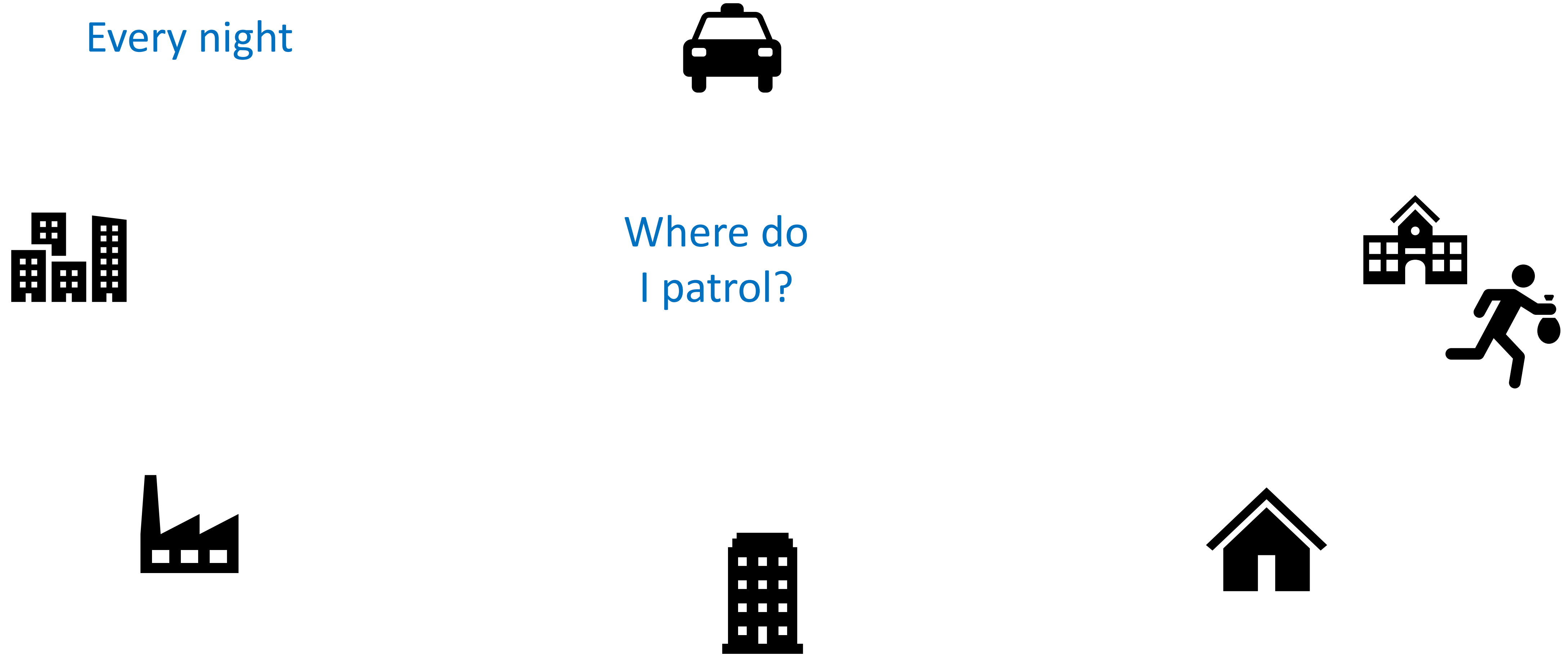
Online Non-Convex Learning: Following Perturbed Leader is Optimal

Arun Sai Suggala
CMU

Praneeth Netrapalli
Microsoft Research India

Example I : Patrolling



Example II : Portfolio selection

Every month

Where do
I invest?

