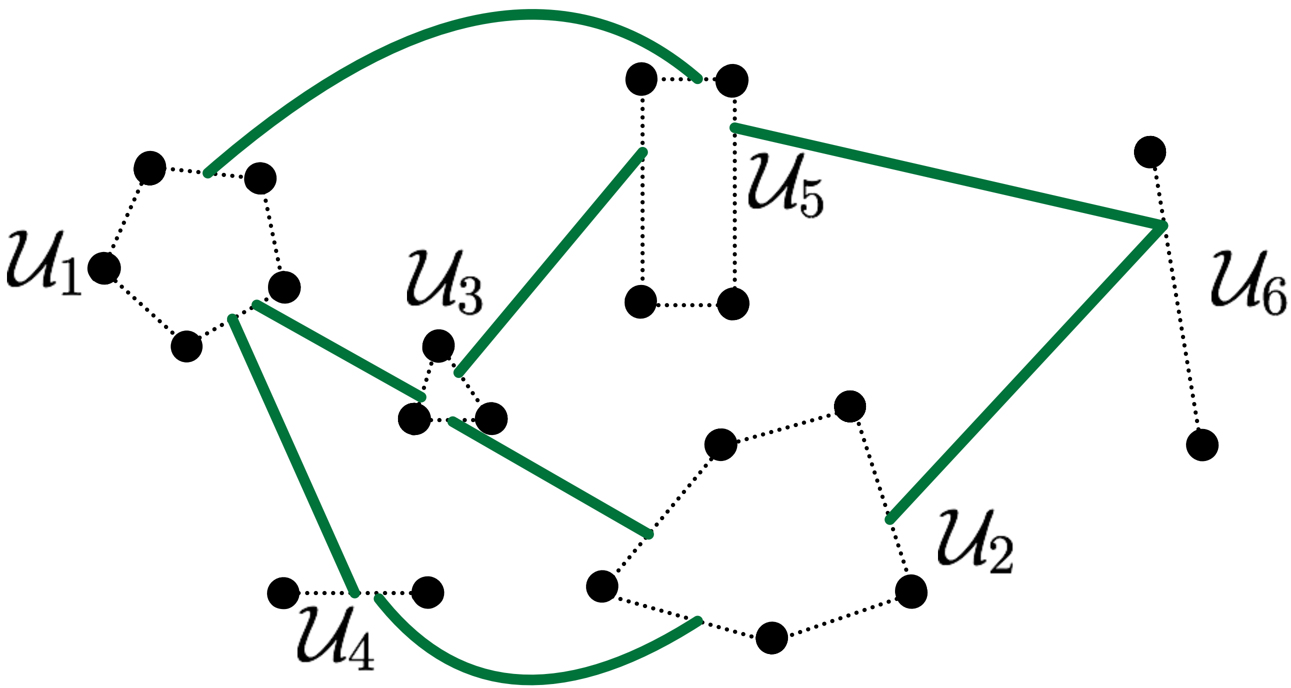
Optimization problems in graphs with locational uncertainty

Marin Bougeret, Jérémy Omer, and Michaël Poss

October 22, 2021

Problem definition

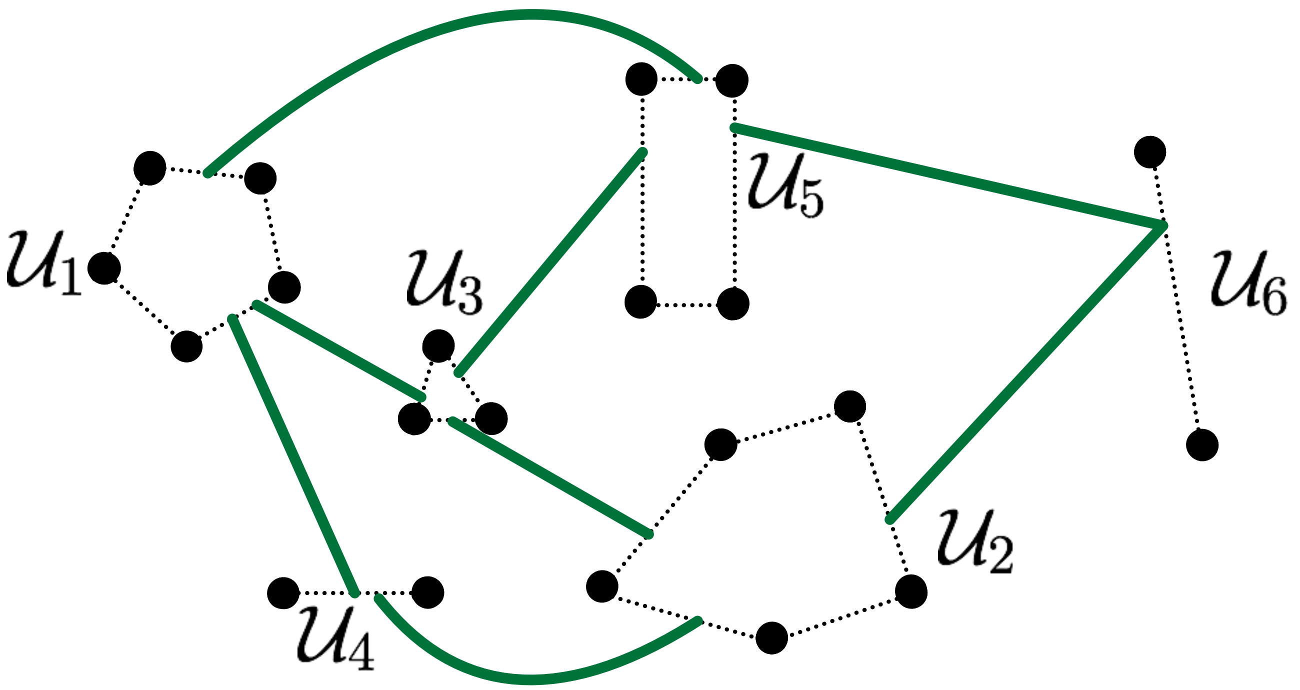
Given:
- $\mathcal{U}_i \subset \mathbb{R}^2$ for $i = 1, \ldots, 6$
- G=(V, E)



Problem definition

Given:
- $\mathcal{U}_i \subset \mathbb{R}^2$ for $i = 1, \ldots, 6$
- G=(V, E)

