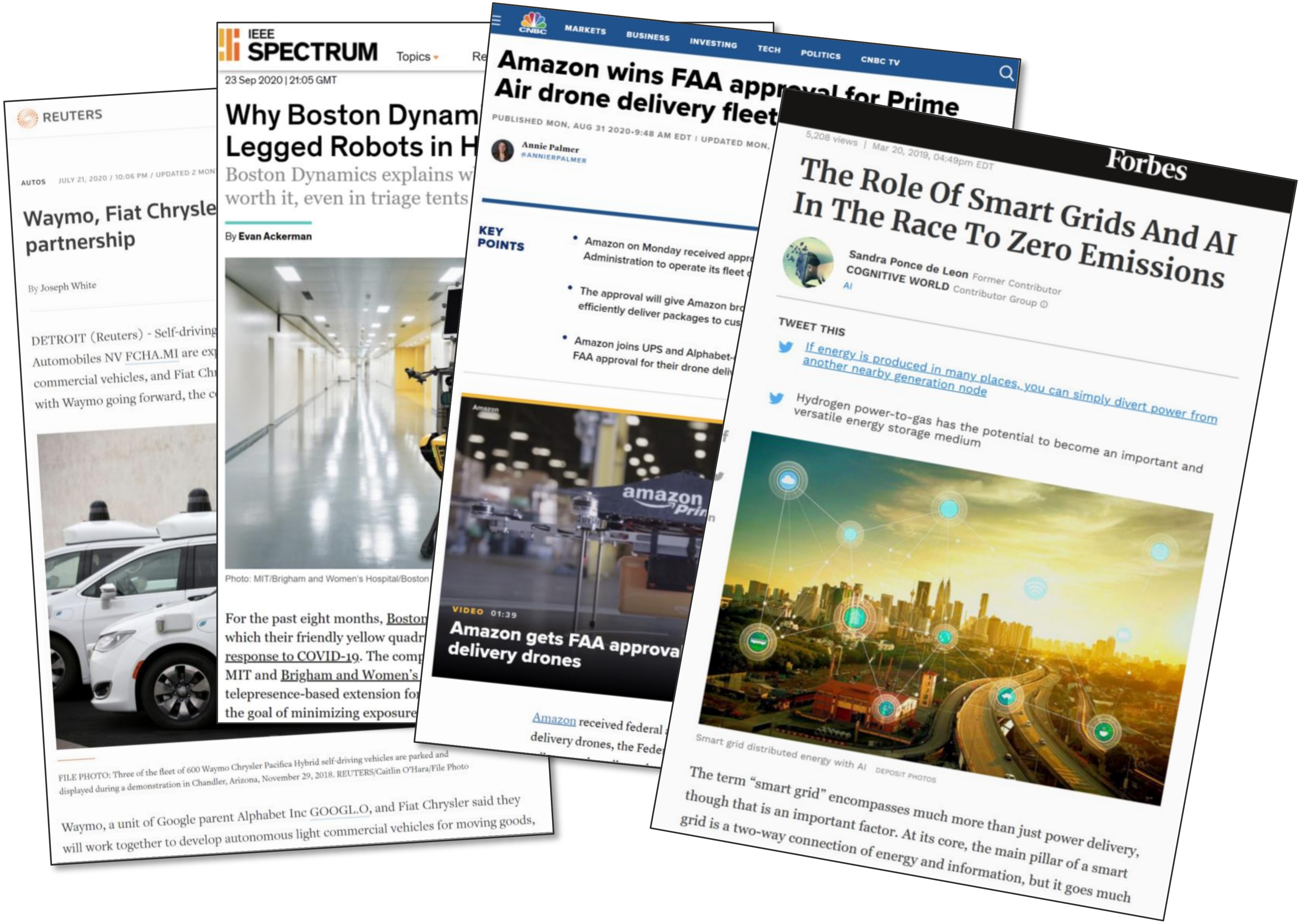
Competitive Control via Online Optimization

Adam Wierman, Caltech

Excitement about the potential of "learning to control" is growing...



REUTERS

AUTOS JULY 21, 2020 / 10:06 PM / UPDATED 2 MON

Waymo, Fiat Chrysle
partnership

By Joseph White

DETROIT (Reuters) - Self-driving
Automobiles NV FCHA.MI are ex
commercial vehicles, and Fiat Ch
with Waymo going forward, the c

FILE PHOTO: Three of the fleet of 600 Waymo Chrysler Pacifica Hybrid self-driving vehicles are parked and displayed during a demonstration in Chandler, Arizona, November 29, 2018. REUTERS/Caitlin O'Hara/File Photo

Waymo, a unit of Google parent Alphabet Inc GOOGL.O, and Fiat Chrysler said they will work together to develop autonomous light commercial vehicles for moving goods,

---

IEEE SPECTRUM Topics Re

23 Sep 2020 | 21:05 GMT

Why Boston Dynam
Legged Robots in H

Boston Dynamics explains w
worth it, even in triage tents

By **Evan Ackerman**

Photo: MIT/Brigham and Women's Hospital/Boston

For the past eight months, Boston
which their friendly yellow quadr
response to COVID-19. The comp
MIT and Brigham and Women's
telepresence-based extension fo
the goal of minimizing exposure

CNBC MARKETS BUSINESS INVESTING TECH POLITICS CNBC TV

Amazon wins FAA appr for Prime
Air drone delivery fleet

PUBLISHED MON, AUG 31 2020•9:48 AM EDT | UPDATED MON,

Annie Palmer
@ANNIERPALMER

KEY POINTS

- Amazon on Monday received appr
Administration to operate its fleet
- The approval will give Amazon br
efficiently deliver packages to cus
- Amazon joins UPS and Alphabet-
FAA approval for their drone deli

VIDEO 01:39

Amazon gets FAA approva
delivery drones

Amazon received federal
delivery drones, the Fede

Forbes

5,208 views | Mar 20, 2019, 04:49pm EDT

The Role Of Smart Grids And AI
In The Race To Zero Emissions

Sandra Ponce de Leon Former Contributor
COGNITIVE WORLD Contributor Group
AI

TWEET THIS

If energy is produced in many places, you can simply divert power from another nearby generation node

Hydrogen power-to-gas has the potential to become an important and versatile energy storage medium

Smart grid distributed energy with AI DEPOSIT PHOTOS

The term "smart grid" encompasses much more than just power delivery, though that is an important factor. At its core, the main pillar of a smart grid is a two-way connection of energy and information, but it goes much

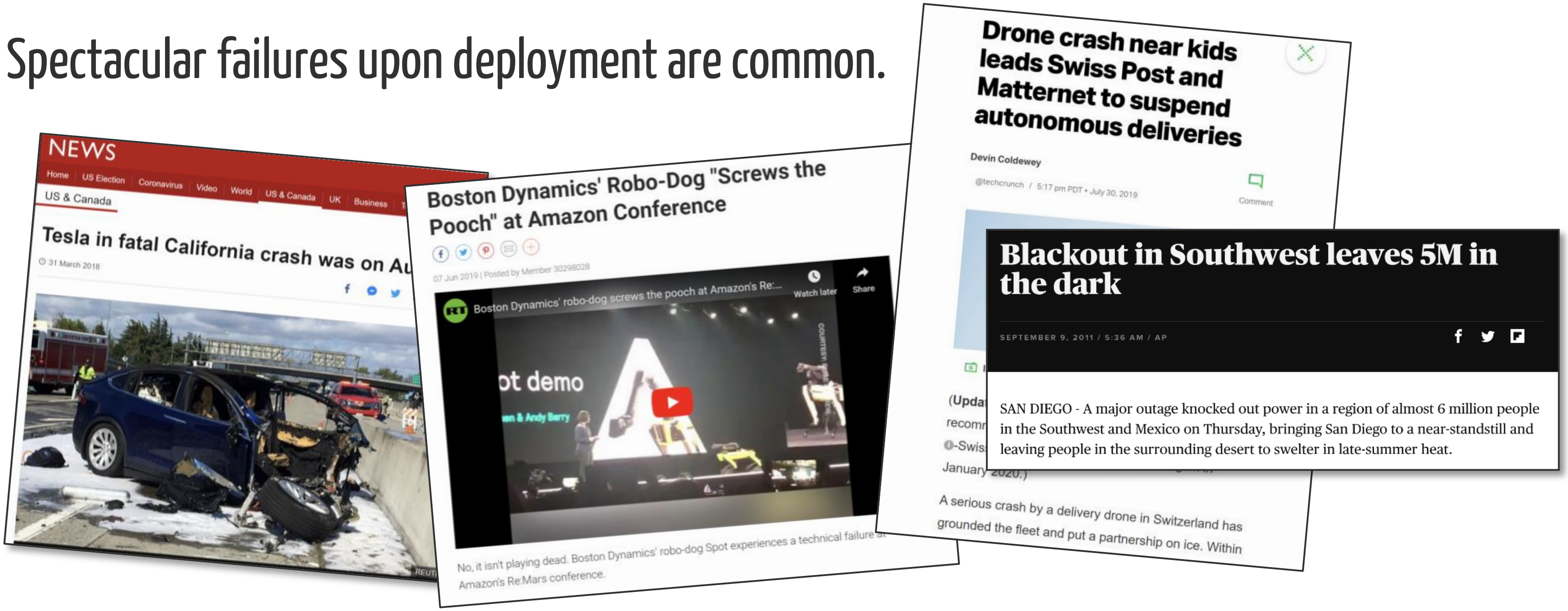
...but realizing the potential is proving to be a challenge.

Predicted arrivals have long passed.

Old Promises Broken, Musk Offers New Pledges on Self-Driving

Ahead of what's expected to be a disappointing earnings report, Tesla CEO Elon Musk predicts 1 million "robotaxis" on the road by next year.

Spectacular failures upon deployment are common.



Tesla in fatal California crash was on Au

31 March 2018

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.

Drone crash near kids leads Swiss Post and Matternet to suspend autonomous deliveries

Devin Coldewey

@techcrunch / 5:17 pm PDT • July 30, 2019

(Upda

recomm

-Swis

January 2020.)

A serious crash by a delivery drone in Switzerland has grounded the fleet and put a partnership on ice. Within

Blackout in Southwest leaves 5M in the dark

SEPTEMBER 9, 2011 / 5:36 AM / AP

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.

Scalable, trustworthy, and predictable control is required.

We’re not there yet.

"Learning & Control" is emerging as a rich, impactful field

LEARNING FOR DYNAMICS & CONTROL (L4DC)

Online June 11-12th, 2020

We're excited to announce that the 3rd L4DC
Further information can be found on the we

CONFERENCE L

L4DC 2020 Livestream (Friday ...

Workshops

Programs > Workshops > Intersections between Control, Learning and Optimization

Intersections between C arning and Optimization

KER LIST

SCHEDULE

ve always been
ays. Optimization
ns in control and
perspectives on
de relationships

Control Meets Learning

Virtual Seminar Series on the Intersection of Control and Learning

The seminar series is broadly focused on the intersection of control and machine learning. It covers a broad range of topics including, but not limited to, learning for dynamical systems, online learning and control, reinforcement learning, control-theoretic perspectives on deep learning, and applications to various real-world systems.

Wednesday 9 a.m. – 10 a.m. (Pacific Time)

Zoom link: Please subscribe to the below Google group to receive Zoom link and future announcements.

Youtube live stream: Sept 30 live streaming

Add Google Calendar

Join Google Group

Core tenet: learning and control have complementary philosophies

Data to action (model-free) ↔ Model to action (model-based)

Mitigate uncertainty by learning from the past ↔ Mitigate uncertainty via feedback

Both philosophies have been successful, **how do we combine them?**

Today

How can ML predictions be integrated into control & autonomy?

Can ML predictions be combined with MPC to improve control in face of time-varying environment, model error, delayed observations, ...

How can ML tools help improve the robustness & efficiency of control?

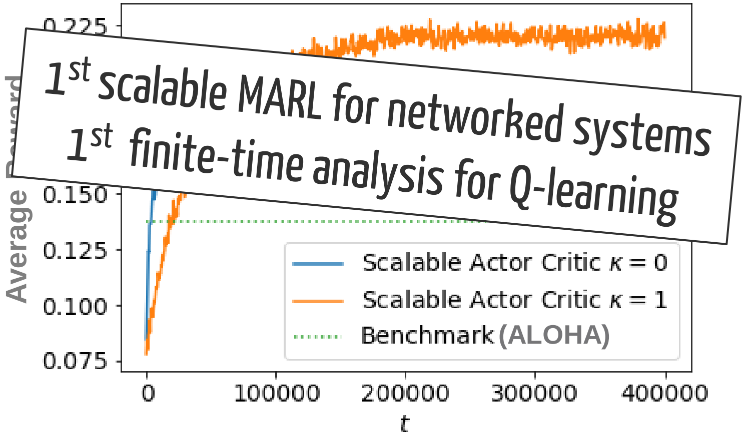
Can tools such as adversarial analysis, finite-time or single trajectory bounds, and general loss functions lead to improvements?

Can model-free and model-based approaches be combined to obtain the best of both worlds?

How much do you need to "understand" about a system to control it? Can we bring scalability and robustness to model-free RL?

Is it possible to exploit network structure to design scalable multi-agent RL (MARL)?

Example: Multi-access Wireless

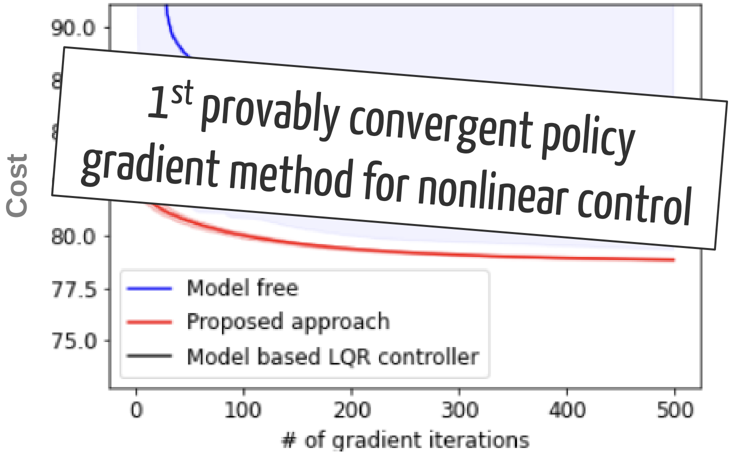


1st scalable MARL for networked systems

1st finite-time analysis for Q-learning

Can LQR be combined with model-free methods to control non-linear systems?

Example: Inverted Pendulum



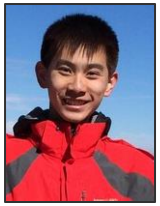
1st provably convergent policy gradient method for nonlinear control

Can model-free and model-based approaches be combined to obtain the best of both worlds? How much do you need to "understand" about a system to control it? Can we bring scalability and robustness to model-free RL?

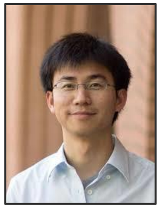
Guannan Qu

Yiheng Lin

Na Li



Chenkai Yu



Longbo Huang

Today

How can ML predictions be integrated into control & autonomy?

