

Fault-Tolerant k -Supplier with Outliers

Deeparnab Chakrabarty, Luc Cote, **Ankita Sarkar**

(Remote) Talk at Purdue Theory Seminar
November 10, 2023

The Problem

Fault-tolerant k -Supplier with Outliers ($FkSO$)

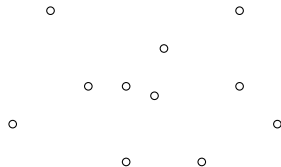
Background: k -Supplier

Input.

Background: k -Supplier

Input.

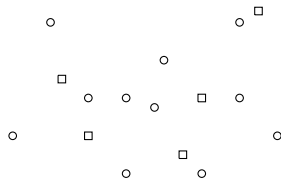
- Clients C , $|C| = n$



Background: k -Supplier

Input.

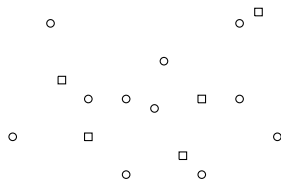
- Clients C , $|C| = n$
- Facilities F



Background: k -Supplier

Input.

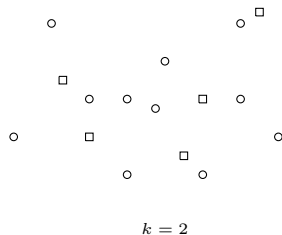
- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$



Background: k -Supplier

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k

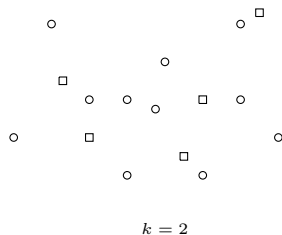


Background: k -Supplier

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k

Output.



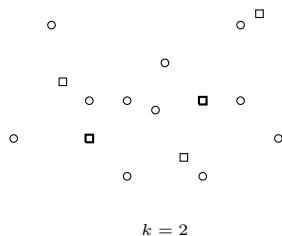
Background: k -Supplier

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k

Output.

- $S \subseteq F$, $|S| \leq k$



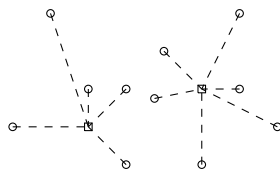
Background: k -Supplier

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k

Output.

- $S \subseteq F$, $|S| \leq k$
- minimize $\max_{v \in C} d(v, S)$



$k = 2$

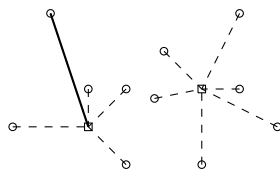
Background: k -Supplier

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k

Output.

- $S \subseteq F$, $|S| \leq k$
- minimize $\max_{v \in C} d(v, S)$



$k = 2$

Background: k -Supplier

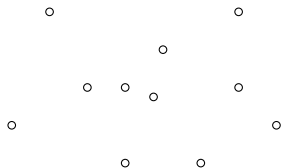
Tight 3-approximation [Hochbaum and Shmoys, 1986].

Background: k -Supplier

Algorithm [HS86]. Guess $r := \text{opt}$.

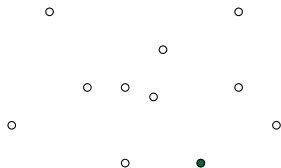
Background: k -Supplier

Algorithm [HS86]. Guess $r := \text{opt}$.



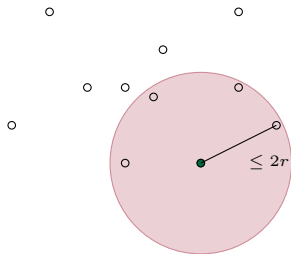
Background: k -Supplier

Algorithm [HS86]. Guess $r := \text{opt}$.



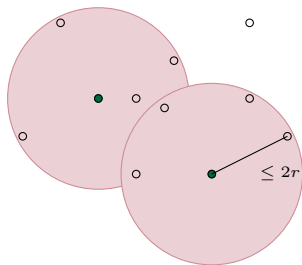
Background: k -Supplier

Algorithm [HS86]. Guess $r := \text{opt}$.



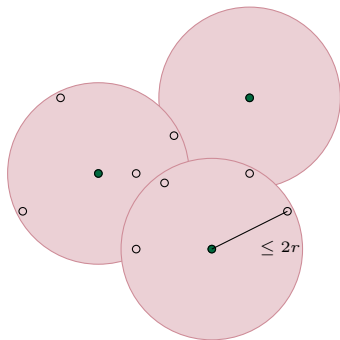
Background: k -Supplier

Algorithm [HS86]. Guess $r := \text{opt}$.