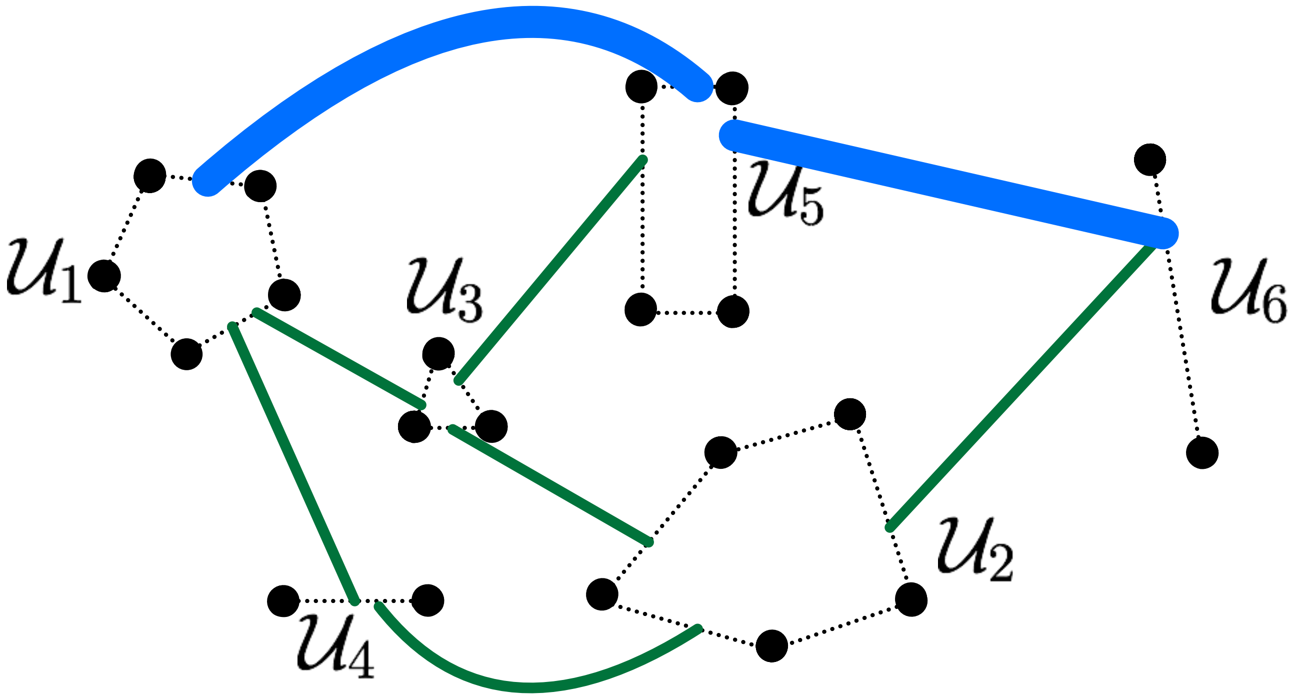
Find: path from $\mathcal{U}_1$ to $\mathcal{U}_6$



Problem definition

Given:
- $\mathcal{U}_i \subset \mathbb{R}^2$ for $i = 1, \ldots, 6$
- G=(V, E)

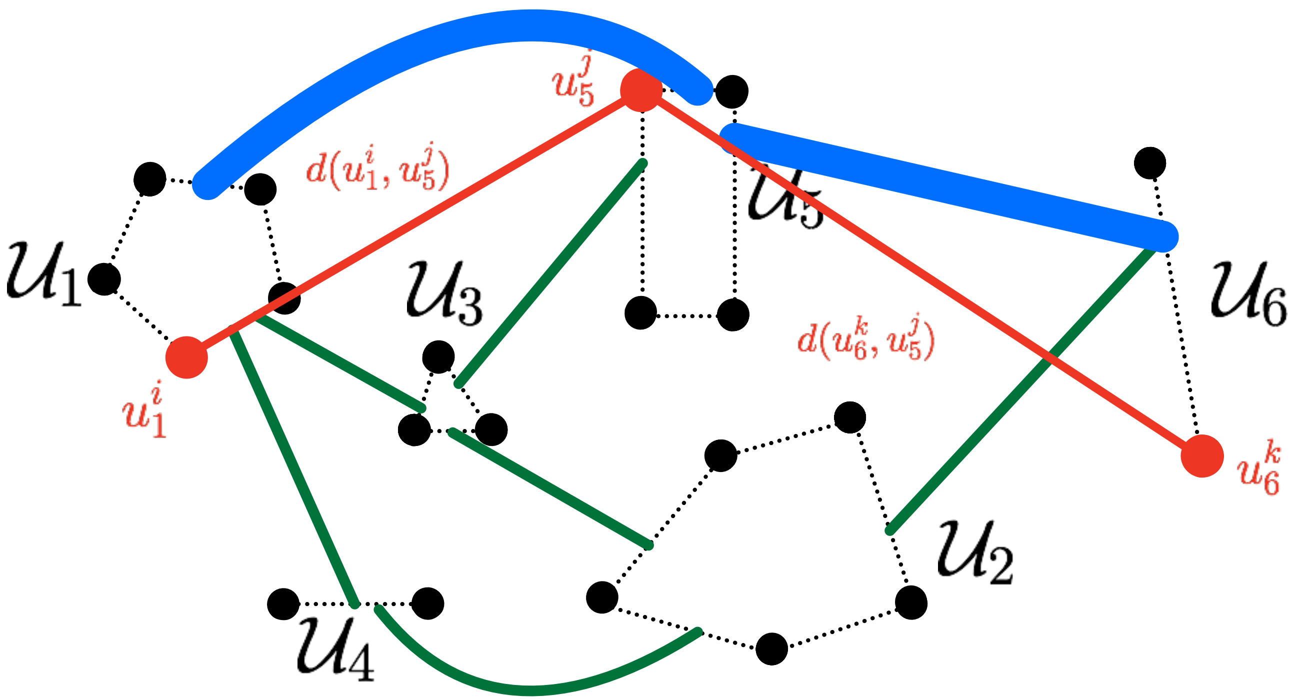
Find: path from $\mathcal{U}_1$ to $\mathcal{U}_6$



Problem definition

Given:
- $\mathcal{U}_i \subset \mathbb{R}^2$ for $i = 1, \ldots, 6$
- G=(V, E)

Find: path from $\mathcal{U}_1$ to $\mathcal{U}_6$



Problem definition

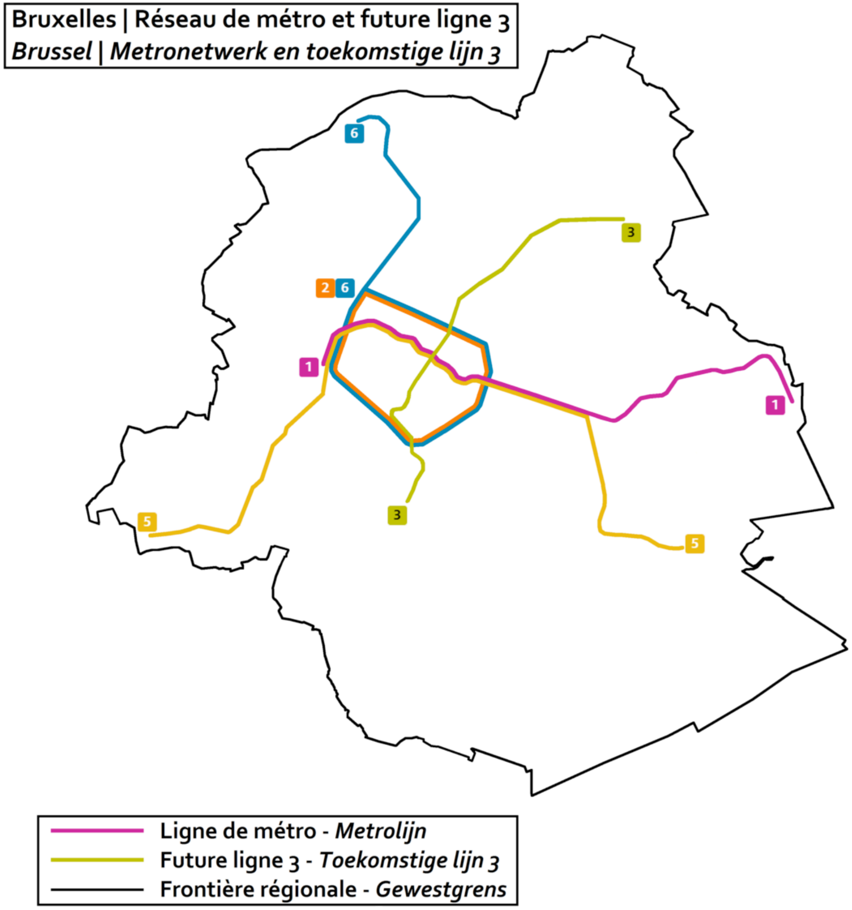
Given

- A graph $G = (V, E)$
- A family of admissible subgraphs $\mathcal{F}$
- Finite uncertainty set $\mathcal{U}_i$ for each $i \in V$
- Distances $d(\boldsymbol{u_i}, \boldsymbol{u_j})$ for each $\boldsymbol{u_i} \in \mathcal{U}_i, \boldsymbol{u_j} \in \mathcal{U}_j$