Can ML predictions by combined with MPC to improve control in face of time-varying environment, model error, delayed observations, ...

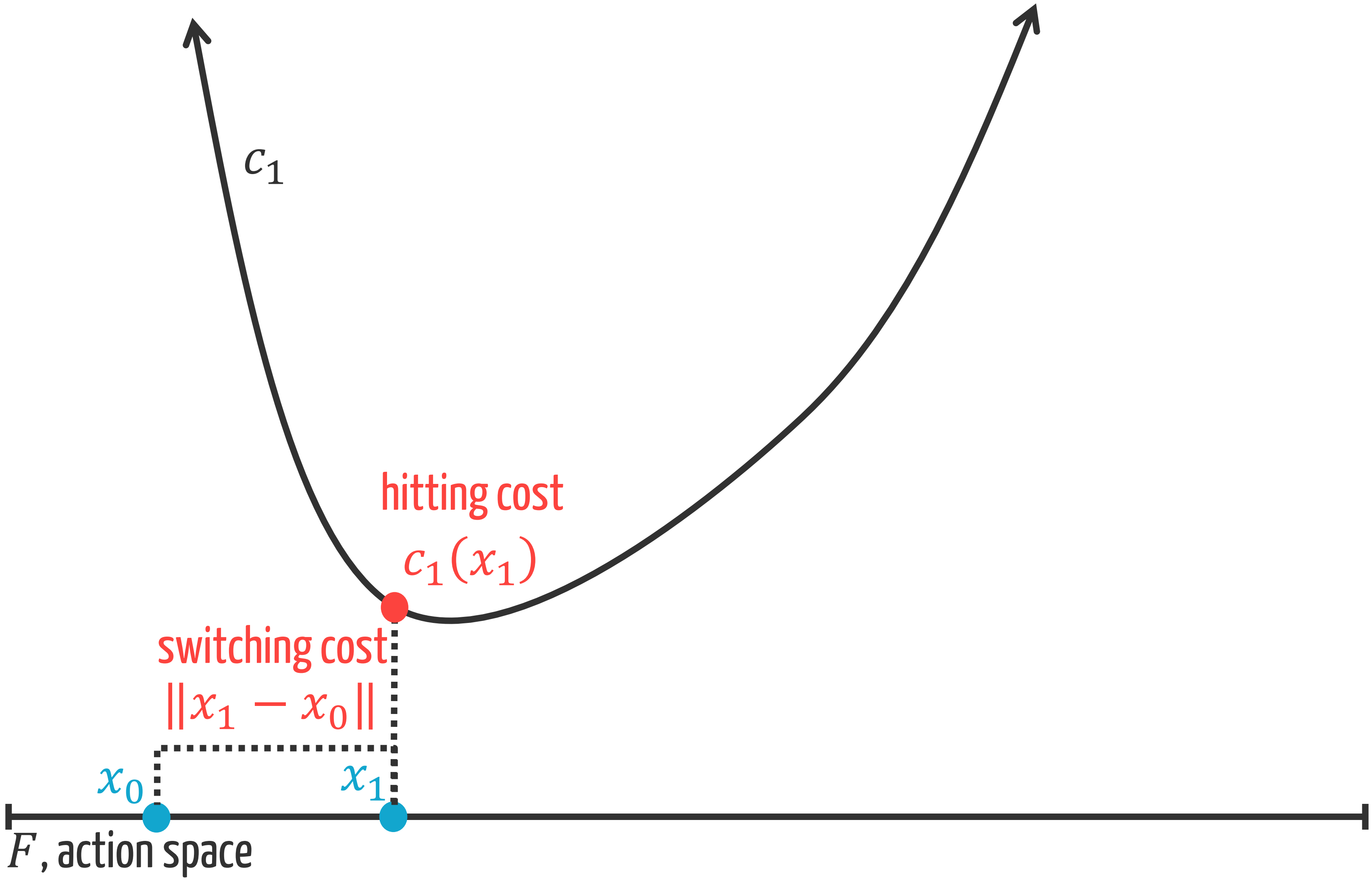
How can ML tools help improve the robustness & efficiency of control?

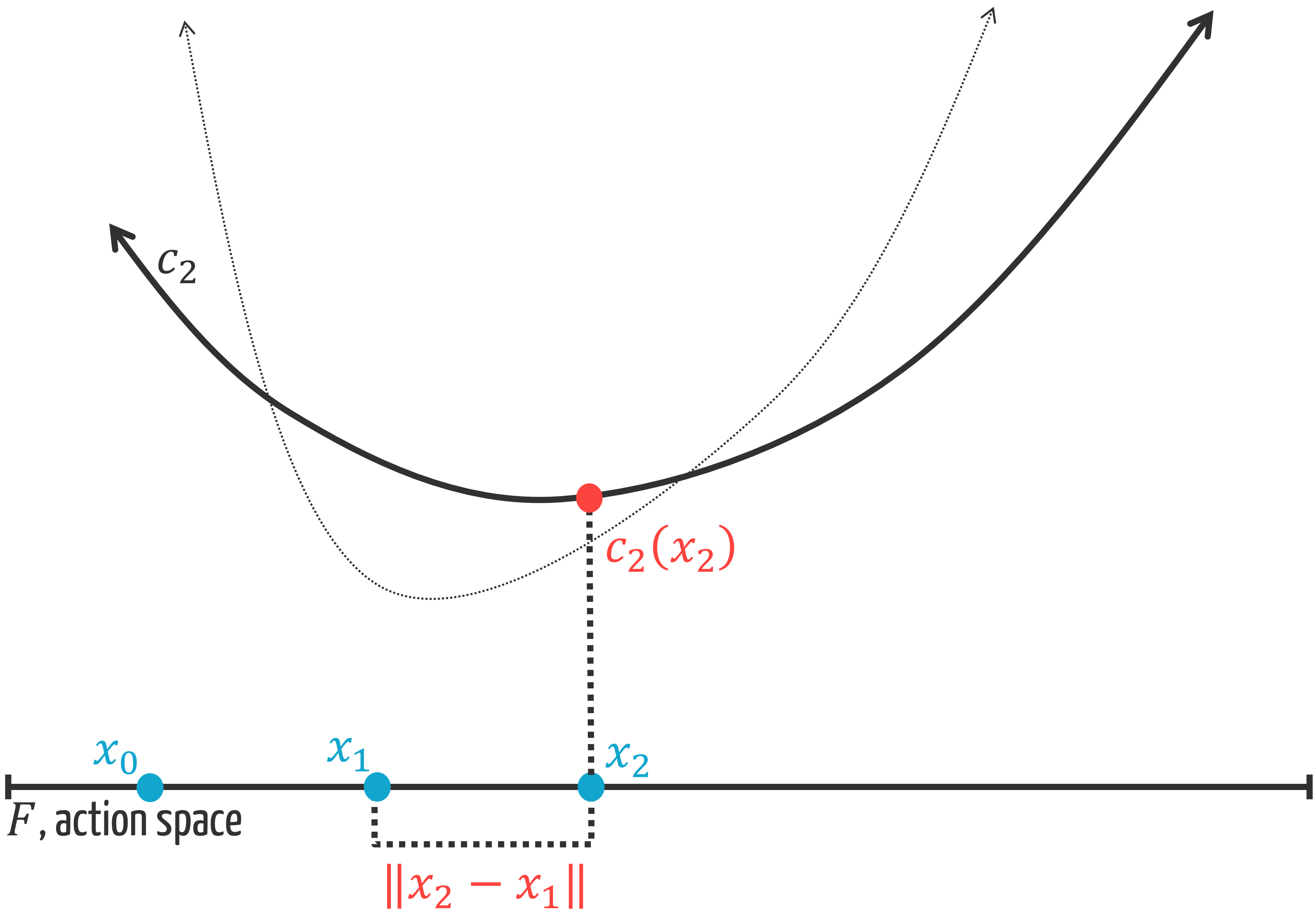
Can tools such as adversarial analysis, finite-time or single trajectory bounds, and general loss functions lead to improvements?

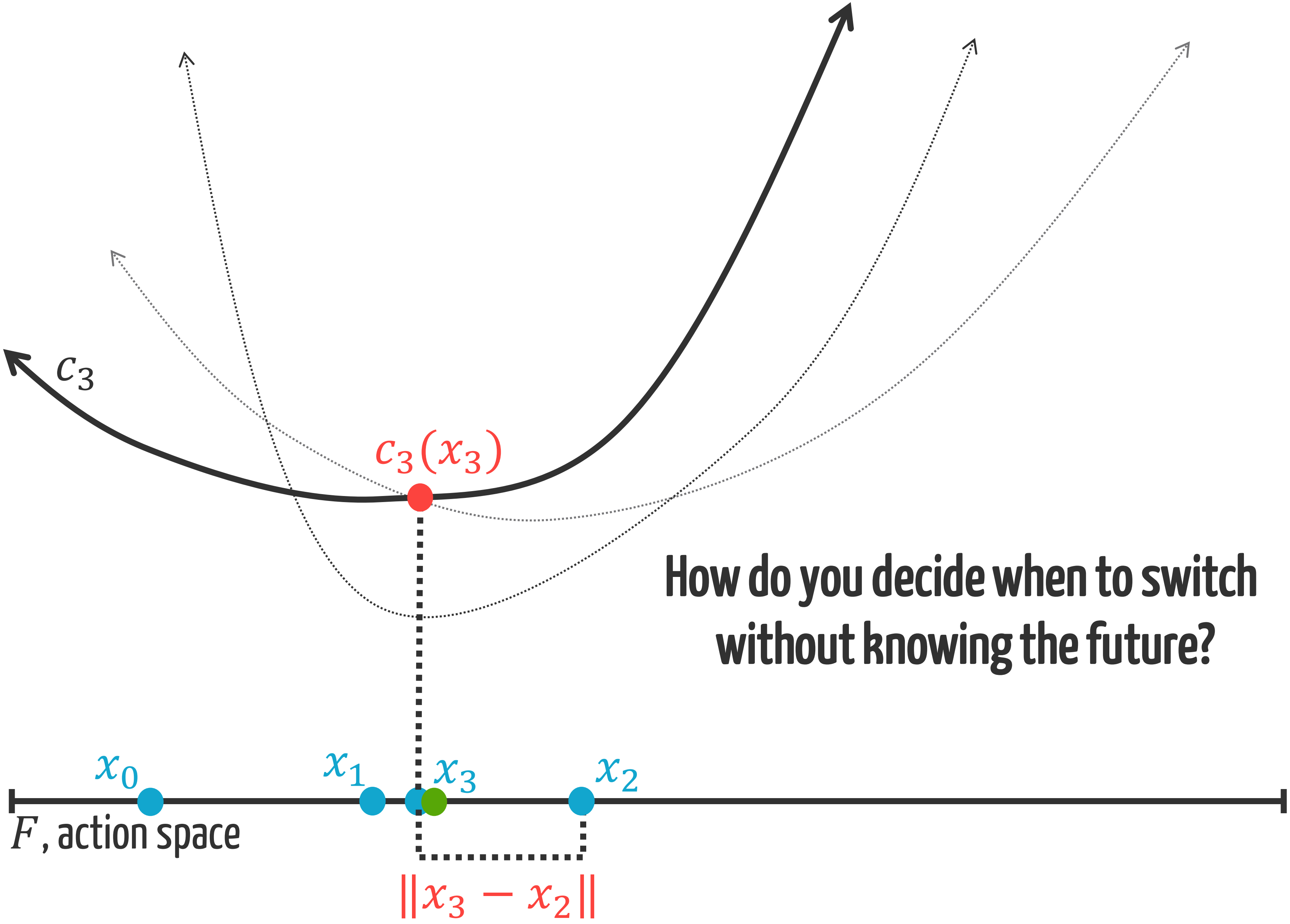
Can model-free and model-based approaches be combined to obtain the best of both worlds?

How much do you need to "understand" about a system to control it? Can we bring scalability and robustness to model-free RL?

Online Optimization & Control


c_1
hitting cost
$c_1(x_1)$
switching cost
$||x_1 - x_0||$
x_0
x_1
F, action space


c_2
$c_2(x_2)$
x_0
x_1
x_2
F, action space
$\|x_2 - x_1\|$



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

$c_1, x_1, c_2, x_2, c_3, x_3$ … ← online

$$\min_{x_t \in F} \sum_t c_t(x_t) + \underbrace{\|x_t - x_{t-1}\|}_{\text{switching cost}}$$

convex

Goal: Algorithms to minimize cost

"Smoothed" Online Convex Optimization
a.k.a. SOCO

$c_1, x_1, c_2, x_2, c_3, x_3 \ldots \leftarrow$ online

$$\min_{x_t \in F} \sum_t c_t(x_t) + \underbrace{\|x_t - x_{t-1}\|}_{\text{switching cost}}$$

convex

Goal: Algorithms to minimize cost

"Smoothed" Online Convex Optimization
a.k.a. SOCO

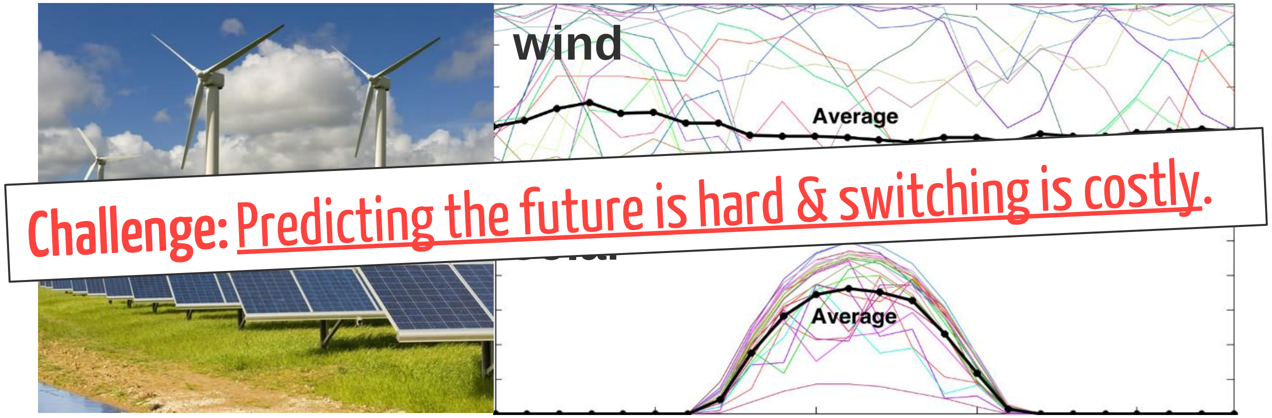
$$\textbf{Competitive ratio(Alg)} = \frac{\text{Cost(Alg)}}{\text{Cost(Offline\_Opt)}}$$

