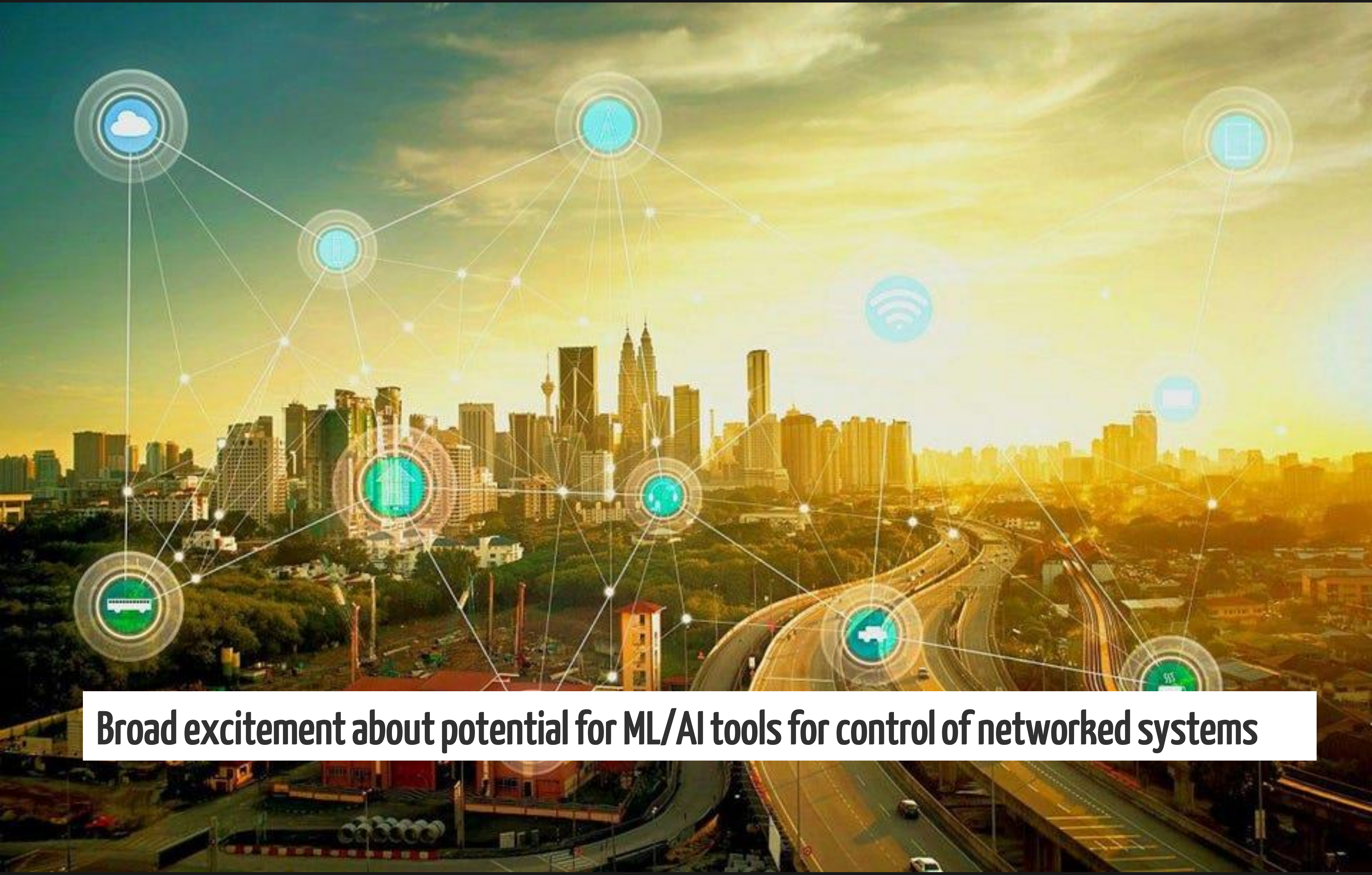
Online Optimization and Control using Black-Box Advice

Adam Wierman, Caltech

Joint work with Nicolas Christianson (Caltech), Tinashe Handina (Caltech), Debankur Mukherjee (Georgia Tech), Daan Rutten (Georgia Tech)

amazon vmware NSF RESNICK INSTITUTE science + energy + sustainability CALIFORNIA INSTITUTE OF TECHNOLOGY BEYOND LIMITS



Broad excitement about potential for ML/AI tools for control of networked systems

But...

=

[Eykholt et al 2018]

But...

NEWS
Home US Election Coronavirus Video World US & Canada
US & Canada
Tesla in fatal California crash w
31 March 2018

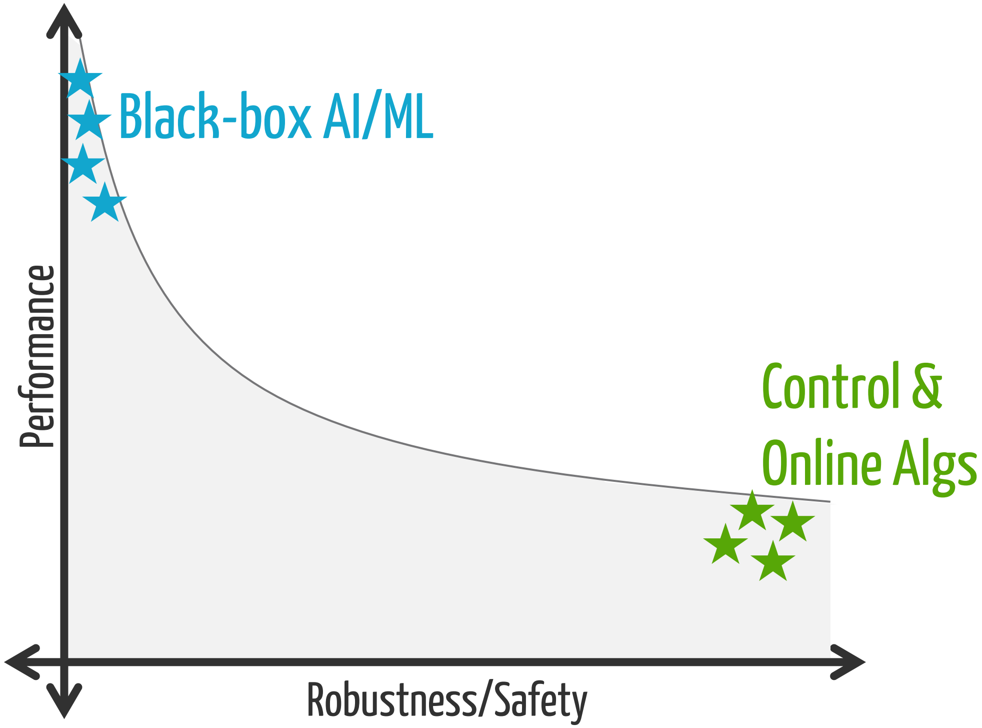
Drone crash near kids leads Swiss Post and Matternet to suspend autonomous deliveries
Devin Coldewey
@techcrunch / 5:17 pm PDT • July 30, 2019

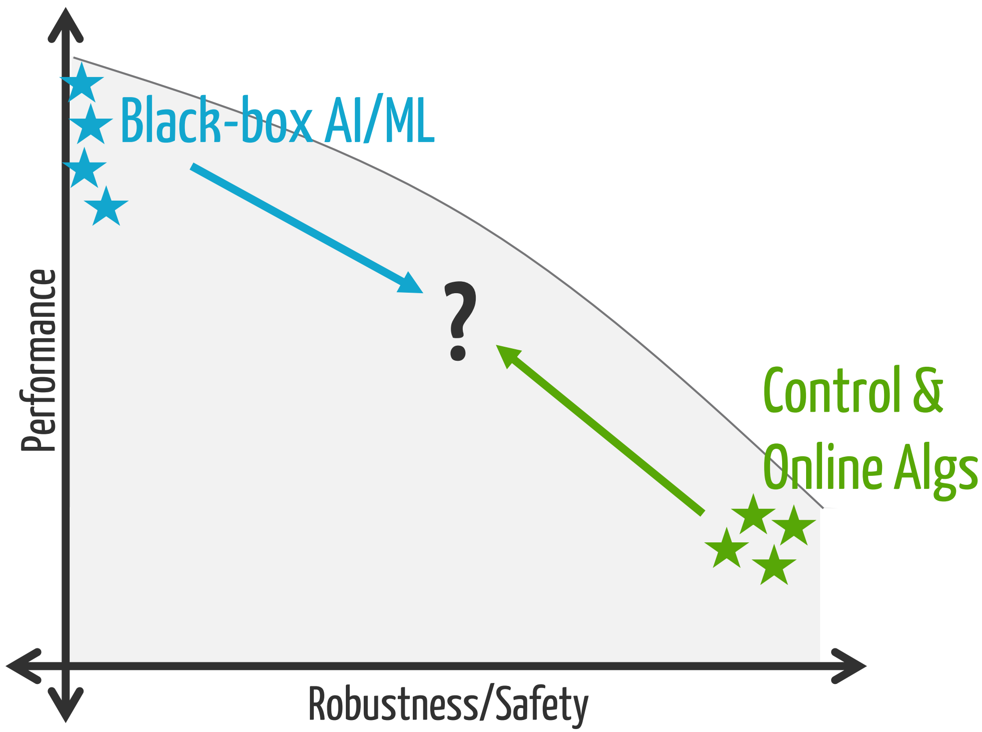
FAST COMPANY
LOGIN
08-10-21
Why high-profile smart cities fail, from Sidewalk's Quayside to Amazon's HQ2 in Queens
Silicon Valley loves to move fast and break things. Cities can't afford such negligence.

Boston Dynamics' Robo-Dog "Screws the Pooch" at Amazon Conference
07 Jun 2019 | Posted by Member 30298028
Boston Dynamics' robo-dog screws the pooch at Amazon's Re:...
No, it isn't playing dead. Boston Dynamics' robo-dog Spot experiences a technical failure at Amazon's Re:Mars conference.

Blackout in Southwest leaves 5M in the dark
SAN DIEGO - A major outage knocked out power in a region of almost 6 million people in the Southwest and Mexico on Thursday, bringing San Diego to a near-standstill and leaving people in the surrounding desert to swelter in late-summer heat.

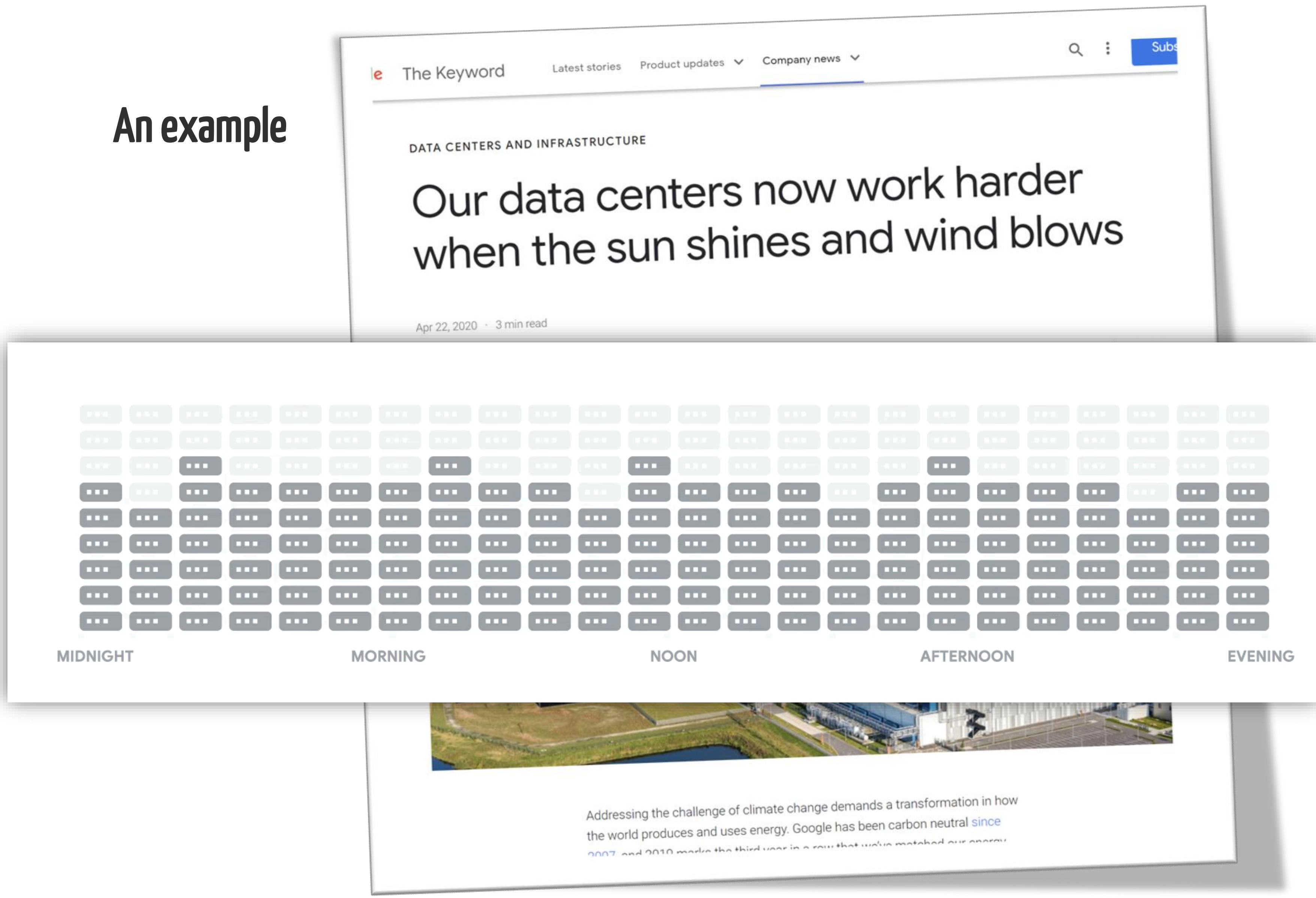
Can guarantees required by safety-critical applications be enforced by AI/ML tools?


