Resilient Monotone Submodular Maximization

Vasileios Tzoumas
with
Konstantinos Gatsis, Ali Jadbabaie, George J. Pappas



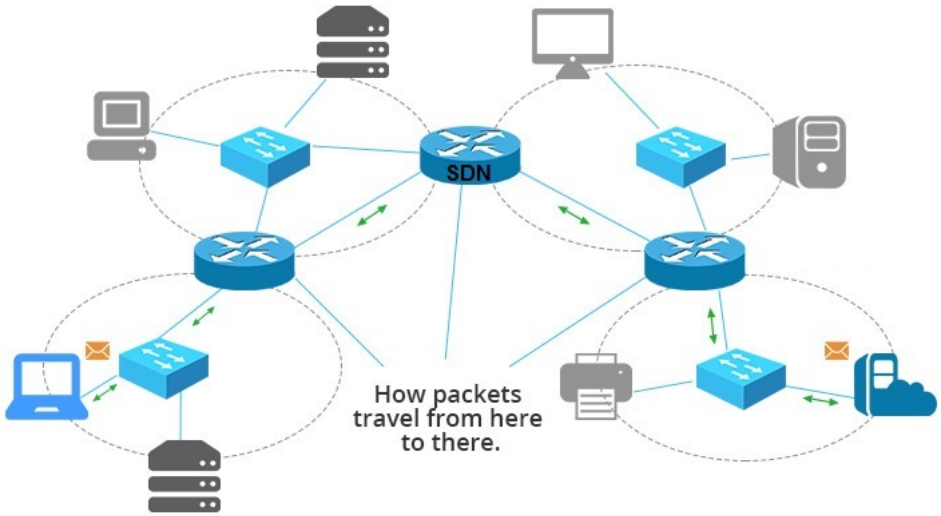
Penn
University of Pennsylvania

Problems in facility location, machine learning, control

Problems in facility location, machine learning, control

Facility location:
Router placement problem

Goal: Maximize controllable traffic flow in internet service provider networks by replacing legacy routers with SDN routers.



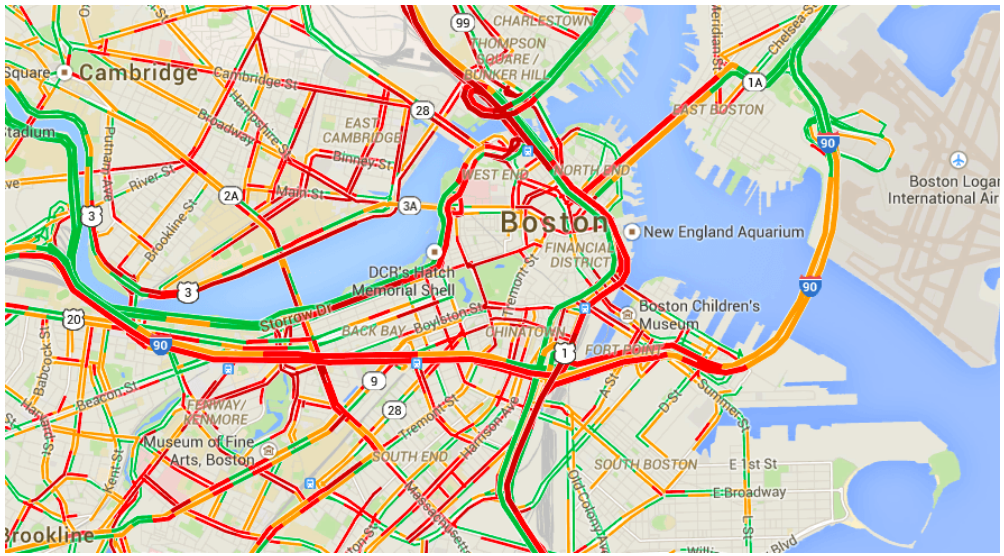
Complication: SDN routers can be expensive.
Problem: Where to place **few** SDNs to achieve goal?[1]

[1] Poularakis et al. '17, IEEE INFOCOM Best paper Award

Problems in facility location, machine learning, control

Machine learning:
Data selection problem

Goal: Maximize prediction accuracy of car traffic by using data collected from cameras and from driver's smart-phones apps.



Complication: Cannot process all data from the data flood.
Problem: What is the **sparsest** data set that achieves goal?[1]

[1] Krause and Guestrin, JMLR '08 (check also: Bilmes at UW; Jegelka at MIT)

Problems in facility location, machine learning, control

Control:
Sensor selection problem

Goal: Maximize quadrotor's localization accuracy by using on-board sensors.

Complication: Quadrotor's battery is limited and sensors need it.
Problem: Which **few** sensors to activate to achieve goal?[1]

---

[1] Tzoumas, Carlone, Pappas, Jadbabaie, arXiv: 1709

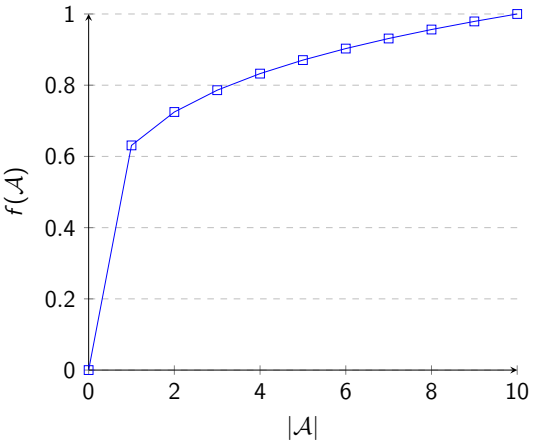
[2] Additional sensor/actuator selection contributors: Bushnell; Bullo; Clark; Cortes; Jovanovic; Krause; Le Ny; Mo; Olshevsky; Pasqualetti; Pequito; Pooven-dran; Roy; Sinopoli; Siami; Smith; Summers; Sundaram; Zampieri; Zhang; ...

Problems in facility location, machine learning, control

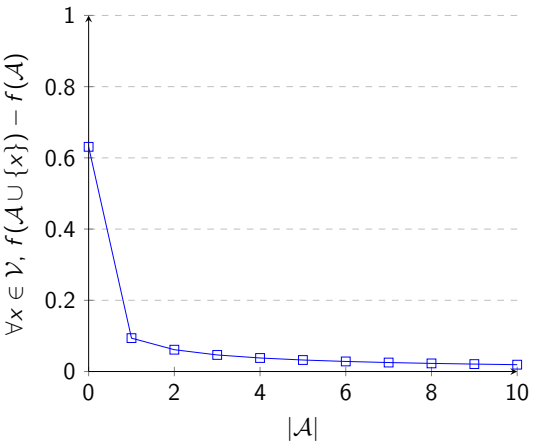
All previous are monotone submodular maximization problems

Monotone submodular maximization:

Given:

- finite ground set $\mathcal{V}$;
- set function f → non-decreasing: 

- budget α,

→ submodular: 

solve:

$$\begin{array}{ll} \max_{\mathcal{A} \subseteq \mathcal{V}} & f(\mathcal{A}) \\ \text{s.t.} & |\mathcal{A}| \leq \alpha. \end{array}$$

Sensors fail; routers get attacked; data get deleted

Denial of Service in Sensor Networks

wireless

Unless their de
sensor networ
to denial-of-se

Sensor netwo
tating large-s
in complex e
applications
military, env
domestic infrastructu

THE WALL STREET JOURNAL.

TECH

New Threats Fuel Fears of Another Global Cyberattack

A new attack hit thousands of computers and a hacking group said it would release more attack softwa

Council of the European Union

Brussels, 11 June 2015
(OR. en)

[...] A natural person should have the right that their personal data are erased and no longer processed [...]

if we pick $\mathcal{A}$ to $\max_{\mathcal{A}\subseteq\mathcal{V},|\mathcal{A}|\leq\alpha} f(\mathcal{A})$ → and later a $\mathcal{B}\subseteq\mathcal{A}$ gets deleted → we end up with $f(\mathcal{A}-\mathcal{B})$

Sensors fail; routers get attacked; data get deleted

Denial of Service in Sensor Networks

Unless their de
sensor networ
to denial-of-se

Sensor netwo
tating large-s
in complex e
applications
military, env
domestic infrastructu

THE WALL STREET JOURNAL.