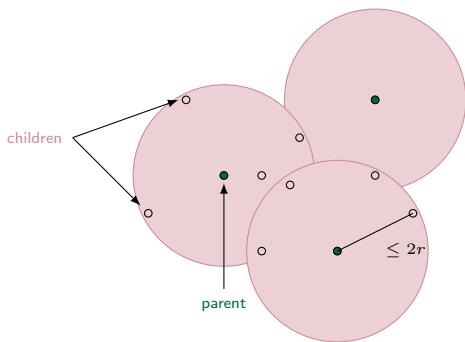
Background: k -Supplier

Algorithm [HS86]. Guess $r := \text{opt}$.



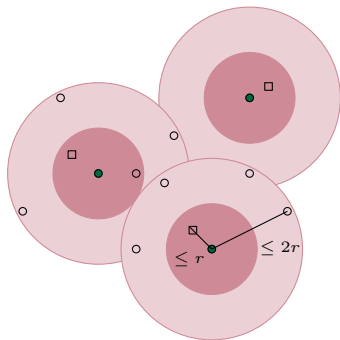
Background: k -Supplier

Algorithm [HS86]. Guess $r := \text{opt}$.



Background: k -Supplier

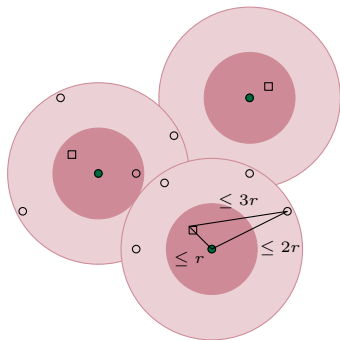
Algorithm [HS86]. Guess $r := \text{opt}$.



- Facility in each $B(j, r)$

Background: k -Supplier

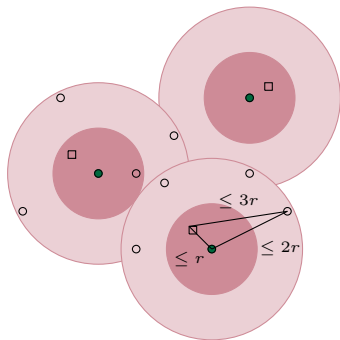
Algorithm [HS86]. Guess $r := \text{opt}$.



- Facility in each $B(j, r)$

Background: k -Supplier

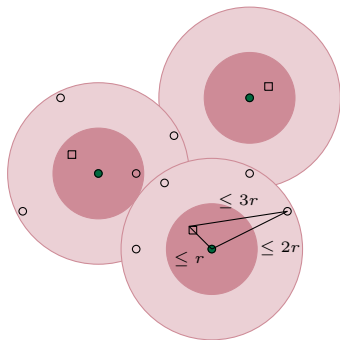
Algorithm [HS86]. Guess $r := \text{opt}$.



- Facility in each $B(j, r)$
- R is **well-separated**

Background: k -Supplier

Algorithm [HS86]. Guess $r := \text{opt}$.



- Facility in each $B(j, r)$
- R is **well-separated**
- $|R| \leq k$

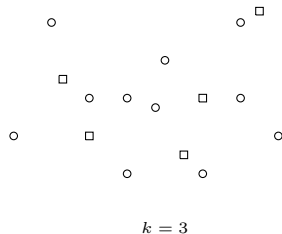
The Problem

Fault-tolerant k -Supplier with Outliers

Background: Fault-tolerant k -Supplier (FkS)

Input.

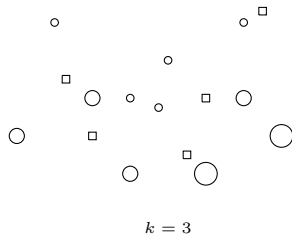
- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k



Background: Fault-tolerant k -Supplier (FkS)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Fault-tolerances $\{\ell_v\}_{v \in C}$



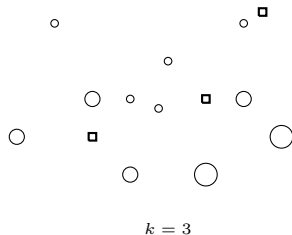
Background: Fault-tolerant k -Supplier (FkS)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Fault-tolerances $\{\ell_v\}_{v \in C}$

Output.

- $S \subseteq F$, $|S| \leq k$



Background: Fault-tolerant k -Supplier (FkS)

Input.

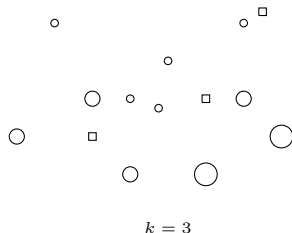
- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Fault-tolerances $\{\ell_v\}_{v \in C}$

Output.