Black-box AI/ML
Control &
Online Algs
Performance
Robustness/Safety



Black-box AI/ML
?
Control &
Online Algs
Performance
Robustness/Safety

An example


The Keyword
Latest stories
Product updates
Company news
Subs

DATA CENTERS AND INFRASTRUCTURE

Our data centers now work harder when the sun shines and wind blows

Apr 22, 2020 · 3 min read

Addressing the challenge of climate change demands a transformation in how the world produces and uses energy. Google has been carbon neutral since

Control Meets Learning
Virtual Seminar Series on the Intersection of Control and Learning
Zoom link
Workshops
Workshops > Intersections between Control, Learning and Optimization
Intersections between Control
and Optimization
FEBRUARY 24 - 28,
SCHEDULE
4th L4DC Conference
HOME
CALL FOR PAPERS
REGISTRATION
INVITED SPEAKERS
PROGRAM
ORGANIZERS & PC
4th Annual Learning for Dynamics & Control Conference
June 23-24, Frances C. Arrillaga Alumni Center
Stanford University, CA, USA

Many exciting results and techniques

- Model-based RL in dynamical systems [Recht 19], [Kakade et al 20], [Simchowitz & Foster 20], [Lale et al 21], ...
- Lyapunov-based policy learning [Chow et al 18], [Richards et al 18], [Chang et al 19], [Jin et al 20], [Shi et al 21], ...
- Model-free policy search [Fazel et al 18], [Malik et al 18], [Bu et al 19], [Mohammadi et al 19], [Li et al 19], [Qu et al 20], ...
- Safe RL [Garcia & Fernandez 15], [Fisac et al 19], [Taylor et al 20], [Hewig et al 20], ...

...and many more

MAD LIBS®

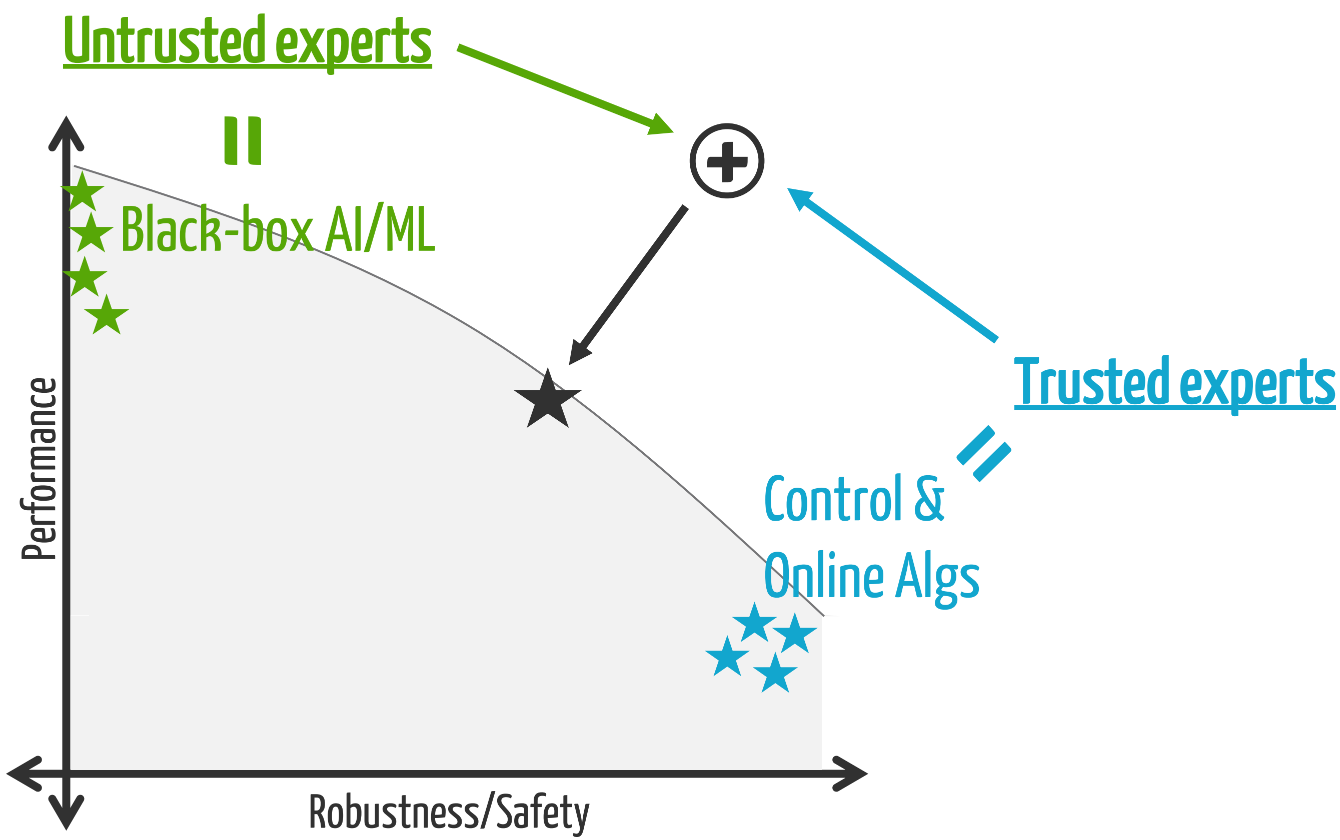
BOOKish Paper Titles

A real title of ours

guarantee	AI/ML algorithm		application
Stability-constrained	RL	for	Voltage Control
Certified Safe	Deep RL	for	Traffic management
Lyapunov-based	Policy learning	for	EV Charging

Today: A "black box" approach for getting the best of both worlds

Today: A "black box" approach for getting the best of both worlds



bicompetitive adversarial guarantee

Goal 1: Consistency

(Nearly) Match the performance of untrusted experts, when they do well.

$$Cost(Alg) \leq (1+\delta)Cost(Untrusted)$$

Goal 2: Robustness

Always provide worst case guarantees that (nearly) match the trusted experts.

$Cost(Alg) \leq \gamma_{Alg}\ Cost(Opt)$, where γ_{Alg} is "close to" $\gamma_{trusted}$

An example: MPC has good consistency, but terrible robustness.

Online decision making with untrusted advice is an emerging framework

Introduced by [Lykouris & Vassilvitskii, 2018] in the context of online caching

Since then, applied in a wide variety of settings:

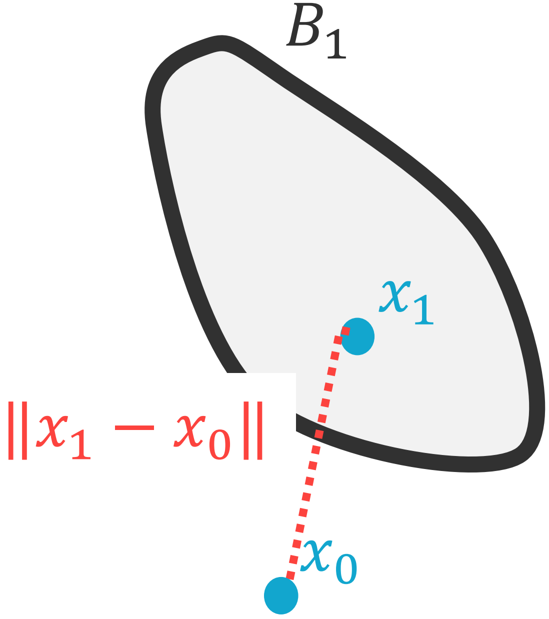
- ski rental [Purohit et al 18] [Angelopoulos et al 19] [Bamas et al 20] [Wei & Zhang 20], ...
- bloom filters [Mitzenmacher 18]
- online set cover [Bamas et al 20]
- online matching [Antoniadis et al 20]
- metrical task systems [Antoniadis et al 20]
- data center capacity [Rutten & Mukherjee 21]
- demand response [Lee et al 21]
- online optimization [Christianson et al 21]
- online conversion problems [Sun et al 21]
- convex body chasing [Christianson et al 21]
- linear quadratic control [Li et al 21]

<u>This talk:</u> Algorithm design & fundamental limits on the use of untrusted advice in online optimization & control

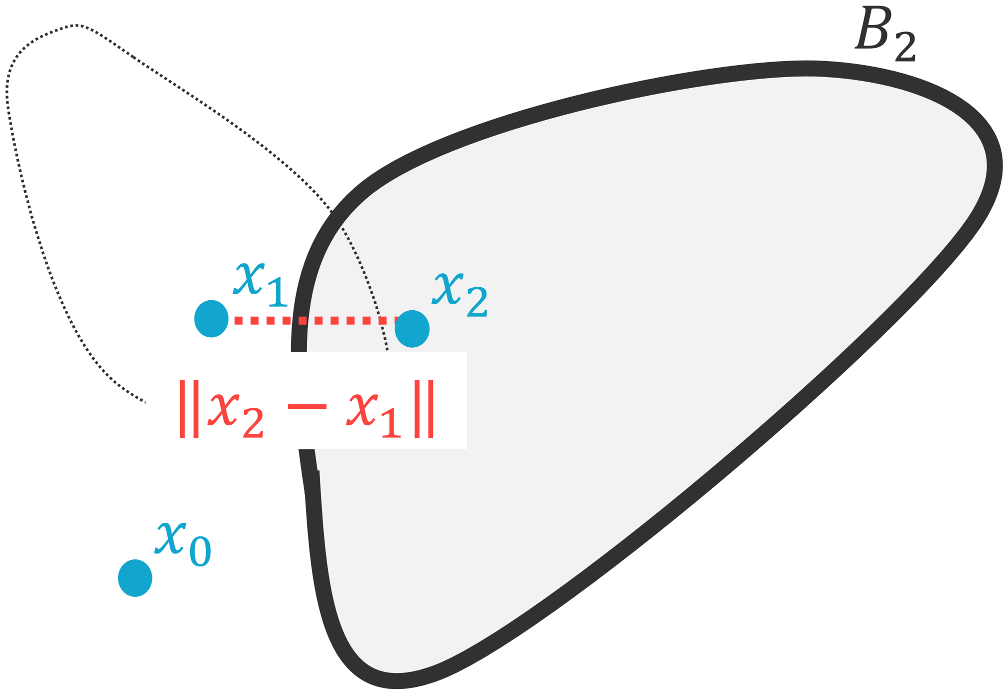
~~Three~~ Two running examples:

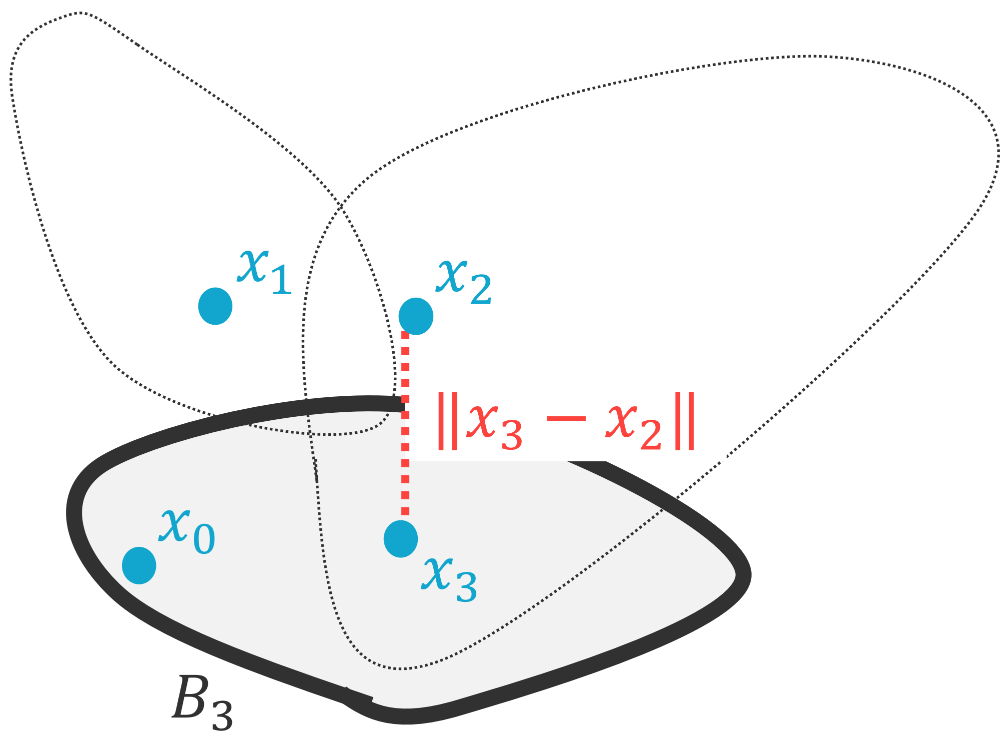
- Convex Body Chasing
- Online Optimization with Switching Costs
- ~~Linear Quadratic Control~~

