# Recent Progress in Approximation Algorithms for the Traveling Salesman Problem

Lecture 3: The $s$-$t$ path TSP

David P. Williamson
Cornell University

July 18-22, 2016
São Paulo School of Advanced Science on Algorithms, Combinatorics, and Optimization

# $s$-$t$ path TSP

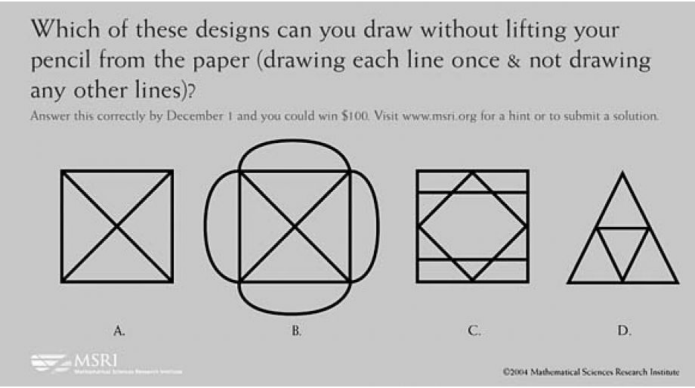
Recall the $s$-$t$ path TSP:
Usual TSP input plus $s, t \in V$, find a min-cost path from $s$ to $t$ visiting all other nodes in between (an *$s$-$t$ Hamiltonian path*).

# Eulerian path

There is an Eulerian path that starts at $s$, ends at $t$, and visits every edge exactly once iff $s$ and $t$ have odd-degree and all other vertices have even degree.

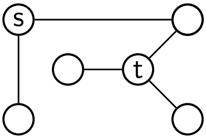
# Eulerian path

There is an Eulerian path that starts at $s$, ends at $t$, and visits every edge exactly once iff $s$ and $t$ have odd-degree and all other vertices have even degree.



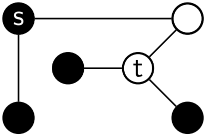
# Hoogeveen's algorithm

Let $F$ be the min-cost spanning tree. Let $T$ be the set of vertices whose *parity needs changing*: $s$ iff $s$ has even degree in $F$, $t$ iff $t$ has even degree in $F$, and $v \neq s, t$ iff $v$ has odd degree. Then find a minimum-cost perfect matching $M$ on the vertices in $T$. Find Eulerian path on $F \cup M$; shortcut to an $s$-$t$ Hamiltonian path.



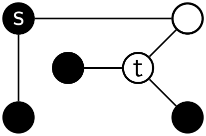
# Hoogeveen's algorithm

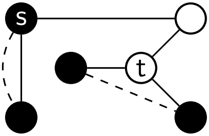
Let $F$ be the min-cost spanning tree. Let $T$ be the set of vertices whose *parity needs changing*: $s$ iff $s$ has even degree in $F$, $t$ iff $t$ has even degree in $F$, and $v \neq s, t$ iff $v$ has odd degree. Then find a minimum-cost perfect matching $M$ on the vertices in $T$. Find Eulerian path on $F \cup M$; shortcut to an $s$-$t$ Hamiltonian path.



# Hoogeveen's algorithm

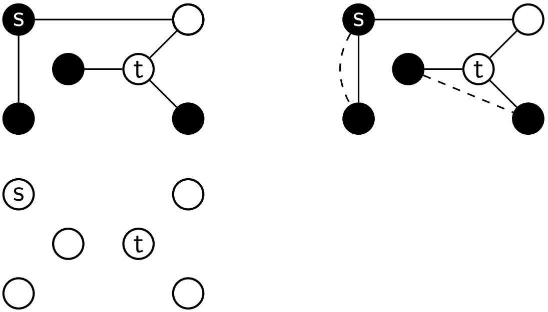
Let $F$ be the min-cost spanning tree. Let $T$ be the set of vertices whose *parity needs changing*: $s$ iff $s$ has even degree in $F$, $t$ iff $t$ has even degree in $F$, and $v \neq s, t$ iff $v$ has odd degree. Then find a minimum-cost perfect matching $M$ on the vertices in $T$. Find Eulerian path on $F \cup M$; shortcut to an $s$-$t$ Hamiltonian path.