Solve

$$\min_{\boldsymbol{F} \in \mathcal{F}} \max_{\boldsymbol{u} \in \mathcal{U}} \sum_{\{i,j\} \in \boldsymbol{F}} d(\boldsymbol{u_i}, \boldsymbol{u_j}) \quad = \quad \min_{\boldsymbol{x} \in \mathcal{X}} \max_{\boldsymbol{u} \in \mathcal{U}} \sum_{\{i,j\} \in E} \boldsymbol{x}_{ij} d(\boldsymbol{u_i}, \boldsymbol{u_j})$$

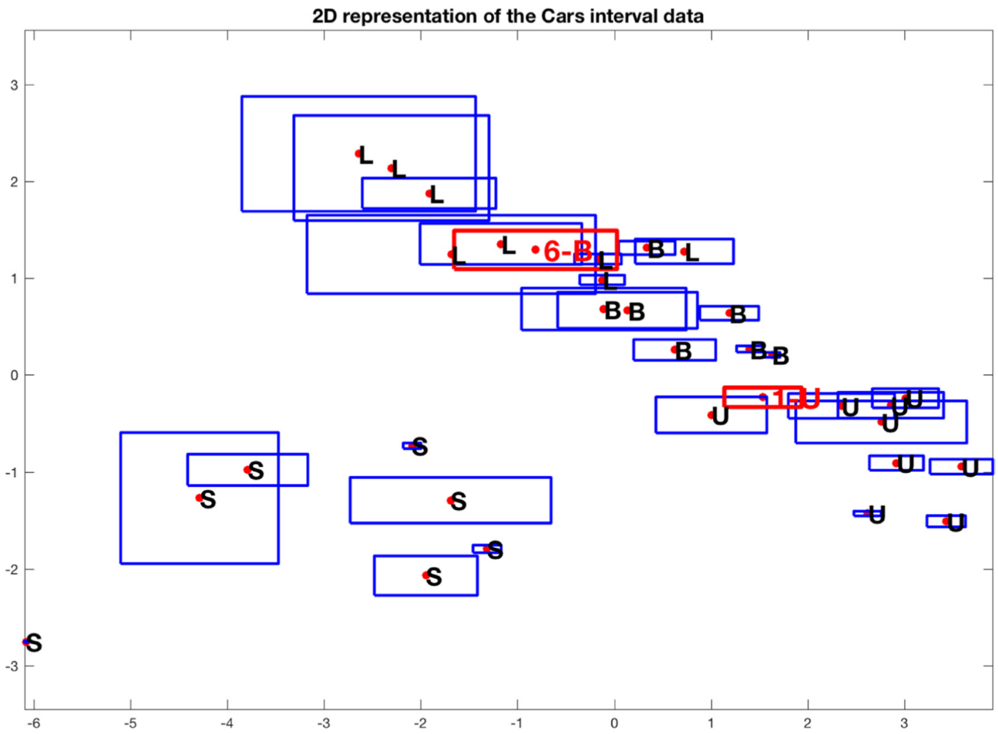
Motivation



Subway network expansion

- Exact locations of the stations hard to predict
- Construction cost proportional to lengths
- $\mathcal{F}$ consists of **paths** (possibly **trees**)

Motivation



Data classification

- Missing entries replaced by intervals
- Different distances may be relevant
- $\mathcal{F}$ consists of unions of **cliques**

Related work I: Affine decision rules [Zhen et al. (2021)]

They consider the special case

$$\min_{\boldsymbol{x}\in\mathcal{X}} \max_{\boldsymbol{u}\in\mathcal{U}} \sum_{\{i,j\}\in E} \boldsymbol{x}_{ij} \|\boldsymbol{u_i} - \boldsymbol{u_j}\|_2$$

1. Replace $\|\cdot\|_2$ by the maximization over ellipsoid $\mathcal{W}^\ell \subset \mathbb{R}^\ell$:

$$\begin{aligned} & \min_{\boldsymbol{x}\in\mathcal{X}} \max_{\boldsymbol{u}\in\mathcal{U}} \sum_{\{i,j\}\in E} \boldsymbol{x}_{ij} \max_{\boldsymbol{w}_{ij}\in\mathcal{W}^\ell} \boldsymbol{w}_{ij}^T(\boldsymbol{u_i} - \boldsymbol{u_j}) \\ = \quad & \min_{\boldsymbol{x}\in\mathcal{X}} \max_{\boldsymbol{w}\in\mathcal{W}} \max_{\boldsymbol{u}\in\mathcal{U}} \sum_{\{i,j\}\in E} \boldsymbol{x}_{ij}\boldsymbol{w}_{ij}^T(\boldsymbol{u_i} - \boldsymbol{u_j}) \\ \overset{\mathcal{U}=\{Au\leq c\}}{=} \quad & \min_{\boldsymbol{x}\in\mathcal{X}} \max_{\boldsymbol{w}\in\mathcal{W}} \min \; c^T\boldsymbol{\lambda} \\ & \qquad \text{s.t. } A_i^T\boldsymbol{\lambda} = \boldsymbol{x}_{ij}\boldsymbol{w}_{ij} - \boldsymbol{x}_{ji}\boldsymbol{w}_{ji}, \; i \in V \\ & \qquad \boldsymbol{\lambda} \geq 0 \end{aligned}$$

2. Use Affine Decision Rules and dualize the robust constraints

Related work II: Maximum pairwise distances [Citovsky et al. (2017)]

$$d_{ij}^{max} = \max_{\boldsymbol{u_i} \in U_i, \boldsymbol{u_j} \in U_j} d(\boldsymbol{u_i}, \boldsymbol{u_j})$$

1. Solve the problem $\min_{\boldsymbol{F} \in \mathcal{F}} \sum_{\{i,j\} \in \boldsymbol{F}} d_{ij}^{max}$
2. Return $\boldsymbol{F}^*$

They prove that for the traveling salesman problem, the returned solution cost is not worse than twice the optimal solution cost.

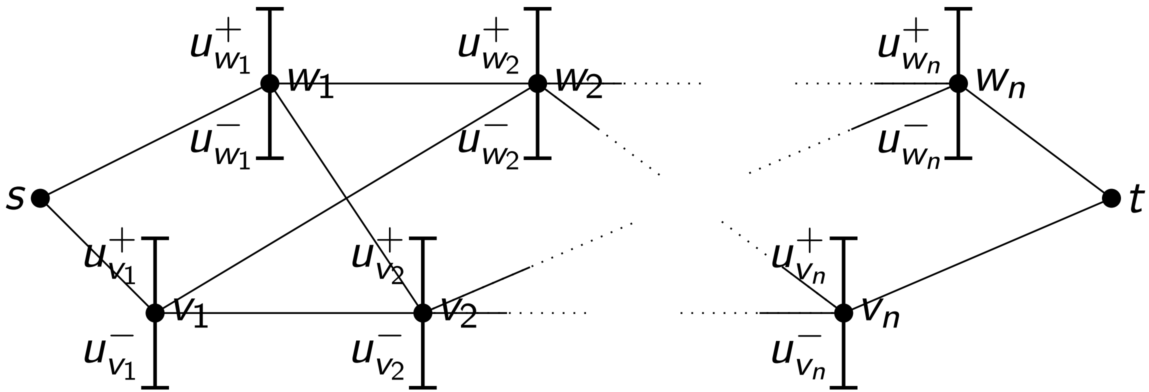
Our contributions

1. Prove the NP-hardness of $\min_{\boldsymbol{F} \in \mathcal{F}} \max_{\boldsymbol{u} \in \mathcal{U}} \sum_{\{i,j\} \in \boldsymbol{F}} d(\boldsymbol{u_i}, \boldsymbol{u_j})$.
2. Provide exact solution algorithms.
3. Study extensions of the approximation guarantee of d^{max} for more general structures and distances.
4. Provide a FPTAS for the shortest path problem.