- $S \subseteq F$, $|S| \leq k$
- minimize

$$\max_{v \in C} \underbrace{d_{\ell_v}(v, S)}$$

distance of v to ℓ_v^{th} open facility



Background: Fault-tolerant k -Supplier (FkS)

Input.

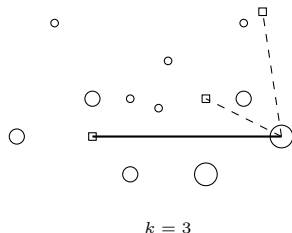
- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Fault-tolerances $\{\ell_v\}_{v \in C}$

Output.

- $S \subseteq F$, $|S| \leq k$
- minimize

$$\max_{v \in C} \underbrace{d_{\ell_v}(v, S)}$$

distance of v to ℓ_v^{th} open facility



Background: Fault-tolerant k -Supplier (FkS)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Fault-tolerances $\{\ell_v\}_{v \in C}$

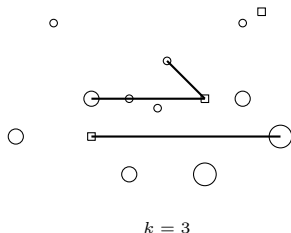
Output.

- $S \subseteq F$, $|S| \leq k$
- minimize

$$\max_{v \in C}$$

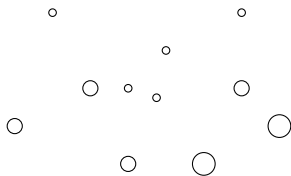
$$\underbrace{d_{\ell_v}(v, S)}$$

distance of v to ℓ_v^{th} open facility



Background: 3-approximation for FkS

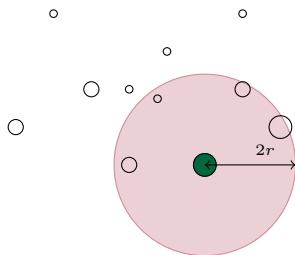
Algorithm. Carve balls in decreasing order of l_v 's.



- l_j facilities in each $B(j, r)$
- $\sum_{j \in R} l_j \leq k$
- Same 3-approximation as before
- **Takeaway: decreasing order of l_v 's.**

Background: 3-approximation for FkS

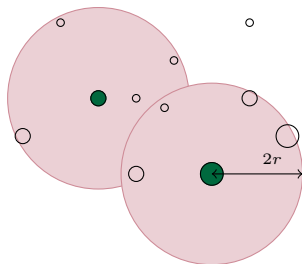
Algorithm. Carve balls in decreasing order of l_v 's.



- l_j facilities in each $B(j, r)$
- $\sum_{j \in R} l_j \leq k$
- Same 3-approximation as before
- **Takeaway: decreasing order of l_v 's.**

Background: 3-approximation for FkS

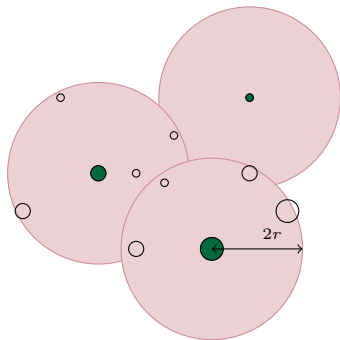
Algorithm. Carve balls in decreasing order of l_v 's.



- l_j facilities in each $B(j, r)$
- $\sum_{j \in R} l_j \leq k$
- Same 3-approximation as before
- **Takeaway: decreasing order of l_v 's.**

Background: 3-approximation for FkS

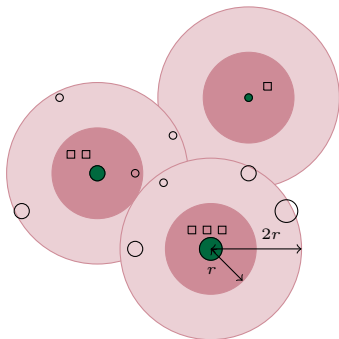
Algorithm. Carve balls in decreasing order of l_v 's.



- l_j facilities in each $B(j, r)$
- $\sum_{j \in R} l_j \leq k$
- Same 3-approximation as before
- **Takeaway: decreasing order of l_v 's.**

Background: 3-approximation for FkS

Algorithm. Carve balls in decreasing order of l_v 's.



- l_j facilities in each $B(j, r)$
- $\sum_{j \in R} l_j \leq k$
- Same 3-approximation as before
- **Takeaway: decreasing order of l_v 's.**

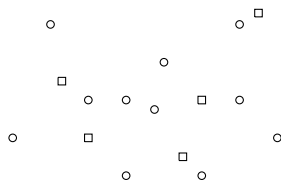
The Problem

Fault-tolerant k -Supplier with Outliers

Background: k -Supplier with Outliers (k SO)

Input.