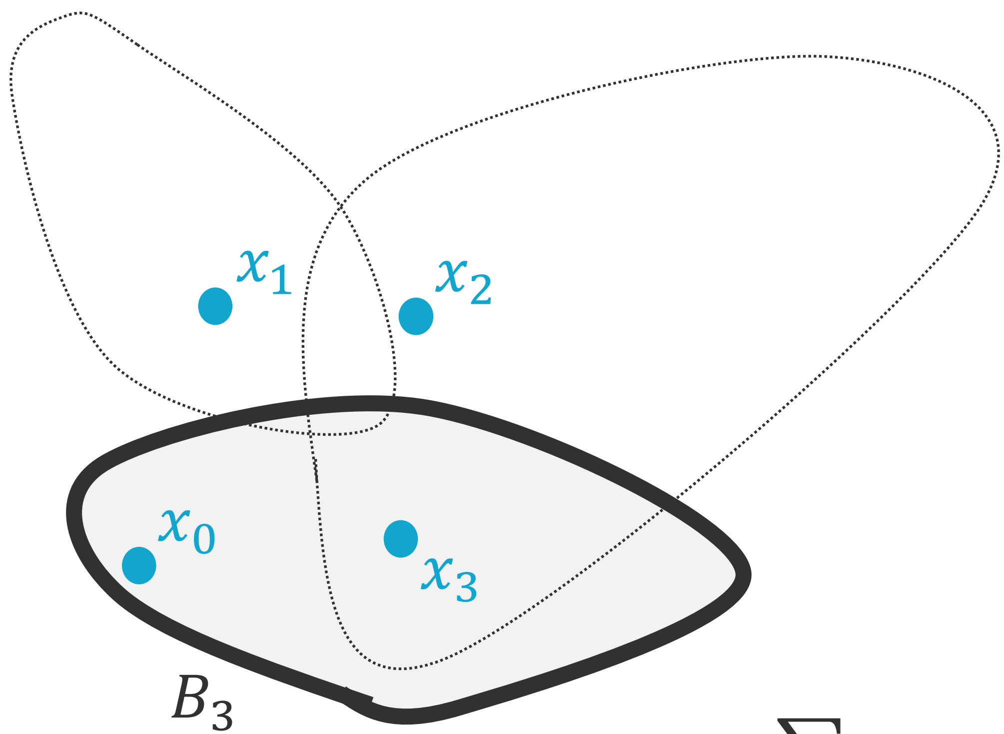
Example 1: Convex Body Chasing



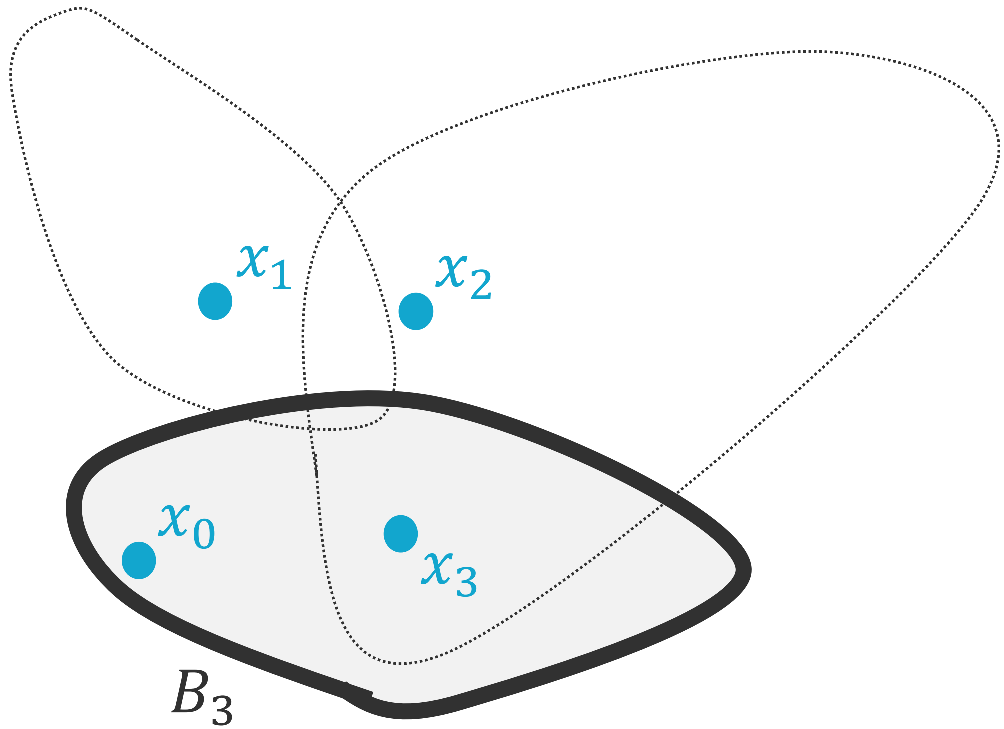
Example 1: Convex Body Chasing




x_1
x_2
$\|x_3 - x_2\|$
x_0
x_3
B_3



$$\min_{x_t \in B_t} \sum_t \|x_t - x_{t-1}\|$$



How do you decide where to move
without knowing the future?

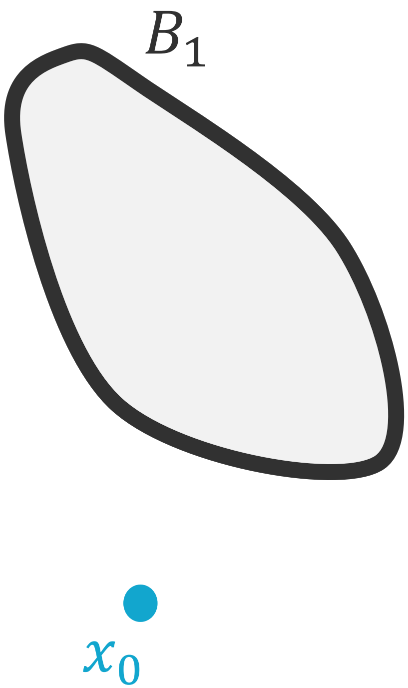
Convex Body Chasing has a long history

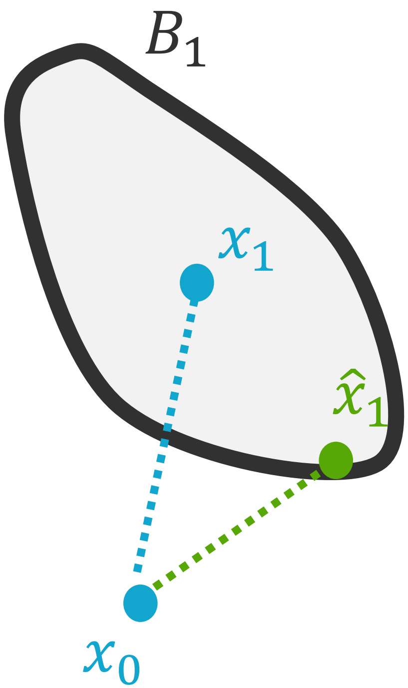
Core problem for safety & stability in control [Ho & Doyle 20], [Yu et al 22], [Yeh et al 22], ...

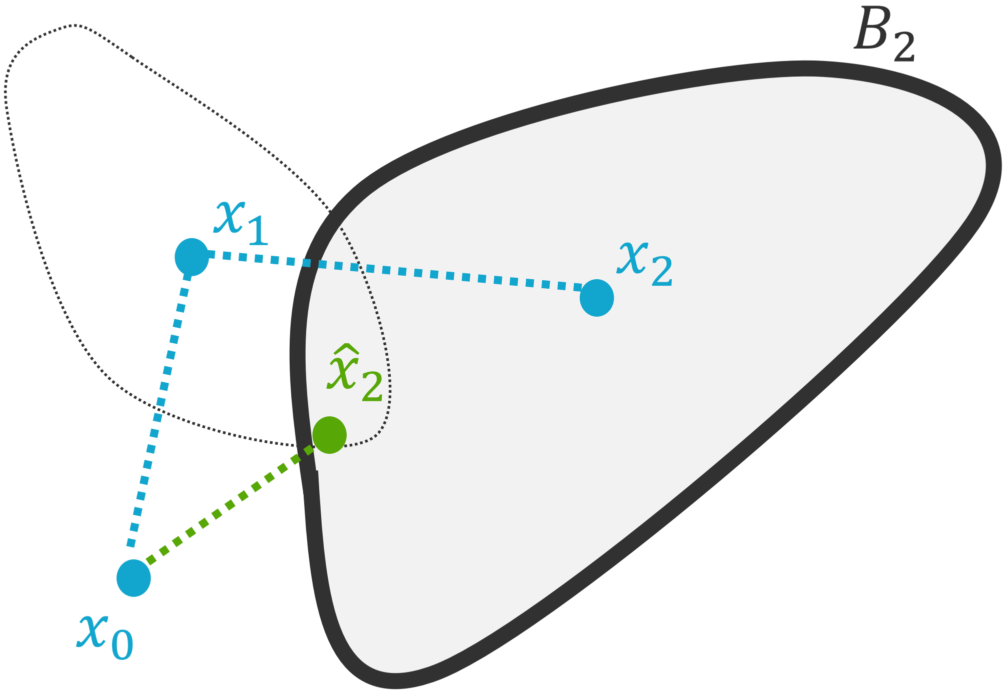
Exciting algorithmic progress in recent years [Abtibuadus 16], [Bansal et al 20], [Bubeck et al 19], [Sellke 20], [Argue 20], [Bubeck et al 20], [Argue 21], ...

Theorem [Bubeck et al 20]. Moving to the Steiner point of the body each round obtains an $O\left(\min\left(d, \sqrt{d \log(T)}\right)\right)$-competitive ratio, and any online algorithms is $\Omega(\sqrt{d})$.

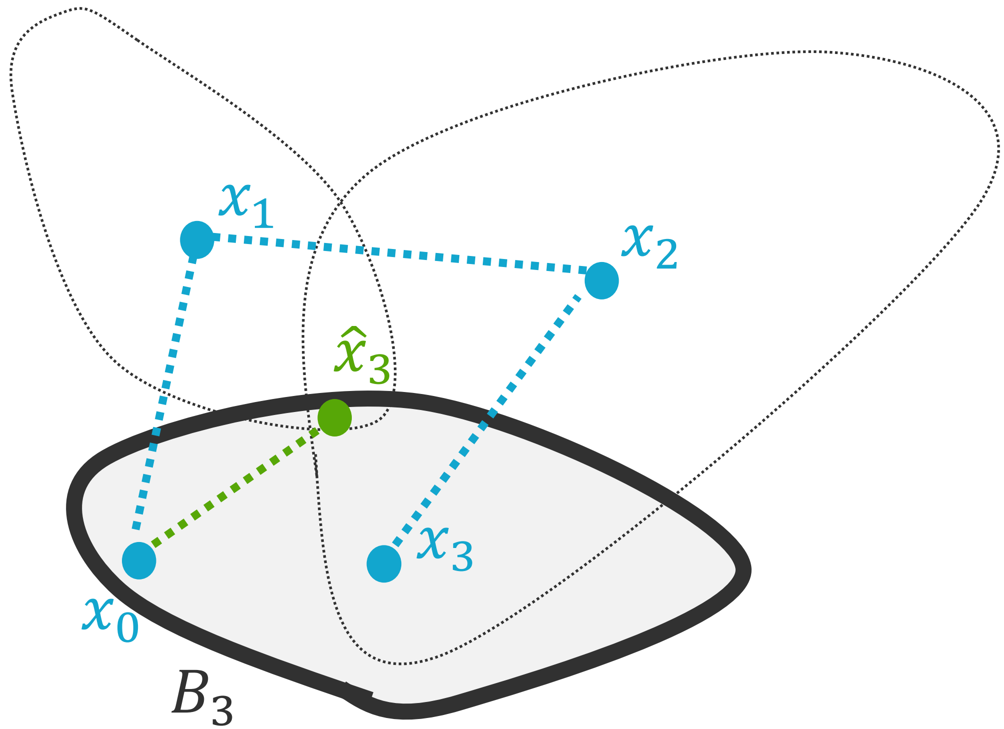
Choices of algorithm are quite conservative. **Advice can help.**


B_1
x_0

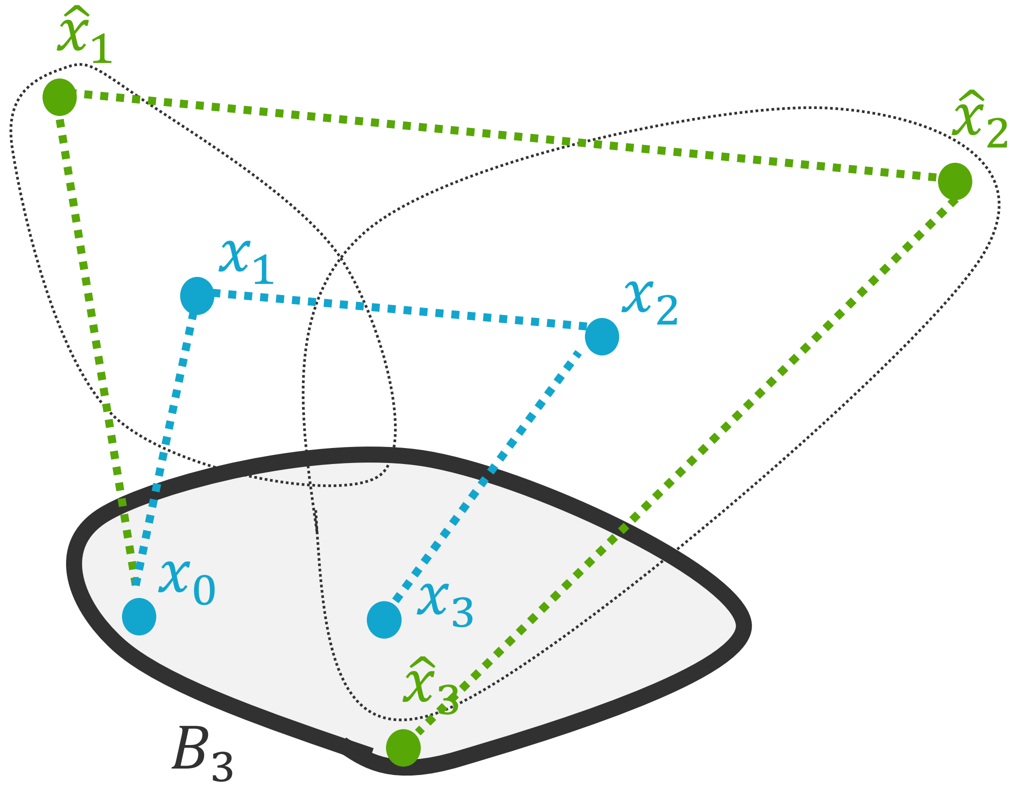

B_1
x_1
$\hat{x}_1$
x_0



B_2
x_1
x_2
$\hat{x}_2$
x_0



But the advice could have been bad...