NP-hardness

Theorem

ROBUST-Δ-SP *is $\mathcal{NP}$-hard, even when $(\mathcal{M}, d)$ is the 1-dimensional Euclidean space.*

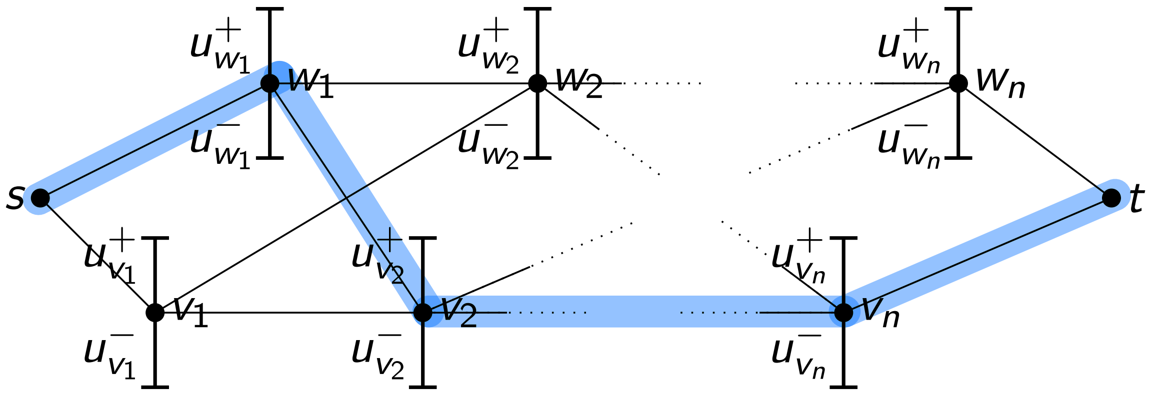


- Reduction from the partition problem for items $a_1, a_2, \ldots, a_n$
- All intervals built around 0
- Intervals length is larger than K >> a_i

NP-hardness

Theorem

ROBUST-Δ-SP *is $\mathcal{NP}$-hard, even when $(\mathcal{M}, d)$ is the 1-dimensional Euclidean space.*

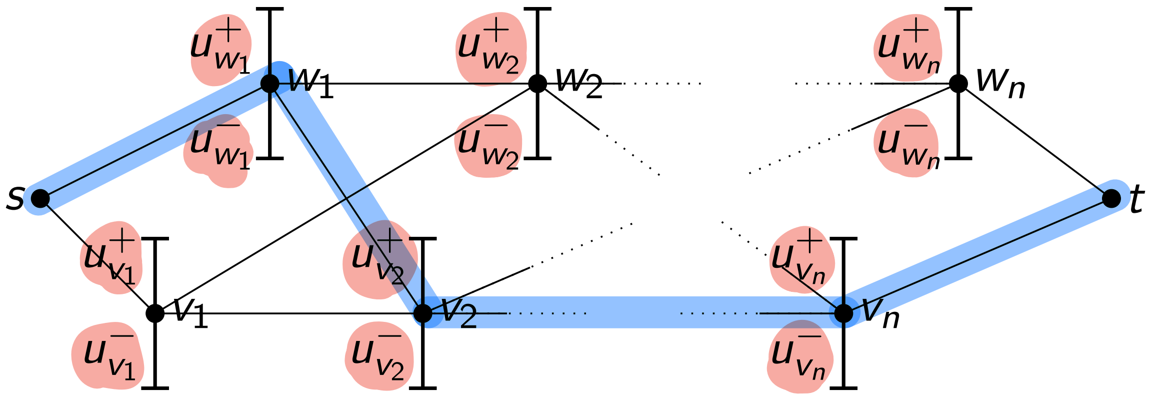


- Reduction from the partition problem for items $a_1, a_2, \ldots, a_n$
- All intervals built around 0
- Intervals length is larger than K >> a_i

NP-hardness

Theorem

ROBUST-Δ-SP is $\mathcal{NP}$-hard, even when $(\mathcal{M}, d)$ is the 1-dimensional Euclidean space.

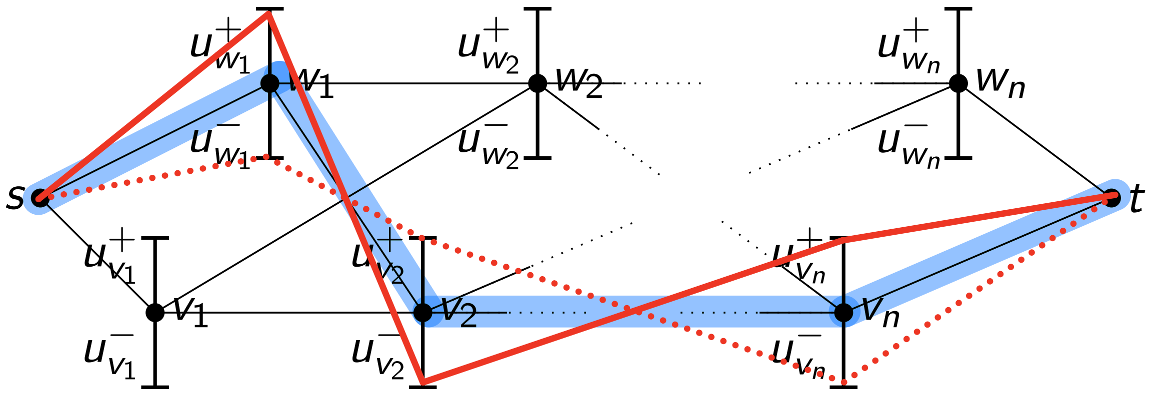


- Reduction from the partition problem for items $a_1, a_2, \ldots, a_n$
- All intervals built around 0
- Intervals length is larger than K >> a_i

NP-hardness

Theorem

ROBUST-Δ-SP *is $\mathcal{NP}$-hard, even when $(\mathcal{M}, d)$ is the 1-dimensional Euclidean space.*



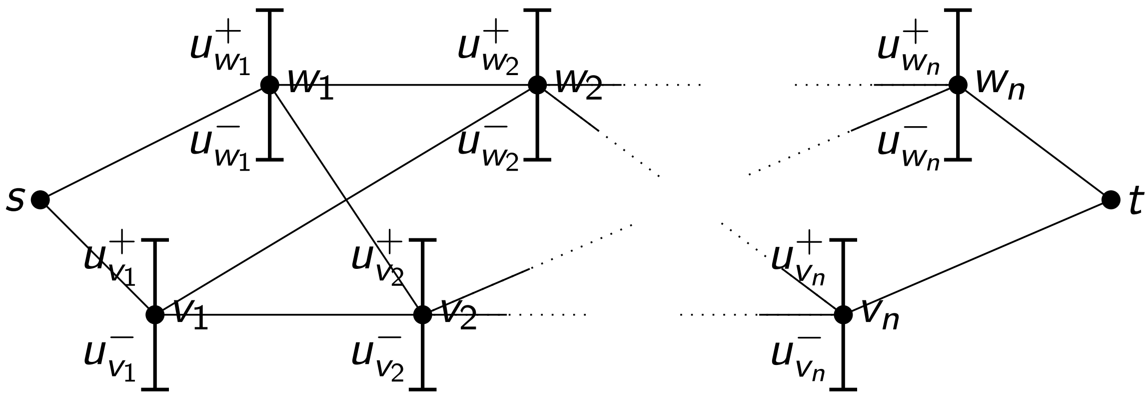
$\Rightarrow$ equivalent to the robust shortest path with 2 scenarios

- Reduction from the partition problem for items $a_1, a_2, \ldots, a_n$
- All intervals built around 0
- Intervals length is larger than K >> a_i

NP-hardness

Theorem

ROBUST-Δ-SP *is $\mathcal{NP}$-hard, even when $(\mathcal{M}, d)$ is the 1-dimensional Euclidean space.*



Similar result for the spanning tree problem.