Stock 1

Stock 2

Stock 3

Online learning

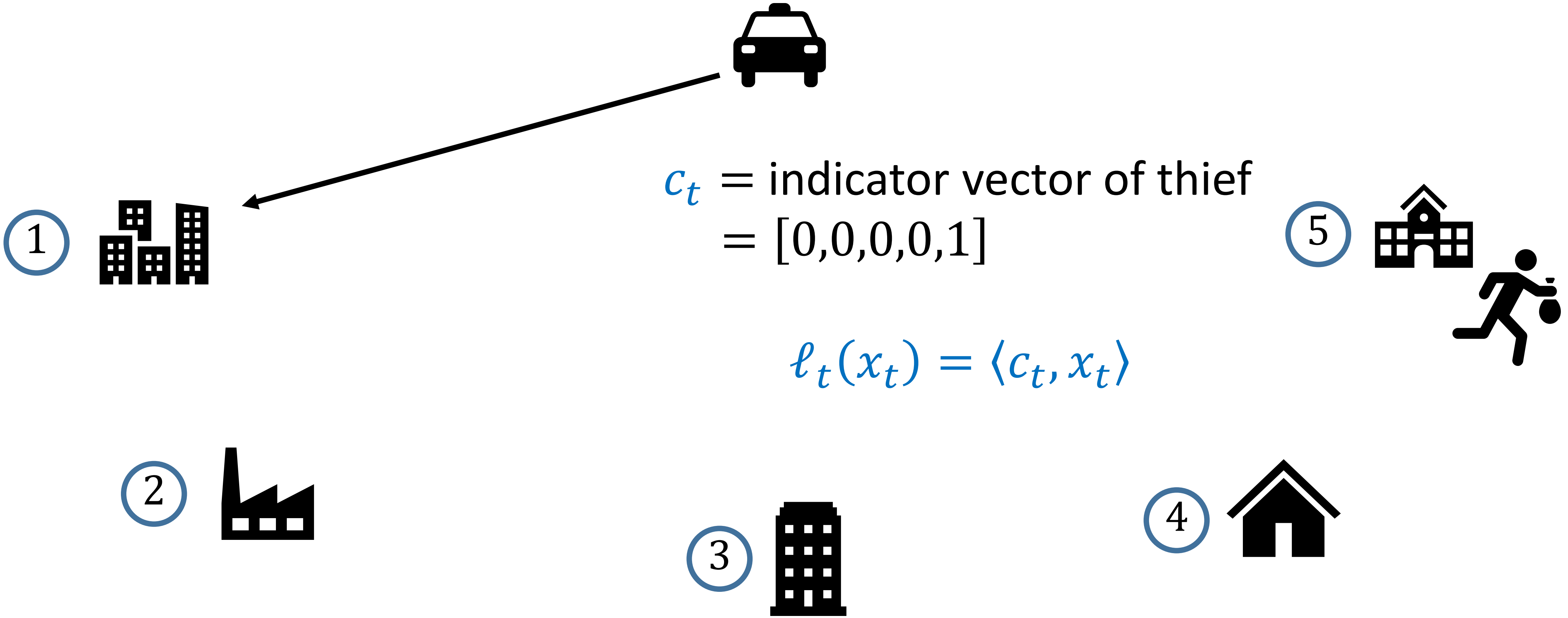
- Time: $1,2,\cdots,t,\cdots,T$
- At time t, predict $x_t \in \mathcal{X}$
- After playing x_t, observe loss function ℓ_t
- **Goal**: minimize cumulative loss $\sum_{t=1}^{T} \ell_t(x_t)$

Example I : Patrolling

x_t = indicator vector of *no* patrol
$= [0,1,1,1,1]$

c_t = indicator vector of thief
$= [0,0,0,0,1]$

$$\ell_t(x_t) = \langle c_t, x_t \rangle$$



Example II : Portfolio selection

x_t = indicator vector of investment
$= [1,0,0]$

c_t = negative yield of different venues
$= -[1.1, 0.9, 1.05]$

$$\ell_t(x_t) = \langle c_t, x_t \rangle$$

Stock 1

Stock 2

Stock 3

Example III : Solving minimax problems

- Minimax problems = Zero sum games

- Widely studied in Economics

- Several recent applications in machine learning: generative adversarial networks (GANs), multi-agent reinforcement learning (RL), adversarial training etc.

- Online learning algorithms can be used to solve such problems

Online learning

- At time t, predict x_t and observe loss function ℓ_t
    - ℓ_t fixed ahead of time
- **Goal**: minimize cumulative loss $\sum_{t=1}^{T} \ell_t(x_t)$
- **Benchmark**: $\min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x)$ — *best fixed policy in hindsight*
- **Regret**: $\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x)$ **Minimize regret**

History

- Online *linear* learning: dates back to [Brown and von Neumann 1950]
- Online *convex* learning: Heavily studied since [Zinkevich 2003]
- **Regret**

$$\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x) \leq O\left(\sqrt{T}\right)$$

Online *nonconvex* learning

- Computationally intractable even if all $\ell_t(\cdot)$ are the same

What can we do?