TECH

New Threats Fuel Fears of Another Global Cyberattack

A new attack hit thousands of computers and a hacking group said it would release more attack softwa

Council of the European Union

Brussels, 11 June 2015 (OR. en)

[...] A natural person should have the right that their personal data are erased and no longer processed [...]

if we pick $\mathcal{A}$ to ~~$\max_{\mathcal{A}\subseteq\mathcal{V},|\mathcal{A}|\leq\alpha} f(\mathcal{A})$~~ → and later a $\mathcal{B}\subseteq\mathcal{A}$ gets deleted → we end up with $f(\mathcal{A}-\mathcal{B})$

Resilient Monotone Submodular Maximization

Problem

Given:

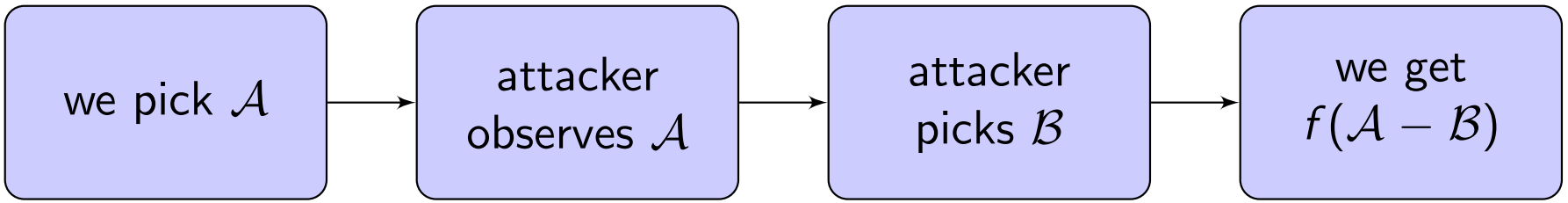
- *finite ground set* $\mathcal{V}$;
- *set function* f *s.t. non-decreasing, submodular,* $f \geq 0$, $f(\emptyset) = 0$;
- *budgets* α, β *s.t.* $0 \leq \beta \leq \alpha \leq |\mathcal{V}|$,

solve:

$$\max_{\mathcal{A} \subseteq \mathcal{V}, |\mathcal{A}| \leq \alpha} \; \min_{\mathcal{B} \subseteq \mathcal{A}, |\mathcal{B}| \leq \beta} \; f(\mathcal{A} - \mathcal{B}).$$

Symbol explanation:

- α: selection budget for resiliency;
- β: maximum number of (future) deletions.



Resilient Monotone Submodular Maximization

Problem

Given:

- *finite ground set* $\mathcal{V}$;
- *set function* f *s.t. non-decreasing, submodular,* $f \geq 0$, $f(\emptyset) = 0$;
- *budgets* α, β *s.t.* $0 \leq \beta \leq \alpha \leq |\mathcal{V}|$,

solve:

$$\max_{\mathcal{A} \subseteq \mathcal{V}, |\mathcal{A}| \leq \alpha} \min_{\mathcal{B} \subseteq \mathcal{A}, |\mathcal{B}| \leq \beta} f(\mathcal{A} - \mathcal{B}).$$

Difficulties:

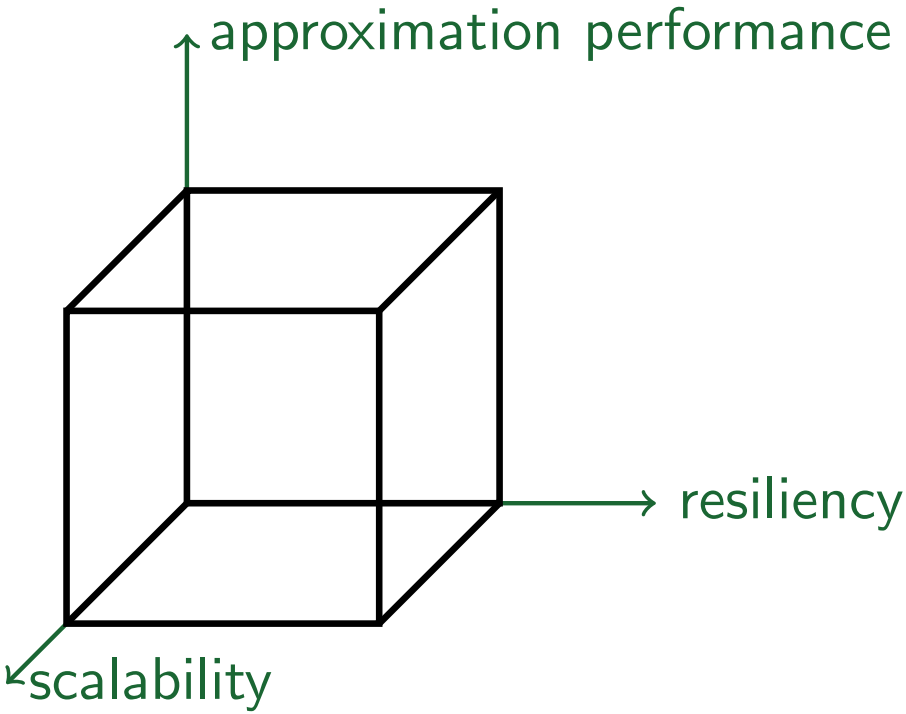
- Problem is NP-hard;[1]
- Function $g(\mathcal{A}) \triangleq \min_{\mathcal{B} \subseteq \mathcal{A}, |\mathcal{B}| \leq \beta} f(\mathcal{A} - \mathcal{B})$ is non-submodular

  $\Rightarrow$ Greedy alg. on $\max_{\mathcal{A} \subseteq \mathcal{V}, |\mathcal{A}| \leq \alpha} g(\mathcal{A})$ can perform arbitrarily bad.[2]