But the advice could have been bad...

When should an algorithm “switch” between the trusted/untrusted advice?

Is it enough to always follow one or the other?

Attempt 1: A switching algorithm

1. Follow the <u>untrusted</u> advice until total distance traveled is r.
2. Follow the <u>trusted</u> advice until total distance traveled is r.
3. Set $r \leftarrow 2r$

Treats advice as black boxes.

Attempt 1: A switching algorithm

1. Follow the <u>untrusted</u> advice until total distance traveled is r.
2. Follow the <u>trusted</u> advice until total distance traveled is r.
3. Set $r \leftarrow 2r$ and repeat.

Optimize to bias toward consistency

<u>Theorem.</u> For nested convex body chasing, the switching algorithm is $(1+\delta)$-consistent & $O(d/\delta)$-robust.

Attempt 1: A switching algorithm

1. Follow the untrusted advice until total distance traveled is r.
2. Follow the trusted advice until total distance traveled is r.
3. Set $r \leftarrow 2r$ and repeat.

Optimize to bias toward consistency

Theorem. For nested convex body chasing, the switching algorithm is $(1+\delta)$-consistent & $O(d/\delta)$-robust.

"Best of both worlds": Black-box AI/ML imbued with robustness guarantee. Constant factor loss in robustness yields near-optimal consistency.

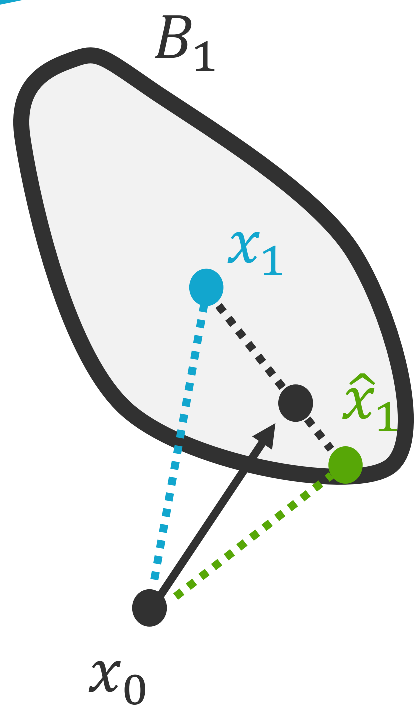
A Fundamental Limit

Theorem. For general convex body chasing, any switching algorithm that is robust must be at least 3-consistent.

Theorem. For nested convex body chasing, the switching algorithm is $(1+\delta)$-consistent & $O(d/\delta)$-robust.

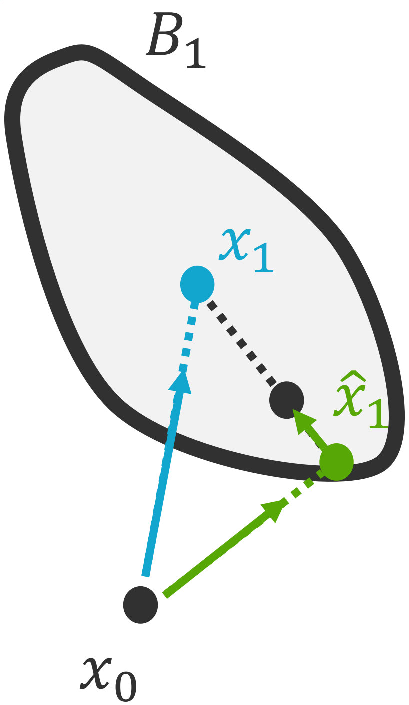
Attempt 2: Beyond switching algorithms

Adaptively choose a convex combination of the two advice points.



Attempt 2: Beyond switching algorithms

Adaptively choose a convex combination of the two advice points.



<u>Bicompetitive Line Chasing</u>

If $Cost(Alg)$ is much better than $Cost(Advice)$
then follow x_1.

Else, take a greedy step from $\hat{x}_1$ toward x_1
with step size depending on $Cost(Alg)$ and $Cost(Advice)$

Attempt 2: Beyond switching algorithms

Adaptively choose a convex combination of the two advice points.

Theorem. For general convex body chasing, the interpolation algorithm is $(\sqrt{2} + \delta)$-consistent & $O(d/\delta^2)$-robust.

Attempt 2: Beyond switching algorithms

Adaptively choose a convex combination of the two advice points.

<u>Theorem</u>. For general convex body chasing, the interpolation algorithm is $(\sqrt{2}+\delta)$-consistent & $O(d/\delta^2)$-robust.

Adding robustness means sacrificing performance of black-box AI.

Is this a fundamental limit?

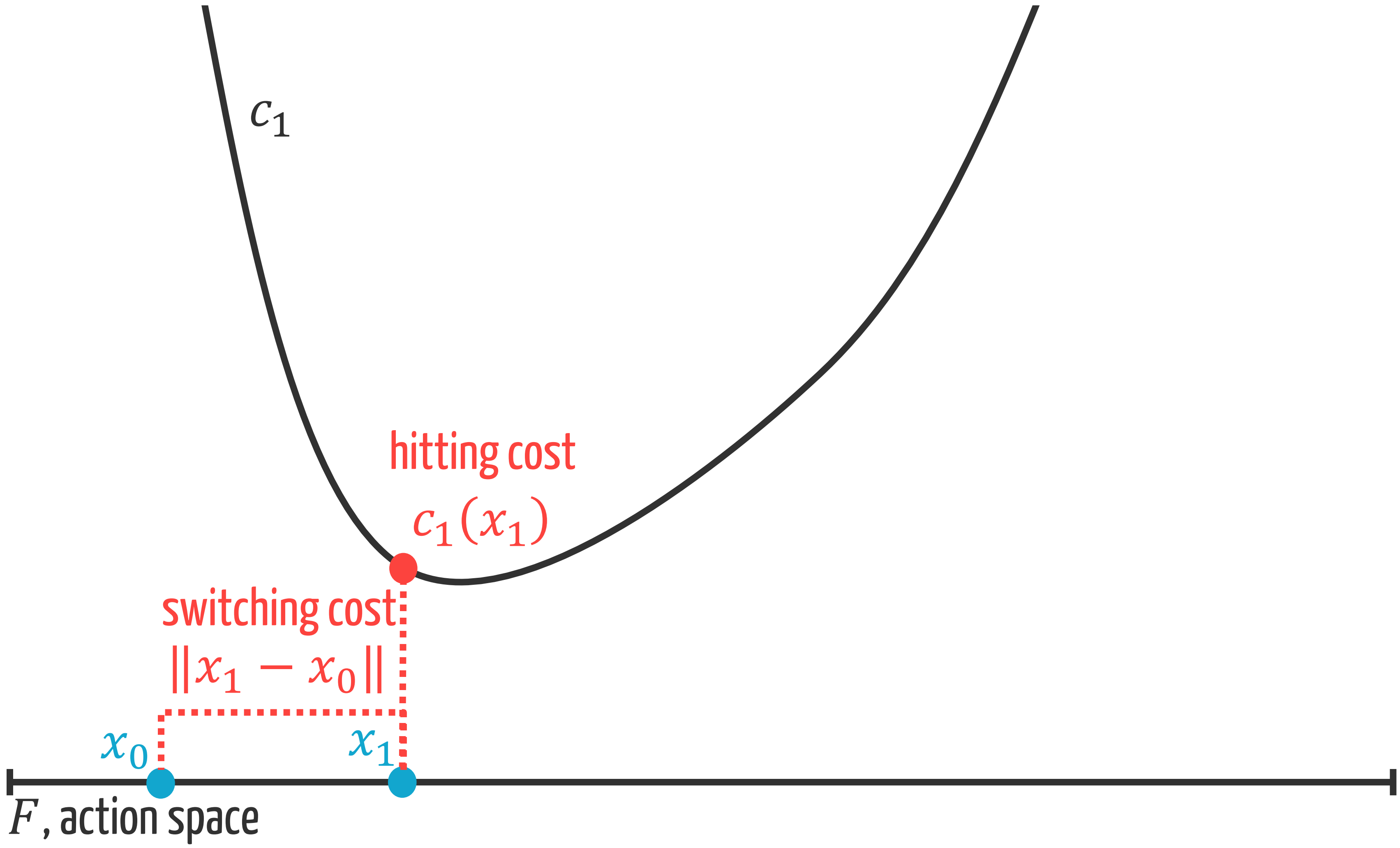
When should an algorithm "switch" between the trusted/untrusted advice?

Is it enough to always follow one or the other?

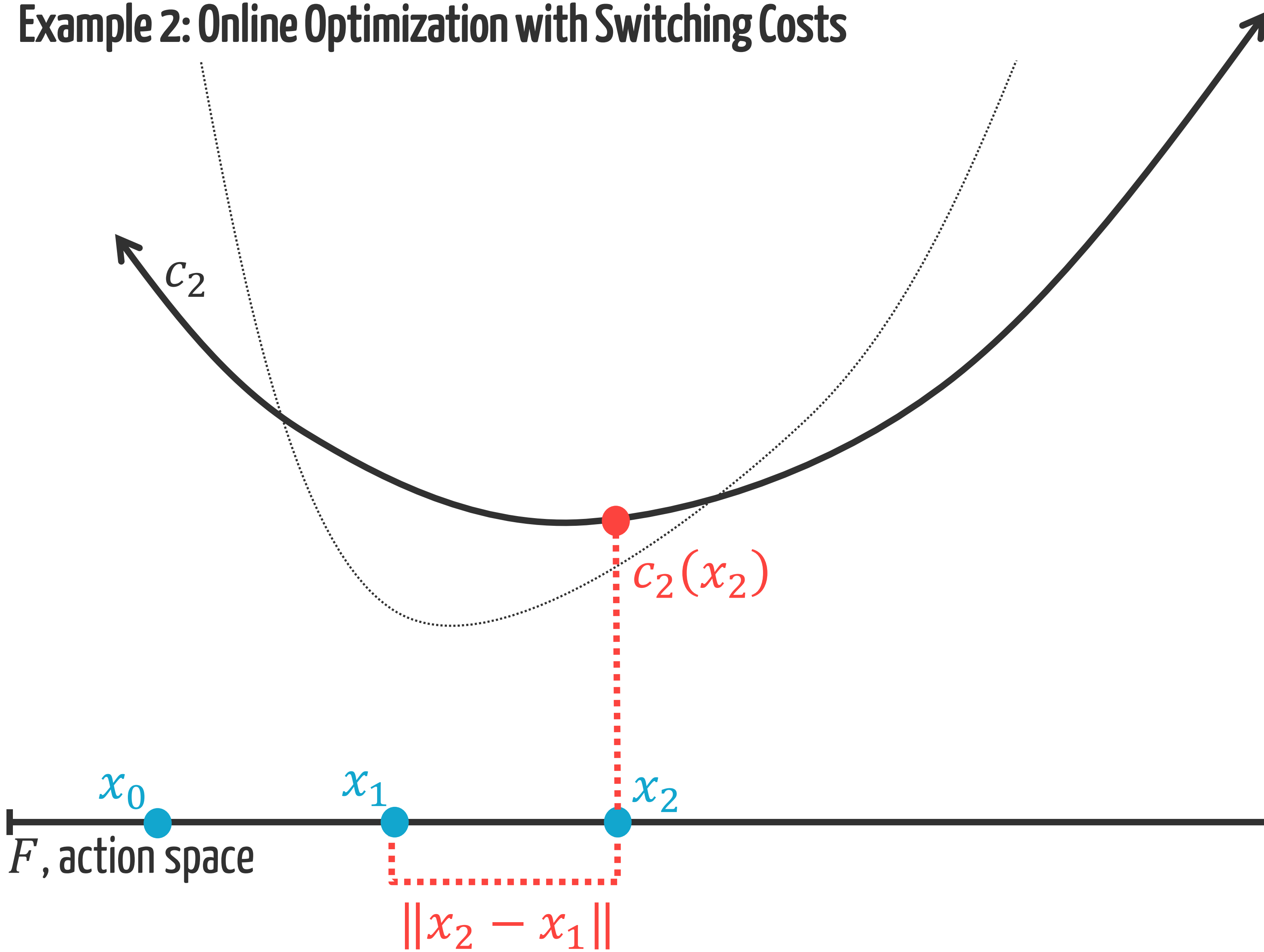
Memory was the key to knowing whether to follow advice or hedge...