Exact solution

Relaxed Master Problem

Let $\widetilde{\mathcal{U}} \subseteq \mathcal{U}$.

$$\begin{aligned} \min \quad & \boldsymbol{\omega} \\ \text{s.t.} \quad & \boldsymbol{\omega} \geq \sum_{\{i,j\} \in E} \boldsymbol{x}_{ij} d(\boldsymbol{u_i}, \boldsymbol{u_j}), \quad \forall \boldsymbol{u} \in \widetilde{\mathcal{U}} \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

This is a Mixed-Integer Linear Program.

Adversarial Separation Problem (EVAL-C)

$$\max_{\boldsymbol{u} \in \mathcal{U}} \sum_{\{i,j\} \in E} \boldsymbol{x}_{ij} d(\boldsymbol{u_i}, \boldsymbol{u_j}) \quad = \quad \max_{\boldsymbol{u} \in \mathcal{U}} \sum_{\{i,j\} \in \boldsymbol{F}} d(\boldsymbol{u_i}, \boldsymbol{u_j})$$

Adversarial separation problem: complexity results

Theorem

Even when $|\mathcal{M}| = 2$*, there is no* $\mathcal{PTAS}$ *for* EVAL-C *unless* $\mathcal{P} = \mathcal{NP}$*.*

Theorem

EVAL-C$/tw + \sigma$ *is* $\mathcal{FPT}$*.*

Adversarial separation problem: ILP

We want to solve

$$\max_{\boldsymbol{u} \in \mathcal{U}} \sum_{\{i,j\} \in E} \boldsymbol{x}_{ij} d(\boldsymbol{u_i}, \boldsymbol{u_j})$$

We use binary variables

$$\boldsymbol{y_i^k} = 1 \Leftrightarrow \boldsymbol{u_i} = u_i^k$$

We obtain a quadratic assignment problem

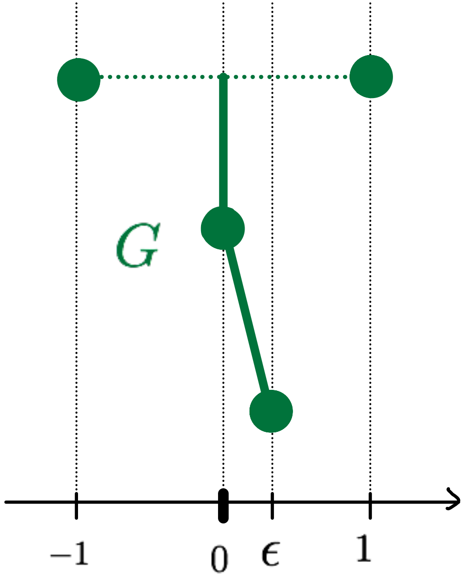
$$\max \sum_{\{i,j\} \in E} \boldsymbol{x}_{ij} \sum_{k=1}^{|\mathcal{U}_i|} \sum_{\ell=1}^{|\mathcal{U}_j|} d(u_i^k, u_j^\ell) \boldsymbol{y_i^k} \boldsymbol{y_j^\ell}$$

$$\text{s.t.} \sum_{k=1}^{|\mathcal{U}_i|} \boldsymbol{y_i^k} = 1, \quad \forall i \in V$$

$$\boldsymbol{y_i} \in \{0,1\}^{|\mathcal{U}_i|}, \quad \forall i \in V$$

The barycenter does not work

$\mathcal{F} = E(G) \implies$ the problem amounts to choosing one edge



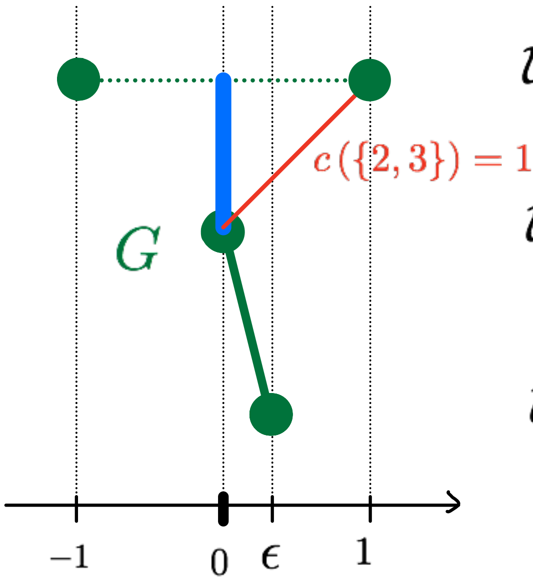
$$\mathcal{U}_3 = \{-1, 1\} \implies u_3^{bc} = 0$$

$$\mathcal{U}_2 = \{0\} \implies u_2^{bc} = 0$$

$$\mathcal{U}_1 = \{\epsilon\} \implies u_1^{bc} = \epsilon$$

The barycenter does not work

$\mathcal{F} = E(G) \implies$ the problem amounts to choosing one edge



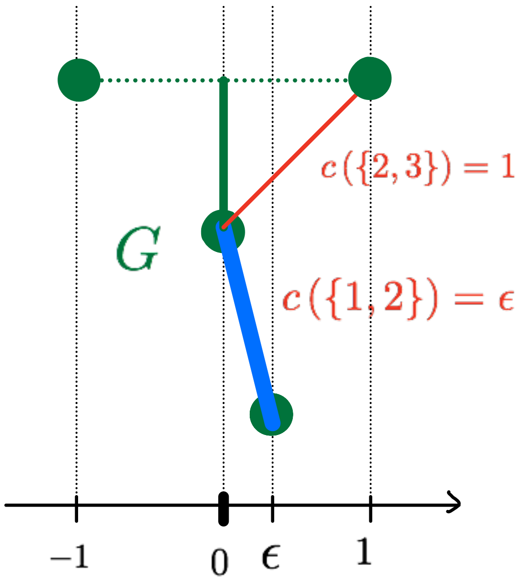
$\mathcal{U}_3 = \{-1, 1\} \implies u_3^{bc} = 0$