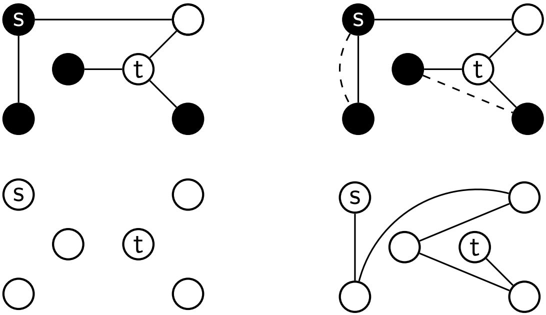
# Hoogeveen's algorithm

Let $F$ be the min-cost spanning tree. Let $T$ be the set of vertices whose *parity needs changing*: $s$ iff $s$ has even degree in $F$, $t$ iff $t$ has even degree in $F$, and $v \neq s, t$ iff $v$ has odd degree. Then find a minimum-cost perfect matching $M$ on the vertices in $T$. Find Eulerian path on $F \cup M$; shortcut to an $s$-$t$ Hamiltonian path.



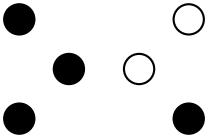
# Hoogeveen's algorithm

Let $F$ be the min-cost spanning tree. Let $T$ be the set of vertices whose *parity needs changing*: $s$ iff $s$ has even degree in $F$, $t$ iff $t$ has even degree in $F$, and $v \neq s, t$ iff $v$ has odd degree. Then find a minimum-cost perfect matching $M$ on the vertices in $T$. Find Eulerian path on $F \cup M$; shortcut to an $s$-$t$ Hamiltonian path.



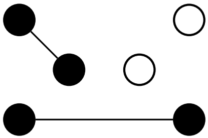
# $T$-joins

Rather than a minimum-cost perfect matching on $T$, will construct a minimum-cost *$T$-join*: a set of edges that has odd degree at every vertex in $T$, even degree at every vertex not in $T$.



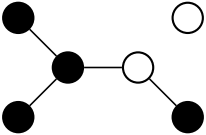
# $T$-joins

Rather than a minimum-cost perfect matching on $T$, will construct a minimum-cost *$T$-join*: a set of edges that has odd degree at every vertex in $T$, even degree at every vertex not in $T$.



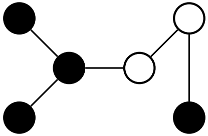
# $T$-joins

Rather than a minimum-cost perfect matching on $T$, will construct a minimum-cost *$T$-join*: a set of edges that has odd degree at every vertex in $T$, even degree at every vertex not in $T$.



# $T$-joins

Rather than a minimum-cost perfect matching on $T$, will construct a minimum-cost *$T$-join*: a set of edges that has odd degree at every vertex in $T$, even degree at every vertex not in $T$.



# Hoogeveen's algorithm

Let $F$ be the min-cost spanning tree. Let $T$ be the set of vertices whose parity needs changing. Then find a minimum-cost $T$-join $J$. Find Eulerian path on $F \cup J$; shortcut to an $s$-$t$ Hamiltonian path.

# Hoogeveen's algorithm

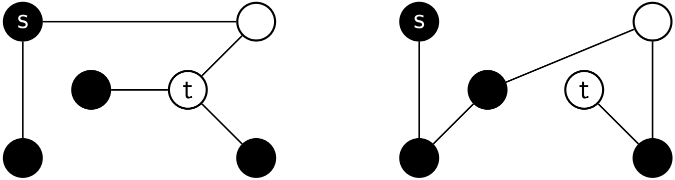
Let $F$ be the min-cost spanning tree. Let $T$ be the set of vertices whose parity needs changing. Then find a minimum-cost $T$-join $J$. Find Eulerian path on $F \cup J$; shortcut to an $s$-$t$ Hamiltonian path.