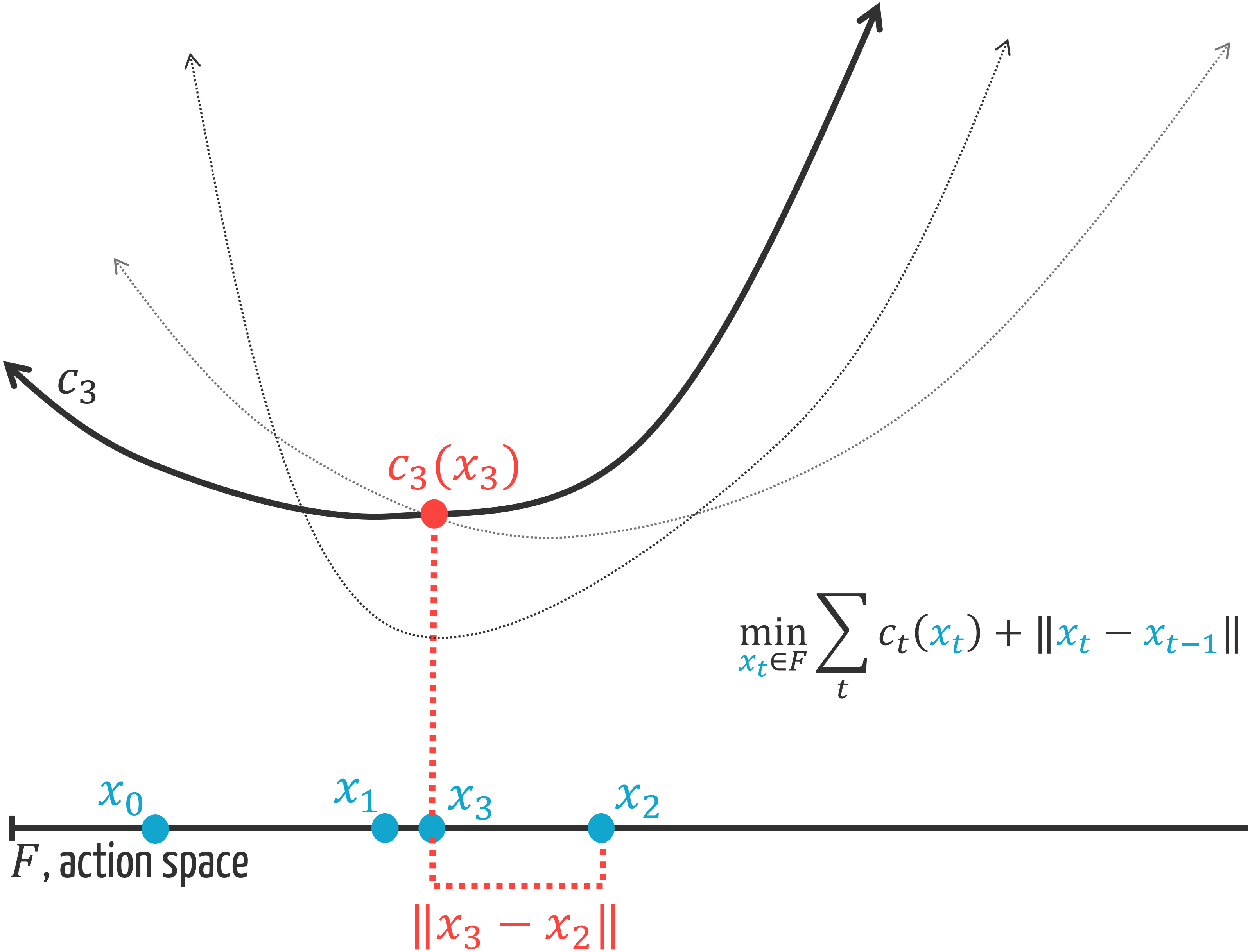
Is memory needed to benefit from untrusted advice?

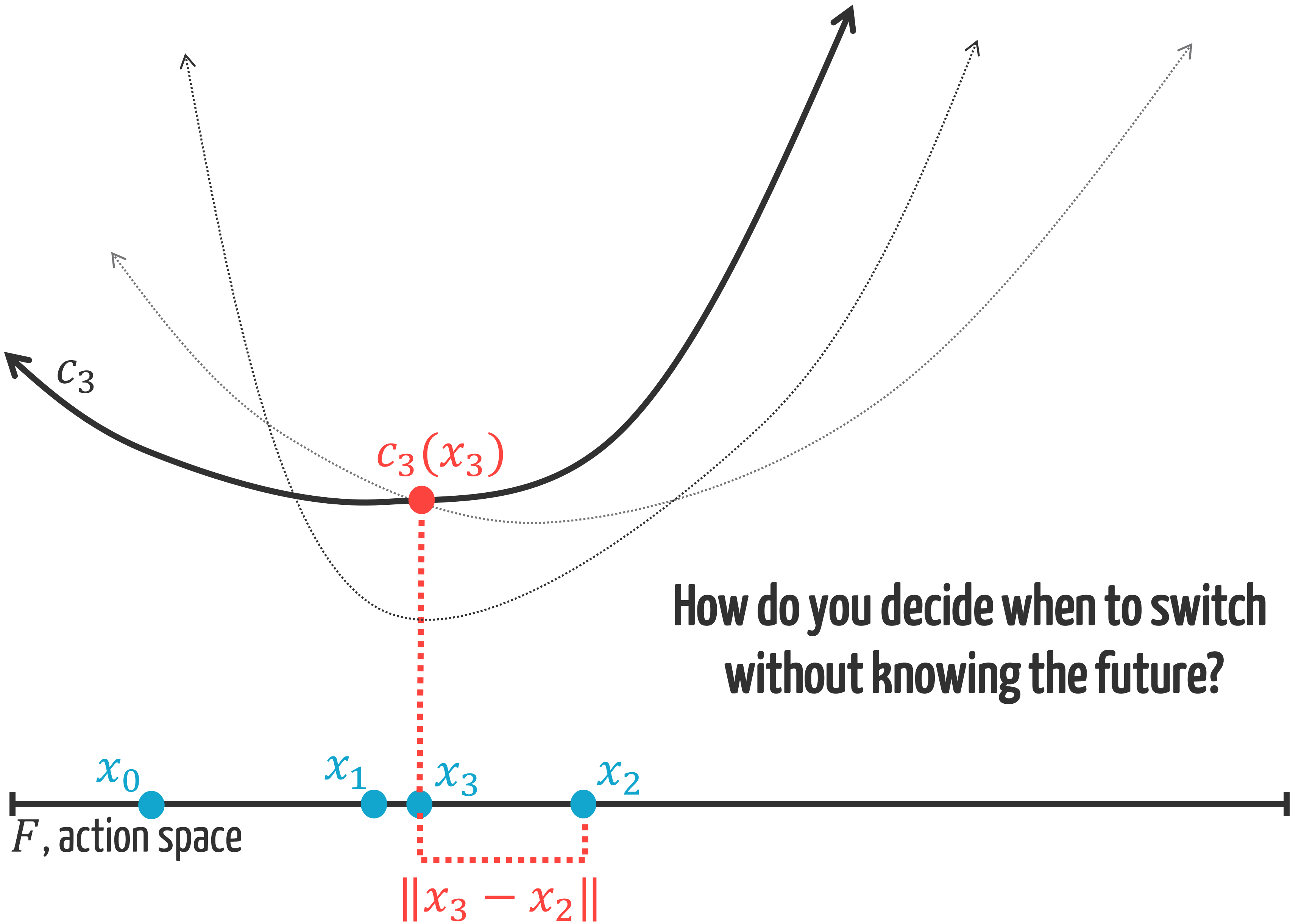
Example 2: Online Optimization with Switching Costs



Example 2: Online Optimization with Switching Costs




c_3
$c_3(x_3)$
$\min_{x_t \in F} \sum_t c_t(x_t) + \|x_t - x_{t-1}\|$
x_0
x_1
x_3
x_2
F, action space
$\|x_3 - x_2\|$



How do you decide when to switch without knowing the future?

Online optimization with switching costs has a huge literature at this point.

Varying applications in data centers, video streaming, EV charging, camera tracking, …

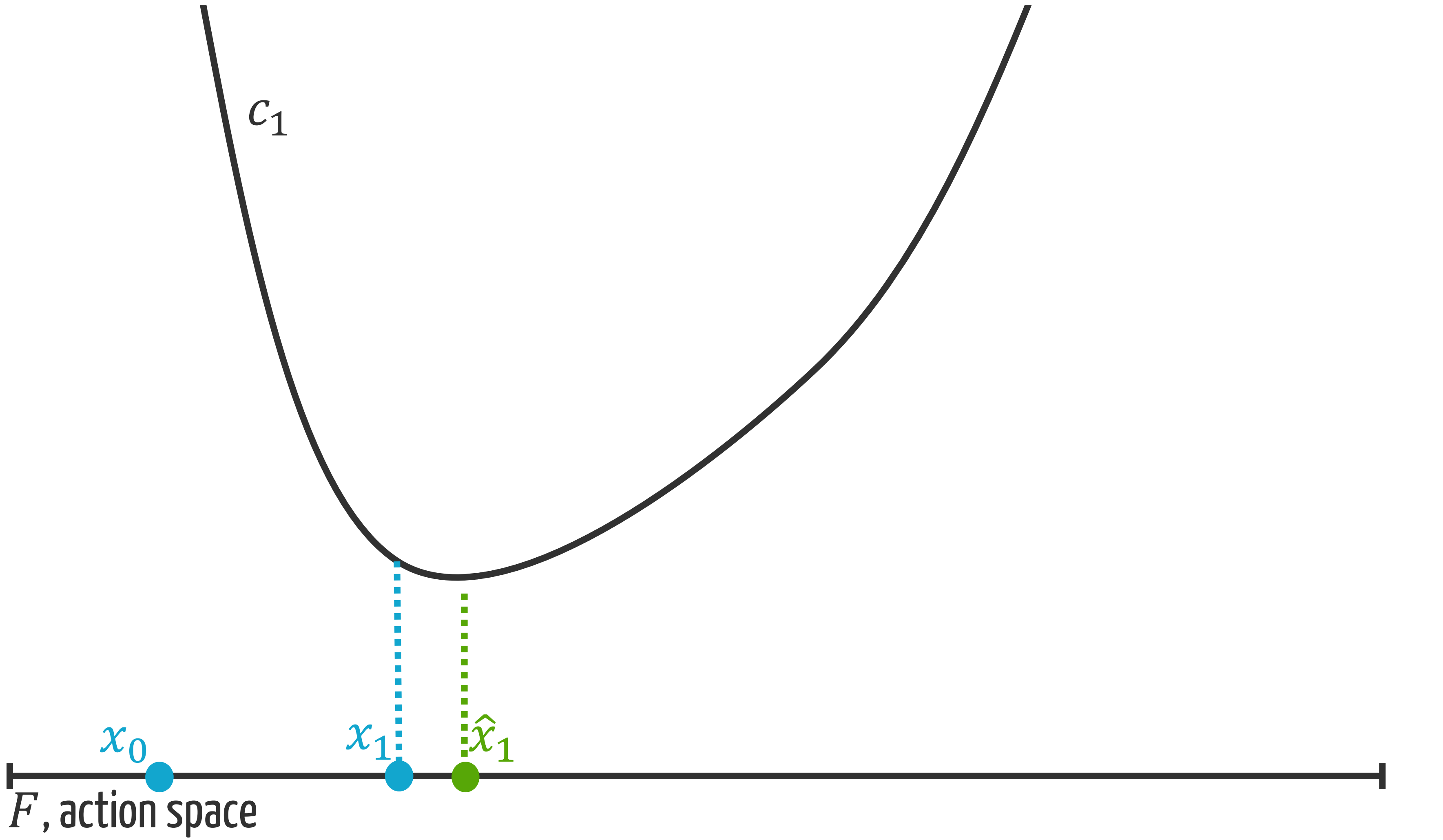
Exciting algorithmic progress in recent years [Chen et al 18], [Goel et al 19], [Goel & Wierman 19], [Li & Li 20], [Lin et al 20], [Zhang et al 21], …

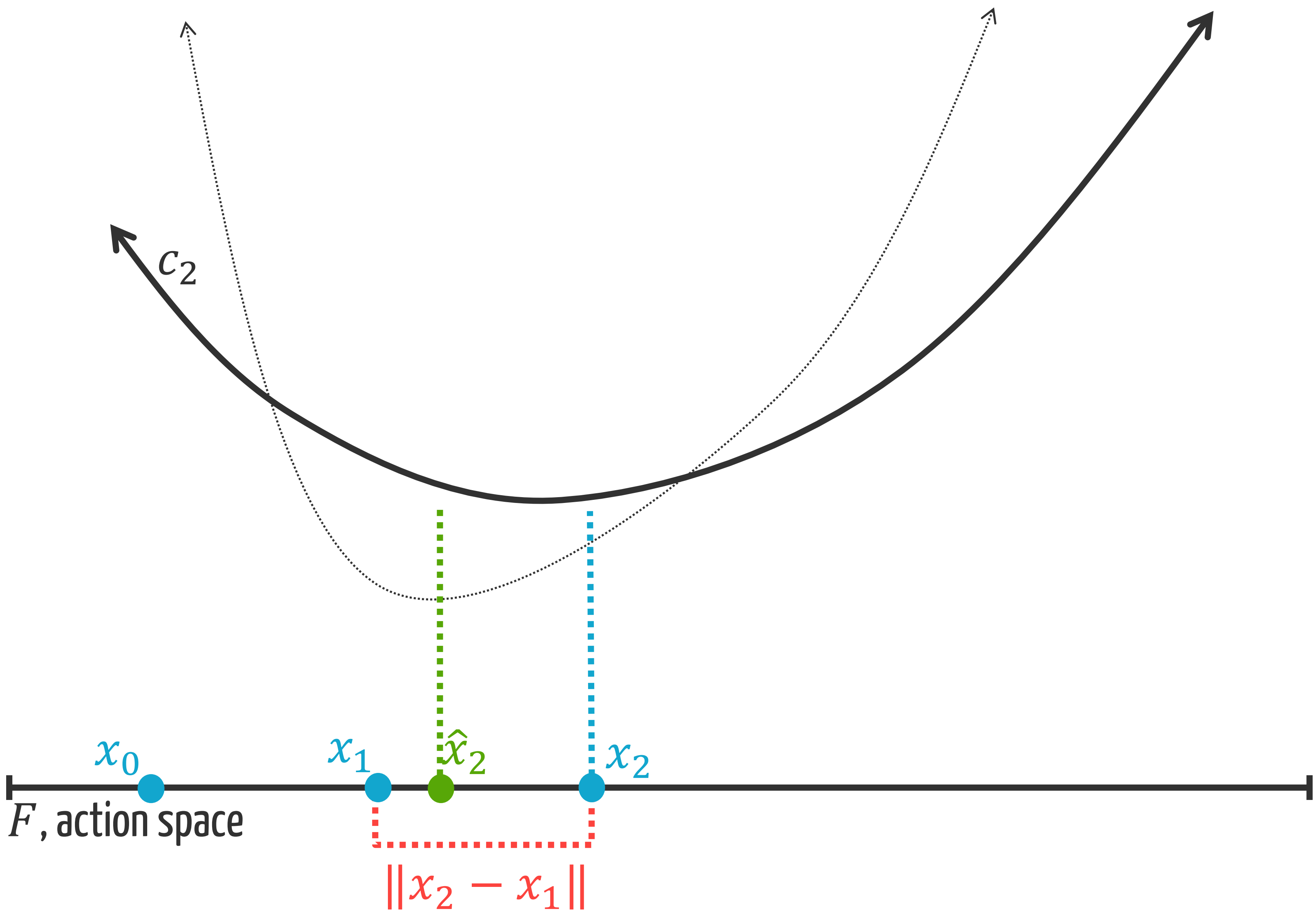
Theorem [Lin et al 20, Zhang et al 21]. Consider α-polyhedral cost functions.