1. Weaker notions of regret (such as stationarity in optimization)
   - [Hazan, Singh and Zhang 2017]

2. Assume access to optimization oracles (only deal with learning)
   - [Agarwal, Gonen and Hazan 2018]

Main result

Setting

- $\ell_t(\cdot)$ is Lipschitz continuous
- $x_t \in \mathcal{X}$ with bounded diameter

Our result

- **Regret**:

$$\sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x) \leq O(\sqrt{T})$$

- Previous best: $O(T^{2/3})$

[Agarwal, Gonen, Hazan 2018]

Algorithm I: Follow the leader

- For any $t \leq T$ leader $\tilde{x}_t \stackrel{\text{def}}{=} \underset{x \in \mathcal{X}}{\operatorname{argmin}} \sum_{i=1}^{t} \ell_i(x)$

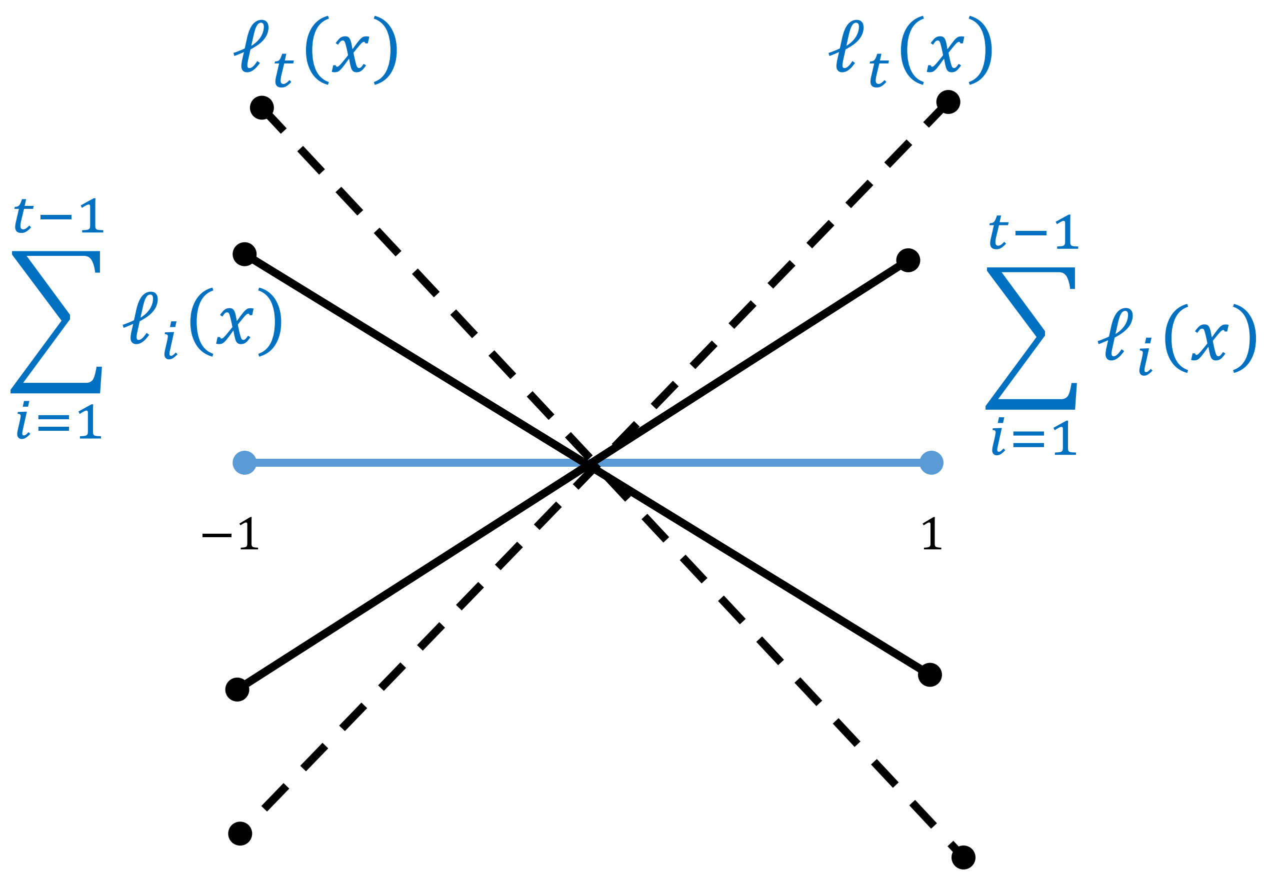
- Cannot compute $\tilde{x}_t$ – do not know $\ell_t(\cdot)$