$$\textbf{Regret(Alg)} = \text{Cost(Alg)} - \text{Cost(Static\_Opt)}$$

Key: Adversarial guarantees

We got interested because of Sustainable Data Centers

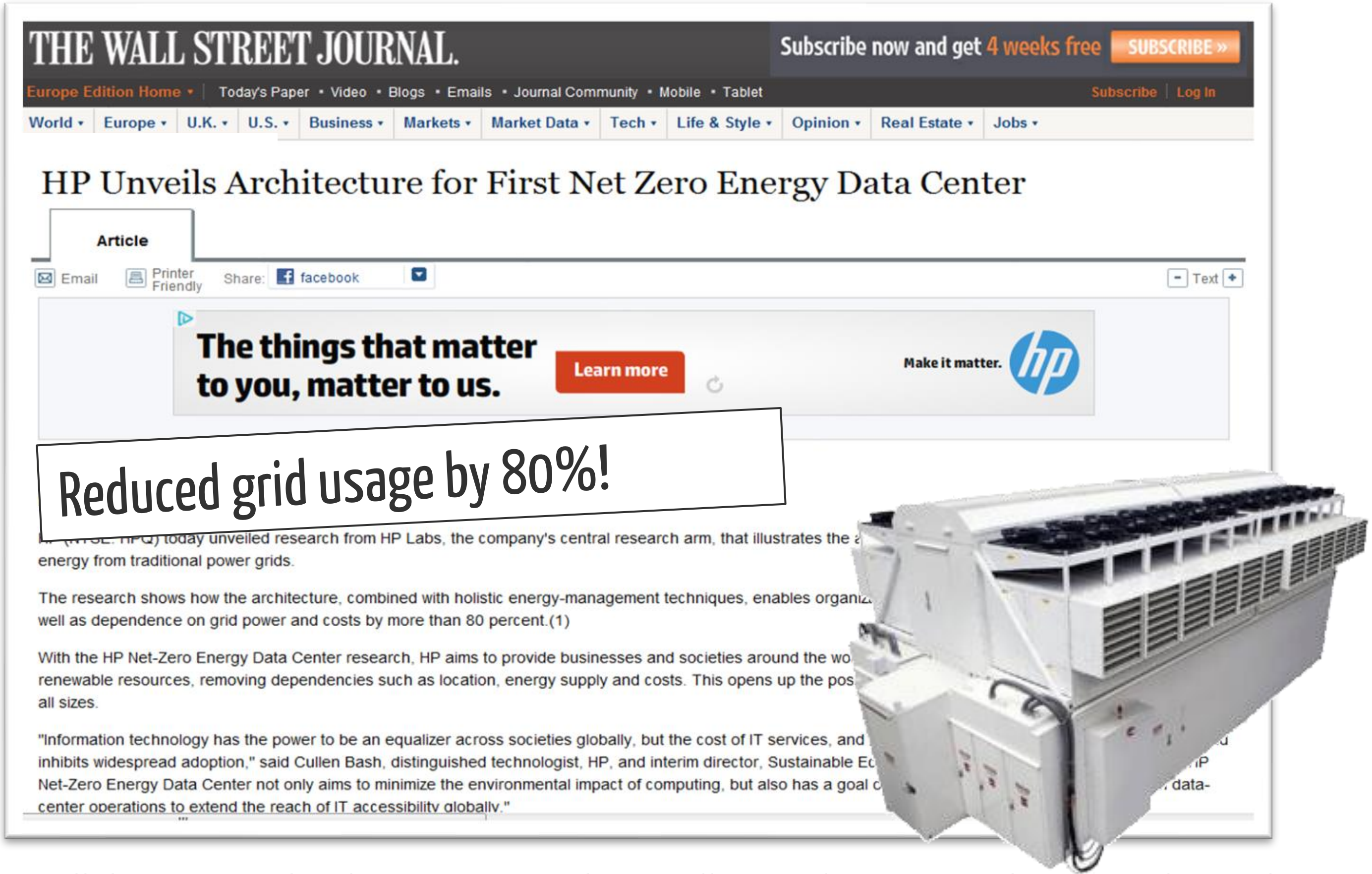
Can a data center run (almost) entirely on renewable sources?



Challenge: Predicting the future is hard & switching is costly.

Requires dynamic rightsizing of capacity and smart deferral of workloads

We got interested because of Sustainable Data Centers



HP Unveils Architecture for First Net Zero Energy Data Center

Reduced grid usage by 80%!

today unveiled research from HP Labs, the company's central research arm, that illustrates the
energy from traditional power grids.

The research shows how the architecture, combined with holistic energy-management techniques, enables organiz
well as dependence on grid power and costs by more than 80 percent.(1)

With the HP Net-Zero Energy Data Center research, HP aims to provide businesses and societies around the wo
renewable resources, removing dependencies such as location, energy supply and costs. This opens up the pos
all sizes.

"Information technology has the power to be an equalizer across societies globally, but the cost of IT services, and
inhibits widespread adoption," said Cullen Bash, distinguished technologist, HP, and interim director, Sustainable Ec
Net-Zero Energy Data Center not only aims to minimize the environmental impact of computing, but also has a goal c
center operations to extend the reach of IT accessibility globally."

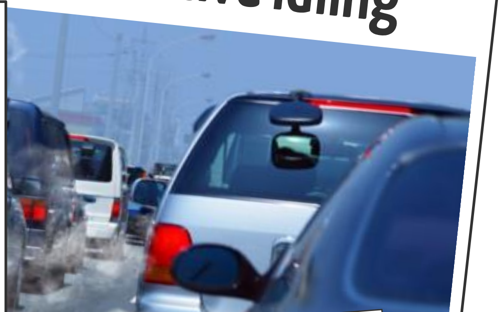
Collaborators: Zhenhua Liu, Yuan Chen, Cullen Bash, Martin Arlitt, Daniel Gmach, Zhikui Wang, Manish Marwah and Chris Hyser

SOCO is now a core model for energy systems...

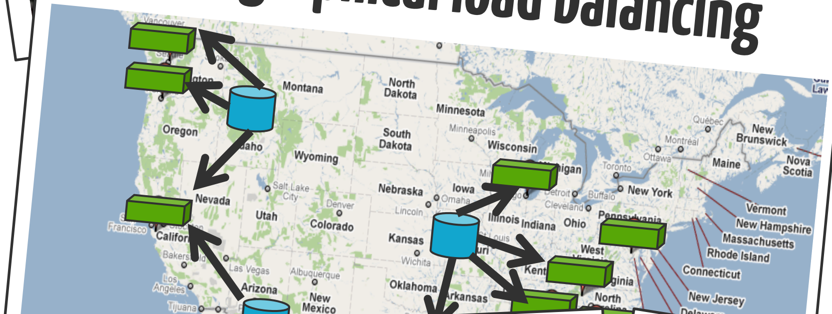
Microgrid control

Residential Demand Response

Automotive idling

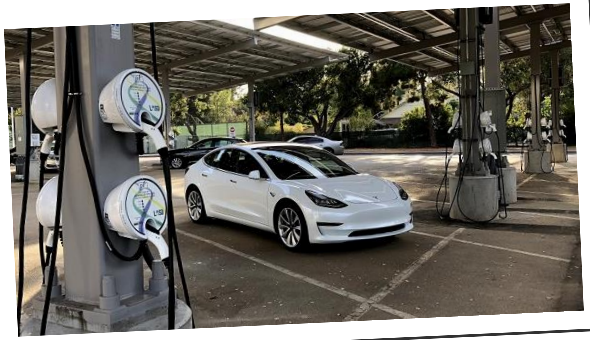


Geographical load balancing



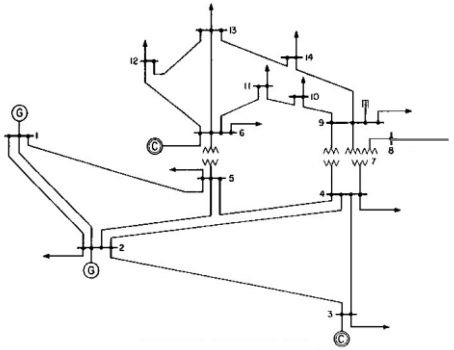
Data Center Demand Response

EV Charging



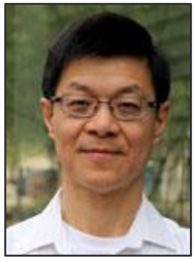
Economic Dispatch

Time-varying OPF

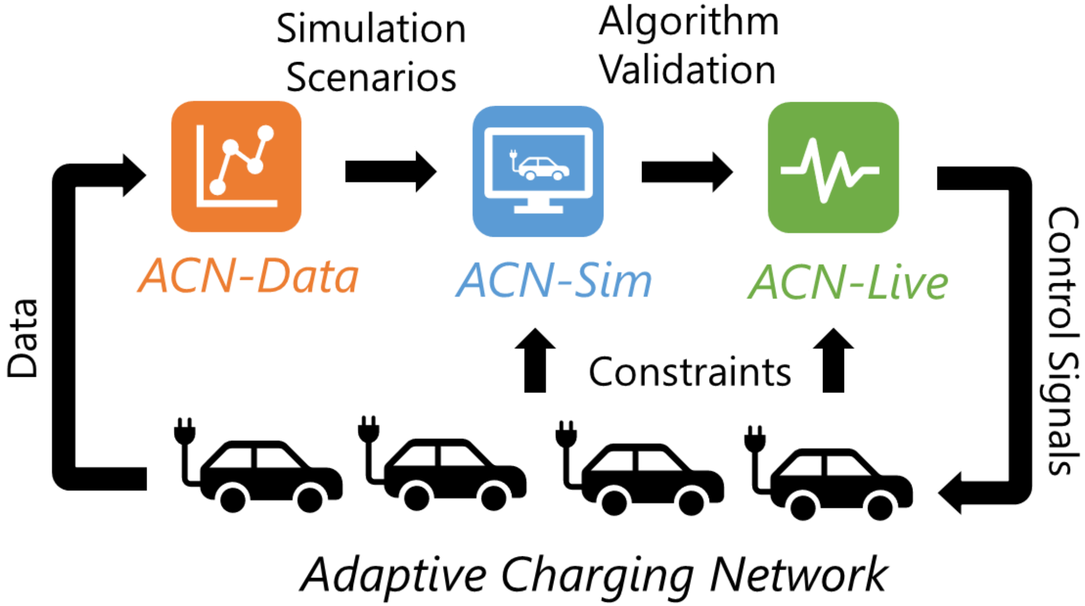


•••

Zach Lee



Steven Low



Simulation Scenarios

Algorithm Validation

ACN-Data

ACN-Sim

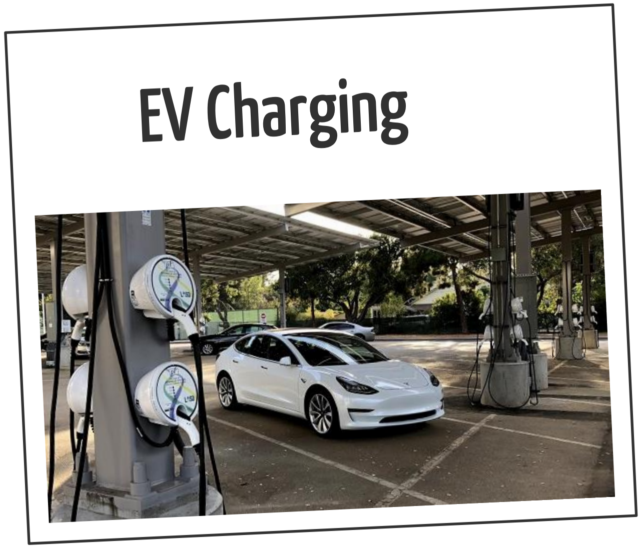
ACN-Live

Data

Constraints

Control Signals

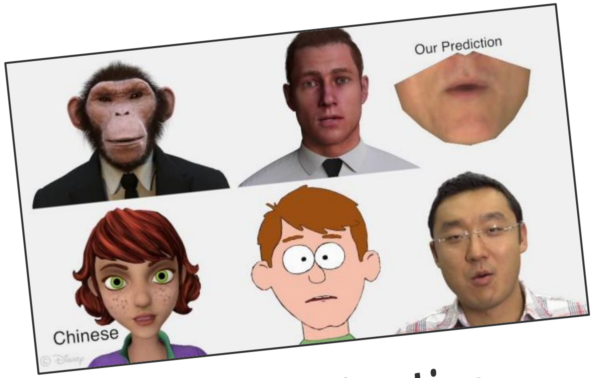
Adaptive Charging Network



EV Charging

powerflex
EDF renewables

...and the applications aren't limited to energy!



Facial animation

Video streaming


Disney using human operators to train automatic cameras for broadcasts

Robotic planning

Portfolio Management

Autonomous vehicles

•••

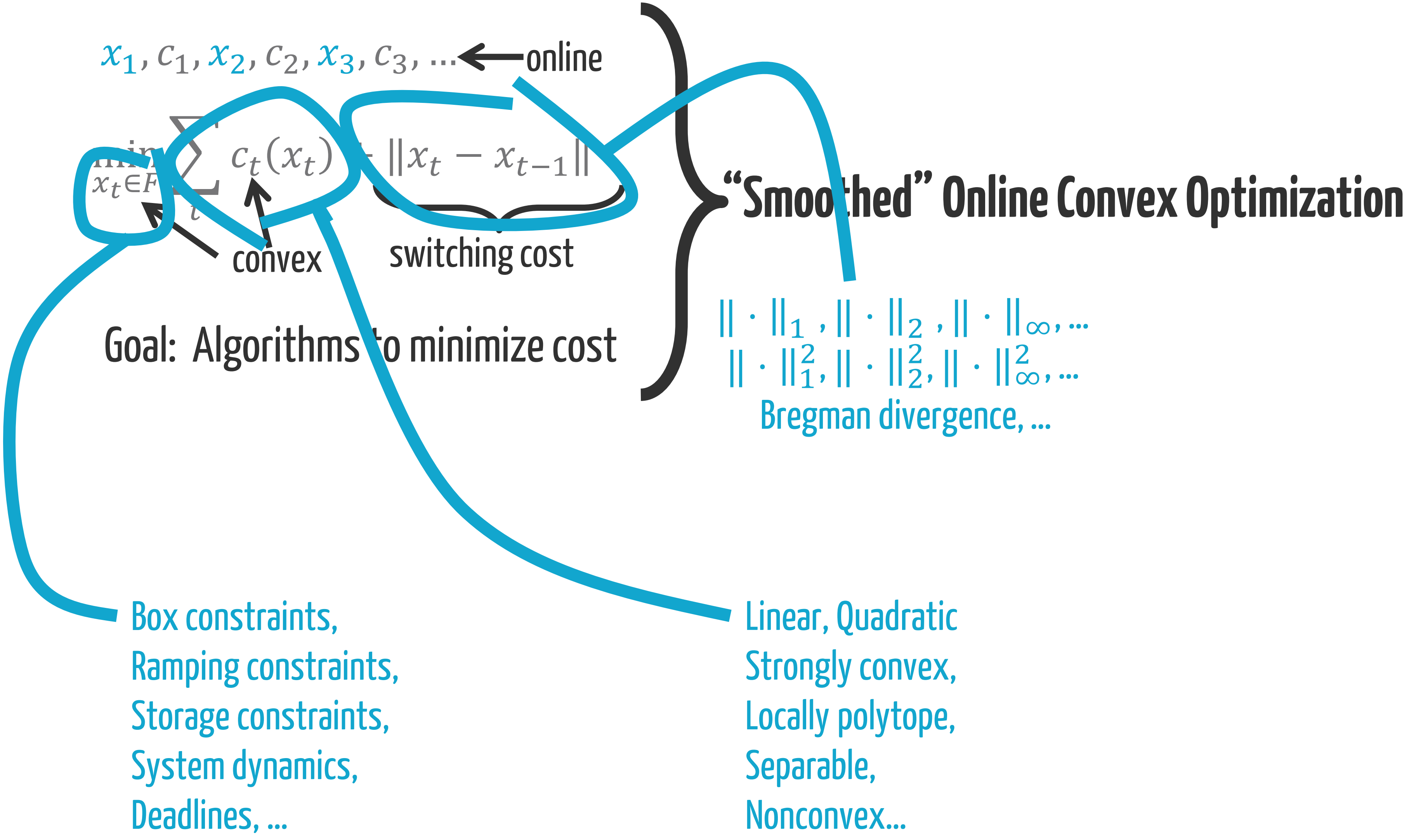
altitude

High-level motion planning

$$\min \sum_t \alpha(x_t - \theta_t)^2 + \beta(x_t - x_{t-1})^2$$

target being tracked
(there may be delay)