**Theorem**

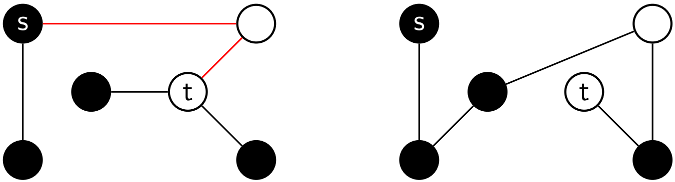
*Hoogeveen's algorithm is a $\frac{5}{3}$-approximation algorithm.*

# Proof of theorem

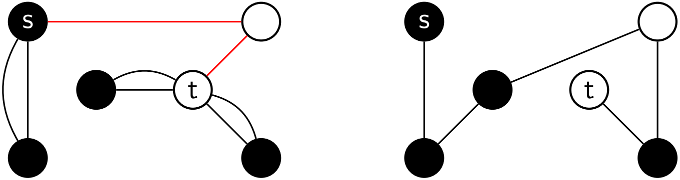
# Proof of theorem



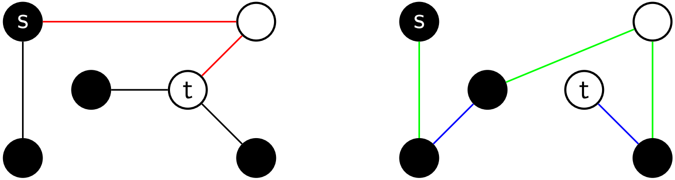
# Proof of theorem



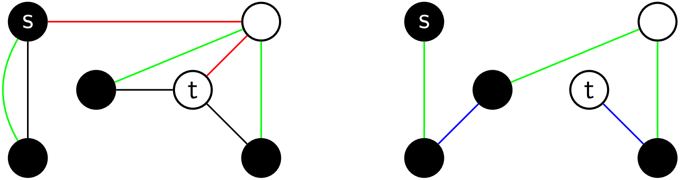
# Proof of theorem



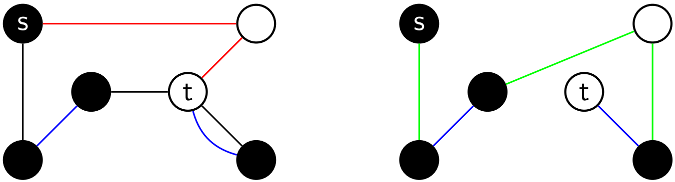
# Proof of theorem



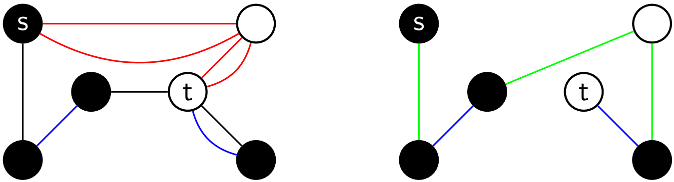
# Proof of theorem



# Proof of theorem

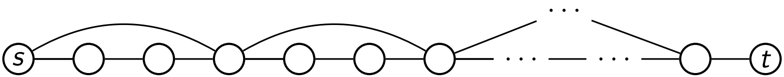


# Proof of theorem



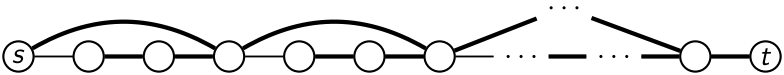
# Tight Example

The analysis is tight. Consider the *graph TSP* instance below: cost $c(e)$ for $e = (i, j)$ is number of edges in shortest $i$-$j$ path in graph.



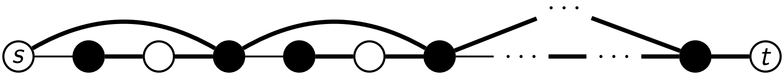
# Tight Example

The analysis is tight. Consider the *graph TSP* instance below: cost $c(e)$ for $e = (i, j)$ is number of edges in shortest $i$-$j$ path in graph.



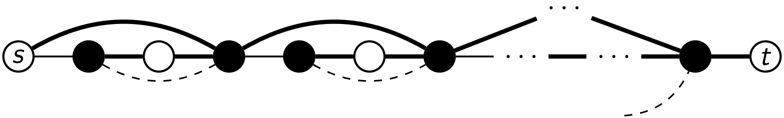
# Tight Example

The analysis is tight. Consider the *graph TSP* instance below: cost $c(e)$ for $e = (i, j)$ is number of edges in shortest $i$-$j$ path in graph.



# Tight Example

The analysis is tight. Consider the *graph TSP* instance below: cost $c(e)$ for $e = (i, j)$ is number of edges in shortest $i$-$j$ path in graph.