- Choose $x_t = \tilde{x}_{t-1}$

Algorithm I: Follow the leader

Regret = $\Omega(T)$

- Choose $x_t \stackrel{\text{def}}{=} \underset{x \in \mathcal{X}}{\operatorname{argmin}} \sum_{i=1}^{t-1} \ell_i(x)$
- Performs poorly!
- $\mathcal{X} = [-1,1]$
- $\ell_1(x) = x$
- $\ell_i(x) = \begin{cases} -2x, & i \text{ is even} \\ 2x, & i \text{ is odd} \end{cases}$

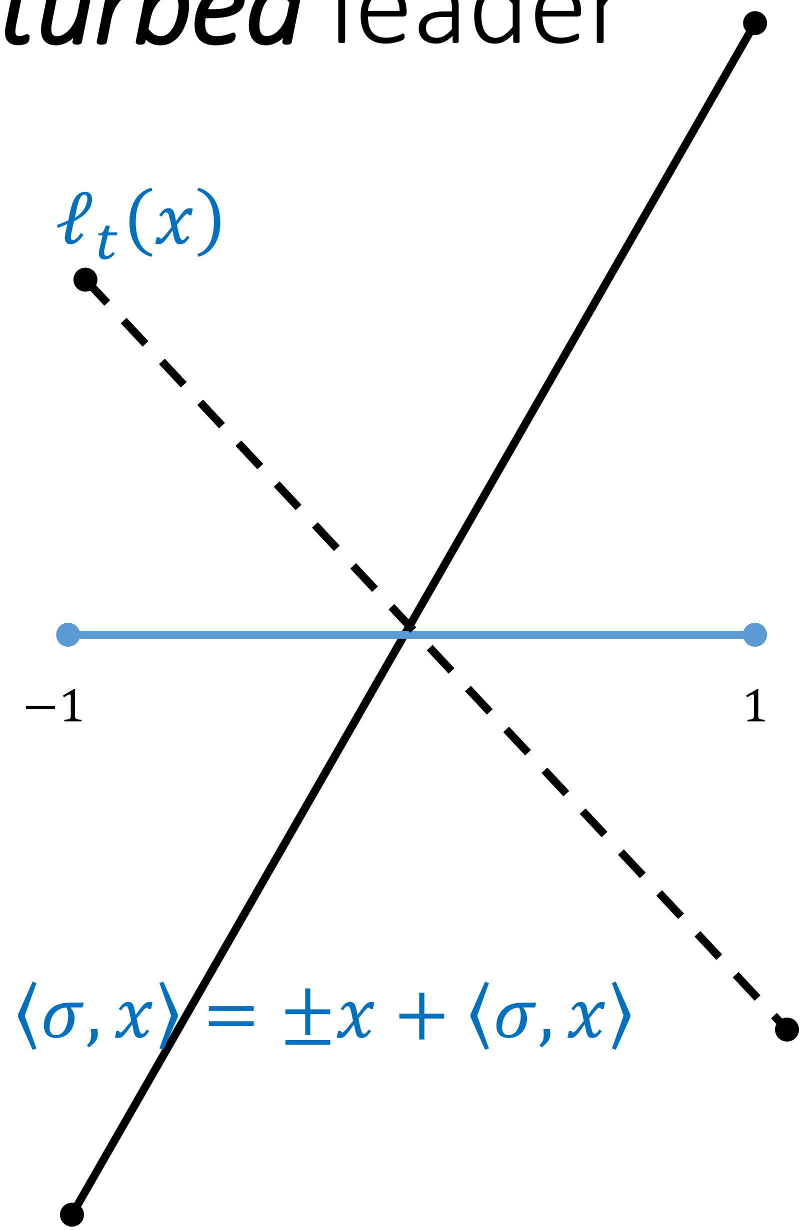


Algorithm I: Follow the *perturbed* leader

- [Hannan 1957], [Kalai, Vempala 2005]