This has led to tons of variations to the model…



$x_1, c_1, x_2, c_2, x_3, c_3, \ldots$ ← online

$$\min_{x_t \in F} \sum_t c_t(x_t) + \underbrace{\|x_t - x_{t-1}\|}_{\text{switching cost}}$$

convex

Goal: Algorithms to minimize cost

"Smoothed" Online Convex Optimization

$\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$, …

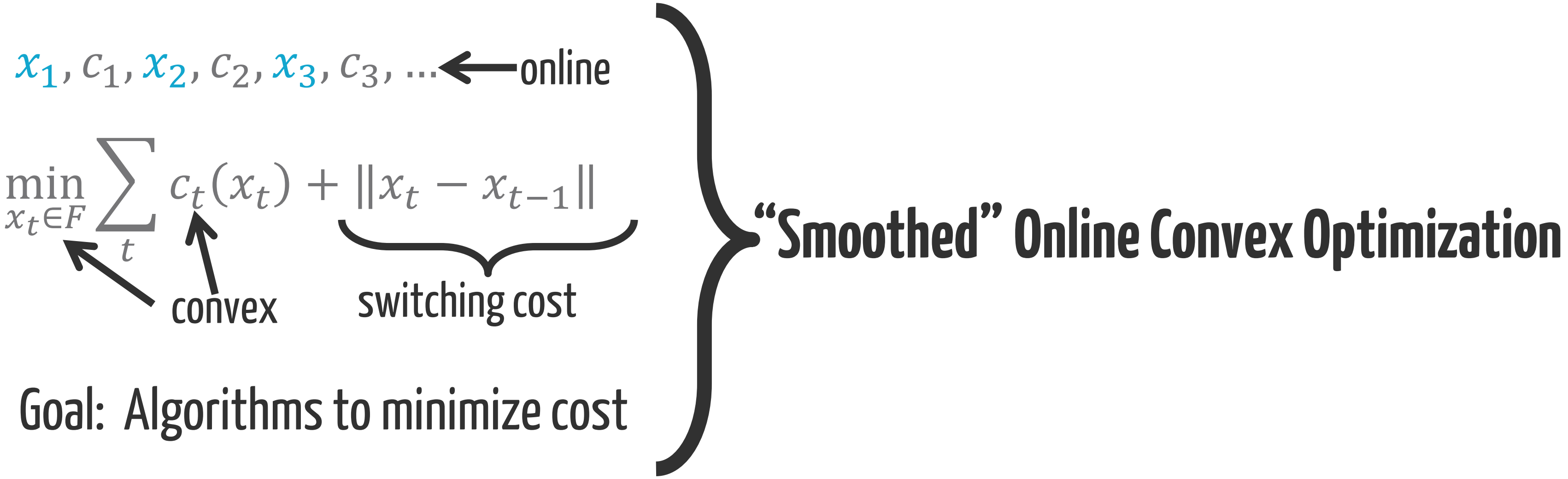
$\|\cdot\|_1^2$, $\|\cdot\|_2^2$, $\|\cdot\|_\infty^2$, …

Bregman divergence, …

Box constraints,
Ramping constraints,
Storage constraints,
System dynamics,
Deadlines, …

Linear, Quadratic
Strongly convex,
Locally polytope,
Separable,
Nonconvex…

This has led to tons of variations to the model…



$x_1, c_1, x_2, c_2, x_3, c_3, \ldots \longleftarrow$ online

$$\min_{x_t \in F} \sum_t c_t(x_t) + \underbrace{\|x_t - x_{t-1}\|}_{\text{switching cost}}$$

convex

Goal: Algorithms to minimize cost

“Smoothed” Online Convex Optimization

…but control has dynamics and SOCO does not!

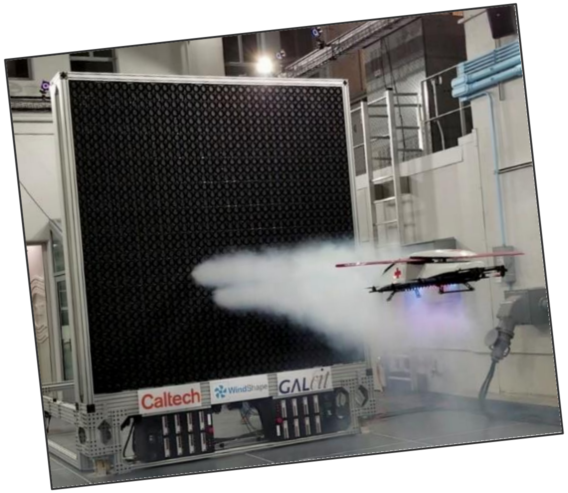
$x_1, c_1, x_2, c_2, x_3, c_3, \dots$ ← online

$$\min_{x_t \in F} \sum_t c_t(x_t) + \underbrace{\|x_t - x_{t-1}\|}_{\text{switching cost}}$$

convex

Goal: Algorithms to minimize cost

"Smoothed" Online Convex Optimization



Low-level physical tracking

x: altitude, $\ddot{x}$: acceleration

$\boxed{\ddot{x} = \gamma u}$ dynamics

$$\sum_t \alpha(x_t - \theta_t)^2 + \beta u_t^2$$

control

target being tracked

$x_1, c_1, x_2, c_2, x_3, c_3, \ldots \leftarrow$ online

$$\min_{x_t \in F} \sum_t c_t(x_t) + \underbrace{\|x_t - x_{t-1}\|}_{\text{switching cost}}$$

convex

Goal: Algorithms to minimize cost

"Smoothed" Online Convex Optimization

Linear Dynamical System (LDS)

dynamics

$$\boxed{x_{t+1} = Ax_t + Bu_t + w_t}$$

$$\min_{u_t} \sum_t q_t \|x_t\|^2 + \|u_t\|^2$$

Goal: Algorithms that stabilize & minimize cost

$x_1, c_1, x_2, c_2, x_3, c_3, \ldots$ ← online

$$\min_{x_t \in F} \sum_t c_t(x_t) + \underbrace{\|x_t - x_{t-1}\|}_{\text{switching cost}}$$

convex

Goal: Algorithms to minimize cost

"Smoothed" Online Convex Optimization

dynamics

$$\boxed{x_{t+1} = Ax_t + Bu_t + w_t}$$

A motivating example:

Frequency regulation with variable inertia

$$\underbrace{\begin{bmatrix}\dot{\theta}\\ \dot{\omega}\end{bmatrix}}_{\dot{\mathbf{x}}} = \underbrace{\begin{bmatrix}0 & I\\ -M_t^{-1}L & -M_t^{-1}D\end{bmatrix}}_{A_t}\begin{bmatrix}\theta\\ \omega\end{bmatrix} + \underbrace{\begin{bmatrix}0\\ M_t^{-1}\end{bmatrix}}_{B_t}\underbrace{p_{in}}_{\mathbf{u_t}}$$

inertia susceptance frequency angle

$x_1, c_1, x_2, c_2, x_3, c_3, \ldots$ ← online

$$\min_{x_t \in F} \sum_t c_t(x_t) + \underbrace{\|x_t - x_{t-1}\|}_{\text{switching cost}}$$

convex

Goal: Algorithms to minimize cost

“Smoothed” Online Convex Optimization

Linear Dynamical System (LDS)

$$x_{t+1} = Ax_t + Bu_t + w_t$$

$$\min_{u_t} \sum_t q_t \|x_t\|^2 + \|u_t\|^2$$

Goal: Algorithms that stabilize & minimize cost

A growing literature focuses on designing no-regret & competitive controllers...

$x_1, c_1, x_2, c_2, x_3, c_3, \ldots$ ← online

$$\min_{x_t \in F} \sum_t c_t(x_t) + \underbrace{\|x_t - x_{t-1}\|}_{\text{switching cost}}$$

convex

Goal: Algorithms to minimize cost

"Smoothed" Online Convex Optimization

Linear Dynamical System (LDS)

stochastic vs. worst case

$$x_{t+1} = Ax_t + Bu_t + w_t$$

$$\min_{u_t} \sum_t q_t\|x_t\|^2 + \|u_t\|^2$$

Goal: Algorithms that stabilize &
minimize cost

quadratic vs. general costs?

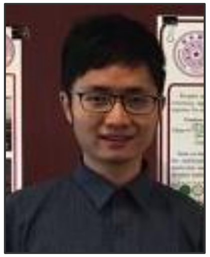
Can we connect SOCO to control?

A deep connection has emerged over the past few years
... [GW19] [ABHKS19] [GLSW19] [GHM 20] [SLCYW20] [PSCYW20] ...

Theorem [SLCYW20] [PSCYW21+]: Any LDS with A,B in canonical form is equivalent to a SOCO problem with memory and delay.


Weici Pan


Yiheng Lin


Guanya Shi


Soon-Jo Chung

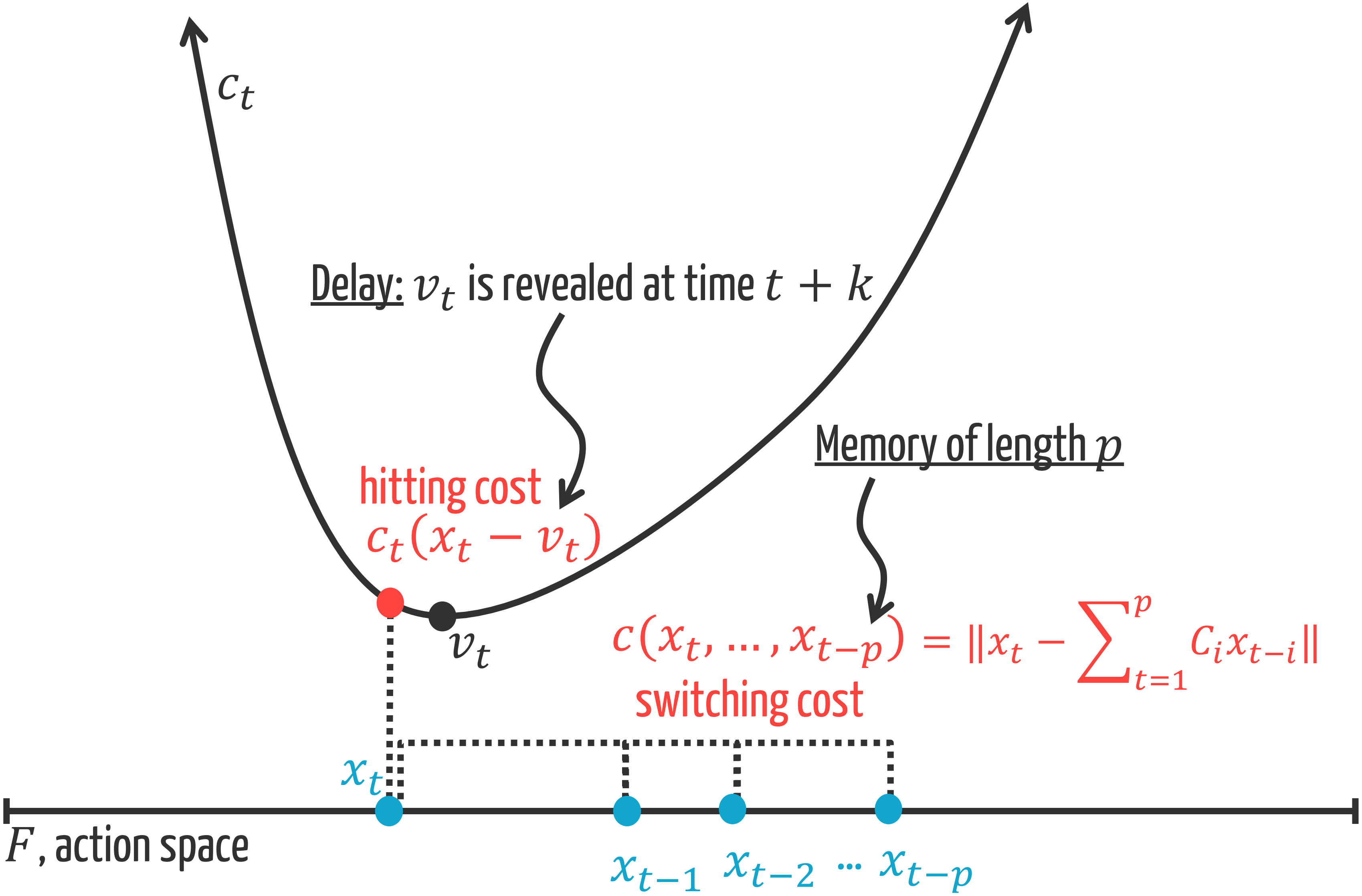

Yisong Yue

Can we connect SOCO to control?

A deep connection has emerged over the past few years
... [GW19] [ABHKS19] [GLSW19] [GHM 20] [SLCYW20] [PSCYW20] ...

Theorem [SLCYW20] [PSCYW21+]: Any LDS with A,B in canonical form is equivalent to a SOCO problem with memory and delay.

Memory first used in OCO in [AHM 2015].
Has importance beyond control too...


c_t
Delay: v_t is revealed at time $t + k$
Memory of length p
hitting cost
$c_t(x_t - v_t)$
v_t
$c(x_t, \dots, x_{t-p}) = \|x_t - \sum_{t=1}^{p} C_i x_{t-i}\|$
switching cost
x_t
F, action space
x_{t-1} x_{t-2} … x_{t-p}

Can we connect SOCO to control?

A deep connection has emerged over the past few years
... [GW19] [ABHKS19] [GLSW19] [GHM 20] [SLCYW20] [PSCYW20] ...

Theorem [SLCYW20] [PSCYW20+]: Any LDS with A,B in canonical form is equivalent to a SOCO problem with memory and delay.