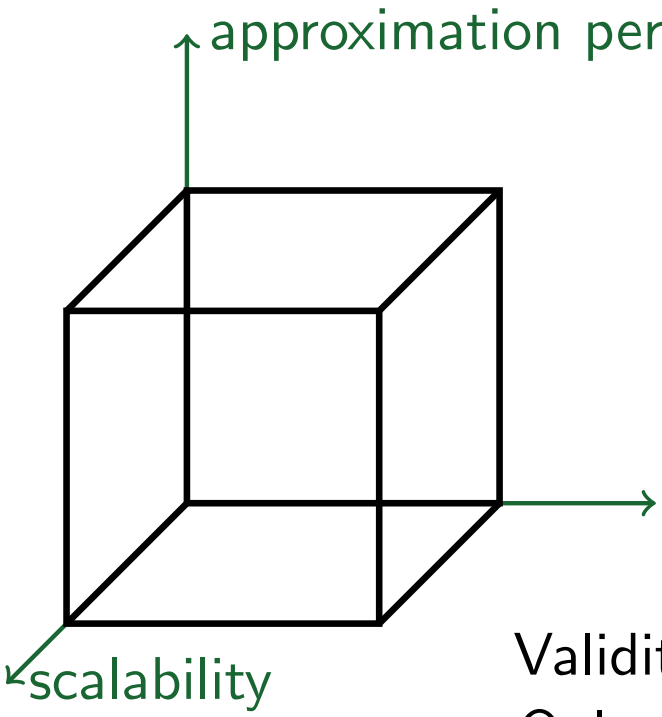
[1] Orlin et al., IPCO '16

[2] Krause et al., JMLR '08

Characteristics of good algorithm



Characteristics of good algorithm

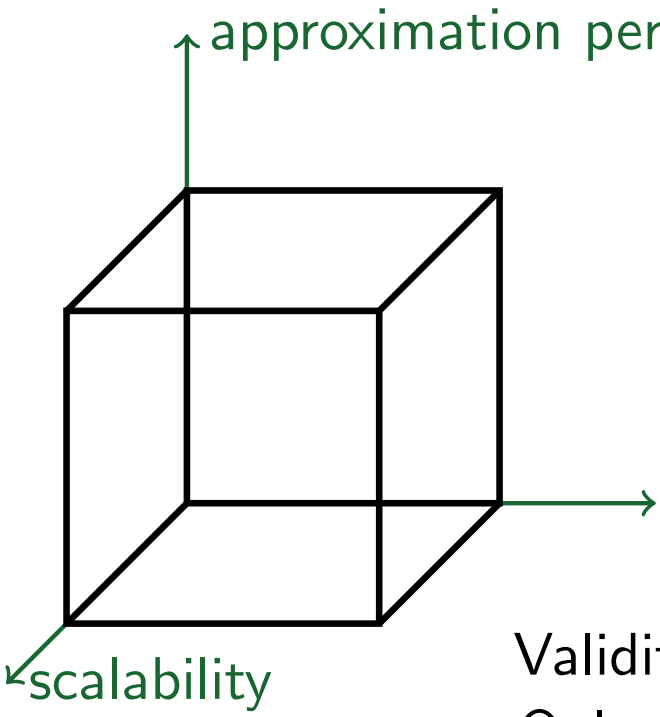


Validity up to what number β of attacks?

Only $\beta \ll \alpha$? Or any $\beta \leq \alpha$?

low resiliency — high resiliency

Characteristics of good algorithm

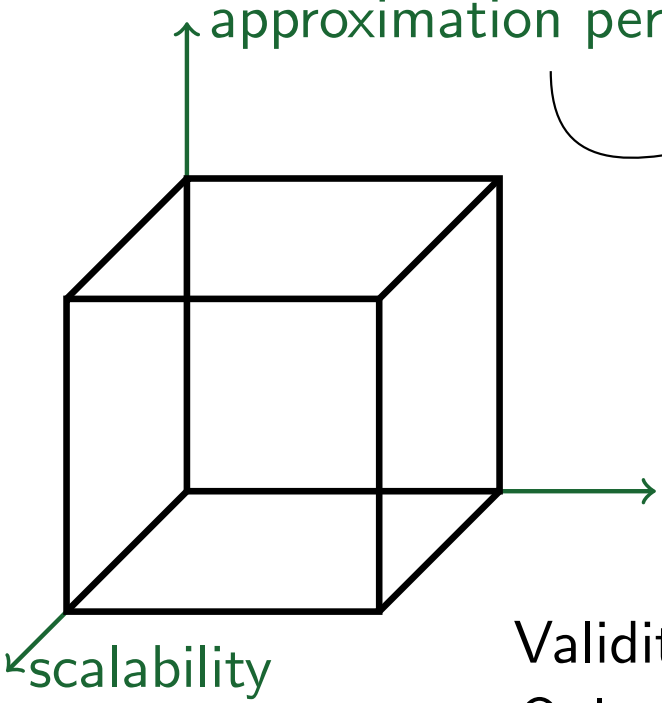


Validity up to what number β of attacks?

Only $\beta \ll \alpha$? (low resiliency) Or any $\beta \leq \alpha$? (high resiliency)

Running time w.r.t. α, β, $|\mathcal{V}|$? Exponential? (slow) Or linear? (fast)

Characteristics of good algorithm



f_{approx} finitely close to f_{opt}?

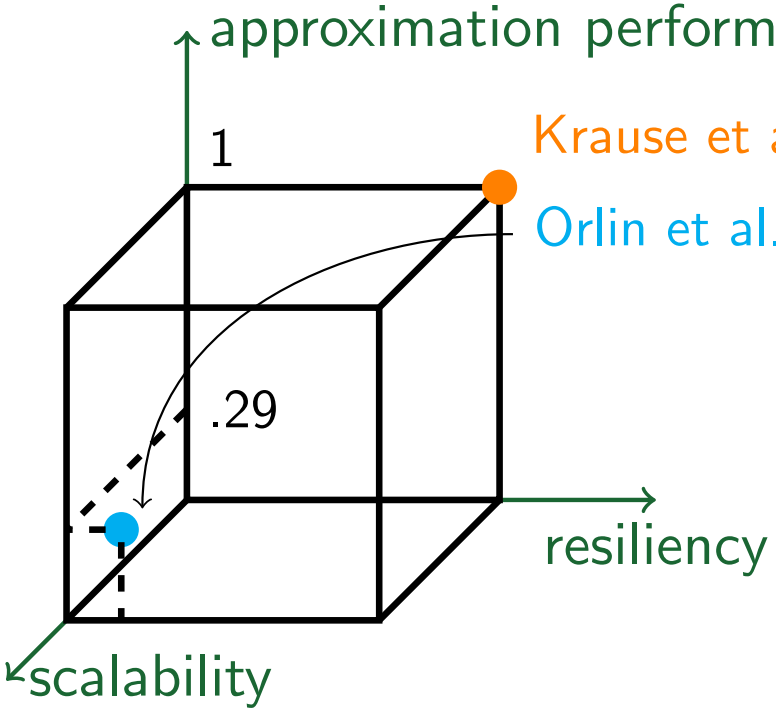
$\underbrace{\frac{f_{\text{approx}}}{f_{\text{opt}}} \geq 0?}_{\text{no suboptimality guarantee}}$ Or $\underbrace{\frac{f_{\text{approx}}}{f_{\text{opt}}} = 1?}_{\text{exact}}$

Validity up to what number β of attacks?

$\underbrace{\text{Only } \beta \ll \alpha?}_{\text{low resiliency}}$ Or $\underbrace{\text{any } \beta \leq \alpha?}_{\text{high resiliency}}$

Running time w.r.t. α, β, $|\mathcal{V}|$? $\underbrace{\text{Exponential?}}_{\text{slow}}$ Or $\underbrace{\text{linear?}}_{\text{fast}}$

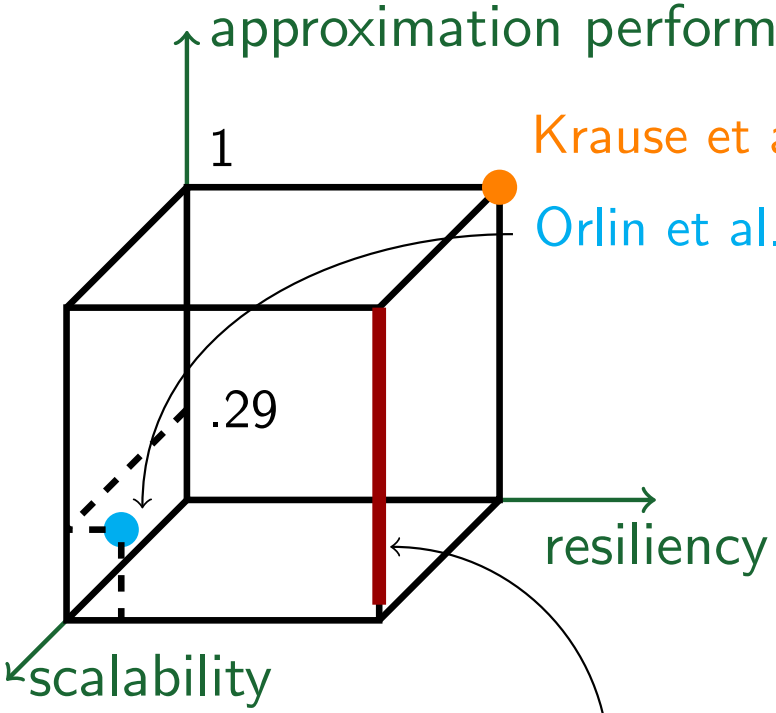
Literature review



Krause et al., JMLR '08 (exponential time: $O(\alpha^\beta)$)

Orlin et al., IPCO '16 (low resiliency: $\beta \leq \sqrt{\alpha}$)

Literature review



Krause et al., JMLR '08 (exponential time: $O(\alpha^\beta)$)