- $\sigma \sim \text{Unif}(0, \sqrt{T})$
- $x_t \stackrel{\text{def}}{=} \underset{x \in \mathcal{X}}{\text{argmin}} \sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle$

- Regret $= O(\sqrt{T})$



$$\sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle = \pm x + \langle \sigma, x \rangle$$

Main intuitions

- Recall: adversary fixes choices ahead of time
- Be the leader lemma
  - Recall, $\boldsymbol{x_t} \stackrel{\text{def}}{=} \underset{x \in \mathcal{X}}{\operatorname{argmin}} \sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle$
  - $\mathrm{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{x_{t+1}})\right] - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x) \leq O\left(\sqrt{T}\right)$ since $\sigma \sim \sqrt{T}$
- Stability
  - $\mathrm{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{x_t})\right] - \mathrm{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{x_{t+1}})\right] \leq T \cdot L \cdot \mathrm{E}[\|\boldsymbol{x_t} - \boldsymbol{x_{t+1}}\|]$

Stability question

- Recall $\boldsymbol{x_t} \overset{\text{def}}{=} \underset{x \in \mathcal{X}}{\operatorname{argmin}} \sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle$

- How large can $\mathrm{E}[\|\boldsymbol{x_t} - \boldsymbol{x_{t+1}}\|]$ be?

Our improvement
$O\left(T^{-1/2}\right)$

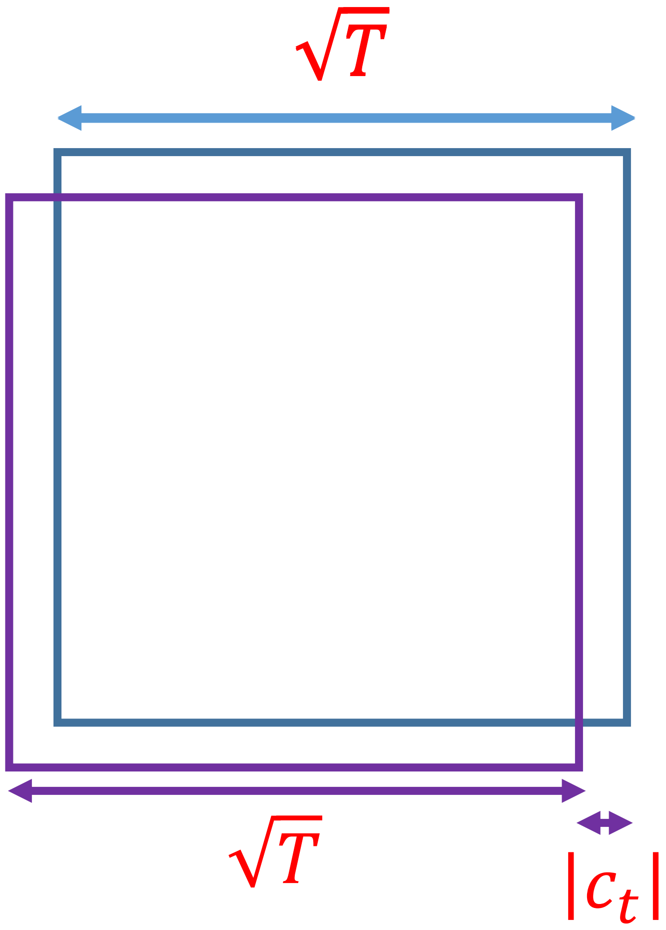
- Agarwal, Gonen, Hazan 2018

$$\mathrm{E}[\|\boldsymbol{x_t} - \boldsymbol{x_{t+1}}\|] \leq O\left(T^{-\frac{1}{3}}\right)$$

Linear case [Kalai and Vempala 2005]