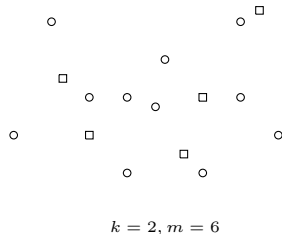
- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k



Background: k -Supplier with Outliers (k SO)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Parameter m



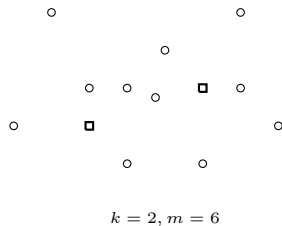
Background: k -Supplier with Outliers (k SO)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Parameter m

Output.

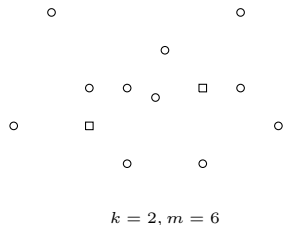
- $S \subseteq F$, $|S| \leq k$



Background: k -Supplier with Outliers (k SO)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Parameter m



Output.

- $S \subseteq F$, $|S| \leq k$
- $\underbrace{T}_{\text{inliers}} \subseteq C$, $|T| \geq m$,
minimize $\max_{v \in T} d(v, S)$

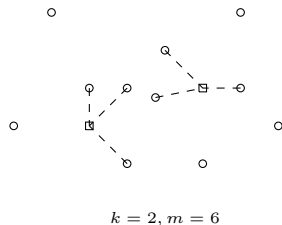
Background: k -Supplier with Outliers (k SO)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Parameter m

Output.

- $S \subseteq F$, $|S| \leq k$
- $\underbrace{T}_{\text{inliers}} \subseteq C$, $|T| \geq m$,
minimize $\max_{v \in T} d(v, S)$



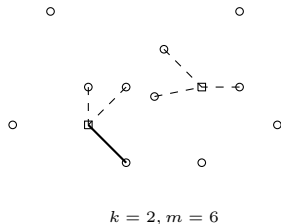
Background: k -Supplier with Outliers (k SO)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Parameter m

Output.

- $S \subseteq F$, $|S| \leq k$
- $\underbrace{T}_{\text{inliers}} \subseteq C$, $|T| \geq m$,
minimize $\max_{v \in T} d(v, S)$



Background: k SO

3-approximation [Chakrabarty, Goyal, and Krishnaswamy, 2016]

Background: k SO

Algorithm [CGK16].

Background: k SO

Algorithm [CGK16].

- Use LP to identify outliers

Background: k SO

Algorithm [CGK16].

- Use LP to identify outliers
- For each $v \in C$, $\text{cov}_v \in [0, 1]$: whether $v \in T$

Background: k SO

Algorithm [CGK16].

- Use LP to identify outliers
- For each $v \in C$, $\text{cov}_v \in [0, 1]$: whether $v \in T$
- For each $i \in F$, $x_i \in [0, 1]$: whether $i \in S$

Background: k SO

Algorithm [CGK16].

- Use LP to identify outliers
- For each $v \in C$, $\text{cov}_v \in [0, 1]$: whether $v \in T$
- For each $i \in F$, $x_i \in [0, 1]$: whether $i \in S$
- Constraints:
 - $\sum_{v \in C} \text{cov}_v \geq m$

Background: k SO

Algorithm [CGK16].

- Use LP to identify outliers
- For each $v \in C$, $\text{cov}_v \in [0, 1]$: whether $v \in T$
- For each $i \in F$, $x_i \in [0, 1]$: whether $i \in S$
- Constraints:
 - $\sum_{v \in C} \text{cov}_v \geq m$
 - $\sum_{i \in F} x_i \leq k$

Background: k SO

Algorithm [CGK16].

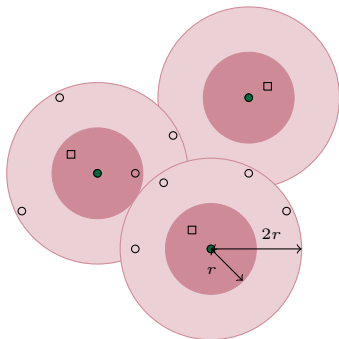
- Use LP to identify outliers
- For each $v \in C$, $\text{cov}_v \in [0, 1]$: whether $v \in T$
- For each $i \in F$, $x_i \in [0, 1]$: whether $i \in S$
- Constraints:
 - $\sum_{v \in C} \text{cov}_v \geq m$
 - $\sum_{i \in F} x_i \leq k$
 - $\forall v \in C, \sum_{i \in B(v,r) \cap F} x_i \geq \text{cov}_v$

Background: k SO

- Carve balls in decreasing order of cov_v 's.