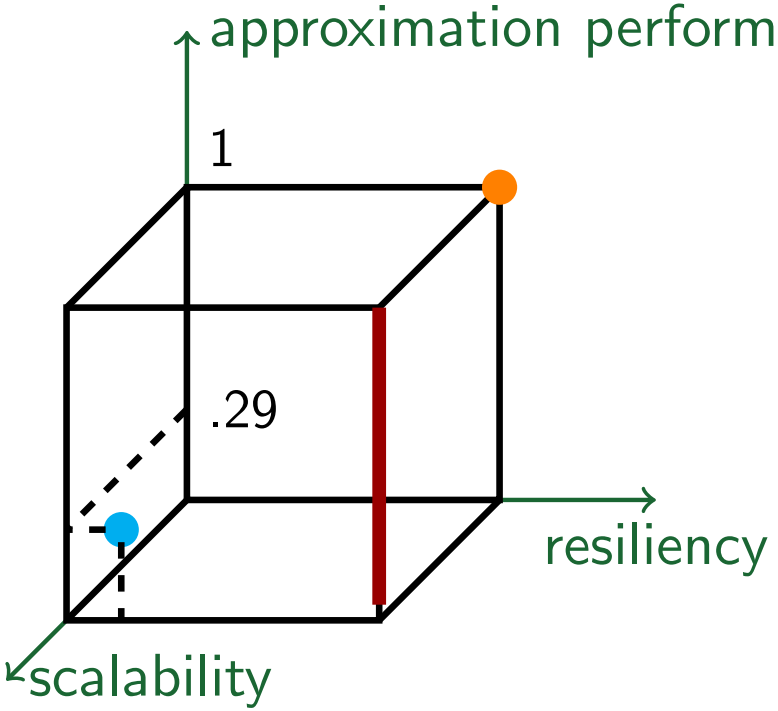
Orlin et al., IPCO '16 (low resiliency: $\beta \leq \sqrt{\alpha}$)

What we ask for: feasibility of red segment

Existence of algorithm that has:

- *High resiliency:* valid for any number β of attacks;
- *High scalability:* running time at most linear in $\alpha, \beta, |\mathcal{V}|$;
- *Provable approximation performance:* non-zero suboptimality guarantee.

Literature review

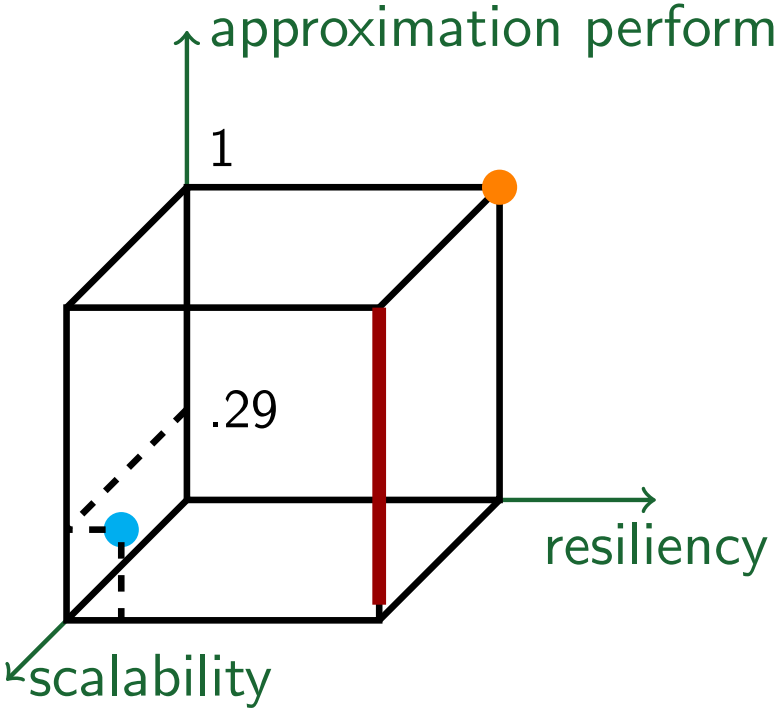


First claim in the CDC paper

First algorithm that has:

- *High resiliency:* valid for any number β of attacks;
- *High scalability:* running time at most linear in $\alpha, \beta, |\mathcal{V}|$;
- *Provable approximation performance:* non-zero suboptimality guarantee.

Literature review

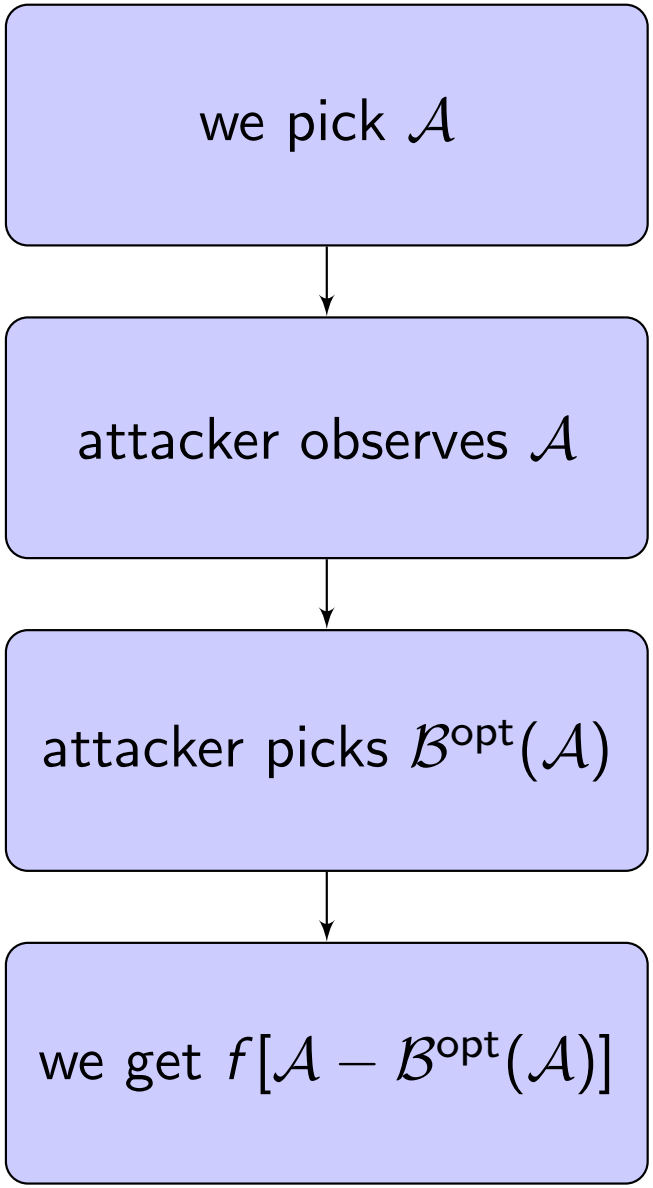


Second claim in the CDC paper

First algorithm that has:

- *High resiliency:* valid for any number β of attacks;
- *High scalability:* running time at most linear in $\alpha, \beta, |\mathcal{V}|$;
- **Superior approximation performance**: For functions f with low curvature, first algorithm with approximation performance $\geq .29$.