Memory → input disturbance, $x_{t+1} = Ax_t + B(u_t + w_t)$

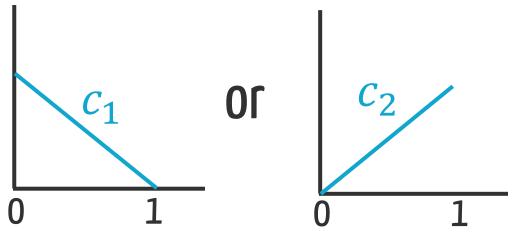
Delay → state disturbance, $x_{t+t} = A(x_t + w_t) + Bu_t$

...and further: non-linear switch costs → non-linear dynamics

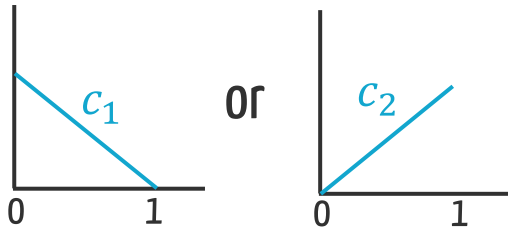
<u>Today</u>: A taste of SOCO &
a new algorithm

We’ll start by ignoring memory & delay…

A warmup: 2 cost functions



A warmup: 2 cost functions

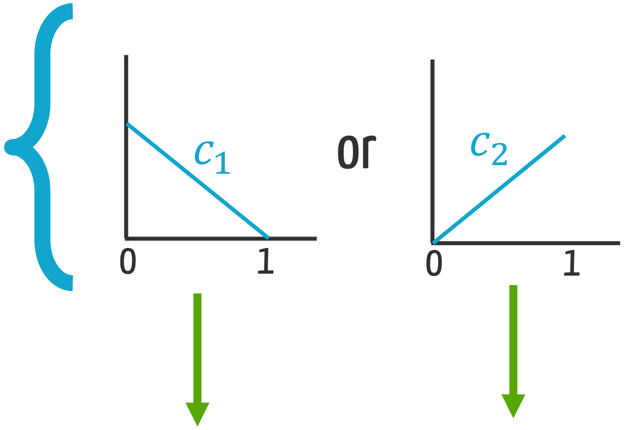


A first attempt: "**Greedy**"

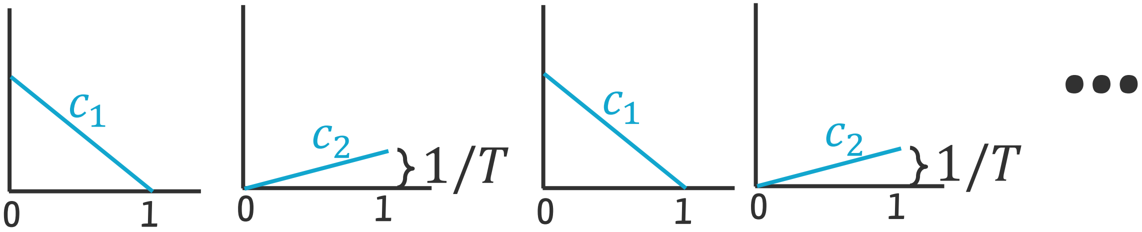
$x_t = 1$

$x_t = 0$

A warmup: 2 cost functions



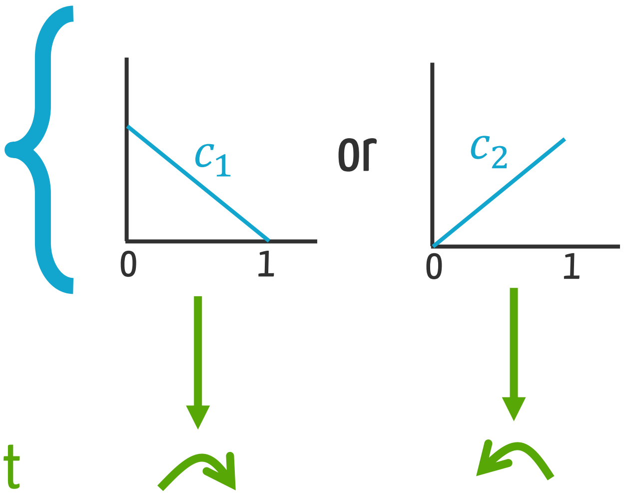
A first attempt: "Greedy" $x_t = 1$ $x_t = 0$



"Greedy:" Cost $= \ T$ "Opt:" Cost $\leq T \times 1/T = 1$

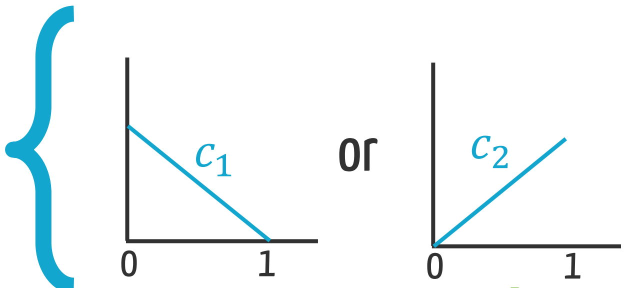
It is crucial to consider whether switching is "worth it"

A warmup: 2 cost functions

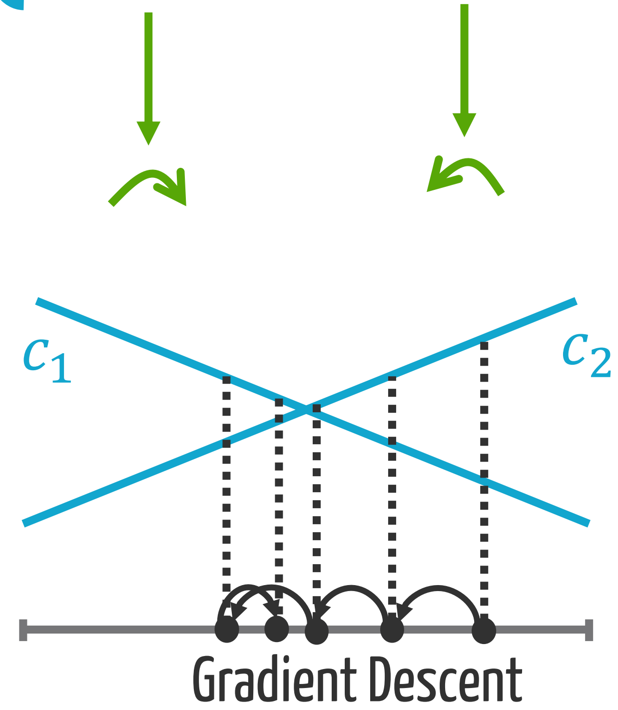


A second attempt: Gradient Descent

A warmup: 2 cost functions

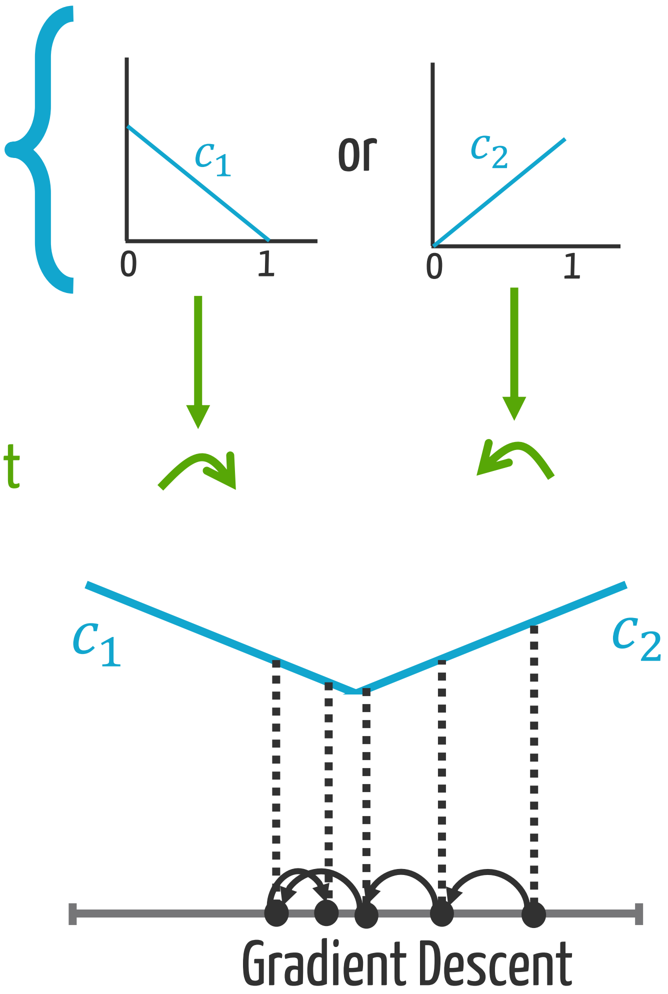


A second attempt: Gradient Descent

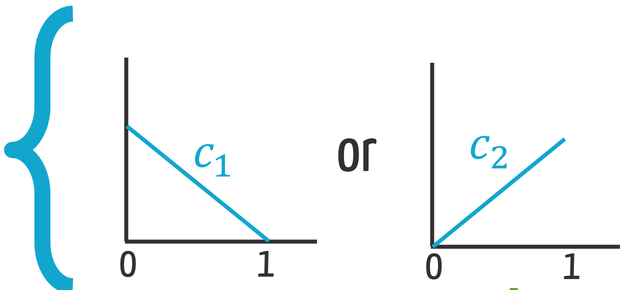


A warmup: 2 cost functions

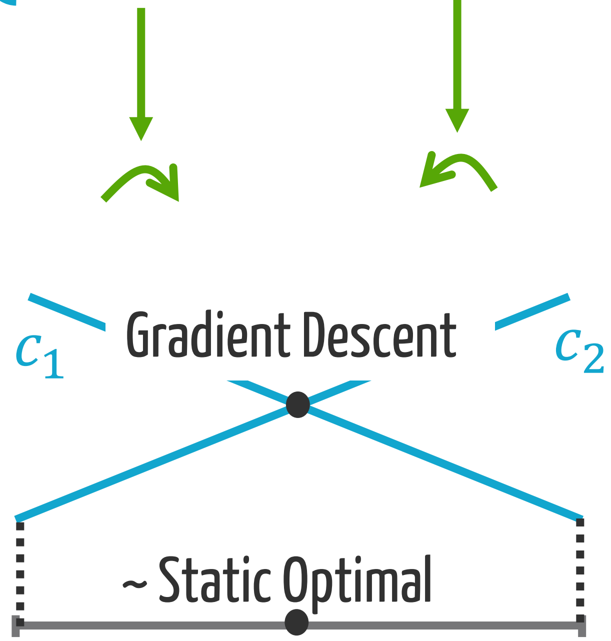
A second attempt: Gradient Descent



A warmup: 2 cost functions

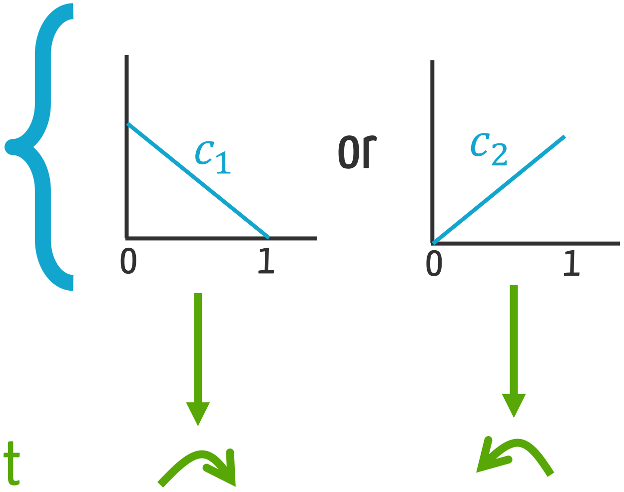


A second attempt: Gradient Descent

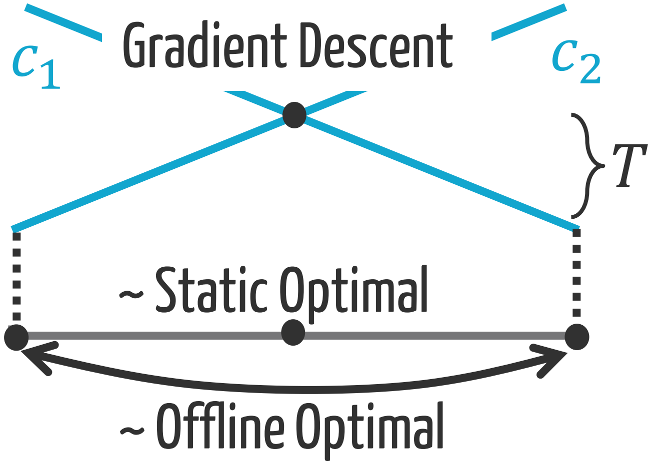


Learns the best static point, but...

A warmup: 2 cost functions

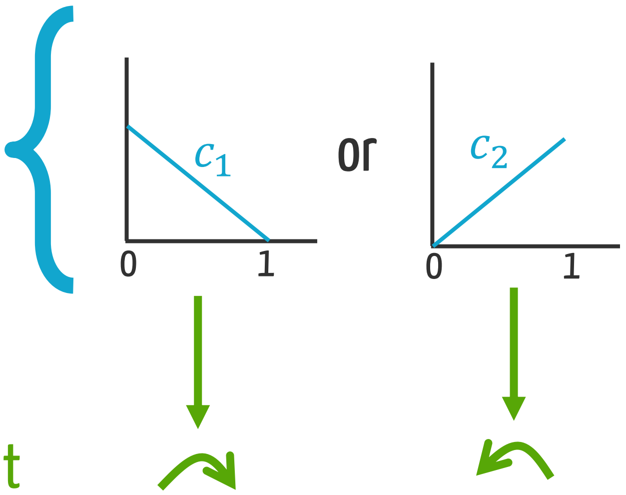


A second attempt: Gradient Descent

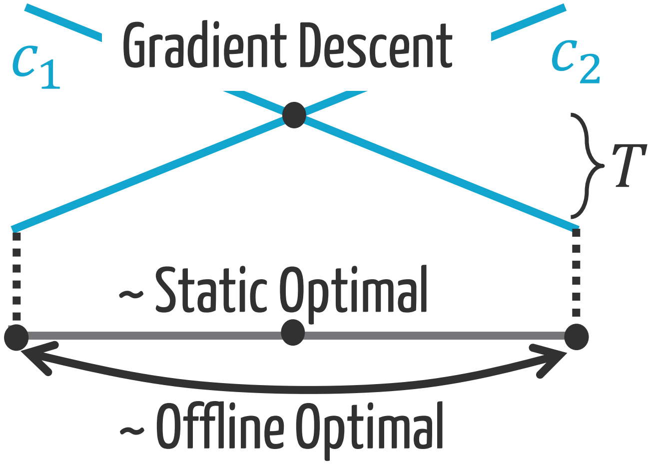


Offline optimal is order-of-magnitude better.