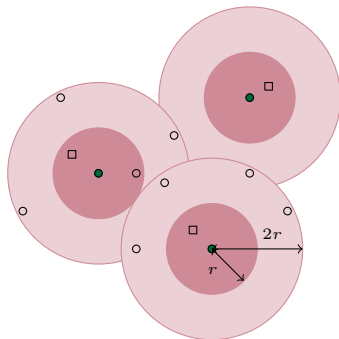
Background: k SO

- Carve balls in decreasing order of cov_v 's.



Background: k SO

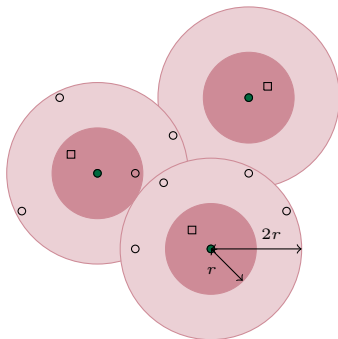
- Carve balls in decreasing order of cov_v 's.



- $\forall j \in R, \forall v \in \text{child}(j), \text{cov}_j \geq \text{cov}_v$

Background: k SO

- Carve balls in decreasing order of cov_v 's.



- $\forall j \in R, \forall v \in \text{child}(j), \text{cov}_j \geq \text{cov}_v$
- Possibly $|R| > k$

Background: k SO

- Enough inliers?

Background: k SO

- Enough inliers?
- Want: k largest $|\text{child}(j)|$'s sum to $\geq m$

Background: k SO

- Enough inliers?
- Want: k largest $|\text{child}(j)|$'s sum to $\geq m$
- $\sum_{j \in R} |\text{child}(j)| \text{cov}_j \geq \sum_{j \in R} \sum_{v \in \text{child}(j)} \text{cov}_v$

Background: k SO

- Enough inliers?
- Want: k largest $|\text{child}(j)|$'s sum to $\geq m$
- $$\sum_{j \in R} |\text{child}(j)| \text{cov}_j \geq \sum_{j \in R} \sum_{v \in \text{child}(j)} \text{cov}_v$$
$$= \sum_{v \in C} \text{cov}_v \geq m$$

Background: k SO

- Enough inliers?
- Want: k largest $|\text{child}(j)|$'s sum to $\geq m$
- $\sum_{j \in R} |\text{child}(j)| \text{cov}_j \geq \sum_{j \in R} \sum_{v \in \text{child}(j)} \text{cov}_v$
 $= \sum_{v \in C} \text{cov}_v \geq m$
- LP and above $\implies |T| \geq m$

Background: k SO

- Enough inliers?
- Want: k largest $|\text{child}(j)|$'s sum to $\geq m$
- $\sum_{j \in R} |\text{child}(j)| \text{cov}_j \geq \sum_{j \in R} \sum_{v \in \text{child}(j)} \text{cov}_v$
 $= \sum_{v \in C} \text{cov}_v \geq m$
- LP and above $\implies |T| \geq m$
- **Takeaway: decreasing order of cov_v 's.**

The Problem

Fault-tolerant k -Supplier with Outliers

Fault-tolerant k -Supplier with Outliers ($FkSO$)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k

Fault-tolerant k -Supplier with Outliers ($FkSO$)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Parameter m
- Fault-tolerances $\{\ell_v\}_{v \in C}$

Fault-tolerant k -Supplier with Outliers ($FkSO$)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Parameter m
- Fault-tolerances $\{\ell_v\}_{v \in C}$

Output.

- $S \subseteq F$, $|S| \leq k$

Fault-tolerant k -Supplier with Outliers ($FkSO$)

Input.

- Clients C , $|C| = n$
- Facilities F
- Metric d on $C \cup F$
- Parameter k
- Parameter m
- Fault-tolerances $\{\ell_v\}_{v \in C}$

Output.

- $S \subseteq F$, $|S| \leq k$
- $\underbrace{T}_{\text{inliers}} \subseteq C$, $|T| \geq m$, minimize $\max_{v \in T} d_{\ell_v}(v, S)$

Fk SO: Natural LP

- $\sum_{v \in C} \text{cov}_v \geq m, \sum_{i \in F} x_i \leq k$

Fk SO: Natural LP

- $\sum_{v \in C} \text{cov}_v \geq m, \sum_{i \in F} x_i \leq k$
- $\forall v \in C, \sum_{i \in B(v,r) \cap F} x_i \geq l_v \text{cov}_v$

F_k SO: Infinite integrality gap!