Algorithm


we pick $\mathcal{A}$
attacker observes $\mathcal{A}$
attacker picks $\mathcal{B}^{\text{opt}}(\mathcal{A})$
we get $f[\mathcal{A} - \mathcal{B}^{\text{opt}}(\mathcal{A})]$

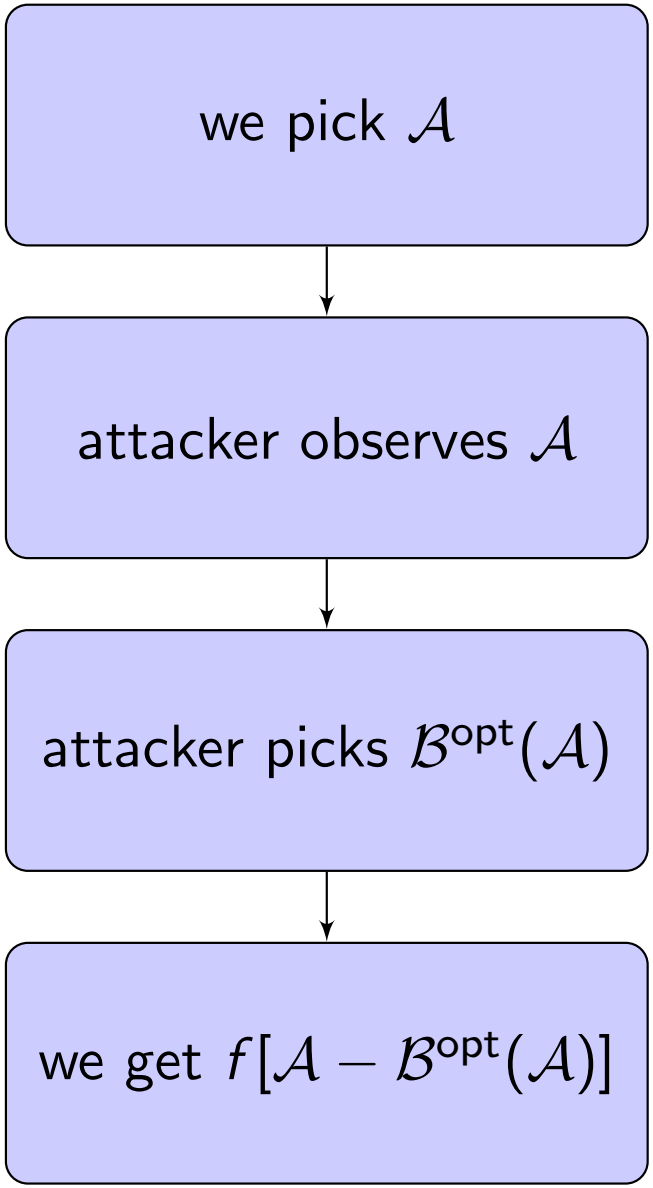
Algorithm

Idea:

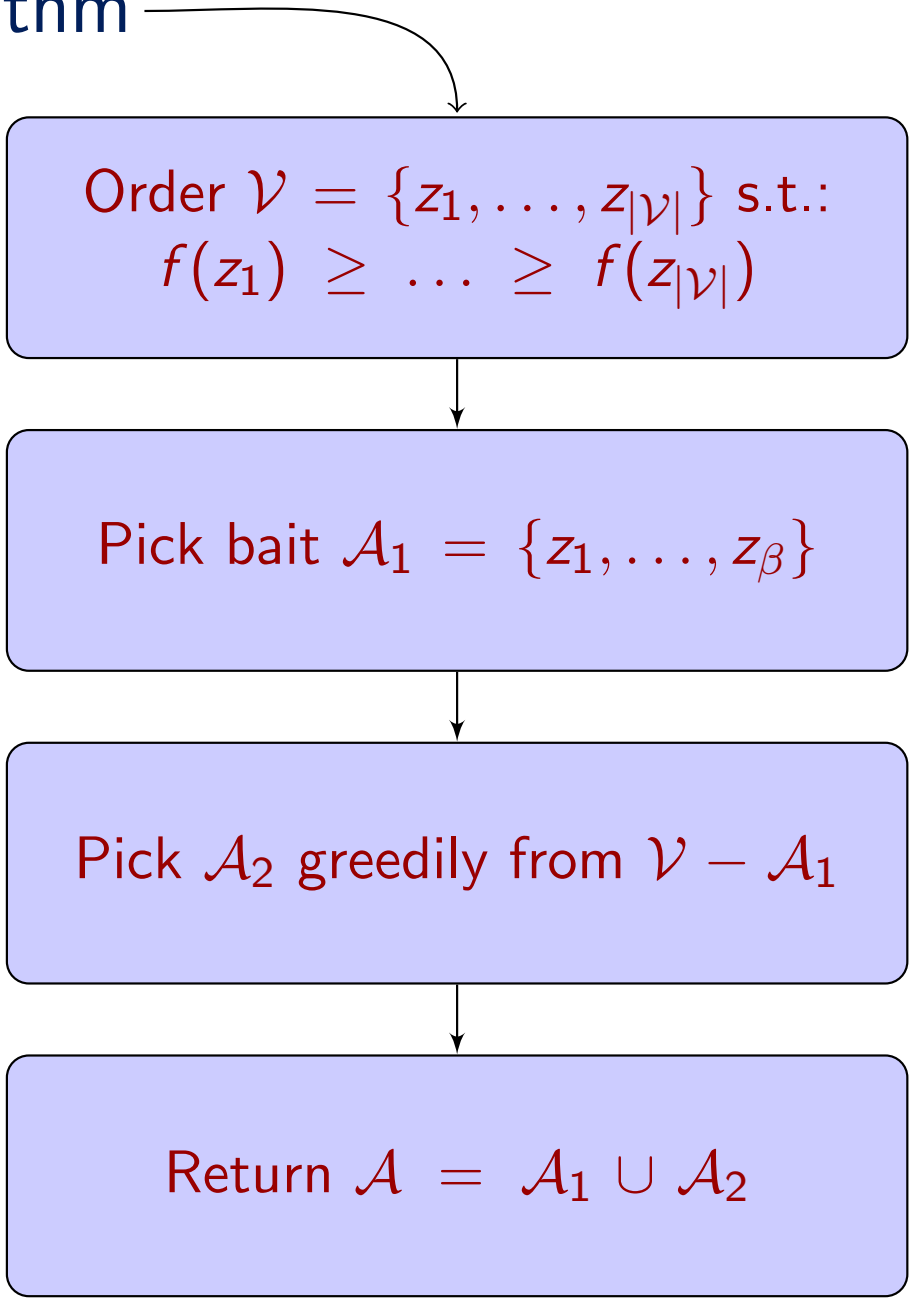
Pick $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$

"bait" $|\mathcal{A}_1| = \beta.$

$\mathcal{A}_2 \subseteq \mathcal{V} - \mathcal{A}_1;$ $|\mathcal{A}_2| = \alpha - \beta.$



Algorithm



Idea:

Pick $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$

"bait"
$|\mathcal{A}_1| = \beta$.

$\mathcal{A}_2 \subseteq \mathcal{V} - \mathcal{A}_1$;
$|\mathcal{A}_2| = \alpha - \beta$.

Algorithm's performance

Definition:

f's curvature is defined as:

$$\kappa_f \triangleq 1 - \min_{z \in \mathcal{V}} \frac{f(\mathcal{V}) - f(\mathcal{V} \setminus \{z\})}{f(z)}.$$

Properties:

- Computable in $O(|\mathcal{V}|)$ time;
- $0 \leq \kappa_f \leq 1$.

Interpretation:

κ_f measures how $\mathcal{V}$'s elements *substitute* each other:

- $\kappa_f = 0 \Leftrightarrow f(\mathcal{A}) = \sum_{z \in \mathcal{A}} f(z)$;
- $\kappa_f = 1 \Leftrightarrow$ there exist $z \in \mathcal{V}$ s. t. $f(\mathcal{V}) = f(\mathcal{V} \setminus \{z\})$.