- For online convex optimization with switching costs, a memoryless algorithm (ROBD) is $O(\sqrt{1/\alpha})$-competitive, and this is tight.
- For online non-convex optimization, a memoryless algorithm (Greedy) is $(1+2/\alpha)$-competitive.

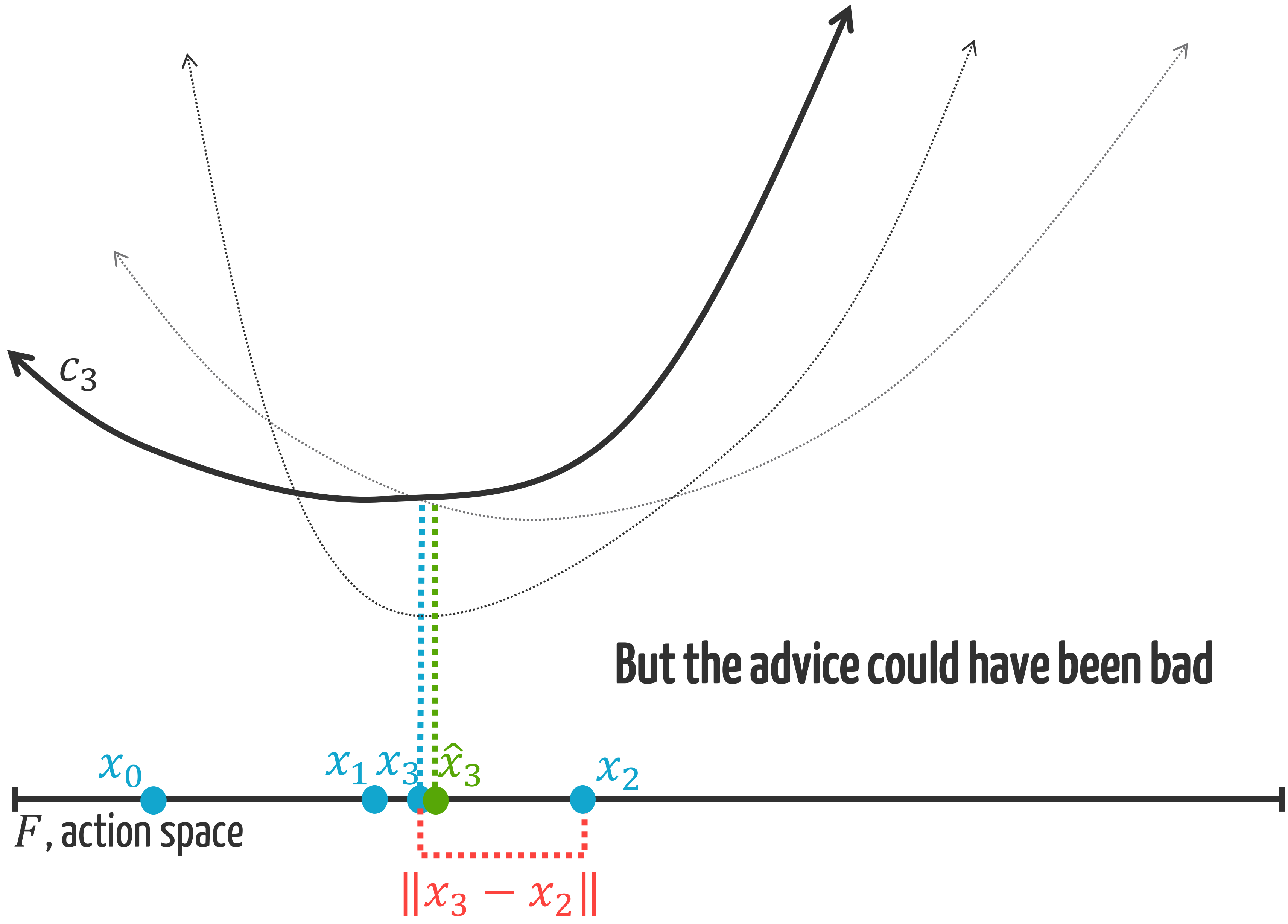
Choices of algorithm are quite conservative. Advice can help.

Example 2: Online Optimization with Switching Costs

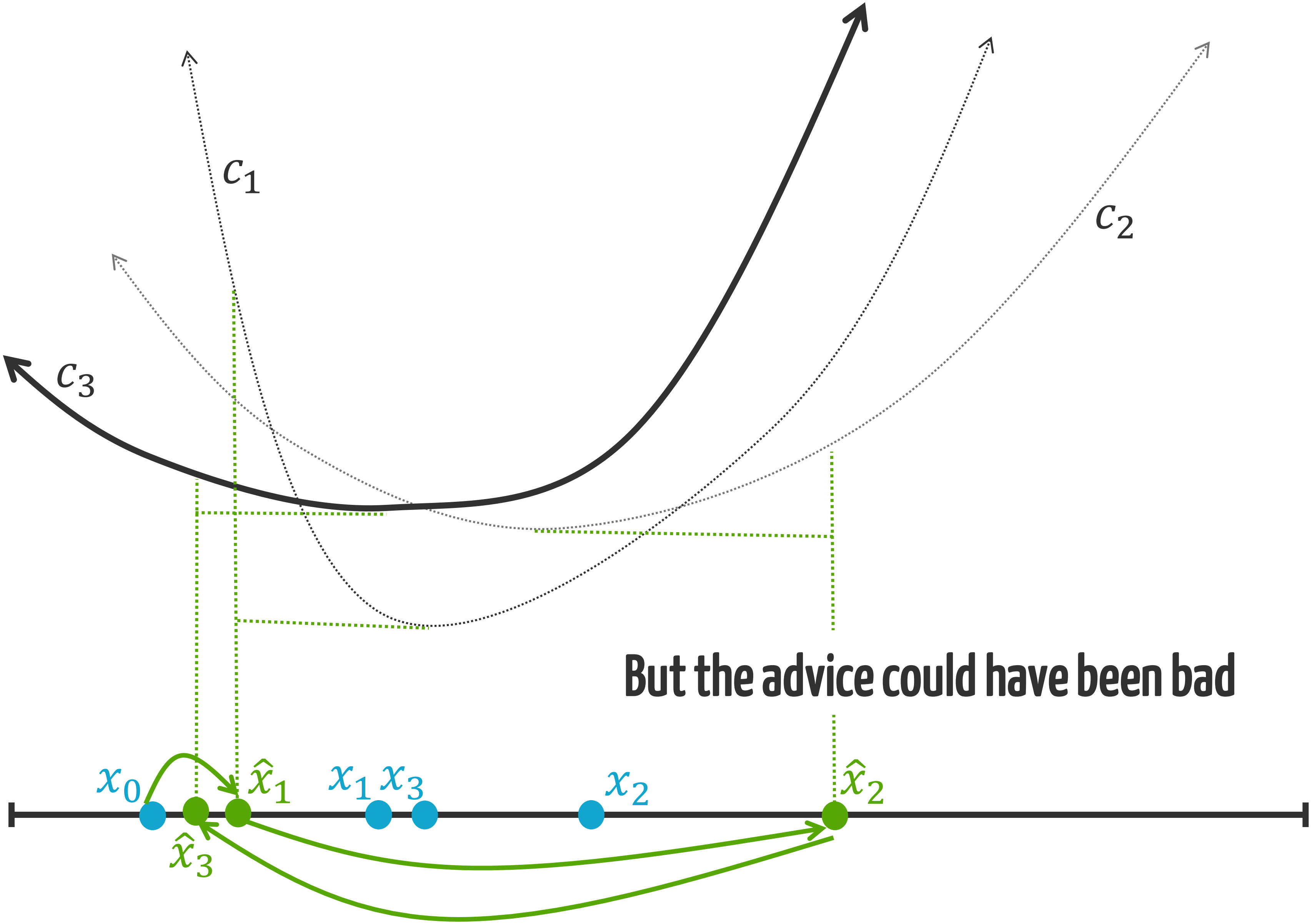




c_2
x_0
x_1
$\hat{x}_2$
x_2
F, action space
$\|x_2 - x_1\|$



c_3
But the advice could have been bad
x_0
x_1
x_3
$\hat{x}_3$
x_2
F, action space
$\|x_3 - x_2\|$



But the advice could have been bad

Attempt 1: A switching algorithm with memory

<u>Theorem.</u> For online non-convex optimization with α-polyhedral costs. Adaptive Online Switching (AOS) is

$$(1+2\delta)\text{-consistent \& } O\left(\frac{1}{\delta}\left(\frac{2}{\alpha+\delta(1+\alpha)}\right)^{\frac{2}{\alpha\delta}}\right)\text{-robust.}$$



Attempt 1: A switching algorithm with memory

Theorem. For online non-convex optimization with α-polyhedral costs. Adaptive Online Switching (AOS) is

$$\underbrace{(1+2\delta)\text{-consistent} \;\&\; O\left(\frac{1}{\delta}\left(\frac{2}{\alpha+\delta(1+\alpha)}\right)^{\frac{2}{\alpha\delta}}\right)\text{-robust.}}_{\text{Black-box AI/ML imbued with robustness guarantee.}}$$

Attempt 1: A switching algorithm with memory

<u>Theorem.</u> For online non-convex optimization with α-polyhedral costs. Adaptive Online Switching (AOS) is

$$(1+2\delta)\text{-consistent } \& \ O\left(\frac{1}{\delta}\left(\frac{2}{\alpha+\delta(1+\alpha)}\right)^{\frac{2}{\alpha\delta}}\right)\text{-robust.}$$

Exponential trade-off is provably necessary for any online algorithm.

Key Challenge: Algorithm sets confidence in advice (δ), which is must be chosen balance robustness & consistency.

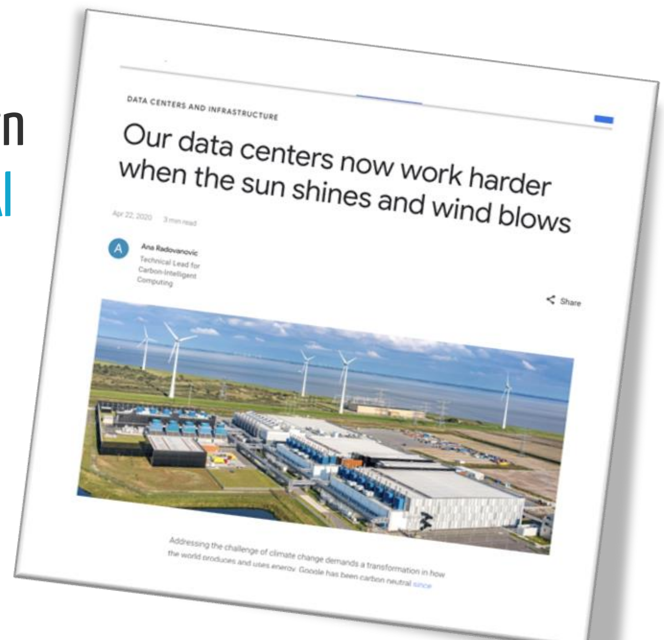
Attempt 1: A switching algorithm with memory

<u>Theorem.</u> For online non-convex optimization with α-polyhedral costs. Adaptive Online Switching (AOS) is

$(1+2\delta)$-consistent & $O\left(\frac{1}{\delta}\left(\frac{2}{\alpha+\delta(1+\alpha)}\right)^{\frac{2}{\alpha\delta}}\right)$-robust.

Example: Sustainable Data Center Design

We'll use AOS to combine a black-box AI with a trusted online algorithm.