# Improvements

No improvement on Hoogeveen's algorithm for *s*-*t* path TSP, until just the last few years.

| | | |
|---|---|---|
| An, Kleinberg, Shmoys | 2012 | 1.618 |
| Sebő | 2013 | 1.6 |
| Vygen | 2015 | 1.599 |
| Gottschalk and Vygen | 2015 | 1.56 |
| Sebő and Van Zuylen | 2016 | 1.52 |

Goal: Understand the An et al. algorithm and analysis; will sketch some of the ideas of the improvements.

# A Linear Programming Relaxation

$$\text{Min} \quad \sum_{e \in E} c_e x_e$$

subject to:

$$x(\delta(v)) = \Big\{$$

$$x(\delta(S)) \geq \Big\{$$

$$0 \leq x_e \leq 1, \qquad \forall e \in E,$$

where $\delta(S)$ is the set of edges with exactly one endpoint in $S$, and $x(E') \equiv \sum_{e \in E'} x_e$.

# A Linear Programming Relaxation

$$\text{Min} \quad \sum_{e \in E} c_e x_e$$

subject to:

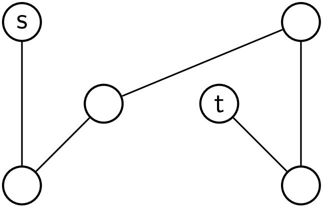
$$x(\delta(v)) = \begin{cases} 1, & v = s, t, \\ 2, & v \neq s, t, \end{cases}$$

$$x(\delta(S)) \geq \begin{cases} 1, & |S \cap \{s, t\}| = 1, \\ 2, & |S \cap \{s, t\}| \neq 1, \end{cases}$$

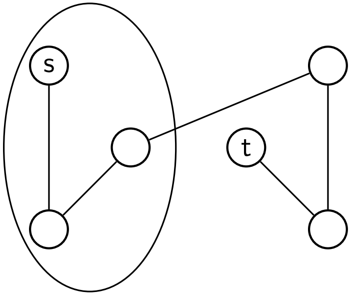
$$0 \leq x_e \leq 1, \qquad \forall e \in E,$$

where $\delta(S)$ is the set of edges with exactly one endpoint in $S$, and $x(E') \equiv \sum_{e \in E'} x_e$.

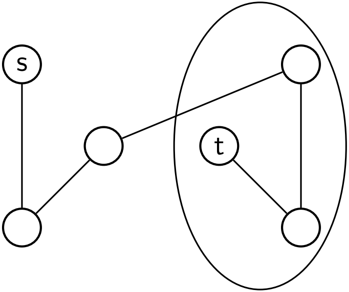
# LP relaxation



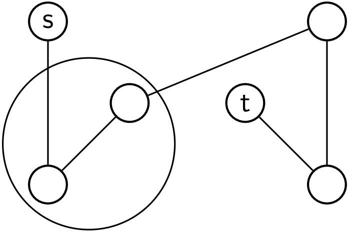
# LP relaxation



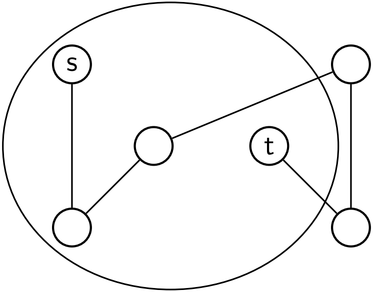
# LP relaxation



# LP relaxation



# LP relaxation

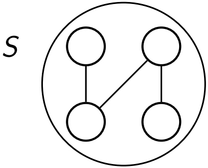


# The spanning tree polytope

The spanning tree polytope (convex hull of all spanning trees) is defined by the following inequalities:

$$\begin{aligned} x(E) &= |V| - 1, & & \\ x(E(S)) &\leq |S| - 1, & & \forall |S| \subseteq V, |S| \geq 2, \\ x(e) &\geq 0, & & \forall e \in E, \end{aligned}$$

where $E(S)$ is the set of all edges with both endpoints in $S$.



# The LP relaxation and spanning trees

**Lemma**

*Any solution $x$ feasible for the $s$-$t$ path TSP LP relaxation is in the spanning tree polytope.*

# Proof

$$x(\delta(v)) = \begin{cases} 1, & v = s,t, \\ 2, & v \neq s,t, \end{cases}$$

$$x(\delta(S)) \geq \begin{cases} 1, & |S \cap \{s,t\}| = 1, \\ 2, & |S \cap \{s,t\}| \neq 1, \end{cases}$$

$$0 \leq x(e) \leq 1, \quad \forall e \in E.$$

$$x(E) = |V| - 1,$$

$$x(E(S)) \leq |S| - 1, \quad \forall |S| \subseteq V, |S| \geq 2,$$

$$x(e) \geq 0, \quad \forall e \in E.$$

# A warmup to the improvements

Let $OPT_{LP}$ be the value of an optimal solution $x^*$ to the LP relaxation.