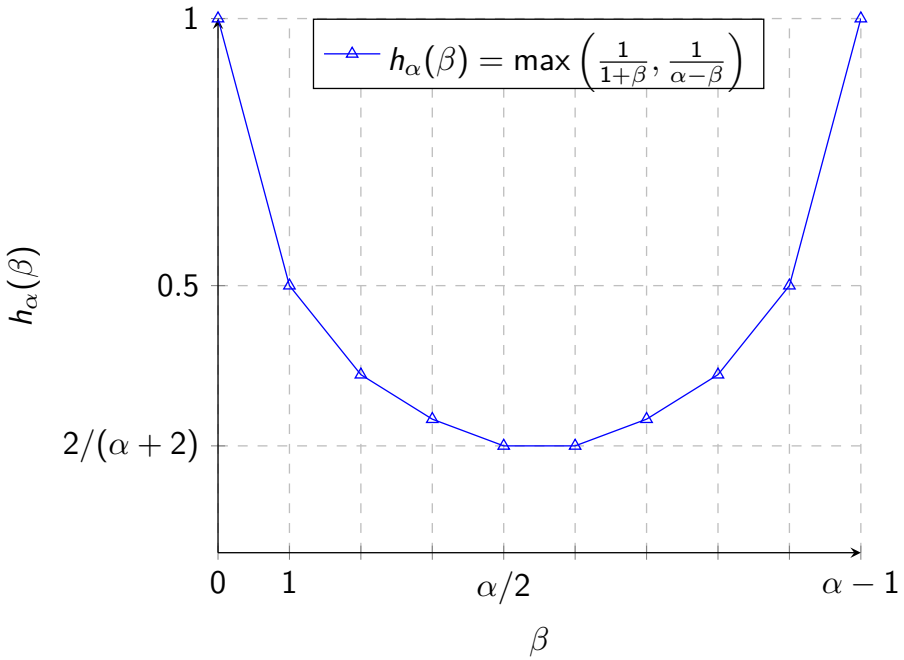
Algorithm's performance

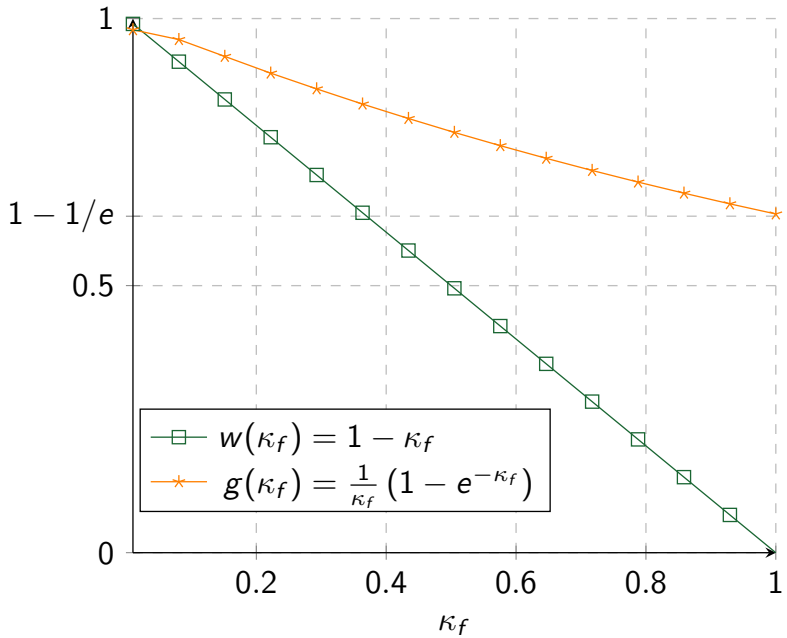
Theorem

Algorithm:

- *(Resiliency)* *is valid for any* $\beta \leq \alpha \leq |\mathcal{V}|$;
- *(Scalability)* *runs in* $O[(\alpha - \beta)|\mathcal{V}|]$ *time;*
- *(Provable approximation performance)* *guarantees:*

$$\frac{f_{approx}}{f_{opt}} \geq \max\left[h_\alpha(\beta), w(\kappa_f)\right] g(\kappa_f).$$





Algorithm's performance

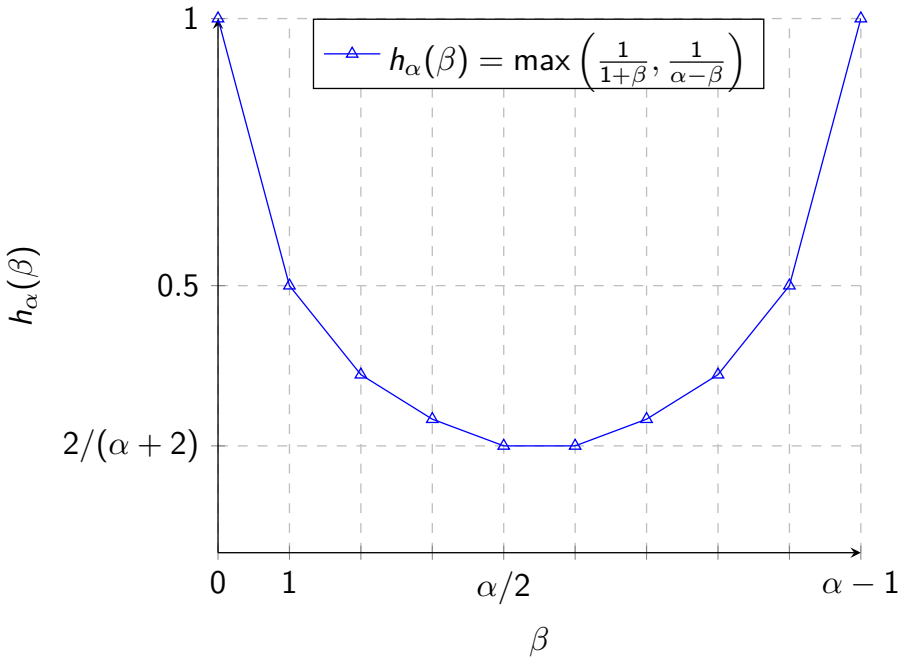
Theorem

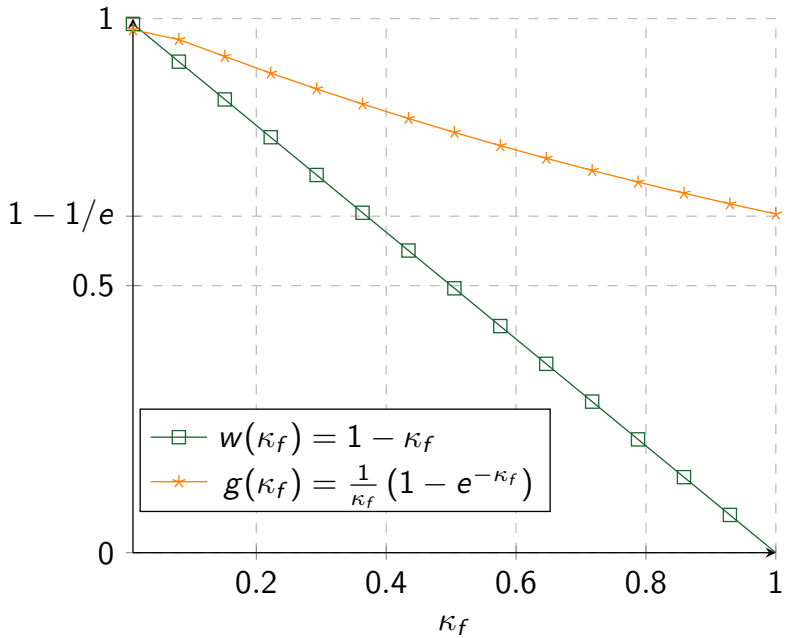
Algorithm:

greedy algorithm's guarantee for $\max_{\mathcal{A}\subseteq\mathcal{V},|\mathcal{A}|\leq\alpha} f(\mathcal{A})$

- *(Resiliency)* *is valid for any* $\beta \leq \alpha \leq |\mathcal{V}|$;
- *(Scalability)* *runs in* $O[(\alpha-\beta)|\mathcal{V}|]$ *time;*
- *(Provable approximation performance)* *guarantees:*

$$\frac{f_{approx}}{f_{opt}} \geq \max\left[h_\alpha(\beta), w(\kappa_f)\right] g(\kappa_f).$$





Classes of functions with $\kappa_f < 1$ and applications

- Concave over modular functions;
  - (Machine learning) Image segmentation, speech processing.
- Functions of the form $f(\mathcal{A}) = \log\det(\sum_{i\in\mathcal{A}} D_i + I)$;
  - (Machine learning) Experiment design; feature, data selection.
  - (Control) Sensor and actuator selection.