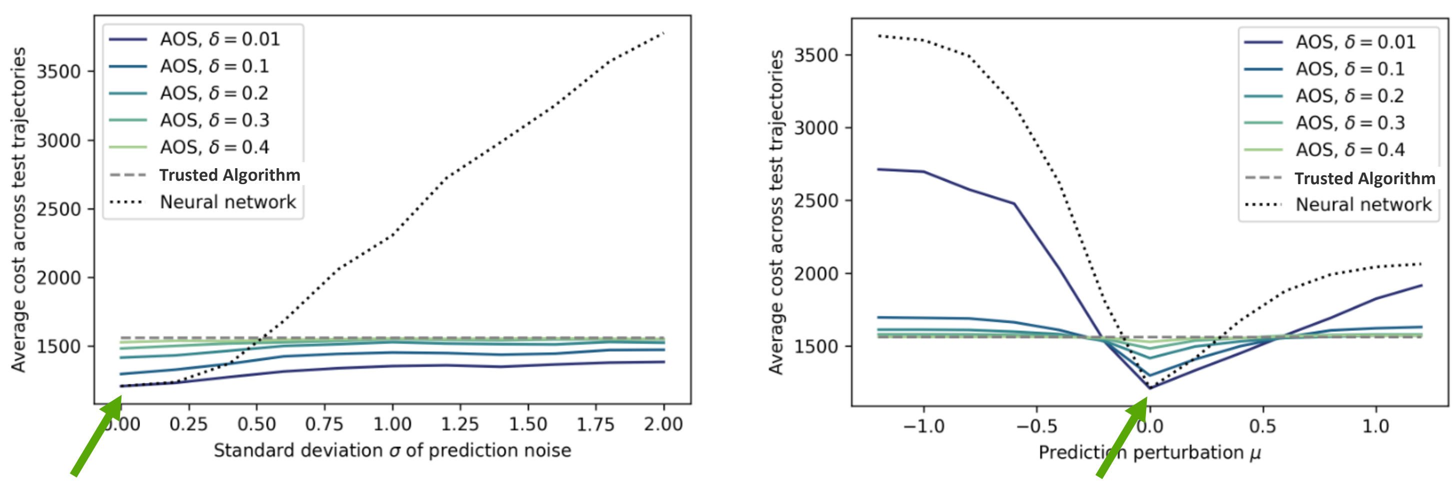


Attempt 1: A switching algorithm with memory

Theorem. For online non-convex optimization with α-polyhedral costs. Adaptive Online Switching (AOS) is

$$(1+2\delta)\text{-consistent \& } O\left(\frac{1}{\delta}\left(\frac{2}{\alpha+\delta(1+\alpha)}\right)^{\frac{2}{\alpha\delta}}\right)\text{-robust.}$$

Example: Sustainable Data Center Design

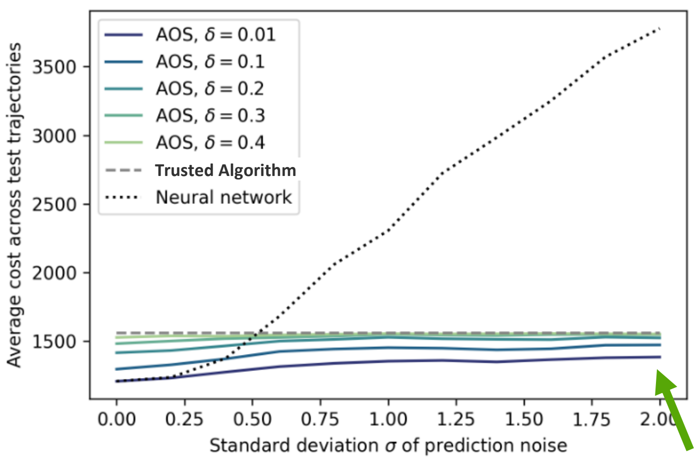


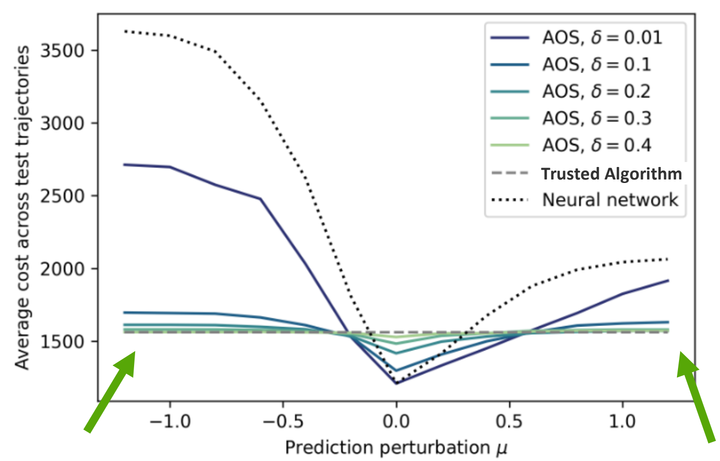
Attempt 1: A switching algorithm with memory

Theorem. For online non-convex optimization with α-polyhedral costs. Adaptive Online Switching (AOS) is

$$(1+2\delta)\text{-consistent \& } O\left(\frac{1}{\delta}\left(\frac{2}{\alpha+\delta(1+\alpha)}\right)^{\frac{2}{\alpha\delta}}\right)\text{-robust.}$$

Example: Sustainable Data Center Design





Attempt 1: A switching algorithm with memory

<u>Theorem.</u> For online non-convex optimization with α-polyhedral costs. Adaptive Online Switching (AOS) is

$$(1+2\delta)\text{-consistent } \& \ O\left(\frac{1}{\delta}\left(\frac{2}{\alpha+\delta(1+\alpha)}\right)^{\frac{2}{\alpha\delta}}\right)\text{-robust.}$$

Open Question: Can we adaptively learn to set confidence in predictions δ?

Example 3: In the Linear Quadratic Control setting we achieve this via "Follow the Leader."

Attempt 1: A switching algorithm with memory

Theorem. For online non-convex optimization with α-polyhedral costs. Adaptive Online Switching (AOS) is

$$(1+2\delta)\text{-consistent \& } O\left(\frac{1}{\delta}\left(\frac{2}{\alpha+\delta(1+\alpha)}\right)^{\frac{2}{\alpha\delta}}\right)\text{-robust.}$$

A Fundamental Limit: Cannot benefit from untrusted advice without memory.

Theorem. Consider $\mathbb{R}^2$. For online (non-)convex optimization with α-polyhedral costs, any memoryless algorithm cannot be both

- γ-robust, for $\gamma < \infty$,
- $(1+\delta)$-consistent for $\delta < \frac{1}{\sqrt{8\alpha}}$.

achievable without advice

When is it possible for a memoryless algorithm to benefit from untrusted advice?

<u>A Fundamental Limit</u>: Cannot benefit from untrusted advice without memory.

<u>Theorem</u>. Consider $\mathbb{R}^2$. For online (non-)convex optimization with α-polyhedral costs, any memoryless algorithm cannot be both

- γ-robust, for $\gamma < \infty$,
- $(1 + \delta)$-consistent for $\delta < \frac{1}{\sqrt{8\alpha}}$.

When is it possible for a memoryless algorithm to benefit from untrusted advice?

<u>Theorem.</u> For online convex optimization in one dimension. A memoryless algorithm, Adaptive Online Balanced Descent, is

$(1+\delta)$-consistent & $O(1/\delta^2)$-robust.



When is it possible for a memoryless algorithm to benefit from untrusted advice?

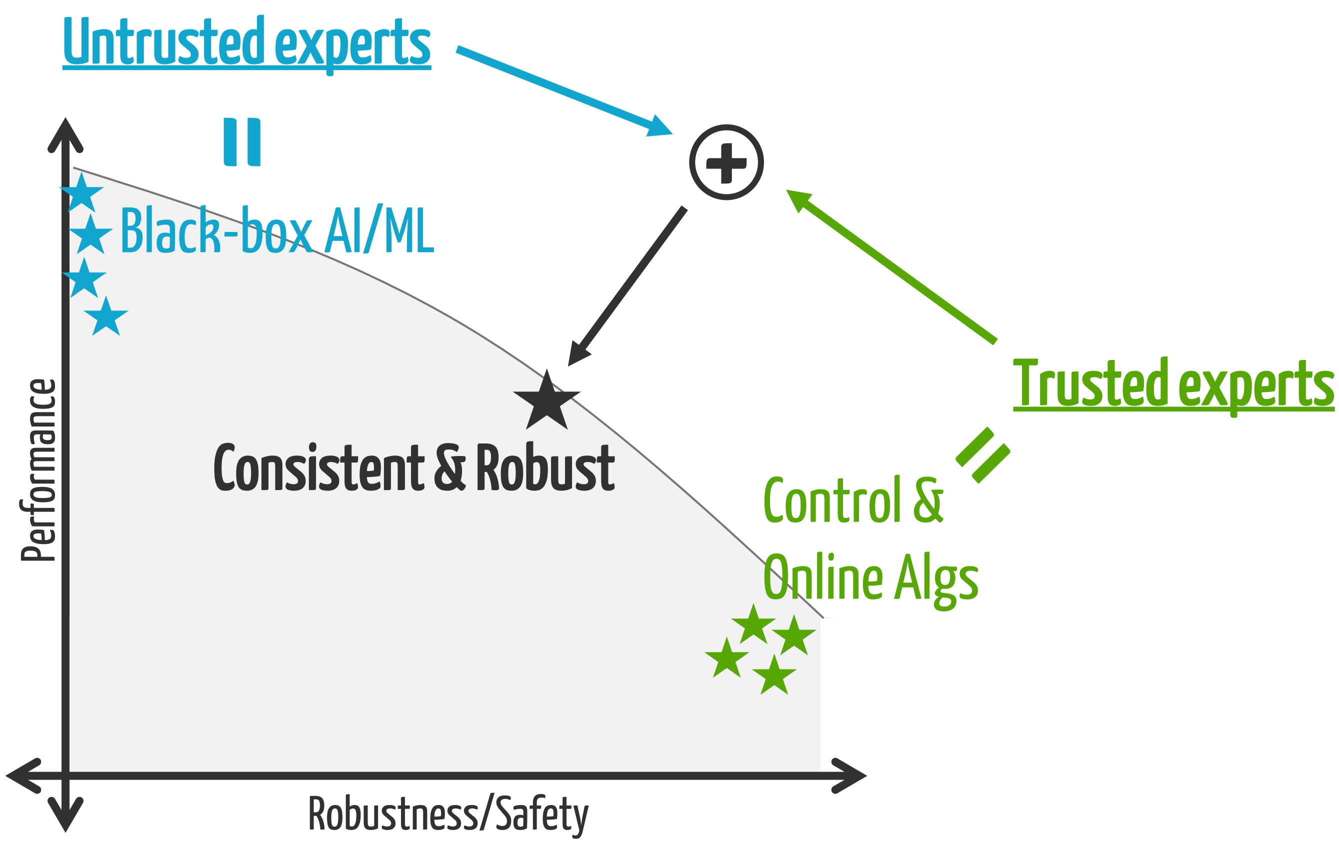
<u>Theorem.</u> For online convex optimization in one dimension. A memoryless algorithm, Adaptive Online Balanced Descent, is

$(1 + \delta)$-consistent & $O(1/\delta^2)$-robust.

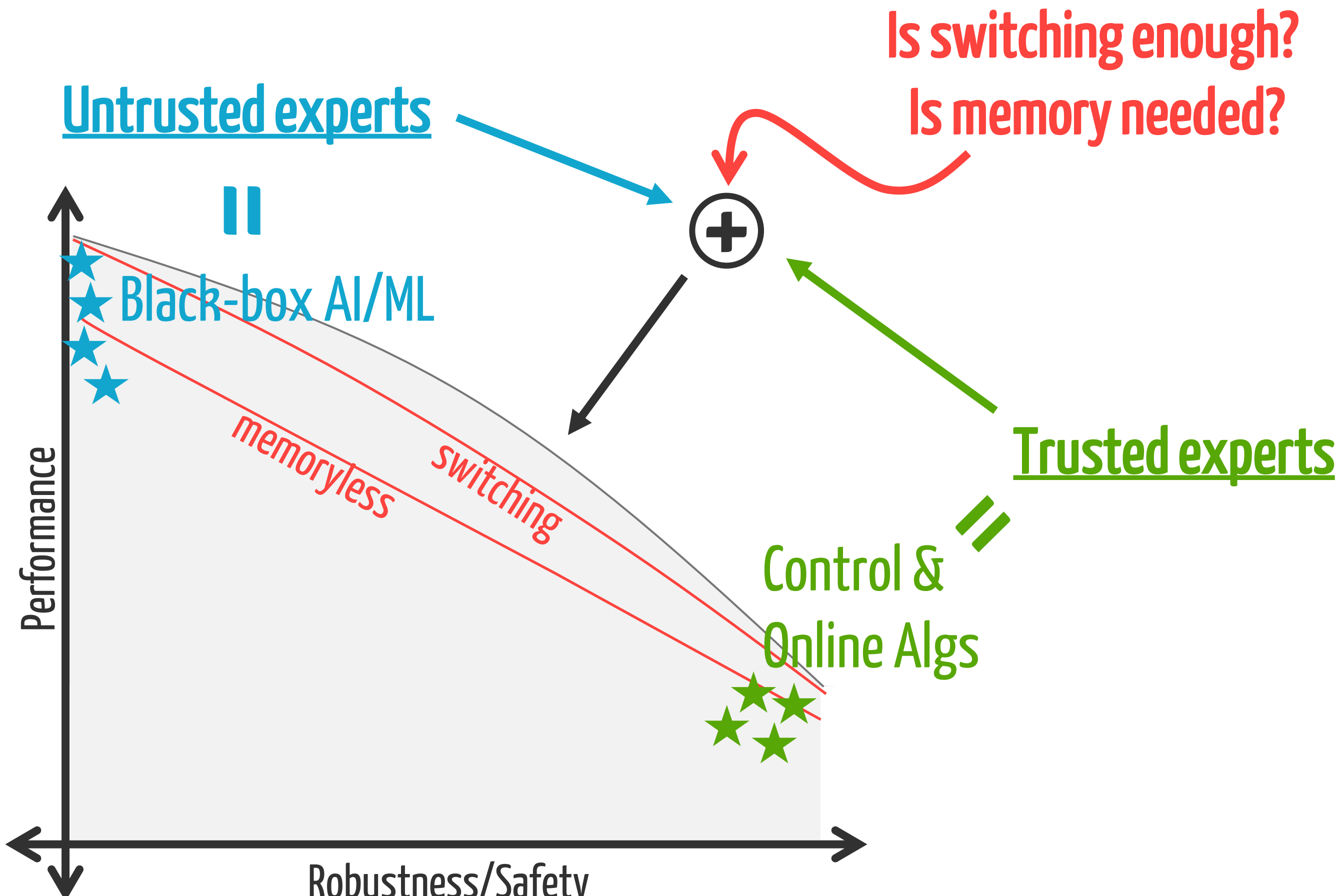
Lower bound: $O(1/\delta)$

Is memory needed to achieve lower bound?