**Theorem (An, Kleinberg, Shmoys (2012))**

*Hoogeveen's algorithm returns a solution of cost at most* $\frac{5}{3}OPT_{LP}$.

# An extremely useful lemma

Let $F$ be a spanning tree, and let $T$ be the vertices whose parity needs fixing in $F$.

**Definition**

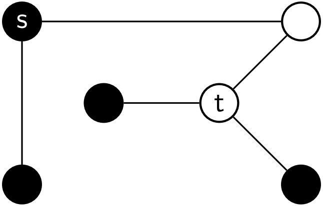
$S$ is an *odd set* if $|S \cap T|$ is odd.

**Lemma**

*Let $S$ be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.*

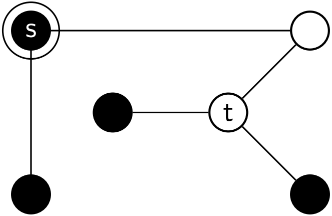
## Lemma

*Let $S$ be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.*



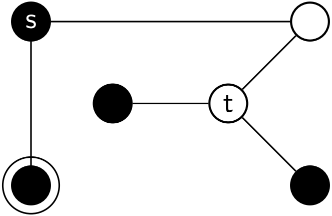
## Lemma

*Let $S$ be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.*



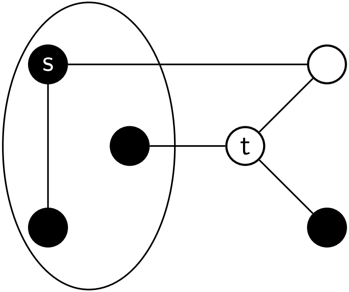
## Lemma

*Let $S$ be an odd set. If $|S \cap \{s, t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s, t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.*



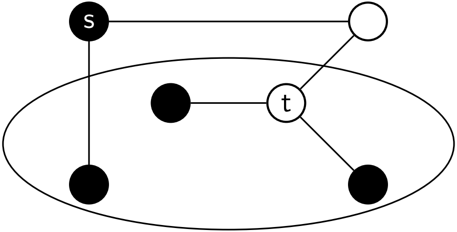
## Lemma

*Let $S$ be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.*

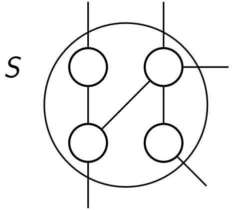


**Lemma**

*Let $S$ be an odd set. If $|S \cap \{s,t\}| = 1$, then $|F \cap \delta(S)|$ is even. If $|S \cap \{s,t\}| \neq 1$, then $|F \cap \delta(S)|$ is odd.*



# Proof of lemma

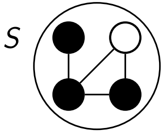



$$\sum_{v \in S} deg_F(v) = 2|E(S) \cap F| + |\delta(S) \cap F|$$

# $T$-join LP

The solution to the following linear program is the minimum-cost $T$-join for costs $c \geq 0$:

$$\begin{array}{lll} \text{Min} & \sum_{e \in E} c_e x_e & \\ \text{subject to:} & x(\delta(S)) \geq 1, & \forall S \subseteq V, |S \cap T| \text{ odd} \\ & x_e \geq 0, & \forall e \in E. \end{array}$$



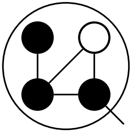

$$\sum_{v \in S} deg_J(v) = 2|E(S) \cap J| + |\delta(S) \cap J|$$

# $T$-join LP

The solution to the following linear program is the minimum-cost $T$-join for costs $c \geq 0$:

$$
\begin{array}{lll}
\text{Min} & \displaystyle\sum_{e \in E} c_e x_e & \\
\text{subject to:} & x(\delta(S)) \geq 1, & \forall S \subseteq V, |S \cap T| \text{ odd} \\
& x_e \geq 0, & \forall e \in E.
\end{array}
$$

$S$



$$\sum_{v \in S} deg_J(v) = 2|E(S) \cap J| + |\delta(S) \cap J|$$

# Proof of theorem

**Theorem (An, Kleinberg, Shmoys (2012))**

*Hoogeveen's algorithm returns a solution of cost at most* $\frac{5}{3} OPT_{LP}$.