A warmup: 2 cost functions

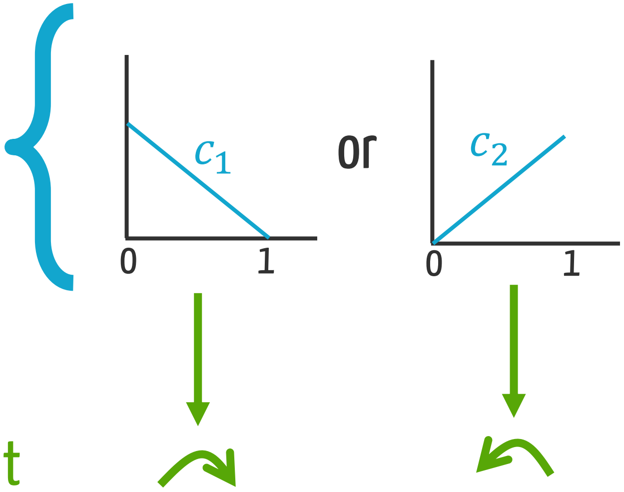


A second attempt: Gradient Descent

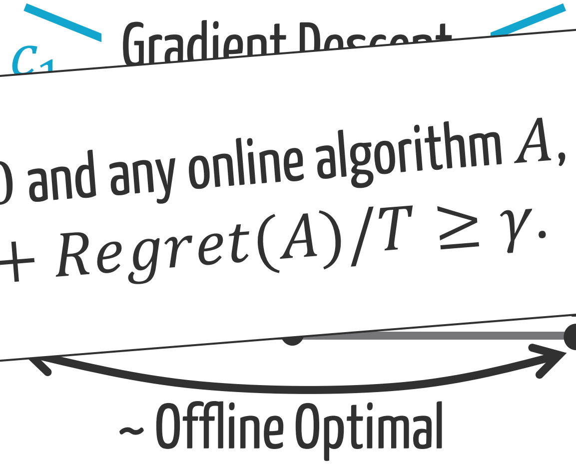


It is crucial to make “big jumps” when needed

A warmup: 2 cost functions



A second attempt: Gradient Descent



Theorem [ABKL+13]: For arbitrary $\gamma > 0$ and any online algorithm A, there exist linear cost functions such that $CR(A) + Regret(A)/T \geq \gamma$.

Many algorithms can be no-regret for SOCO but…

Can an algorithm be constant competitive?

Yes! …but it took a long time to get there.

The starting point

<u>Theorem [LWAT13]:</u> Lazy Capacity Provisioning (LCP) is 3-competitive in 1-dimension.

[BGK+ 2015] give a 2-competitive algorithm.
[AS 2018] show 2 is the best possible.

Complex!

"Memoryless" algorithms can't be better than 3 competitive.
[ABKL+ 2013] & [BGK+2015] give 3-competitive memoryless algorithms.

The starting point

Theorem [LWAT13]: Lazy Capacity Provisioning (LCP) is 3-competitive in 1-dimension.

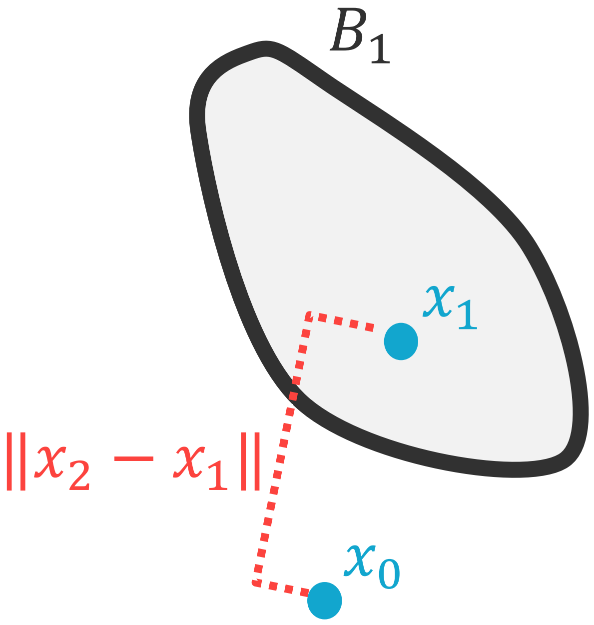
Years passed with no progress outside of 1 dimension.

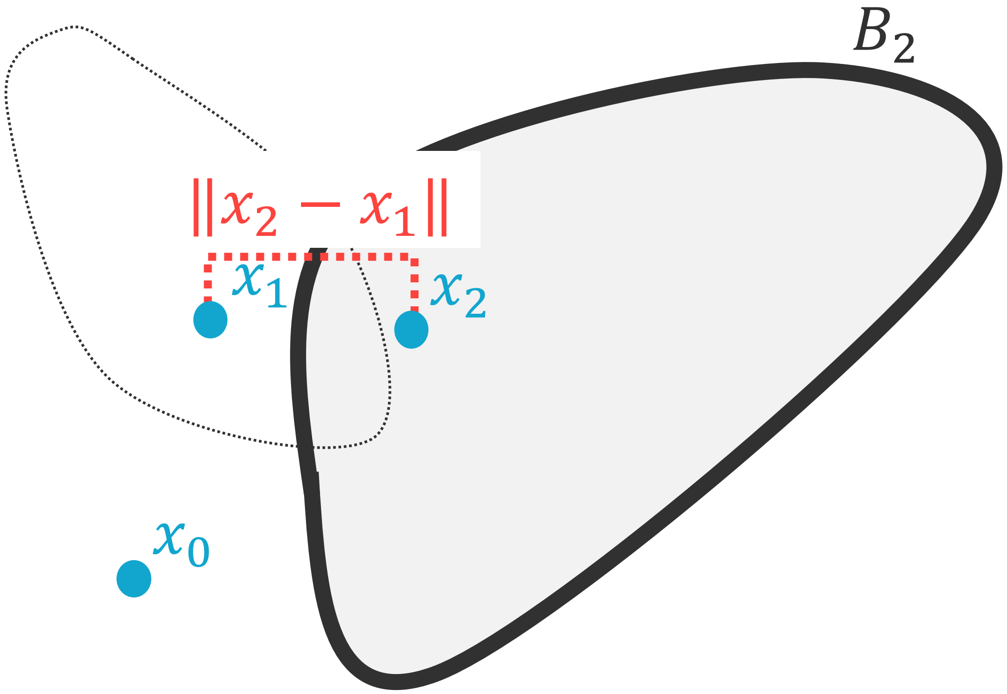
We now understand why...

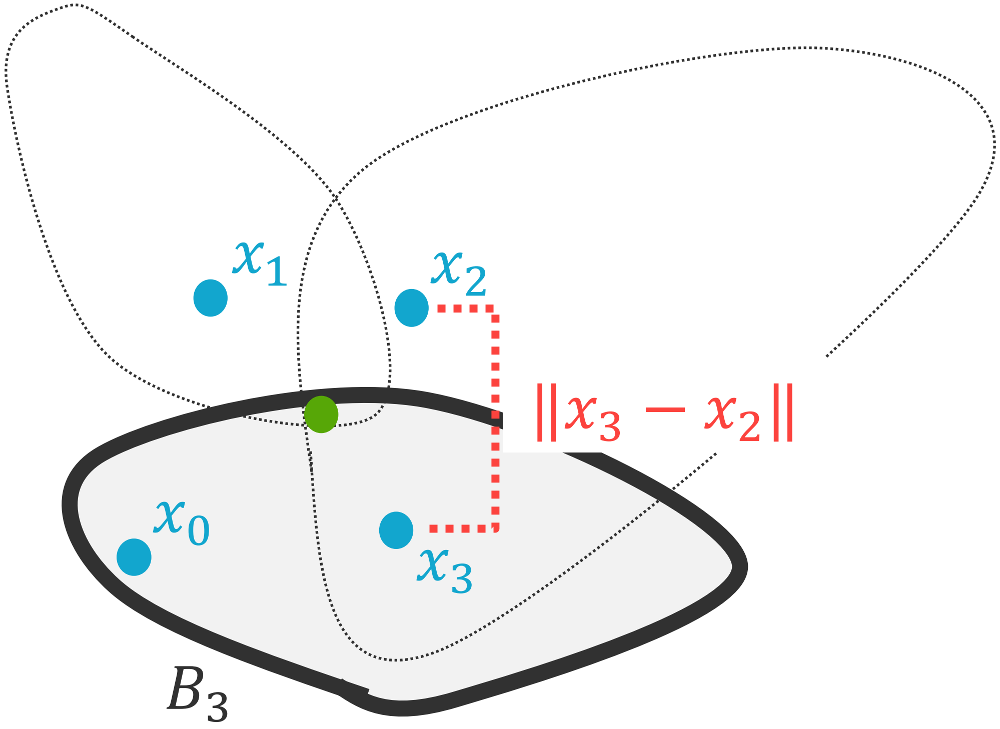
Theorem [BLLS19] [LGW20]: In d-dimensional convex body chasing (CBC) problem, any online algorithm is $\Omega(\sqrt{d})$-competitive , even when the algorithm can perfectly forecast the next w steps bodies.

Yiheng Lin

Gautam Goel


B_1
x_1
$\|x_2 - x_1\|$
x_0

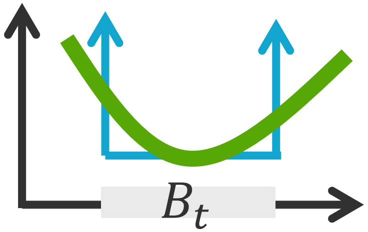

B_2
$\|x_2 - x_1\|$
x_1
x_2
x_0



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

We now understand why:

Theorem [BLLS19] [LGW20]: In d-dimensional convex body chasing (CBC) problem, any online algorithm is $\Omega(\sqrt{d})$-competitive , even when the algorithm can perfectly forecast the next w steps bodies.



But we usually have some structure!

A breakthrough

Theorem [GLSW19]: Regularized Online Balanced Descent (ROBD) is $1 + O(1/\sqrt{m})$-competitive for m-strongly convex hitting costs and switching costs that are either the squared-L_2 norm or Bregman divergence.

Gautam Goel

Yiheng Lin

Haoyuan Sun

Theorem [GLSW19]: Regularized Online Balanced Descent (ROBD) is $1 + O(1/\sqrt{m})$-competitive for m-strongly convex hitting costs and switching costs that are either the squared-L_2 norm or Bregman divergence.

No dependence on the dimension d!

What is achievable?

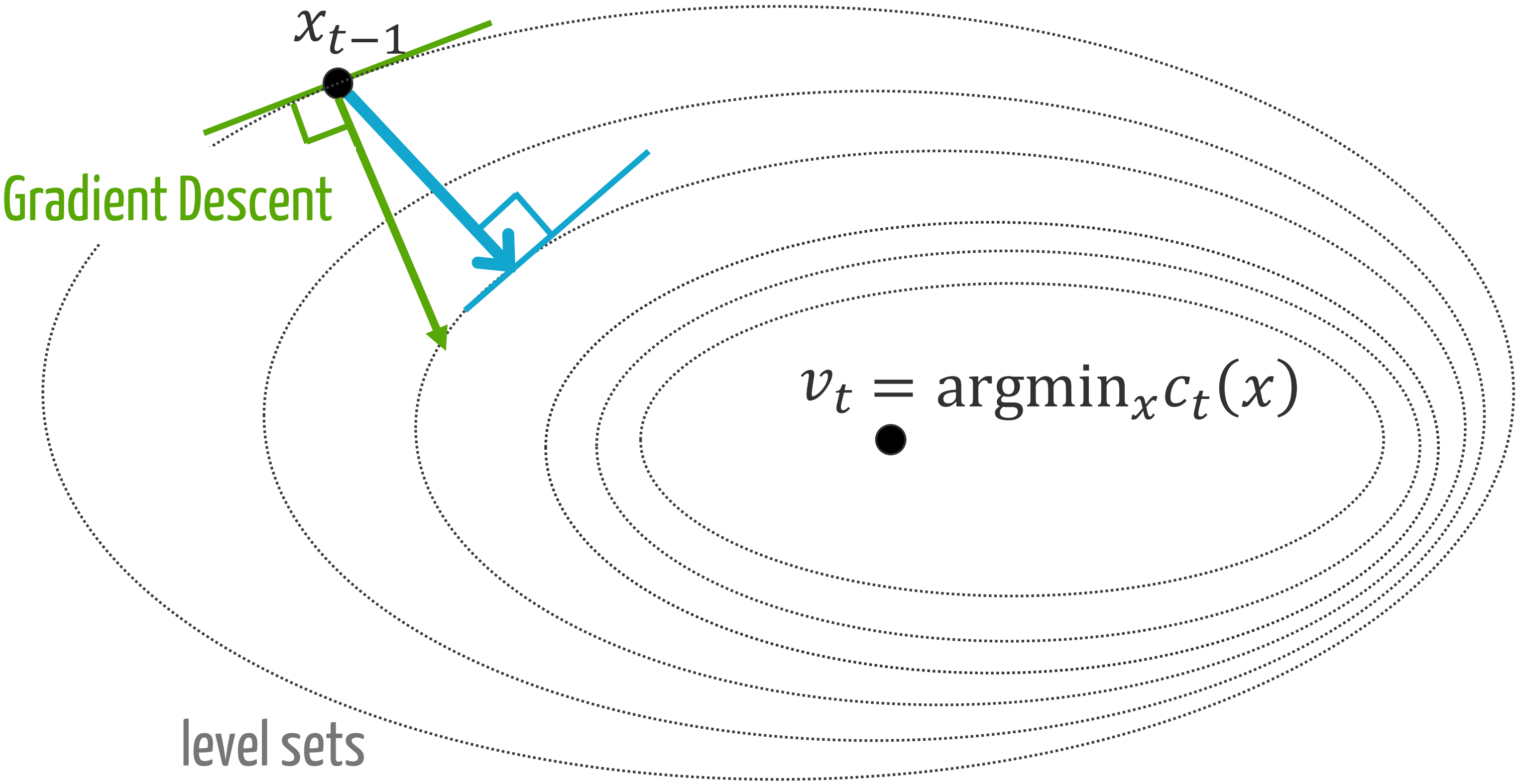
Theorem [GLSW19]: Regularized Online Balanced Descent (ROBD) is $1 + O(1/\sqrt{m})$-competitive for m-strongly convex hitting costs and switching costs that are either the squared-L_2 norm or Bregman divergence.

Theorem [GLSW19]: All online algorithms have competitive ratio $\geq \frac{1}{2}\left(1 + \sqrt{1 + 4/m}\right)$ for m-strongly convex hitting costs and squared L_2 switching costs.

Theorem [GLSW19]: Regularized Online Balanced Descent (ROBD) is $\frac{1}{2}\left(1+\sqrt{1+4/m}\right)$-competitive for m-strongly convex hitting costs and switching costs that are either the squared-L_2 norm or Bregman divergence.