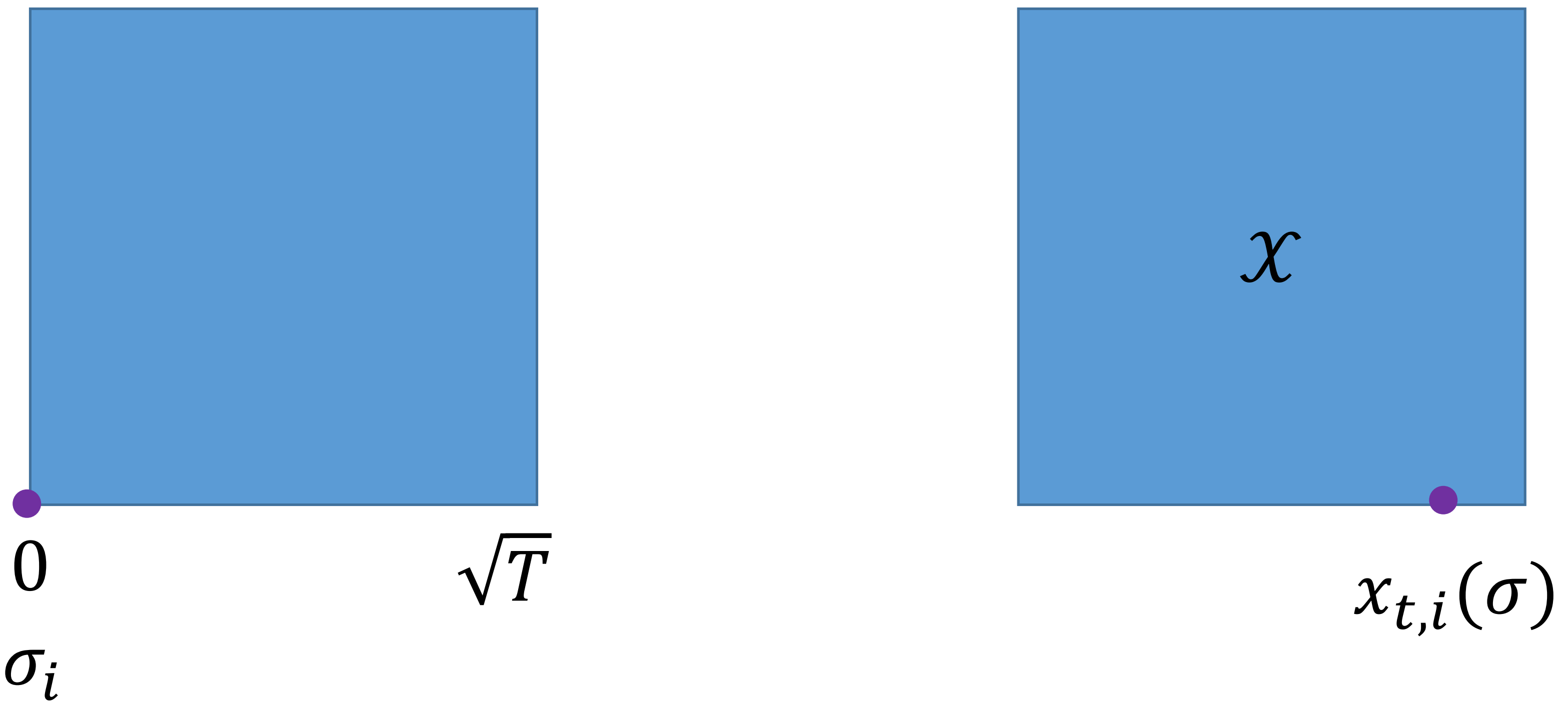
- $\ell_i(x) = \langle c_i, x\rangle$; $\ell_i(\cdot)$ Lipschitz $\Rightarrow c_i$ bounded
- $\sigma \sim \text{Unif}(0, \sqrt{T})$
- $\sum_{i=1}^{t} \ell_i(x) + \langle \sigma, x\rangle = \left\langle \sigma + \sum_{i=1}^{t} c_i \, , x \right\rangle$

- **Key idea**:
- $\sigma + \sum_{i=1}^{t-1} c_i \sim \sigma + \sum_{i=1}^{t} c_i$
- $x_t \sim x_{t+1}$



