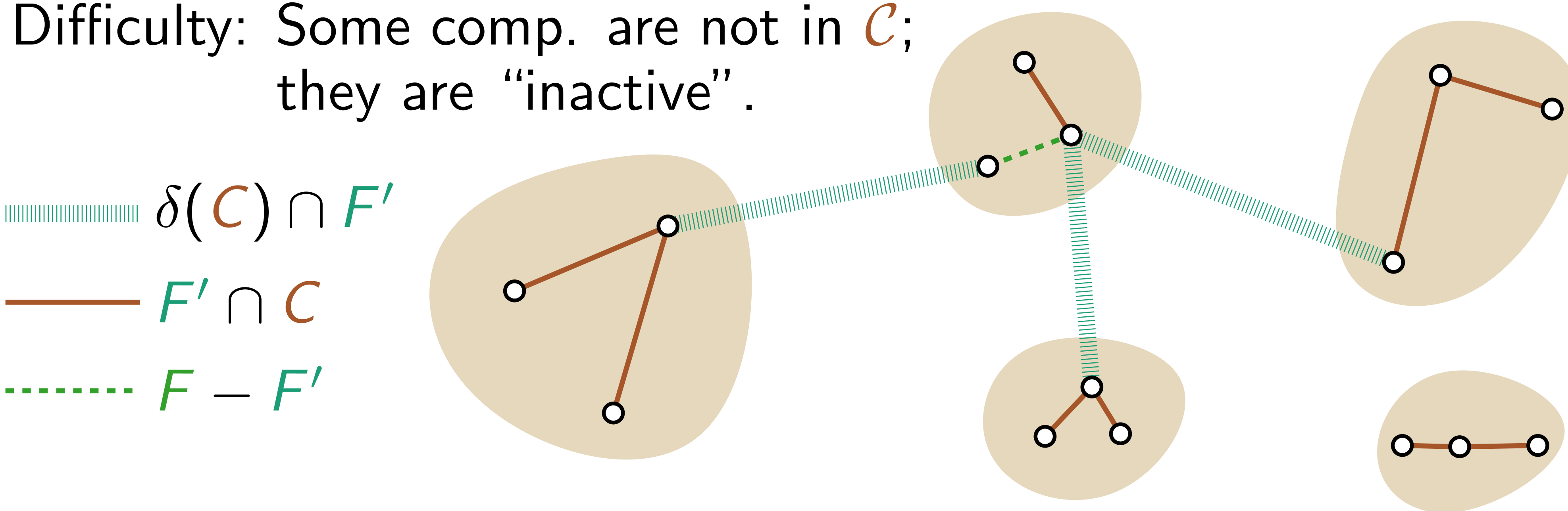
Part V:
Structure Lemma

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".



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

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

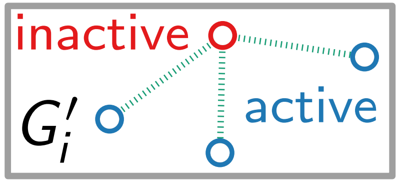
Contract every component C of G_i in G_i^* to a single vertex $\rightsquigarrow G_i'$.

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

Claim. G_i' is a forest.

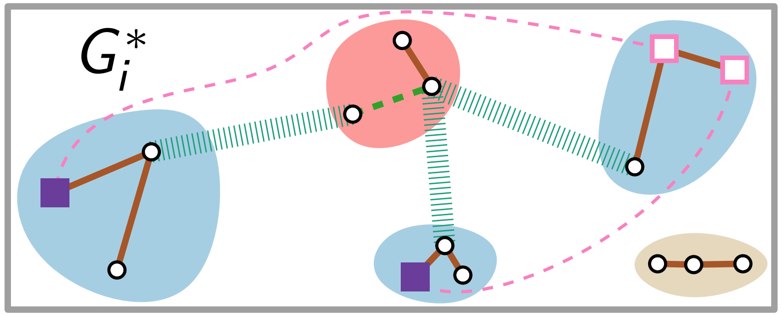
Note: $\sum_{C \text{ comp.}} |\delta(C) \cap F'| = \sum_{v \in V(G_i')} \deg_{G_i'}(v)$

$= 2|E(G_i')| < 2|V(G_i')|$



Claim. Inactive vertices have degree ≥ 2.

$\Rightarrow \sum_{v \text{ active}} \deg_{G_i'}(v) \leq$
$2 \cdot |V(G_i')| - 2 \cdot \#(\text{inactive}) = 2|\mathcal{C}|$. $\square$



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

$$\sum_{S} |\delta(S) \cap F'| \cdot y_S \leq 2 \sum_{S} y_S . \qquad (*)$$

From that, the claim of the theorem follows.

Analysis

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

Proof.

$$\sum_{S} |\delta(S) \cap F'| \cdot y_S \leq 2 \sum_{S} y_S. \qquad (*)$$

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 $\varepsilon \geq 0$.

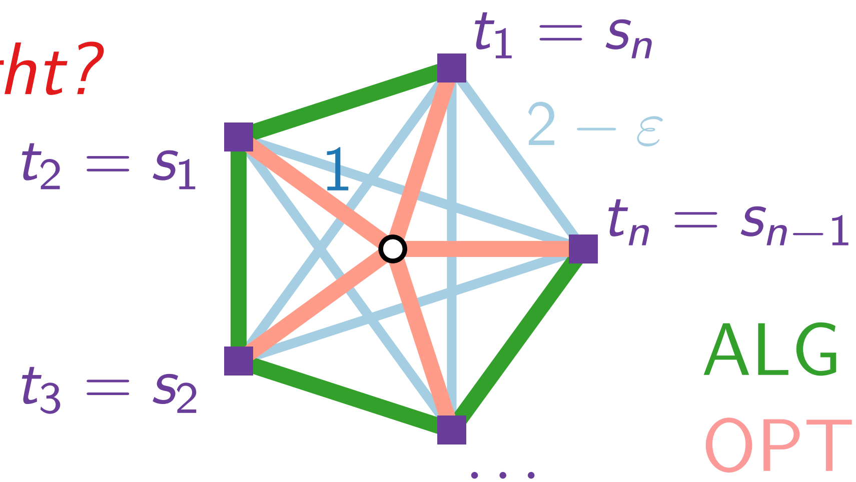
This increases the left side of $(*)$ by $\varepsilon \cdot \sum_{C \in \mathcal{C}} |\delta(C) \cap F'|$ and the right side by $\varepsilon \cdot 2|\mathcal{C}|$.

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?



ALG $= (2 - \varepsilon)(n - 1)$
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]