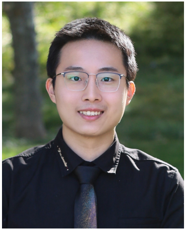
Dynamic Allocation of Reusable Resources to Strategic Agents under Long-Term Constraints

(NeurIPS'25; **Winner** of ACM Student Research @ SIGMETRICS'25)

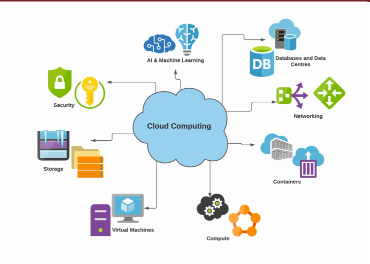
Yan Dai Negin Golrezaei Patrick Jaillet
Massachusetts Institute of Technology



Today

Motivating Examples

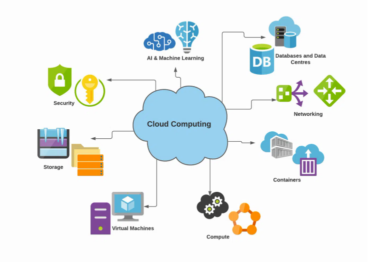
GPU Allocation



- **Resource:** Reusable GPU
- **Agents:** Research groups
- **Constr:** Energy & budget

Motivating Examples

GPU Allocation

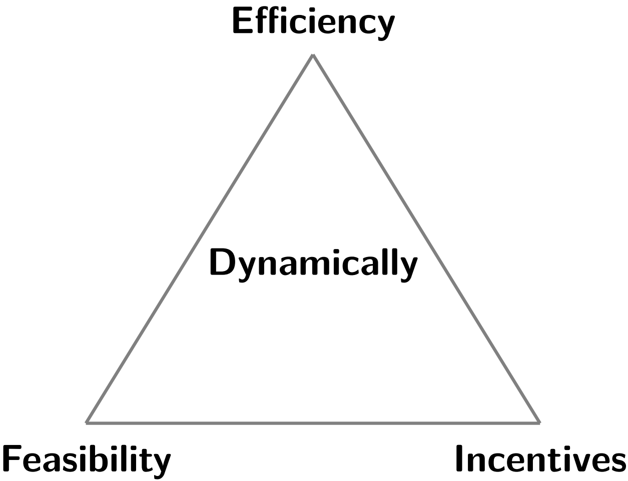


- **Resource:** Reusable GPU
- **Agents:** Research groups
- **Constr:** Energy & budget

Mobile Health Unit

- **Resource:** MHU
- **Agents:** Remote regions
- **Constr:** Staffing & budget

Efficiency-Feasibility-Incentives Trilemma



Efficiency-Feasibility-Incentives Trilemma

Efficiency

- T rounds, K agents, value $v_{t,i} \sim$ **unknown** $\mathcal{V}_i$

$$\boxed{\text{Max value: } \sum_t v_{t,i_t}}$$

Efficiency-Feasibility-Incentives Trilemma

Efficiency

- T rounds, K agents, value $v_{t,i} \sim$ **unknown** $\mathcal{V}_i$

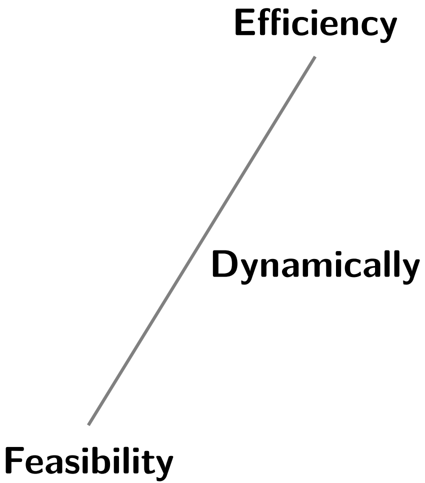
$$\boxed{\text{Max value: } \sum_t v_{t,i_t}}$$