**Lemma**

*Let* $S$ *be an odd set. If* $|S \cap \{s,t\}| = 1$, *then* $|F \cap \delta(S)|$ *is even. If* $|S \cap \{s,t\}| \neq 1$, *then* $|F \cap \delta(S)|$ *is odd.*

$$
\begin{array}{lll}
\text{Min} & \displaystyle\sum_{e \in E} c(e)x(e) & \\
 & x(\delta(S)) \geq 1, & \forall S \subseteq V, |S \cap T| \text{ odd} \\
 & x(e) \geq 0, & \forall e \in E.
\end{array}
$$

# Convex combination

Let $x^*$ be an optimal LP solution. Let $\chi_F$ be the *characteristic vector* of a set of edges $F$, so that

$$\chi_F(e) = \begin{cases} 1 & e \in F \\ 0 & e \notin F \end{cases}$$

Since $x^*$ is in the spanning tree polytope, can write $x^*$ as a convex combination of spanning trees $F_1, \ldots, F_k$:

$$x^* = \sum_{i=1}^{k} \lambda_i \chi_{F_i},$$

such that $\sum_{i=1}^{k} \lambda_i = 1$, $\lambda_i \geq 0$.

# Best-of-Many Christofides' Algorithm

An, Kleinberg, Shmoys (2012) propose the *Best-of-Many Christofides'* algorithm: given optimal LP solution $x^*$, compute convex combination of spanning trees

$$x^* = \sum_{i=1}^{k} \lambda_i \chi_{F_i}.$$

For each spanning tree $F_i$, let $T_i$ be the set of vertices whose parity needs fixing, let $J_i$ be the minimum-cost $T_i$-join. Find $s$-$t$ Hamiltonian path by shortcutting $F_i \cup J_i$. Return the shortest path found over all $i$.

# Best-of-Many Christofides' Algorithm

$$x^* = \sum_{i=1}^{k} \lambda_i \chi_{F_i}.$$

For each spanning tree $F_i$, let $T_i$ be the set of vertices whose parity needs fixing, $J_i$ be the minimum-cost $T_i$-join. Find $s$-$t$ Hamiltonian path by shortcutting $F_i \cup J_i$. Return the shortest path found over all $i$.

**Theorem**

*The Best-of-Many Christofides' algorithm returns a solution of cost at most $\frac{5}{3} OPT_{LP}$.*

# Proof

# Improvement?

To do better, we need to improve the analysis for the costs of the $T_i$-joins; recall that we use that

$$y_i = \frac{1}{3}\chi_{F_i} + \frac{1}{3}x^*$$

is feasible for the $T_i$-join LP.

Consider

$$y_i = \alpha\chi_{F_i} + \beta x^*.$$

Then the cost of the best $s$-$t$ Hamiltonian path is at most

$$(1 + \alpha + \beta) OPT_{LP}.$$

# Improvement?

Proof that $y_i$ feasible for $T_i$-join LP had two cases. Assume $S$ odd ($|S \cap T_i|$ odd).

If $|S \cap \{s, t\}| \neq 1$, then

$$y_i(\delta(S)) = \alpha |F_i \cap \delta(S)| + \beta x^*(\delta(S)) \geq \alpha + 2\beta.$$

We will want $\alpha + 2\beta \geq 1$, so the $T_i$-join LP constraint is satisfied.

# Improvement?

If $|S \cap \{s,t\}| = 1$, then

$$y_i(\delta(S)) = \alpha |F_i \cap \delta(S)| + \beta x^*(\delta(S)) \geq 2\alpha + \beta x^*(\delta(S)).$$

# Improvement?

If $|S \cap \{s, t\}| = 1$, then

$$y_i(\delta(S)) = \alpha|F_i \cap \delta(S)| + \beta x^*(\delta(S)) \geq 2\alpha + \beta x^*(\delta(S)).$$

Since we assume $\alpha + 2\beta \geq 1$, we only run into problems if

$$x^*(\delta(S)) < \frac{1 - 2\alpha}{\beta}.$$

Note that $\alpha = 0$, $\beta = \frac{1}{2}$ works if $x^*(\delta(S)) \geq 2$ for all $S \subset V$, and gives a tour of cost at most $\frac{3}{2} OPT_{LP}$.