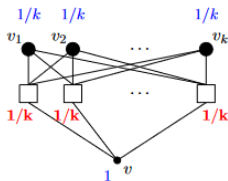


Figure 1: One of the k identical gadgets in the gap example, showing LP values in **red** (x values) and **blue** (cov values). The “edges” represent distance 1, and all other distances are determined by making triangle inequalities tight. The fault-tolerances are $\ell_{v_1} = \ell_{v_2} = \dots = \ell_{v_k} = k$, and $\ell_v = 1$.

$FkSO$: Stronger LP

$FkSO$: Stronger LP

- Exponential-sized LP [Chakrabarty and Negahbani, 2019]

$FkSO$: Stronger LP

- Exponential-sized LP [Chakrabarty and Negahbani, 2019]
- Find solutions via round-or-cut

$FkSO$: Stronger LP

- Exponential-sized LP [Chakrabarty and Negahbani, 2019]
- Find solutions via round-or-cut
- cov_v 's as before

FkSO: cov_v 's vs l_v 's

FkSO: cov_v 's vs l_v 's

- FkS: $l_j \geq l_v$

FkSO: cov_v 's vs ℓ_v 's

- FkS: $\ell_j \geq \ell_v$
- kSO: $\text{cov}_j \geq \text{cov}_v$

FkSO: cov_v 's vs l_v 's

- FkS: $l_j \geq l_v$
- kSO: $\text{cov}_j \geq \text{cov}_v$
- FkSO: need $\text{cov}_j \geq \text{cov}_v$ and $l_j \geq l_v$

FkSO: cov_v 's vs l_v 's

- FkS: $l_j \geq l_v$
- kSO: $\text{cov}_j \geq \text{cov}_v$
- FkSO: need $\text{cov}_j \geq \text{cov}_v$ and $l_j \geq l_v$
- Orders can be opposite!

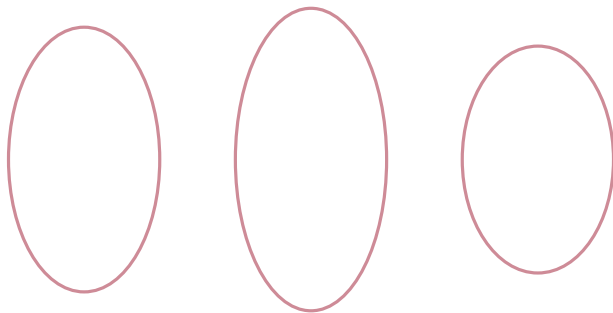
FkSO: cov_v 's vs l_v 's

- FkS: $l_j \geq l_v$
- kSO: $\text{cov}_j \geq \text{cov}_v$
- FkSO: need $\text{cov}_j \geq \text{cov}_v$ and $l_j \geq l_v$
- Orders can be opposite!
- Orders likely to be opposite

FkSO: cov_v 's vs l_v 's

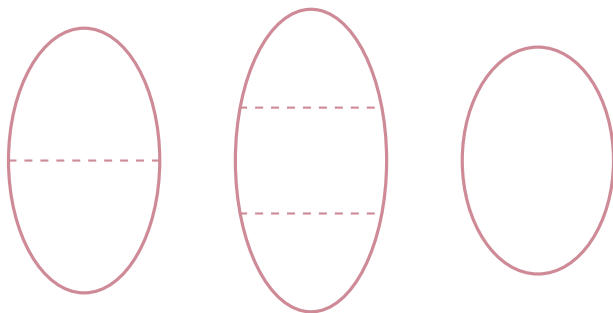
- FkS: $l_j \geq l_v$
- kSO: $\text{cov}_j \geq \text{cov}_v$
- FkSO: need $\text{cov}_j \geq \text{cov}_v$ and $l_j \geq l_v$
- Orders can be opposite!
- Orders likely to be opposite
- Gap example for weaker LP had opposite orders

Good partition



A partition \mathcal{P} is (ρ, cov) -good if:

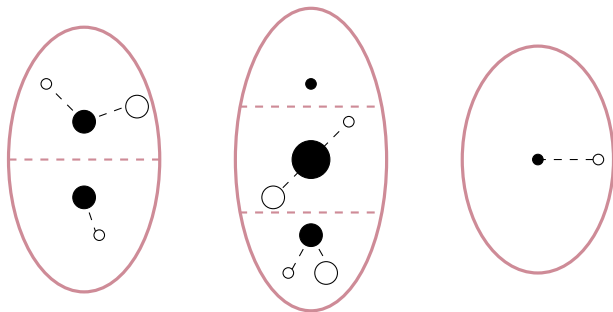
Good partition



A partition \mathcal{P} is (ρ, cov) -good if:

- 1 child(j)'s refine \mathcal{P}

Good partition

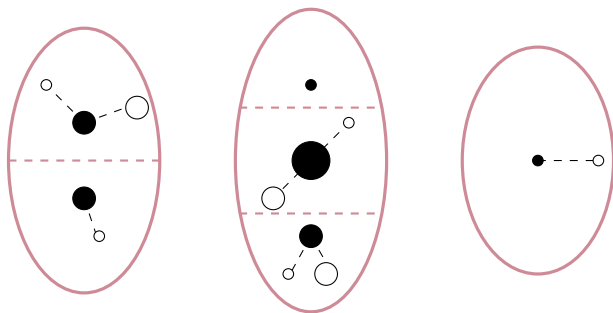


A partition \mathcal{P} is (ρ, cov) -good if:

① $\text{child}(j)$'s refine \mathcal{P}

$$\text{cov}_j \geq \text{cov}_v \text{ and } \ell_j \geq \ell_v$$

Good partition



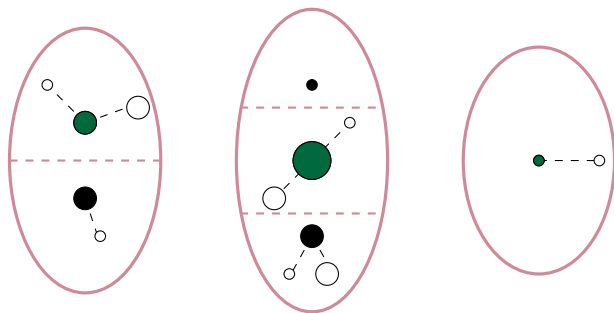
A partition \mathcal{P} is (ρ, cov) -good if:

- 1 child(j)'s refine \mathcal{P}

$$\text{cov}_j \geq \text{cov}_v \text{ and } \ell_j \geq \ell_v$$

- 2 Every $j, j' \in R$ in different P, P' are well-separated

Good partition



A partition \mathcal{P} is (ρ, cov) -good if:

- 1 child(j)'s refine \mathcal{P}

$$\text{cov}_j \geq \text{cov}_v \text{ and } \ell_j \geq \ell_v$$

- 2 Every $j, j' \in R$ in different P, P' are well-separated
- 3 Every $P \in \mathcal{P}$ has radius ρr around $j_P := \text{argmax}_{v \in P} \ell_v$

$(\rho + 1)$ -approximation

Theorem

Given cov_v 's from the LP, if there exists a (ρ, cov) -good partition, then there is a $(\rho + 1)$ -approximation.

$(\rho + 1)$ -approximation

Theorem

Given cov_v 's from the LP, if there exists a (ρ, cov) -good partition, then there is a $(\rho + 1)$ -approximation.

Our main result. $(4t - 2, \text{cov})$ -good partition

$(\rho + 1)$ -approximation

Theorem

Given cov_v 's from the LP, if there exists a (ρ, cov) -good partition, then there is a $(\rho + 1)$ -approximation.

Our main result. $(4t - 2, \text{cov})$ -good partition
where $t = \#$ distinct ℓ_v values.

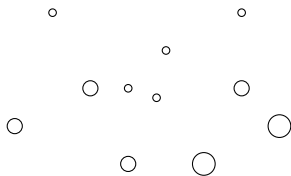
$(\rho + 1)$ -approximation

Theorem

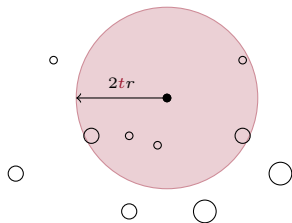
Given cov_v 's from the LP, if there exists a (ρ, cov) -good partition, then there is a $(\rho + 1)$ -approximation.

Our main result. $(4t - 2, \text{cov})$ -good partition
where $t = \#$ distinct ℓ_v values.
Hence $(4t - 1)$ -approximation.