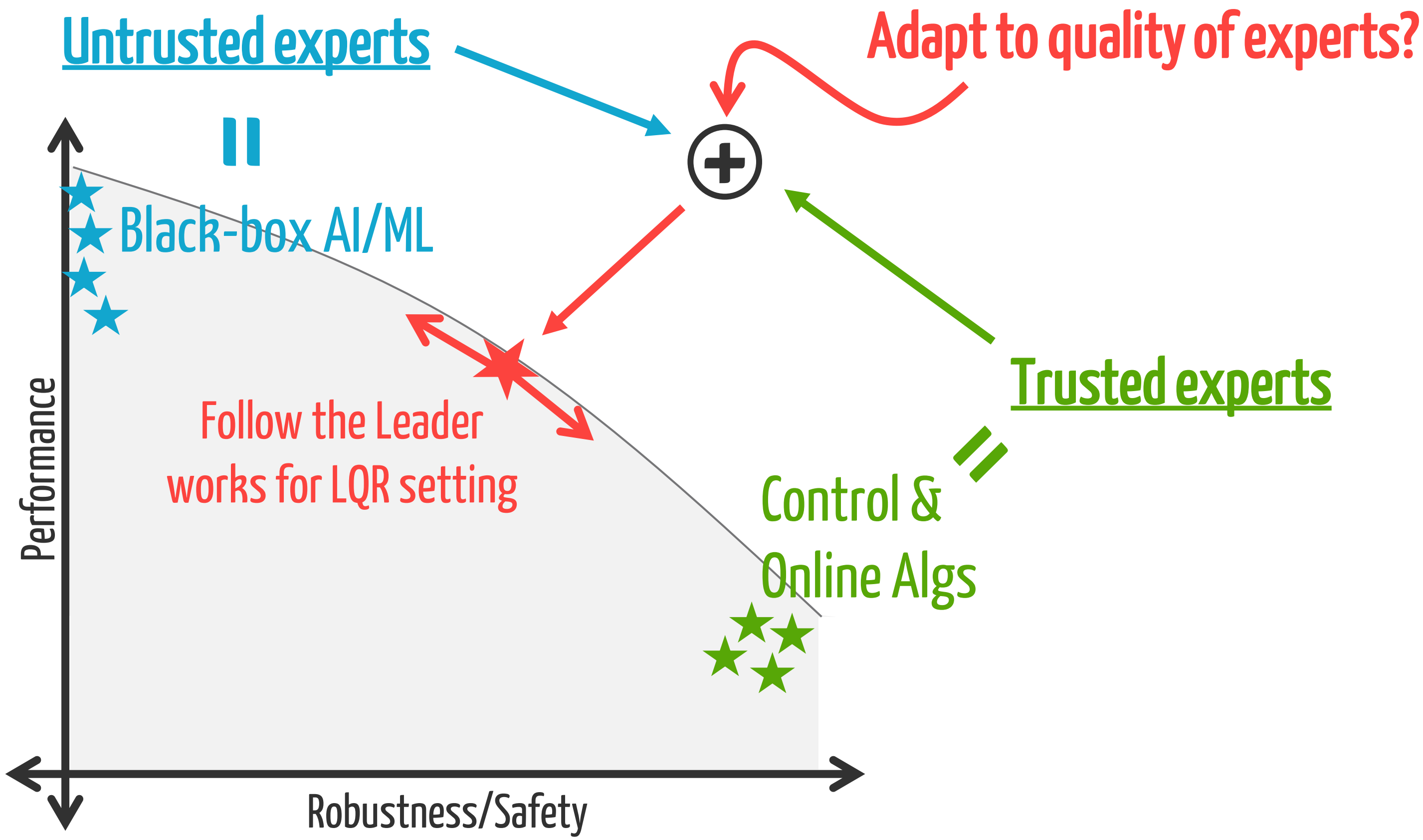
Today: A "black box" approach for getting the best of both worlds



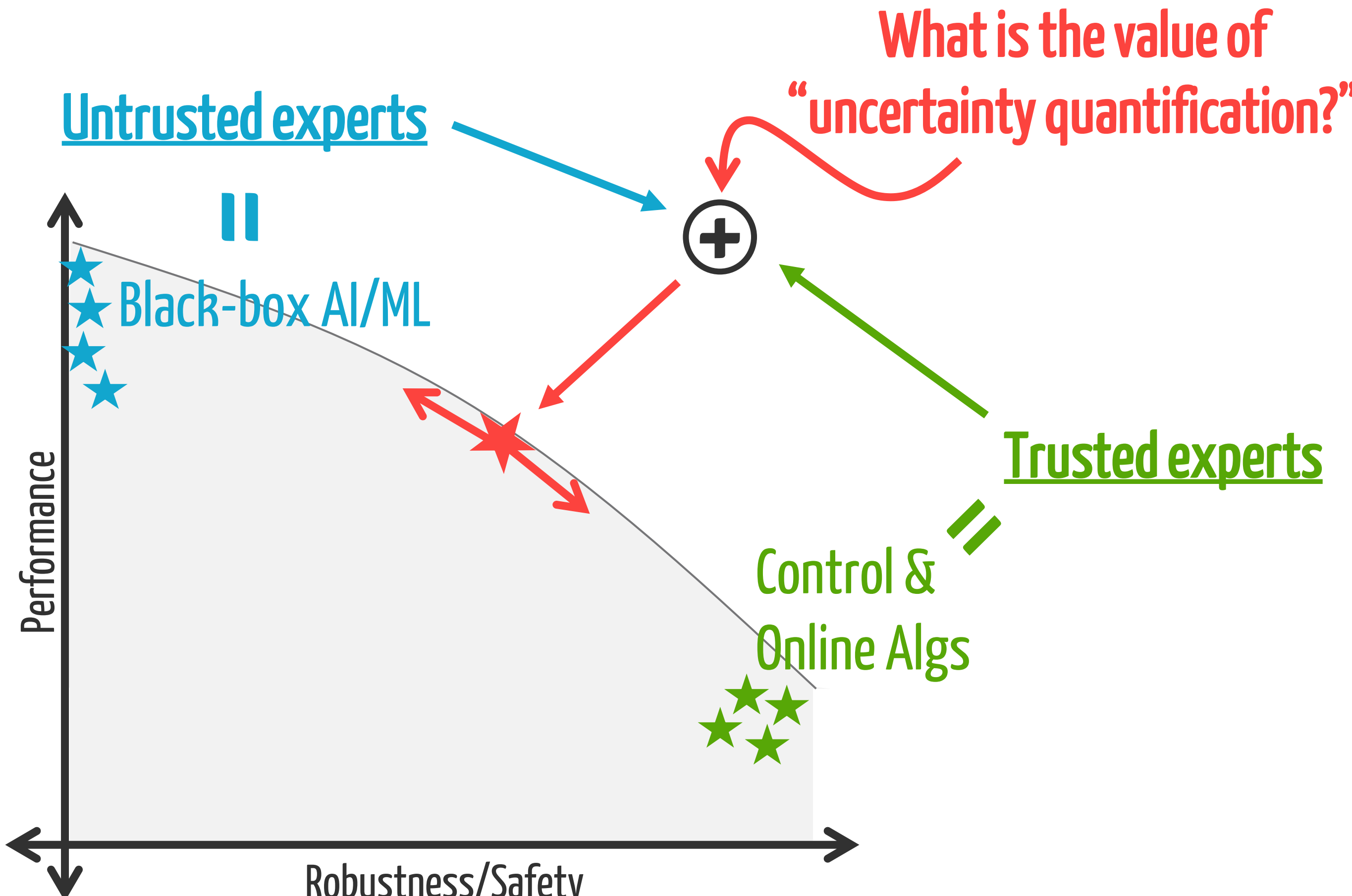
Today: A "black box" approach for getting the best of both worlds



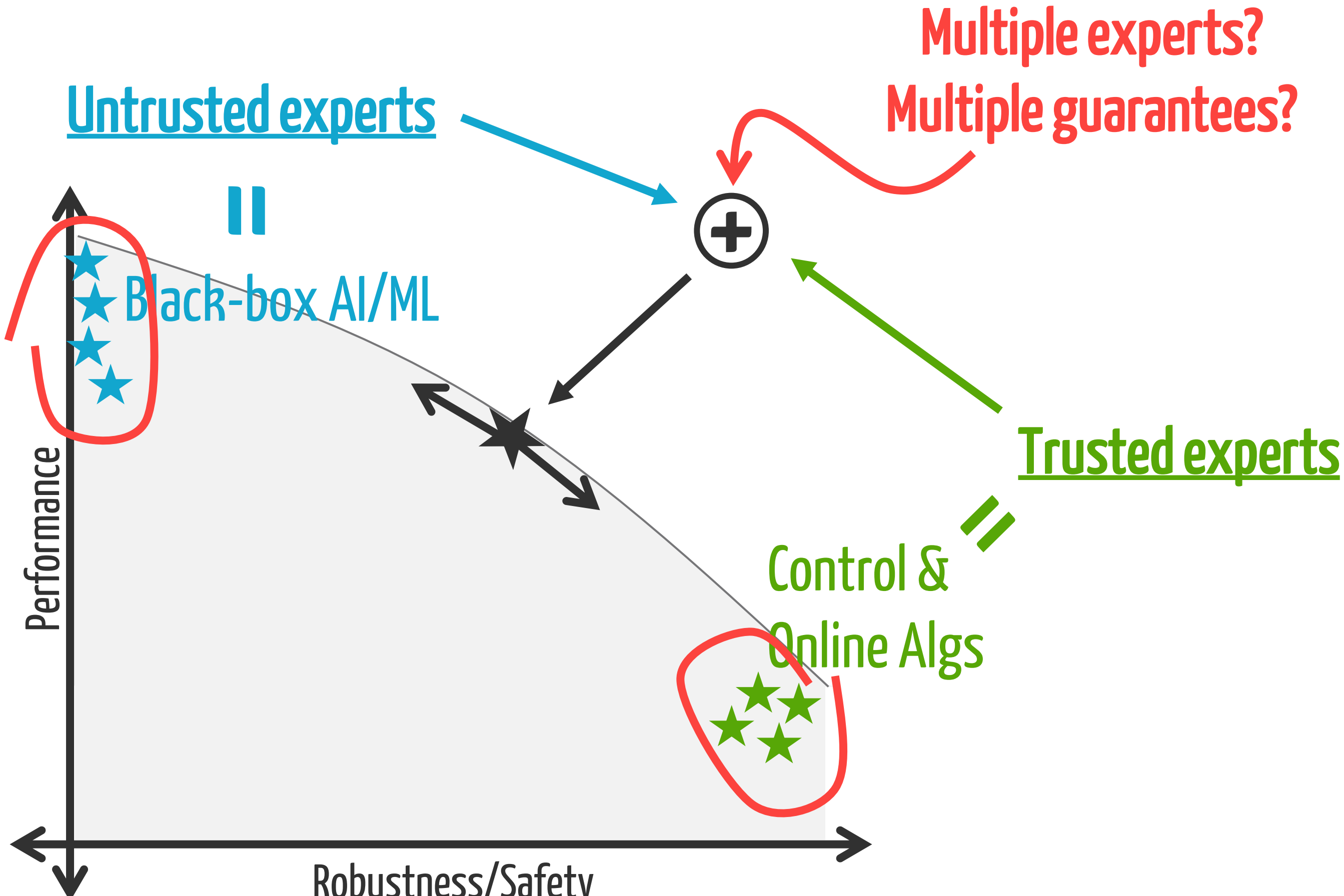
Today: A "black box" approach for getting the best of both worlds



Today: A "black box" approach for getting the best of both worlds



Today: A "black box" approach for getting the best of both worlds



Online decision making with untrusted advice is a promising framework with many exciting open problems!

~~Three~~ Two running examples:

- Convex Body Chasing
- Online Optimization with Switching Costs
- ~~Linear Quadratic Control~~

Many other applications are of interest!

A general view of design/analysis is just now emerging.

Online Optimization and Control using Black-Box Advice

Adam Wierman, Caltech

Papers from our group on “learning augmented” online decision making:

- A Zeynali, B Sun, M Hajiesmaili, A Wierman. Data-driven Competitive Algorithms for Online Knapsack and Set Cover. AAAI 2021.
- B Sun, R Lee, M Hajiesmaili, A Wierman, D. Tsang. Pareto-Optimal Learning-Augmented Algorithms for Online Conversion Problems. NeurIPS 2021.
- Y Su, J Yu, V Anand, A Wierman. Learning-Augmented Energy Aware Scheduling of Precedence-Constrained Tasks. MAMA workshop at Sigmetrics 2021.
- T Li, R Yang, G Qu, G Shi, C Yu, A Wierman, S Low. Robustness and Consistency in Linear Quadratic Control with Untrusted Predictions. Sigmetrics 2022
- D Rutten, N Christianson, D Mukherjee, A Wierman. Online Non-convex Optimization with Untrusted Advice. Under submission.
- N Christianson, T Handina, A Wierman. Chasing Convex Bodies and Functions with Black-Box Advice. Under submission.