$\mathcal{U}_2 = \{0\} \implies u_2^{bc} = 0$

$\mathcal{U}_1 = \{\epsilon\} \implies u_1^{bc} = \epsilon$

The barycenter does not work

$$\mathcal{F} = E(G) \implies \text{the problem amounts to choosing one edge}$$



$$\mathcal{U}_3 = \{-1, 1\} \implies u_3^{bc} = 0$$

$$\mathcal{U}_2 = \{0\} \implies u_2^{bc} = 0$$

$$\mathcal{U}_1 = \{\epsilon\} \implies u_1^{bc} = \epsilon$$

Choosing the cheapest solution according to barycenters can be arbitrarily bad

Approximation algorithms based on d^{max}

Notations

$$c(\boldsymbol{F}) = \max_{\boldsymbol{u} \in \mathcal{U}} \sum_{\{i,j\} \in \boldsymbol{F}} d(\boldsymbol{u_i}, \boldsymbol{u_j}), \qquad c^{max}(\boldsymbol{F}) = \sum_{\{i,j\} \in \boldsymbol{F}} \overbrace{\max_{\boldsymbol{u_i} \in U_i, \boldsymbol{u_j} \in U_j} d(\boldsymbol{u_i}, \boldsymbol{u_j})}^{d_{ij}^{max}}$$

Algorithm

- Let $\boldsymbol{F^{max}}$ be a solution to $\min_{\boldsymbol{F} \in \mathcal{F}} c^{max}(\boldsymbol{F})$
- return $\boldsymbol{F^{max}}$

Theorem

$c^{max}(\boldsymbol{F}) \leq \rho \cdot c(\boldsymbol{F})$ *for any* $\boldsymbol{F} \in \mathcal{F} \implies c(\boldsymbol{F^{max}}) \leq \rho \cdot \mathrm{OPT}.$

Proof.

- $\boldsymbol{F^{max}}$ be the output of the algorithm
- $\boldsymbol{F^*}$ be an optimal solution ($c(\boldsymbol{F^*}) = \mathrm{OPT}$)
- $c(\boldsymbol{F^{max}}) \leq c^{max}(\boldsymbol{F^{max}}) \leq c^{max}(\boldsymbol{F^*}) \leq \rho \cdot c(\boldsymbol{F^*})$

Example of bound $c^{max}(F) \leq \rho \, c(F)$

Theorem

Let $(\mathcal{M}, d)$ be a Ptolemaic metric space. Then, $c^{max}(F) \leq 4 \, c(F)$.

think of it as Euclidean

Proof.

Example of bound $c^{max}(F) \leq \rho\ c(F)$

Theorem

Let $(\mathcal{M}, d)$ be a Ptolemaic metric space. Then, $c^{max}(F) \leq 4\, c(F)$.

Proof.

- Define $\tilde{\mathcal{U}} \subseteq \mathcal{U}$
- We have $\tilde{c}(F) = \max_{u \in \tilde{\mathcal{U}}} \sum_{\{i,j\} \in F} d(u_i, u_j) \geq \mathbb{E}\left[\sum_{\{i,j\} \in F} d(\tilde{u}_i, \tilde{u}_j)\right] = \sum_{\{i,j\} \in F} \mathbb{E}\left[d(\tilde{u}_i, \tilde{u}_j)\right] \geq \sum_{\{i,j\} \in F} \frac{d_{ij}^{max}}{4}$

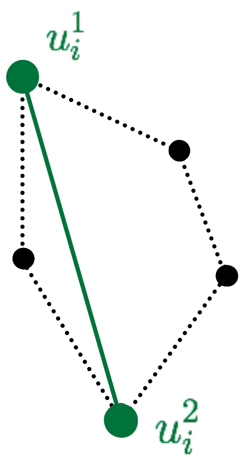
Example of bound $c^{max}(F) \leq \rho\, c(F)$

Theorem

Let $(\mathcal{M}, d)$ be a Ptolemaic metric space. Then, $c^{max}(F) \leq 4\, c(F)$.

Proof.

- Define $\tilde{\mathcal{U}} \subseteq \mathcal{U}$
- We have $\tilde{c}(F) = \max_{u \in \tilde{\mathcal{U}}} \sum_{\{i,j\} \in F} d(u_i, u_j) \geq \mathbb{E}\left[\sum_{\{i,j\} \in F} d(\tilde{u}_i, \tilde{u}_j)\right] = \sum_{\{i,j\} \in F} \mathbb{E}\left[d(\tilde{u}_i, \tilde{u}_j)\right] \geq \sum_{\{i,j\} \in F} \frac{d_{ij}^{max}}{4}$



$$\tilde{\mathcal{U}}_i = \{u_i^1, u_i^2\}$$

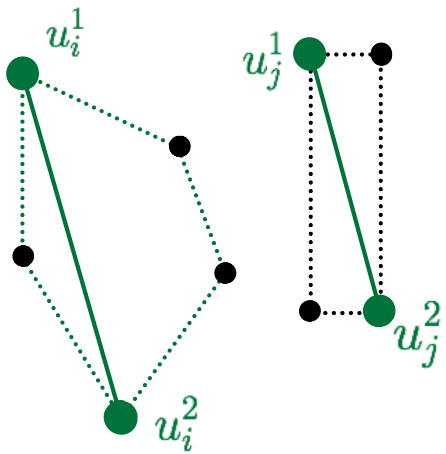
Example of bound $c^{max}(F) \leq \rho\; c(F)$

Theorem

Let $(\mathcal{M}, d)$ be a Ptolemaic metric space. Then, $c^{max}(F) \leq 4\, c(F)$.

Proof.

- Define $\tilde{\mathcal{U}} \subseteq \mathcal{U}$
- We have $\tilde{c}(F) = \max_{u \in \tilde{\mathcal{U}}} \sum_{\{i,j\} \in F} d(u_i, u_j) \geq \mathbb{E}\left[\sum_{\{i,j\} \in F} d(\tilde{u}_i, \tilde{u}_j)\right] = \sum_{\{i,j\} \in F} \mathbb{E}\left[d(\tilde{u}_i, \tilde{u}_j)\right] \geq \sum_{\{i,j\} \in F} \frac{d_{ij}^{max}}{4}$



$\tilde{\mathcal{U}}_i = \{u_i^1, u_i^2\}$

Example of bound $c^{max}(F) \leq \rho\, c(F)$

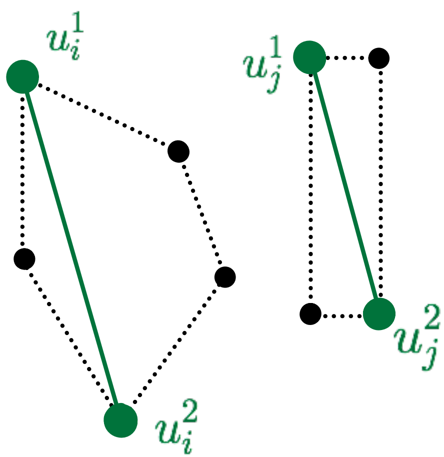
Theorem

Let $(\mathcal{M}, d)$ be a Ptolemaic metric space. Then, $c^{max}(F) \leq 4\, c(F)$.

Proof.

- Define $\tilde{\mathcal{U}} \subseteq \mathcal{U}$
- We have $\tilde{c}(F) = \max_{u \in \tilde{\mathcal{U}}} \sum_{\{i,j\} \in F} d(u_i, u_j) \geq \mathbb{E}\left[\sum_{\{i,j\} \in F} d(\tilde{u}_i, \tilde{u}_j)\right] = \sum_{\{i,j\} \in F} \mathbb{E}\left[d(\tilde{u}_i, \tilde{u}_j)\right] \geq \sum_{\{i,j\} \in F} \frac{d_{ij}^{max}}{4}$

$$\frac{1}{4}\left(d(u_i^1, u_j^1) + d(u_i^1, u_j^2) + d(u_i^2, u_j^1) + d(u_i^2, u_j^2)\right)$$



$$\tilde{\mathcal{U}}_i = \{u_i^1, u_i^2\}$$

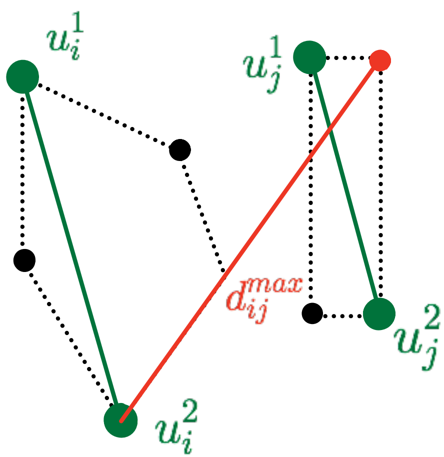
Example of bound $c^{max}(F) \leq \rho\ c(F)$

Theorem

Let $(\mathcal{M}, d)$ be a Ptolemaic metric space. Then, $c^{max}(F) \leq 4\, c(F)$.