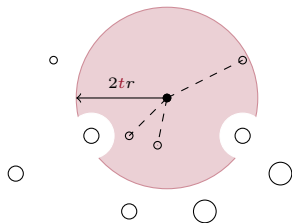
Constructing \mathcal{P} : New child-sets



Constructing \mathcal{P} : New child-sets

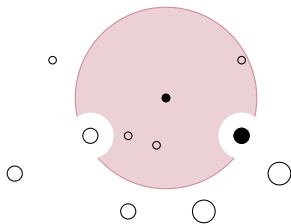


Constructing \mathcal{P} : New child-sets



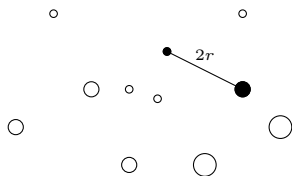
✓ $\text{cov}_j \geq \text{cov}_v, l_j \geq l_v$

Constructing \mathcal{P} : New child-sets



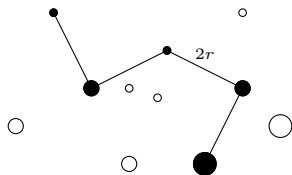
✓ $\text{cov}_j \geq \text{cov}_v, l_j \geq l_v$

Constructing \mathcal{P} : New child-sets



- ✓ $\text{cov}_j \geq \text{cov}_v, l_j \geq l_v$
 R is no longer well-separated!

Constructing \mathcal{P} : New child-sets



- ✓ $\text{cov}_j \geq \text{cov}_v, l_j \geq l_v$
 R is no longer well-separated!

Constructing \mathcal{P} : Coarsening

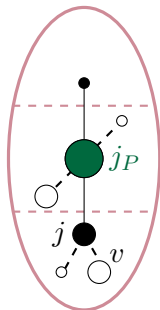
- **Claim.** Shortest-paths in G have $\leq t$ vertices

Constructing \mathcal{P} : Coarsening

- **Claim.** Shortest-paths in G have $\leq t$ vertices
- Diameter of connected components: $2(t - 1)r$

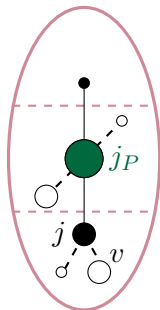
Constructing \mathcal{P} : Coarsening

- **Claim.** Shortest-paths in G have $\leq t$ vertices
- Diameter of connected components: $2(t-1)r$
- $P \in \mathcal{P}$: connected component \cup children



Constructing \mathcal{P} : Coarsening

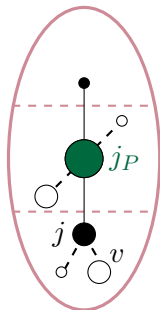
- **Claim.** Shortest-paths in G have $\leq t$ vertices
- Diameter of connected components: $2(t-1)r$
- $P \in \mathcal{P}$: connected component \cup children



- $j_P := \operatorname{argmax}_{v \in P} \ell_v$

Constructing \mathcal{P} : Coarsening

- **Claim.** Shortest-paths in G have $\leq t$ vertices
- Diameter of connected components: $2(t-1)r$
- $P \in \mathcal{P}$: connected component \cup children



- $j_P := \operatorname{argmax}_{v \in P} \ell_v$
- $d(v, j_P) \leq \underbrace{d(v, j)}_{2tr \because v \in \text{child}(j)} + \underbrace{d(j, j_P)}_{2(t-1)r \text{ by diam}} \leq \underbrace{(4t-2)r}_{=\rho}$

Which facilities to open?

Which facilities to open?

- Opening k_P facilities near j_P gives at least $\sum_{j \in R \cap P: \ell_j \leq k_P} |\text{child}(j)|$ inliers

Which facilities to open?

- Opening k_P facilities near j_P gives at least $\sum_{j \in R \cap P: \ell_j \leq k_P} |\text{child}(j)|$ inliers
- Determine k_P 's by DP

Which facilities to open?

- Opening k_P facilities near j_P gives at least $\sum_{j \in R \cap P: \ell_j \leq k_P} |\text{child}(j)|$ inliers
- Determine k_P 's by DP
- If $\text{opt}_{\text{DP}} \geq m$, good!

Which facilities to open?

- Opening k_P facilities near j_P gives at least $\sum_{j \in R \cap P: \ell_j \leq k_P} |\text{child}(j)|$ inliers
- Determine k_P 's by DP
- If $\text{opt}_{\text{DP}} \geq m$, good!
- Else: violated LP constraint.

Takeaways

Takeaways

- Tension between cov_v 's and l_v 's

Takeaways

- Tension between cov_v 's and ℓ_v 's
- (ρ, cov) -good partition yields $(\rho + 1)$ -approximation

Takeaways

- Tension between cov_v 's and ℓ_v 's
- (ρ, cov) -good partition yields $(\rho + 1)$ -approximation
- $\rho = 4t - 2$ possible

Takeaways

- Tension between cov_v 's and ℓ_v 's
- (ρ, cov) -good partition yields $(\rho + 1)$ -approximation
- $\rho = 4t - 2$ possible
- Limitation of good partition: $\rho = o(t)$ **not** possible.