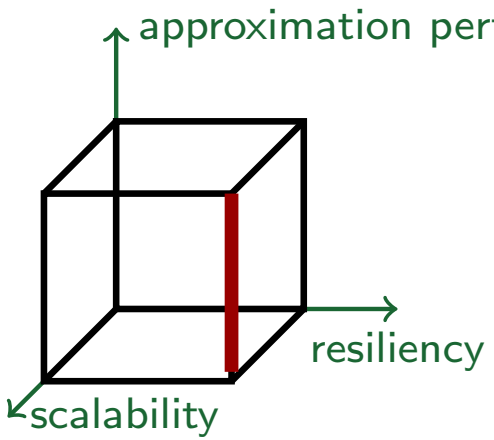
  Example: Gaussian processes with RBF kernels:[3] $\kappa_f \simeq 0$.[4]
- Functions of the form $f(\mathcal{A}) = \text{trace}[(\sum_{i\in\mathcal{A}} D_i + I)^{-1}]$.
  - (Machine learning) Experiment design; feature, data selection.
  - (Control) Sensor and actuator selection.

[3] RBF kernels model physical phenomena such as temperature in buildings.
[4] Sharma et al., ICML '15.

Summary of results

First algorithm that has:

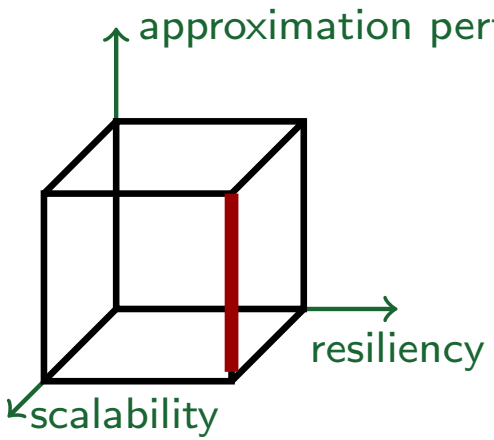
- *High resiliency:* valid for any number β of attacks;
- *High scalability:* running time at most linear in $\alpha, \beta, |\mathcal{V}|$;
- *Provable approximation performance:* suboptimality guarantees.
- *Superior approximation performance:* For curvature values $\kappa_f \leq .71$, first algorithm with approximation performance $\geq .29$.



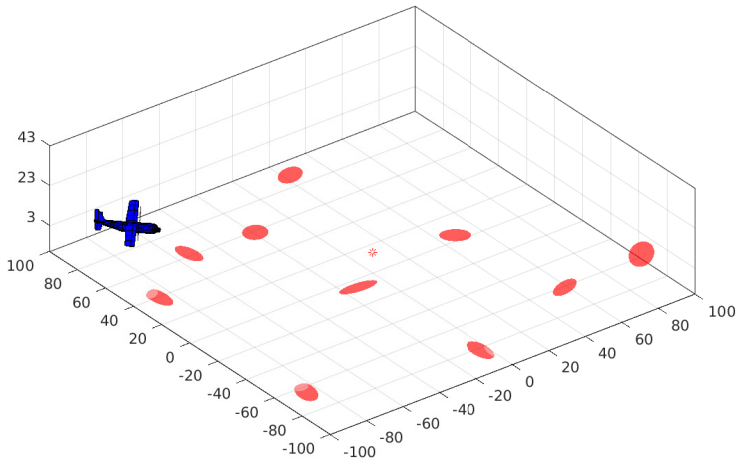
Summary of results

First algorithm that has:

- *High resiliency:* valid for any number β of attacks;
- *High scalability:* running time at most linear in $\alpha, \beta, |\mathcal{V}|$;
- *Provable approximation performance:* suboptimality guarantees.
- *Superior approximation performance:* For curvature values $\kappa_f \leq .71$, first algorithm with approximation performance $\geq .29$.



Simulations: robot localization scenario

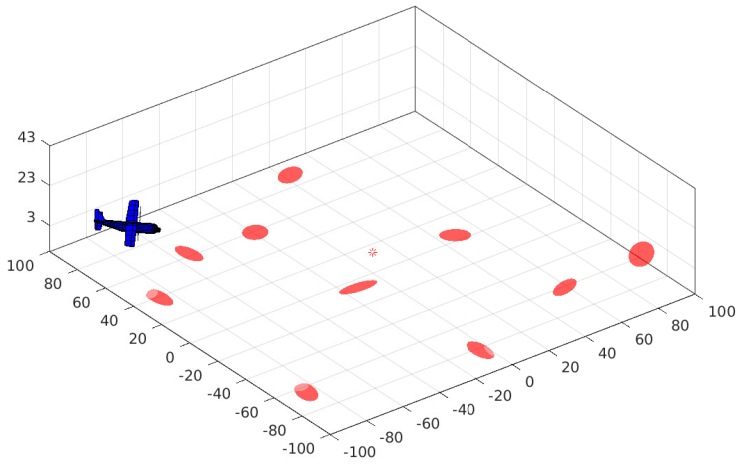


Scenario: UAV moves in a 3D space.

UAV's model: double-integrator, with state $x_k = [p_k, v_k]^\top$ s.t. $p_k =$ position; $v_k =$ velocity. ← disturbed with process noise

UAV's objective: UAV wants to land at position [0, 0, 0] with 0 velocity, by controlling its acceleration.

Simulations: robot localization scenario



Available sensors and landmarks for localization:

- 1 GPS (measuring position);
- 1 altimeter;
- 1 stereo camera;

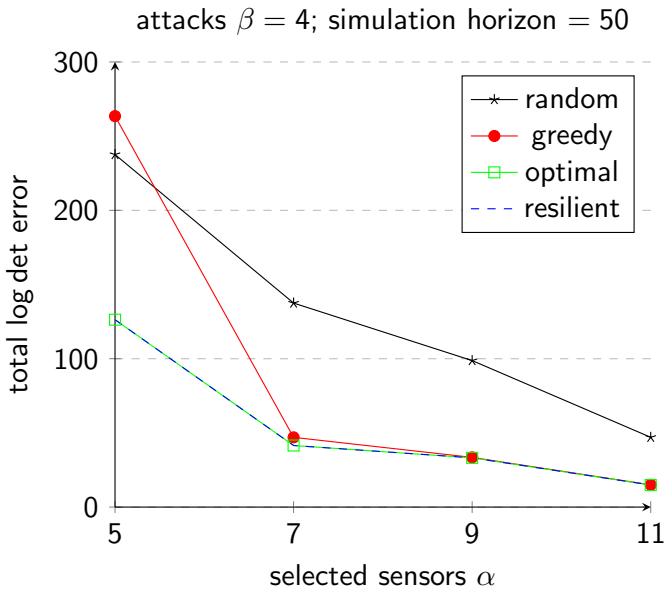
} corrupted with measurement noise

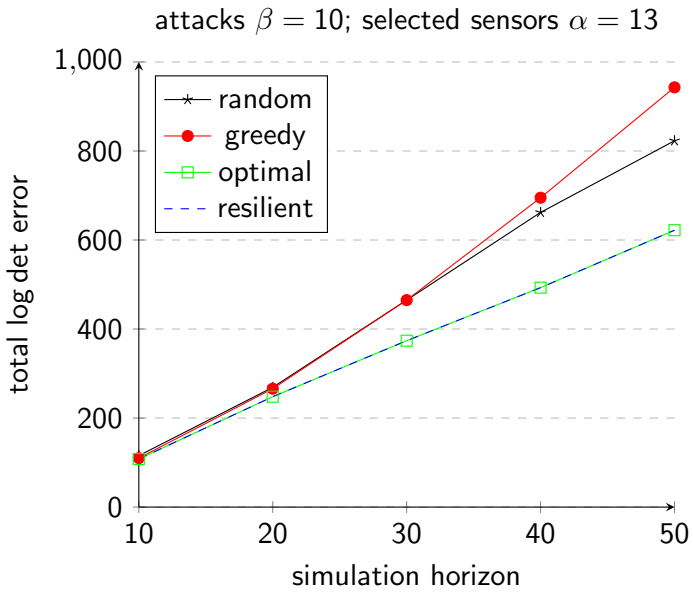
- 10 landmarks on the ground; — noisy knowledge of their position