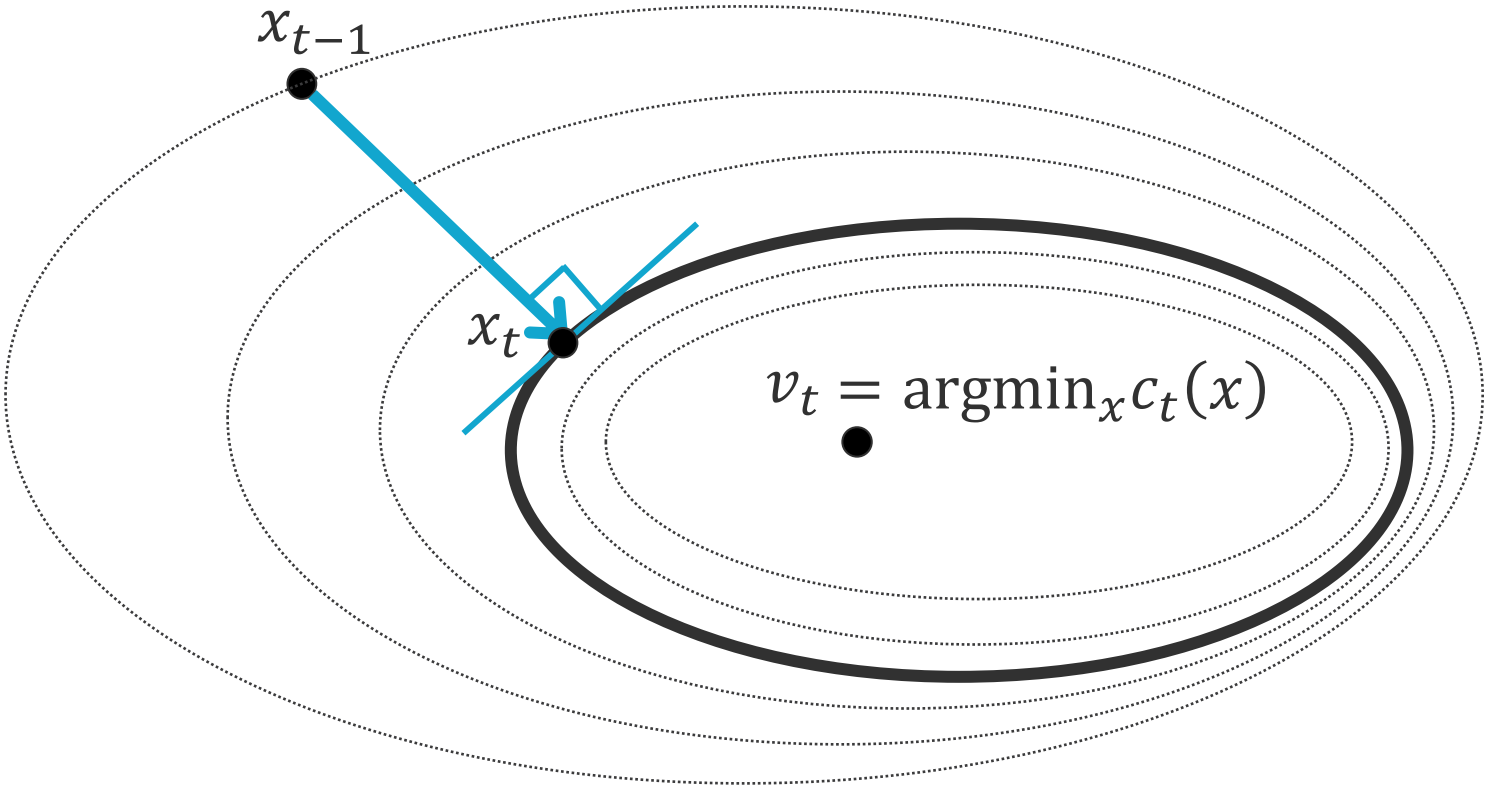
Theorem [GLSW19]: All online algorithms have competitive ratio $\geq \frac{1}{2}\left(1+\sqrt{1+4/m}\right)$ for m-strongly convex hitting costs and squared L_2 switching costs.

~~Regularized~~ Online Balanced Descent (OBD)



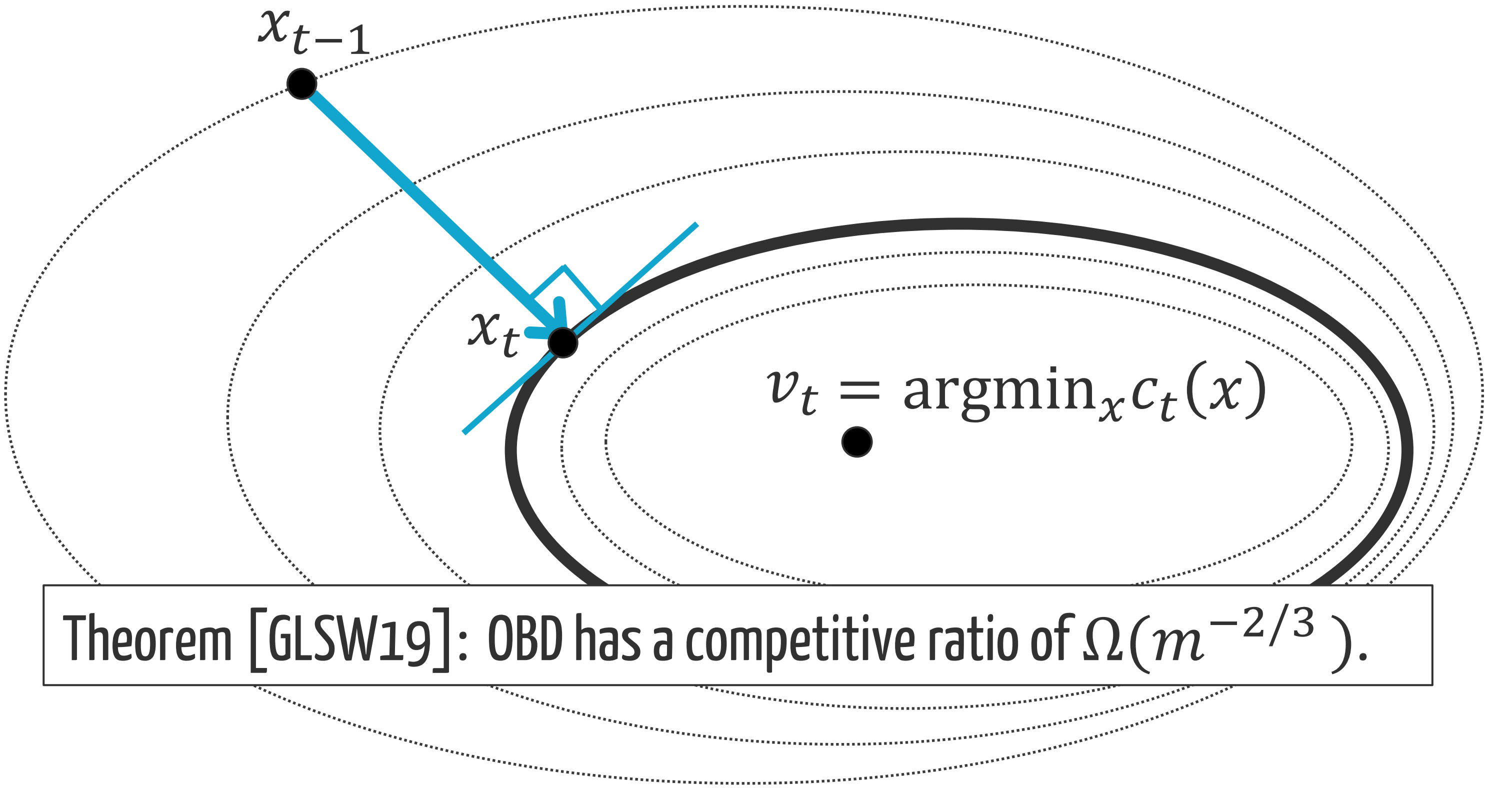
~~Regularized~~ Online Balanced Descent (OBD)

Project onto a level set $K_l = \{\, x | c_t(x) \leq l\}$
where l balances costs, i.e., $\|x(l) - x_{t-1}\| = \beta l.$



~~Regularized~~ Online Balanced Descent (OBD)

Project onto a level set $K_l = \{ x | c_t(x) \leq l \}$ where l balances costs, i.e., $\|x(l) - x_{t-1}\| = \beta l$.



Theorem [GLSW19]: OBD has a competitive ratio of $\Omega(m^{-2/3})$.

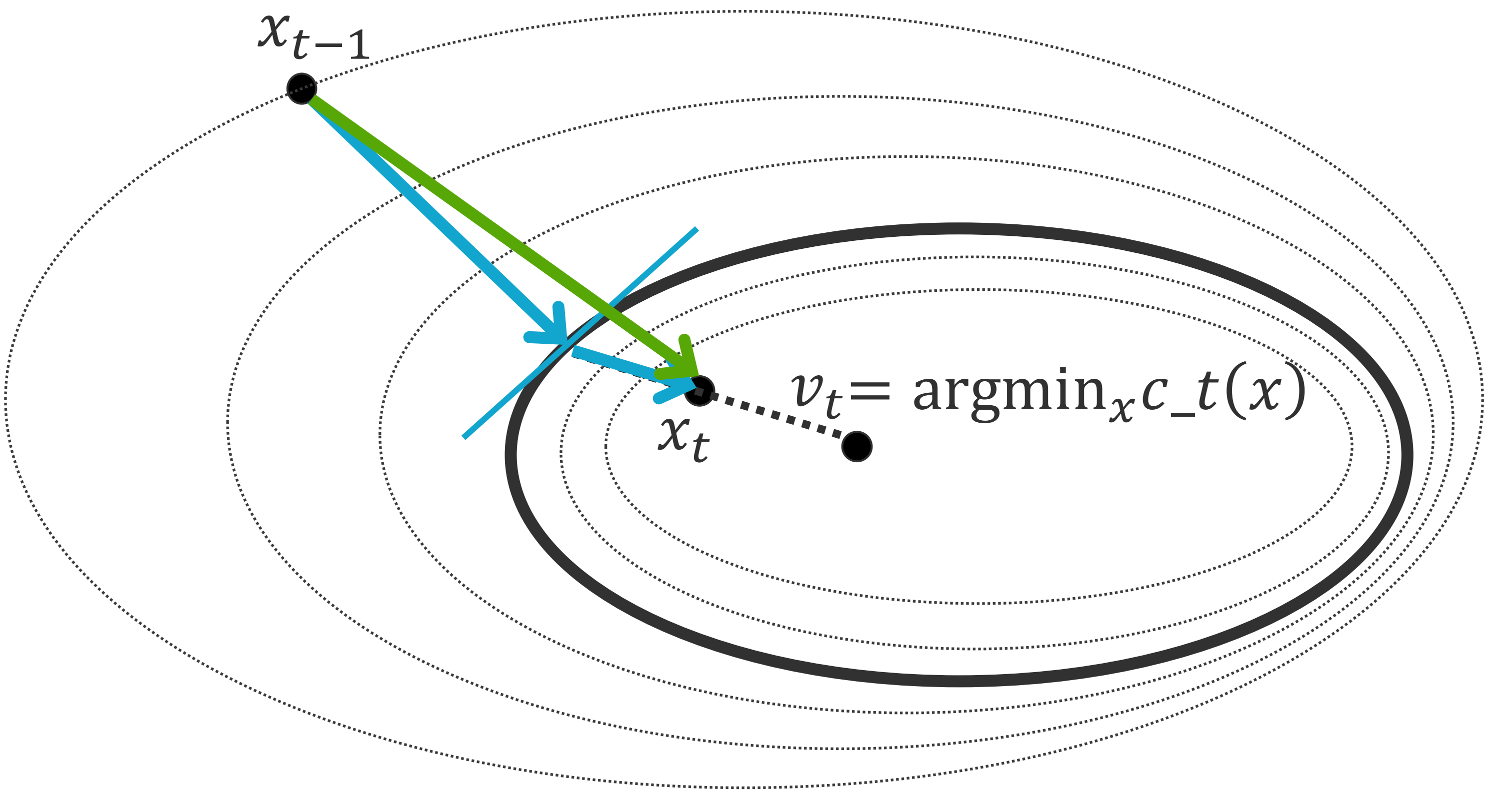
Key idea: OBD is not “greedy” enough.

Theorem [GLSW19]: OBD has a competitive ratio of $\Omega(m^{-2/3})$.

Greedy Online Balanced Descent (G-OBD)

"Geometric view"

Project onto a level set $K_l = \{ x | c_t(x) \leq l\}$ where l balances costs, i.e., $||x(l) - x_{t-1}|| = \beta l$. Then, take an $O\left(\sqrt{m}\right)$-size step toward the minimizer.



Greedy Online Balanced Descent (G-OBD)

"Geometric view"

Project onto a level set $K_l = \{ x | c_t(x) \leq l\}$
where l balances costs, i.e., $\|x(l) - x_{t-1}\| = \beta l$.
Then, take an $O\left(\sqrt{m}\right)$-size step toward the minimizer.

"Local view"

$$x_t = \text{argmin}_x \ \underbrace{c_t(x) + \frac{\lambda_1}{2}\|x - x_{t-1}\|^2}_{} + \lambda_2 c(x, v_t)$$

Project onto a level set!

$$\nabla c_t(x) + \lambda(x - x_{t-1}) = 0$$

Greedy Online Balanced Descent (G-OBD)

“Geometric view”

Project onto a level set $K_l = \{ x | c_t(x) \leq l \}$ where l balances costs, i.e., $||x(l) - x_{t-1}|| = \beta l$. Then, take an $O\left(\sqrt{m}\right)$-size step toward the minimizer.

“Local view”

$$x_t = \mathrm{argmin}_x \;\; c_t(x) + \frac{\lambda_1}{2} ||x - x_{t-1}||^2 + \lambda_2 c(x, v_t)$$

Regularized Online Balanced Descent (R-OBD)

Computationally easier and…

Theorem [GLSW19]: Regularized Online Balanced Descent (ROBD) is $\frac{1}{2}\left(1+\sqrt{1+4/m}\right)$-competitive for m-strongly convex hitting costs and switching costs that are either the squared-L_2 norm or Bregman divergence.

Regularized Online Balanced Descent (R-OBD)

Computationally easier and obtains the optimal competitive ratio!

Can an algorithm be constant competitive?

Yes, with a little structure!

...and the results extend to settings with delay & memory

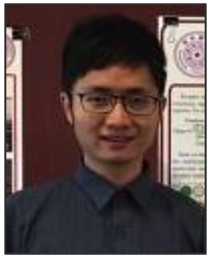
Theorem [SLCYW20] [PSCYW20+]: Optimistic ROBD(λ) is

$$O(l+\alpha)^k \max\{\frac{1}{\lambda}, \frac{\lambda+m}{m+(1-\alpha^2)\lambda}\}\text{-competitive}$$

for l-smooth, m-strongly convex hitting costs with delay k and memory of length p with $\alpha = \sum_{i=1}^{p} ||C_i||$.