Proof.

- Define $\tilde{\mathcal{U}} \subseteq \mathcal{U}$
- We have $\tilde{c}(F) = \max_{u \in \tilde{\mathcal{U}}} \sum_{\{i,j\} \in F} d(u_i, u_j) \geq \mathbb{E}\left[\sum_{\{i,j\} \in F} d(\tilde{u}_i, \tilde{u}_j)\right] = \sum_{\{i,j\} \in F} \mathbb{E}\left[d(\tilde{u}_i, \tilde{u}_j)\right] \geq \sum_{\{i,j\} \in F} \frac{d_{ij}^{max}}{4}$



$$\tilde{\mathcal{U}}_i = \{u_i^1, u_i^2\}$$

$$\mathbb{E}\left[d(\tilde{u}_i, \tilde{u}_j)\right] = \frac{1}{4}\left(d(u_i^1, u_j^1) + d(u_i^1, u_j^2) + d(u_i^2, u_j^1) + d(u_i^2, u_j^2)\right)$$

Applying Ptolemy's inequality ...

$$\geq \frac{d_{ij}^{max}}{4}$$

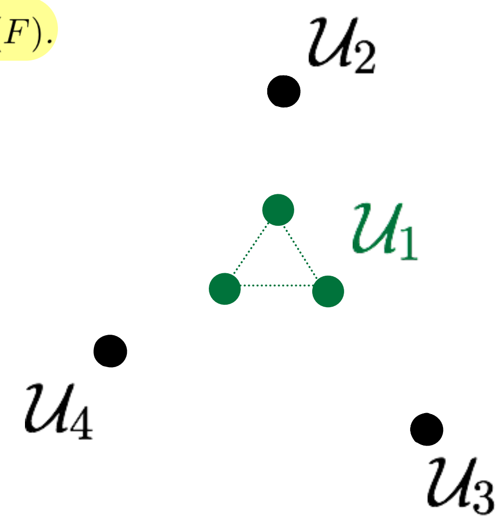
Example of bound $c^{max}(F) \leq \rho\ c(F)$

THEOREM 5. *For any metric space* $(\mathcal{M}, d)$, $c^{max}(F) \leq 9\, c(F)$.

PROPOSITION 10. *Let* F *be a clique. Then,* $c^{max}(F) \leq 2\, c(F)$.

COROLLARY 4. *Let* F *be a star. Then,* $c^{max}(F) \leq 3\, c(F)$.

PROPOSITION 13. *Let* F *be a star on* n *vertices. There is an uncertainty set* $\mathcal{U}$ *such that* $c^{max}(F) = \frac{3(n-1)}{n+1}\, c(F)$.



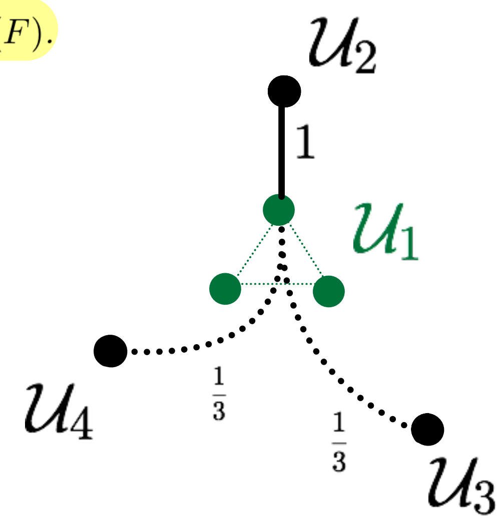
Example of bound $c^{max}(F) \leq \rho\ c(F)$

THEOREM 5. *For any metric space* $(\mathcal{M}, d)$, $c^{max}(F) \leq 9\, c(F)$.

PROPOSITION 10. *Let F be a clique. Then,* $c^{max}(F) \leq 2\, c(F)$.

COROLLARY 4. *Let F be a star. Then,* $c^{max}(F) \leq 3\, c(F)$.

PROPOSITION 13. *Let F be a star on n vertices. There is an uncertainty set $\mathcal{U}$ such that* $c^{max}(F) = \frac{3(n-1)}{n+1}\, c(F)$.



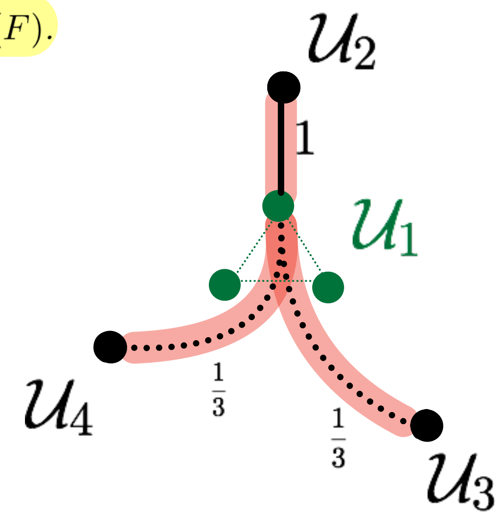
Example of bound $c^{max}(F) \leq \rho\ c(F)$

THEOREM 5. *For any metric space* $(\mathcal{M}, d)$, $c^{max}(F) \leq 9\, c(F)$.

PROPOSITION 10. *Let F be a clique. Then,* $c^{max}(F) \leq 2\, c(F)$.

COROLLARY 4. *Let F be a star. Then,* $c^{max}(F) \leq 3\, c(F)$.

PROPOSITION 13. *Let F be a star on n vertices. There is an uncertainty set $\mathcal{U}$ such that* $c^{max}(F) = \frac{3(n-1)}{n+1}\, c(F)$.



$$c(F) = 1 + \frac{2}{3} = \frac{5}{3}$$

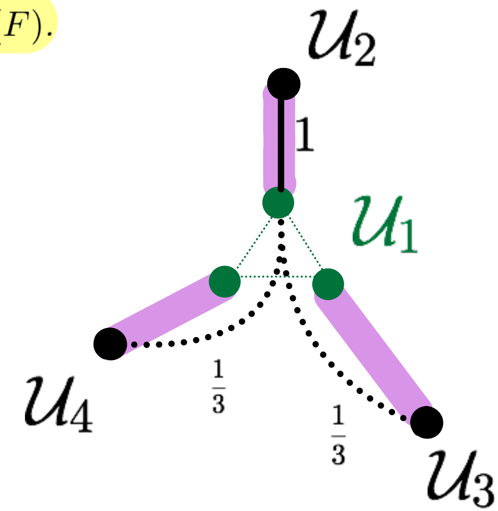
Example of bound $c^{max}(F) \leq \rho\ c(F)$

Theorem 5. *For any metric space* $(\mathcal{M}, d)$, $c^{max}(F) \leq 9\, c(F)$.

Proposition 10. *Let F be a clique. Then,* $c^{max}(F) \leq 2\, c(F)$.