# Improvement?

If $|S \cap \{s,t\}| = 1$, then

$$y_i(\delta(S)) = \alpha|F_i \cap \delta(S)| + \beta x^*(\delta(S)) \geq 2\alpha + \beta x^*(\delta(S)).$$

Since we assume $\alpha + 2\beta \geq 1$, we only run into problems if

$$x^*(\delta(S)) < \frac{1-2\alpha}{\beta}.$$

Note that $\alpha = 0$, $\beta = \frac{1}{2}$ works if $x^*(\delta(S)) \geq 2$ for all $S \subset V$, and gives a tour of cost at most $\frac{3}{2} OPT_{LP}$.

So focus on cuts for which $x^*(\delta(S)) < 2$, and add an extra "correction" term to $y_i$ to handle these cuts.

# $\tau$-Narrow Cuts

**Definition**

$S$ is *$\tau$-narrow* if $x^*(\delta(S)) < 1 + \tau$ for fixed $\tau \leq 1$.

Only $S$ such that $|S \cap \{s, t\}| = 1$ are $\tau$-narrow.

**Definition**

Let $\mathcal{C}_\tau$ be all $\tau$-narrow cuts $S$ with $s \in S$.

# $\tau$-Narrow Cuts

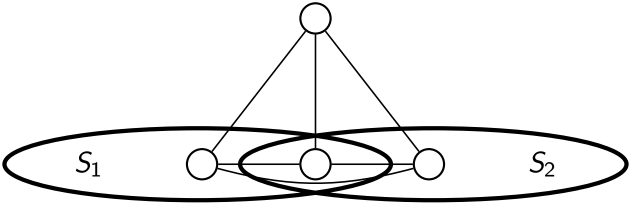
The $\tau$-narrow cuts in $\mathcal{C}_\tau$ have a nice structure.

**Theorem (An, Kleinberg, Shmoys (2012))**

*If $S_1, S_2 \in \mathcal{C}_\tau$, $S_1 \neq S_2$, then either $S_1 \subset S_2$ or $S_2 \subset S_1$.*

First need to show that

$$x^*(\delta(S_1)) + x^*(\delta(S_2)) \geq x^*(\delta(S_1 - S_2)) + x^*(\delta(S_2 - S_1)).$$



# Proof of theorem

**Theorem (An, Kleinberg, Shmoys (2012))**

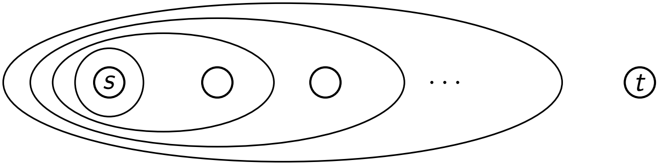
*If $S_1, S_2 \in \mathcal{C}_\tau$, $S_1 \neq S_2$, then either $S_1 \subset S_2$ or $S_2 \subset S_1$.*

# Proof of theorem

**Theorem (An, Kleinberg, Shmoys (2012))**

*If $S_1, S_2 \in \mathcal{C}_\tau$, $S_1 \neq S_2$, then either $S_1 \subset S_2$ or $S_2 \subset S_1$.*

So the $\tau$-narrow cuts look like $s \in Q_1 \subset Q_2 \subset \cdots \subset Q_k \subset V$.



# Correction Factor

Let $e_Q$ be the minimum-cost edge in $\delta(Q)$. Then consider the following (from Gao (2014)):

$$y_i = \alpha \chi_{F_i} + \beta x^* + \sum_{Q \in \mathcal{C}_\tau, |Q \cap T_i| \text{ odd}} (1 - 2\alpha - \beta x^*(\delta(Q)))\, \chi_{e_Q}$$

for $\alpha, \beta, \tau \geq 0$ such that

$$\alpha + 2\beta = 1 \qquad \text{and} \qquad \tau = \frac{1 - 2\alpha}{\beta} - 1.$$

**Theorem**

*$y_i$ is feasible for the $T_i$-join LP.*

# Proof

# Two Lemmas

Recall $x^* = \sum_{i=1}^{k} \lambda_i \chi_{F_i}$, with $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_i \geq 0$. So $\lambda_i$ is a probability distribution on the trees $F_i$; probability of $F_i$ is $\lambda_i$.