The general nonconvex case

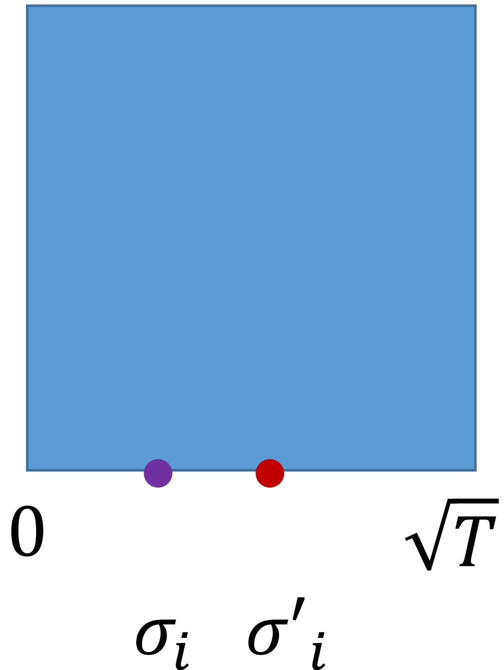
- $x_t(\sigma) \stackrel{\text{def}}{=} \underset{x \in \mathcal{X}}{\operatorname{argmin}} \sum_{i=1}^{t-1} \ell_i(x) + \langle \sigma, x \rangle$
- *Weak monotonicity property:* $x_{t,i}(\sigma + c e_i) \leq x_{t,i}(\sigma) \;\; \forall \;\; c \geq 0$

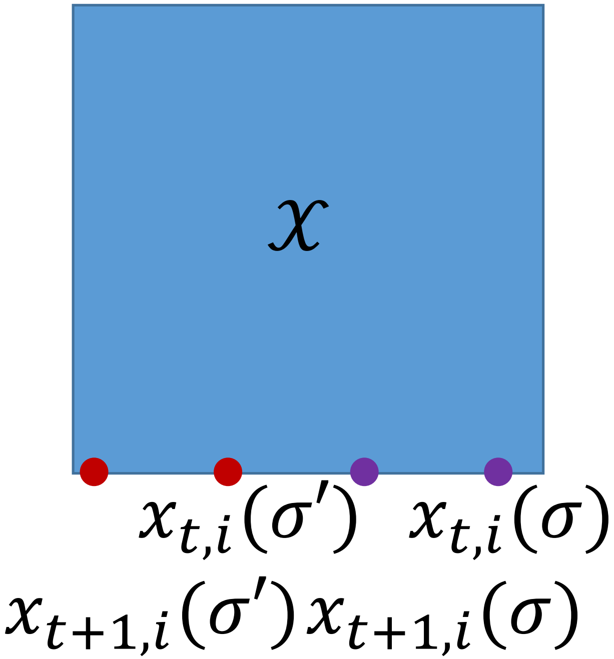


Strong monotonicity property

- Suppose $\|x_t(\sigma) - x_{t+1}(\sigma)\|_1 \leq 10d \cdot |x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|$
- Then for $\sigma' = \sigma + 100Lde_i$,

$$\max\left(x_{t,i}(\sigma'), x_{t+1,i}(\sigma')\right) \leq \max\left(x_{t,i}(\sigma), x_{t+1,i}(\sigma)\right) - \frac{9}{10}|x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|$$





Strong monotonicity property

- Suppose $\|x_t(\sigma) - x_{t+1}(\sigma)\|_1 \leq 10d \cdot |x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|$
- Then for $\sigma' = \sigma + 100Lde_i$,

$$\max\left(x_{t,i}(\sigma'), x_{t+1,i}(\sigma')\right) \leq \max\left(x_{t,i}(\sigma), x_{t+1,i}(\sigma)\right) - \frac{9}{10}|x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|$$

$$\mathrm{E}\left[|x_{t,i}(\sigma) - x_{t+1,i}(\sigma)|\right] \leq \frac{1}{10d}\mathrm{E}[\|x_t(\sigma) - x_{t+1}(\sigma)\|_1] + O(d\|\mathcal{X}\|_\infty)$$

$$\mathrm{E}[\|x_t(\sigma) - x_{t+1}(\sigma)\|_1] \leq \frac{1}{10}\mathrm{E}[\|x_t(\sigma) - x_{t+1}(\sigma)\|_1] + O(d^2\|\mathcal{X}\|_\infty)$$

Recap

- Follow the perturbed leader
- Be the leader lemma: playing $\boldsymbol{x_{t+1}}$ at time $\boldsymbol{t}$ is very good
- Stability: With perturbations, $\|\boldsymbol{x_t} - \boldsymbol{x_{t+1}}\|$ very small
- Key technical results: Tight monotonicity lemmas

Summary

- Online learning a powerful and interesting framework
- Most work so far in convex setting
- Nonconvex setting quite important in practice – work has just begun
- Perturbation based strategies are useful in this context as well