Weici Pan


Yiheng Lin


Guanya Shi


Soon-Jo Chung

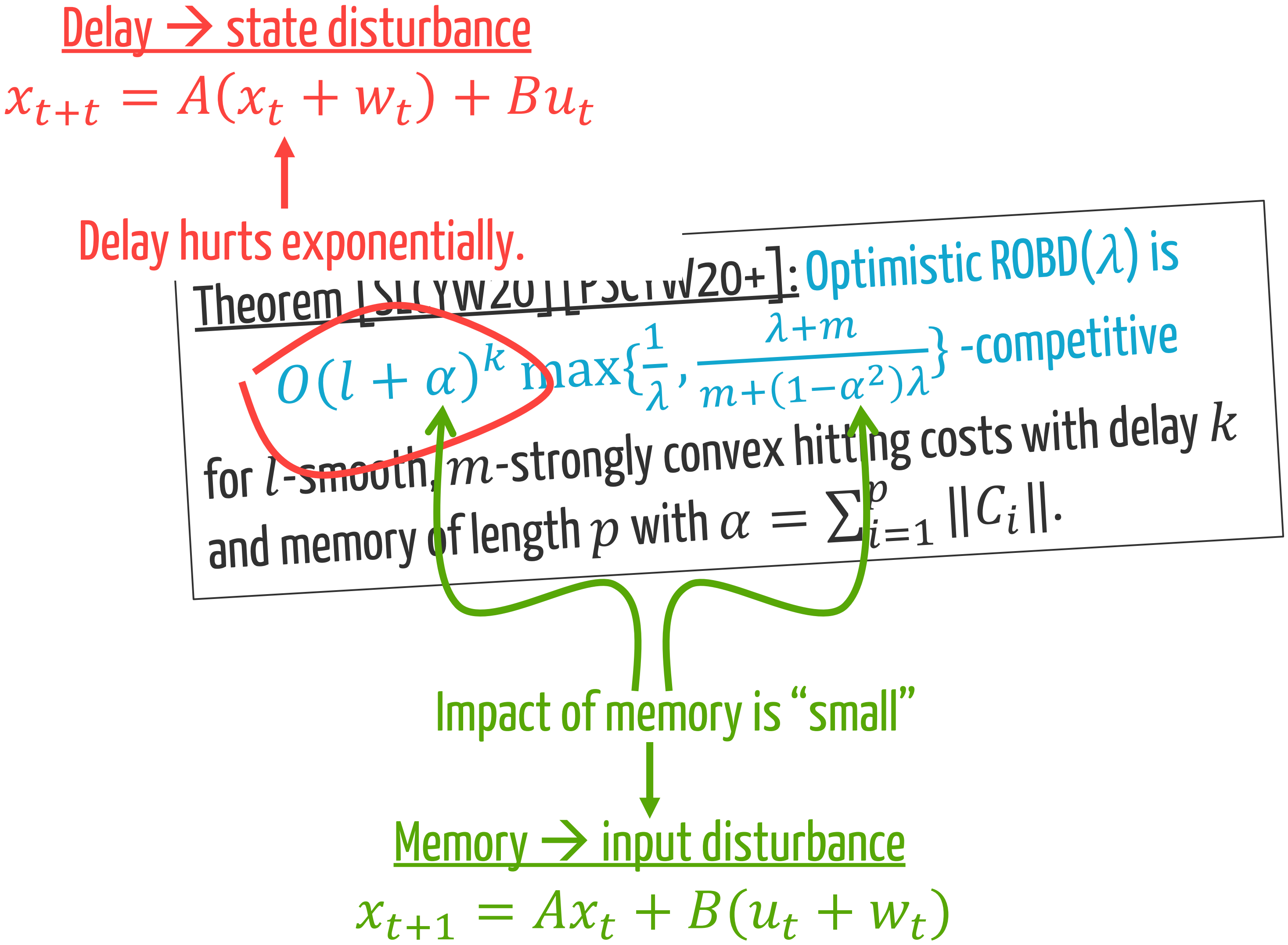

Yisong Yue

Theorem [SLCYW20][PSCYW20+]: Optimistic ROBD(λ) is

$$O(l+\alpha)^k \max\{\frac{1}{\lambda}, \frac{\lambda+m}{m+(1-\alpha^2)\lambda}\}$$ -competitive

for l-smooth, m-strongly convex hitting costs with delay k and memory of length p with $\alpha = \sum_{i=1}^{p} ||C_i||$.

Implies the first constant competitive policy for LDS in the case of adversarial noise.



Delay → state disturbance

$$x_{t+t} = A(x_t + w_t) + Bu_t$$

Delay hurts exponentially.

Theorem [SLCYW20][PSCYW20+]: Optimistic ROBD(λ) is $O(l + \alpha)^k \max\{\frac{1}{\lambda}, \frac{\lambda+m}{m+(1-\alpha^2)\lambda}\}$ -competitive for l-smooth, m-strongly convex hitting costs with delay k and memory of length p with $\alpha = \sum_{i=1}^{p} ||C_i||$.

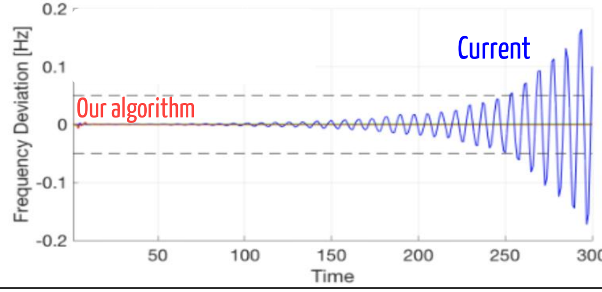
Impact of memory is "small"

Memory → input disturbance

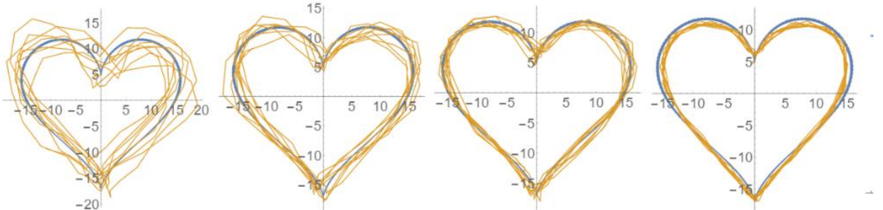
$$x_{t+1} = Ax_t + B(u_t + w_t)$$

Theorem [SLCYW20] [PSCYW20+]: Optimistic ROBD(λ) is $O(l+\alpha)^k \max\{\frac{1}{\lambda}, \frac{\lambda+m}{m+(1-\alpha^2)\lambda}\}$ -competitive for l-smooth, m-strongly convex hitting costs with delay k and memory of length p with $\alpha = \sum_{i=1}^{p} ||C_i||$.

Example: Frequency regulation with time-varying inertia



Example: Trajectory tracking with double integrator dynamics



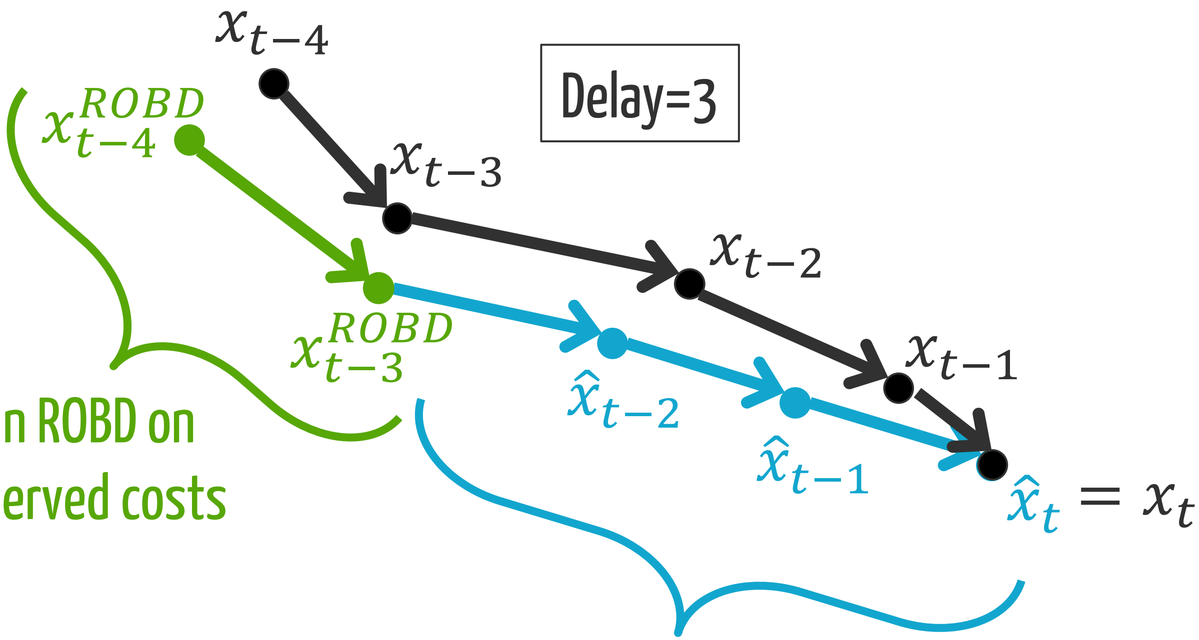
delay=5 delay=3 delay=0 5 predictions

Optimistic ROBD

Key idea: "optimistically" track the ROBD (full information) trajectory.

Optimistic ROBD

Key idea: “optimistically” track the ROBD (full information) trajectory.

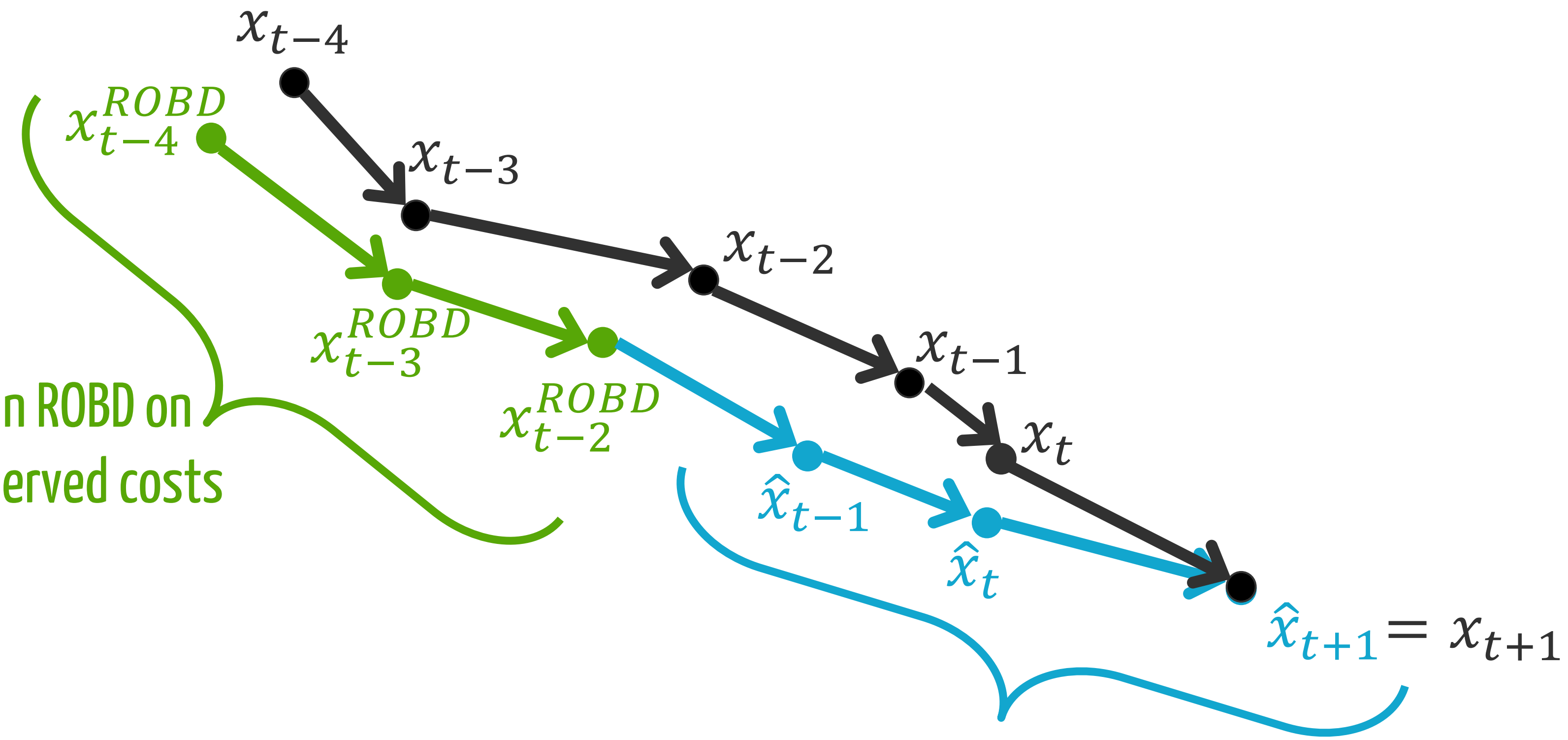


Run ROBD on observed costs

Run ROBD on cost function that minimizes cost experienced to ROBD $(\mathrm{i.e.}, v_{t-i}^*)$

Optimistic ROBD

Key idea: “optimistically” track the ROBD (full information) trajectory.



Run ROBD on observed costs

Run ROBD on cost function (i. e., v_{t-i}) that minimizes cost experienced to ROBD

Can an algorithm be constant competitive?

Yes, with a little structure!

...and the results extend to settings with delay & memory

& storage/inventory constraints

& non-convex costs

& predictions

Competitive Control via Online Optimization

Huge progress!

No predictions: LCP → RBG → ... → ROBD → Optimistic ROBD...

Predictions: MPC → AFHC → CHC → RHGD → SFHC ...

Can we deepen the connection between SOCO and control?

Many successful applications!

Sustainable Data Centers,
EV charging, video streaming,
CDNs, Microgrids, ...

...still lots of open questions remain!

Competitive Control via Online Optimization

Papers introducing Online Balanced Descent and its variations:

- N Chen, G Goel, A Wierman. Smoothed Online Convex Optimization in High Dimensions via Online Balanced Descent. Conference on Learning Theory (COLT) 2018
- G Goel, A Wierman. An Online Algorithm for Smoothed Regression and LQR Control. Conference on Artificial Intelligence and Statistics (AISTATS) 2019
- G Goel, Y Lin, H Sun, A Wierman. Beyond Online Balanced Descent: An Optimal Algorithm for Smoothed Online Optimization. NeurIPS 2019, oral spotlight.
- G Shi, Y Lin, S Chung, Y Yue, and A Wierman. Online Optimization with Memory and Competitive Control. NeurIPS 2020.
- Y Lin, G Goel, A Wierman. Online Optimization with Predictions and Non-Convex Losses. Sigmetrics 2020.
- W Pan, G Shi, S Chung, Y Yue, and A Wierman. Competitive Control via Online Optimization. Under Preparation.

Papers connecting online optimization with distributed optimization:

- P London, S Vardi, A Wierman, and H Yi. A parallizable acceleration framework for packing linear programs. AAAI 2018.
- P London, N Chen, S Vardi, and A Wierman. Logarithmic Communication for Distributed Optimization in Multi-Agent Systems. Sigmetrics 2020.

Papers bridging model-based & model-free control:

- G Qu, A Wierman. Finite-time Analysis of Asynchronous Stochastic Approximation and Q-learning. COLT 2020.
- G Qu, A Wierman, N Li. Scalable Reinforcement Learning of Localized Policies for Multi-Agent Networked Systems. L4DC 2020, oral spotlight.
- G Qu, Y Lin, A Wierman, N Li. Scalable Multi-Agent Reinforcement Learning for Networked Systems with Average Reward. NeurIPS 2020.
- G Qu, C Yu, S Low, A Wierman. Combining Model-Based and Model-Free Methods for Nonlinear Control: A Provably Convergent Policy Gradient Approach. Under submission