**Lemma**

*Let $\mathcal{F}$ be a randomly sampled tree $F_i$, and $\mathcal{T}$ the corresponding vertices $T_i$. Let $Q \in \mathcal{C}_\tau$ be a $\tau$-narrow cut. Then*

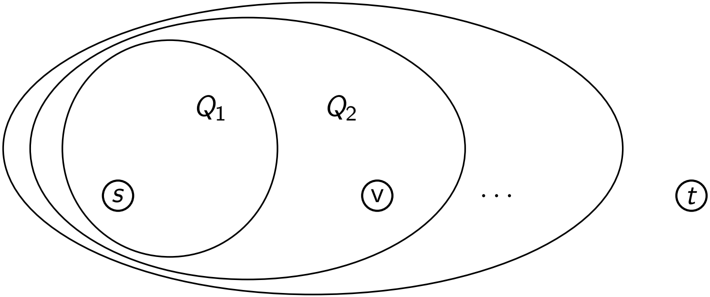
$$\Pr[|\delta(Q) \cap \mathcal{F}| = 1] \geq 2 - x^*(\delta(Q))$$
$$\Pr[|Q \cap \mathcal{T}| \ \mathit{odd}] \leq x^*(\delta(Q)) - 1.$$

# Two Lemmas

Recall $e_Q$ is the cheapest edge crossing a $\tau$-narrow cut $Q \in \mathcal{C}_\tau$.

**Lemma**

$$\sum_{Q \in \mathcal{C}_\tau} c_{e_Q} \leq \sum_{e \in E} c_e x_e^*.$$



# An-Kleinberg-Shmoys

**Theorem (An, Kleinberg, and Shmoys (2012))**

*Best-of-Many Christofides' is a* $\frac{1+\sqrt{5}}{2}$*-approximation algorithm for s-t path TSP.*

# Proof of AKS

For the proof, recall that $e_Q$ is min-cost edge in $\delta(Q)$, $\mathcal{C}_\tau$ are the cuts $Q$ with $x^*(\delta(Q)) < 1 + \tau$,

$$y_i = \alpha\chi_{F_i} + \beta x^* + \sum_{Q \in \mathcal{C}_\tau, |Q \cap T_i| \text{ odd}} (1 - 2\alpha - \beta x^*(\delta(Q)))\, \chi_{e_Q}$$

is feasible for the $T_i$-join LP, and

**Lemma**

*Let $\mathcal{F}$ be a randomly sampled tree $F_i$, and $\mathcal{T}$ the corresponding vertices $T_i$. Let $Q \in \mathcal{C}_\tau$ be a $\tau$-narrow cut. Then*

$$\Pr[|\delta(Q) \cap \mathcal{F}| = 1] \geq 2 - x^*(\delta(Q))$$
$$\Pr[|Q \cap \mathcal{T}| \text{ odd}] \leq x^*(\delta(Q)) - 1.$$

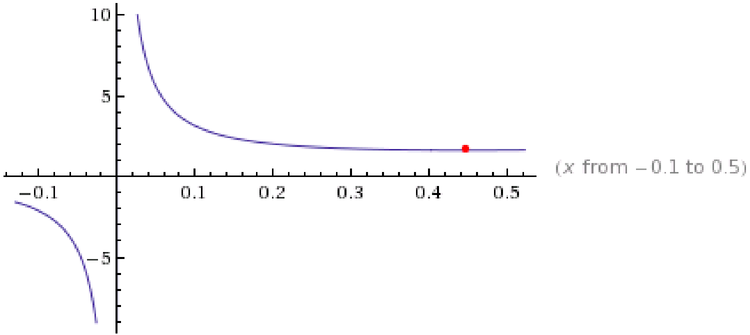
**Lemma**

$$\sum_{Q \in \mathcal{C}_\tau} c_{e_Q} \leq \sum_{e \in E} c_e x^*_e.$$

Global minimum:

$$\min\left\{2 - x + \frac{(3x-1)^2}{4x} \;\middle|\; x \geq 0\right\} = \frac{1}{2}\left(1+\sqrt{5}\right) \text{ at } x = \frac{1}{\sqrt{5}}$$

Plot:



# Next time

A $\frac{3}{2}$-approximation algorithm for $s$-$t$ TSP path in graph TSP instances.