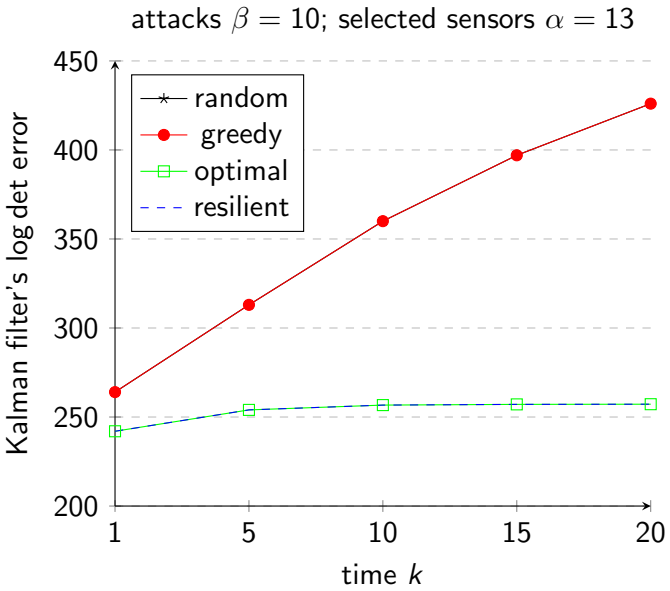
Sensor selection metric: Anticipation error $\log\det[\Sigma(x_k, \ldots, x_{k+20})]$.

minimum mean square error covariance

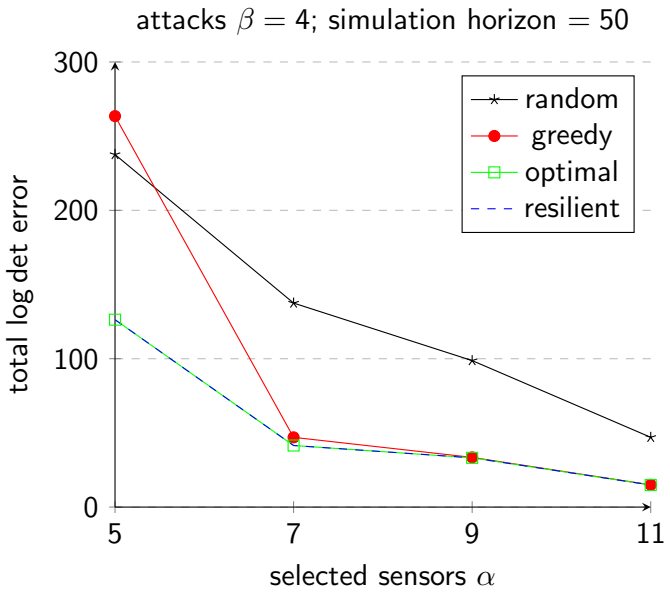
Simulation results

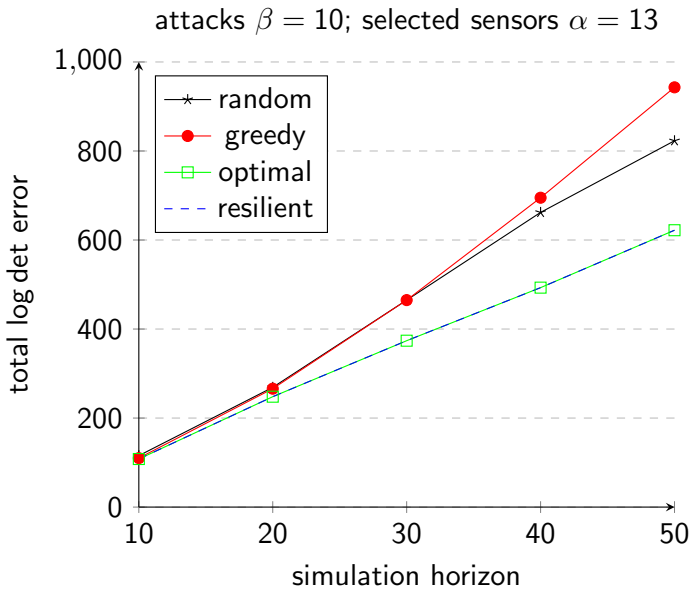


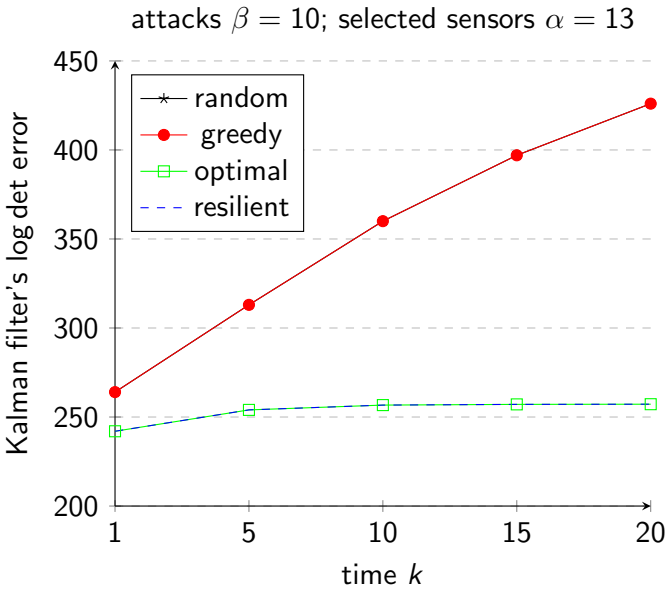


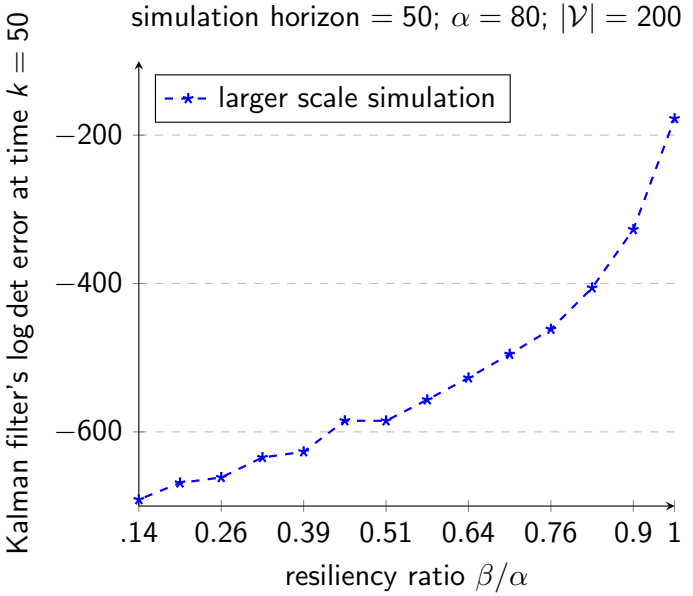


Simulation results





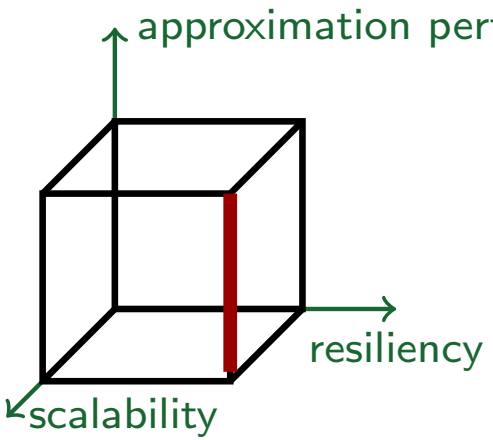




Summary and extensions

First algorithm that has:

- *High resiliency:* valid for any number β of attacks;
- *High scalability:* running time at most linear in $\alpha, \beta, |\mathcal{V}|$;
- *Provable approximation performance:* suboptimality guarantees.
- *Superior approximation performance:* For curvature values $\kappa_f \leq .71$, first algorithm with approximation performance $\geq .29$.



Extensions:

- Matroid constraints;
- Approximatelly submodular functions.