Corollary 4. *Let F be a star. Then,* $c^{max}(F) \leq 3\, c(F)$.

Proposition 13. *Let F be a star on n vertices. There is an uncertainty set $\mathcal{U}$ such that* $c^{max}(F) = \frac{3(n-1)}{n+1}\, c(F)$.



$$c\,(F) = 1 + \frac{2}{3} = \frac{5}{3}$$

$$c^{max}\,(F) = 3$$

$$\Rightarrow \frac{c^{max}\,(F)}{c\,(F)} = \frac{9}{5}$$

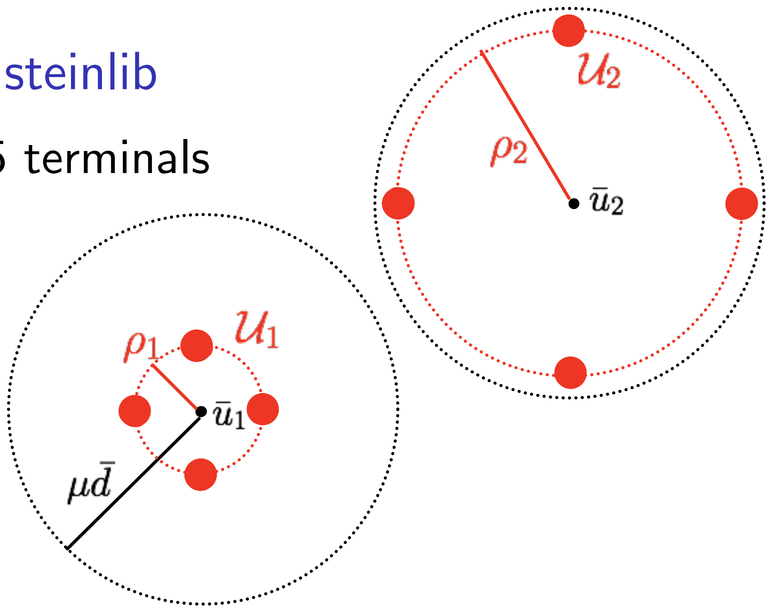
Numerical experiment on Steiner Tree Problem

Algorithms

- **exact**: Cutting-plane algorithm
- **adr**: Affine-decision-rule bases heuristic from Zhen et al. 2021
- **center**: Solving the nominal counterpart based on barycenters
- **dmax**: Solving the nominal counterpart using d^{max}

p619, p620, and 621 from steinlib

- 100 nodes, 180 edges, 5 terminals
- $\bar{d}$ = average distance
- random $\rho_i \in [0, \mu\bar{d}]$



Numerical experiments: solution times (in seconds)

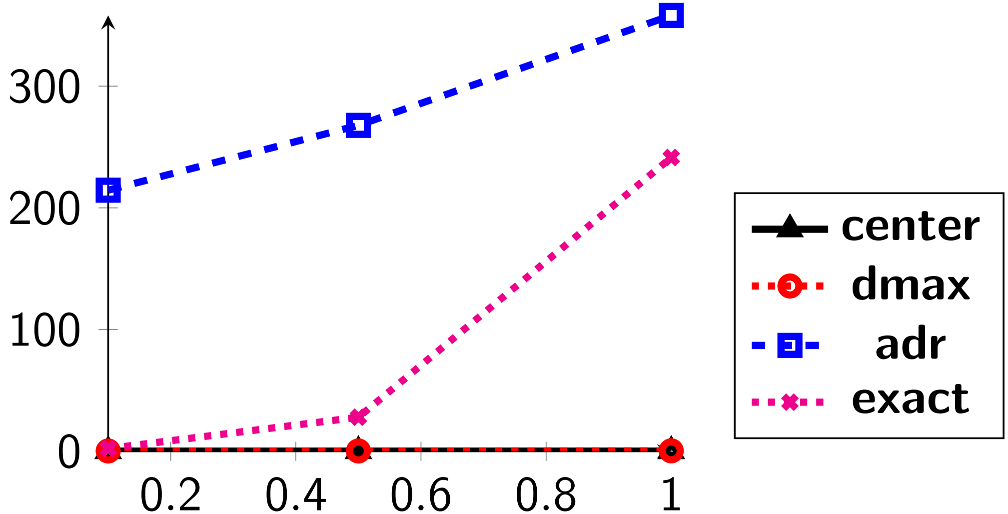
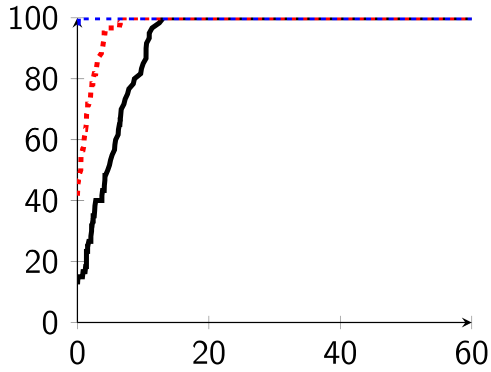
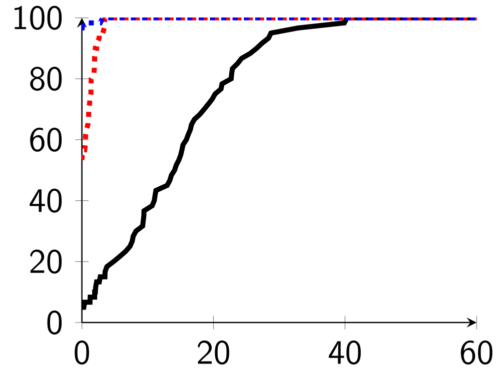


Figure: Value of μ.

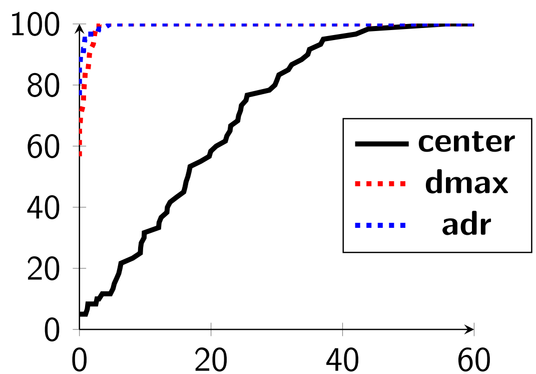
Numerical experiments: solution quality



(a) P6E, $\mu = 0.1$



(b) P6E, $\mu = 0.5$



(c) P6E, $\mu = 1$

Figure: % of instances for which the additional relative cost is less than x.

Open Journal of Mathematical Optimization (OJMO)

- With classical publishers, either
  - papers are behind an (expensive) paywall;
  - or authors pay ($\pm$2k) for Open Access (the so-called **gold** OA)
- OJMO provides a **free** OA alternative (thanks to Mersenne)
- Papers have doi, indexed in DBLP, zbMATH, Crossref, ...

Area Editors

- **Continuous Optimization** - David Russell Luke
- **Discrete Optimization** - Sebastian Pokutta
- **Optimization under Uncertainty** - Guzin Bayraksan
- **Computational aspects and applications** - Bernard Gendron

10 papers published already, and one prize!

2021 Beale — Orchard-Hays Prize Citation

Alberto Costa and Giacomo Nannicini
"RBFOpt: an open-source library for black-box optimization with costly function evaluations"
Mathematical Programming Computation 10 (2018) 597-629.
"On the implementation of a global optimization method for mixed-variable problems"
Open Journal of Mathematical Optimization 2 (2021).

Visit https://ojmo.centre-mersenne.org/