- Alloc cost $\boldsymbol{c}_{t,i}$ (d-dim), **iid** $\sim$ **unknown** $\mathcal{C}_i$

$$\boxed{\text{Constr: } \sum_t \boldsymbol{c}_{t,i_t} \le T\boldsymbol{\rho}}$$

Feasibility

Efficiency-Feasibility-Incentives Trilemma



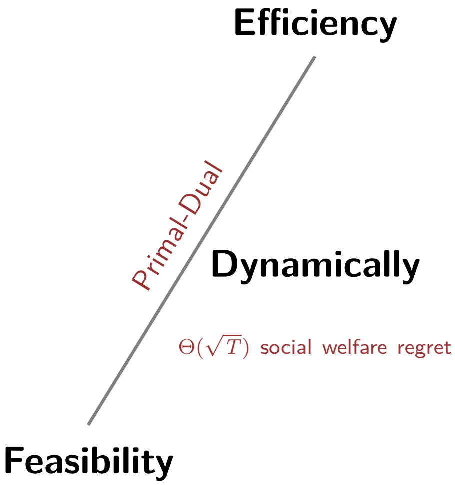
- T rounds, K agents, value $v_{t,i} \sim$ **unknown** $\mathcal{V}_i$

$$\boxed{\text{Max value: } \sum_t v_{t,i_t}}$$

- Alloc cost $\boldsymbol{c}_{t,i}$ (d-dim), **iid** $\sim$ **unknown** $\mathcal{C}_i$

$$\boxed{\text{Constr: } \sum_t \boldsymbol{c}_{t,i_t} \leq T\boldsymbol{\rho}}$$

Efficiency-Feasibility-Incentives Trilemma



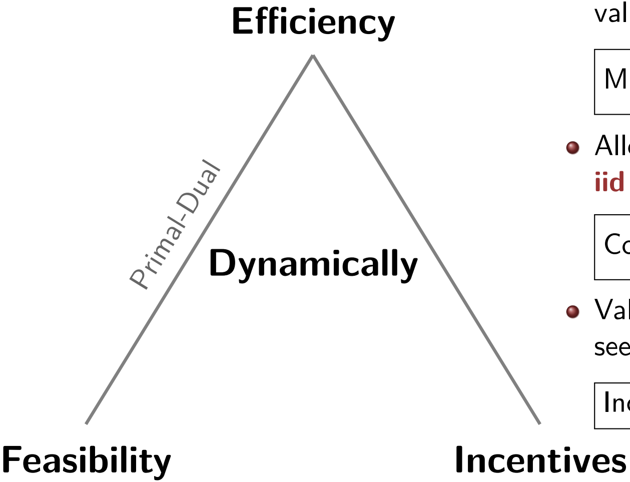
- T rounds, K agents, value $v_{t,i} \sim$ **unknown** $\mathcal{V}_i$

$$\boxed{\text{Max value: } \sum_t v_{t,i_t}}$$

- Alloc cost $\boldsymbol{c}_{t,i}$ (d-dim), **iid** $\sim$ **unknown** $\mathcal{C}_i$

$$\boxed{\text{Constr: } \sum_t \boldsymbol{c}_{t,i_t} \le T\boldsymbol{\rho}}$$

Efficiency-Feasibility-Incentives Trilemma



- T rounds, K agents, value $v_{t,i} \sim$ **unknown $\mathcal{V}_i$**

$$\boxed{\text{Max value: } \sum_t v_{t,i_t}}$$

- Alloc cost $\boldsymbol{c}_{t,i}$ (d-dim), **iid $\sim$ unknown $\mathcal{C}_i$**

$$\boxed{\text{Constr: } \sum_t \boldsymbol{c}_{t,i_t} \le T\boldsymbol{\rho}}$$

- Value $v_{t,i}$ private; only see **strategic report $u_{t,i}$**

$$\boxed{\text{Incentivize } u_{t,i} \approx v_{t,i}}$$

Efficiency-Feasibility-Incentives Trilemma

Efficiency

Dynamically

Incentives

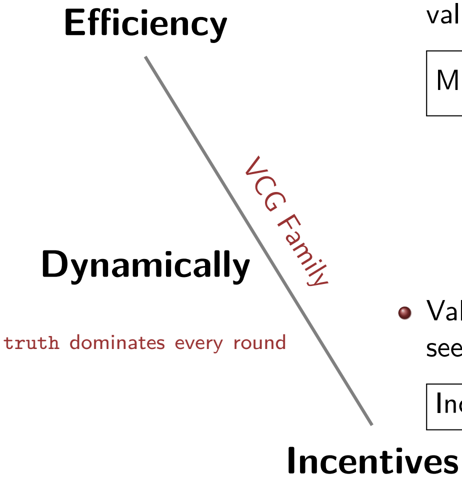
- T rounds, K agents, value $v_{t,i} \sim$ **unknown** $\mathcal{V}_i$

$\boxed{\text{Max value: } \sum_t v_{t,i_t}}$

- Value $v_{t,i}$ private; only see **strategic report** $u_{t,i}$

$\boxed{\text{Incentivize } u_{t,i} \approx v_{t,i}}$

Efficiency-Feasibility-Incentives Trilemma



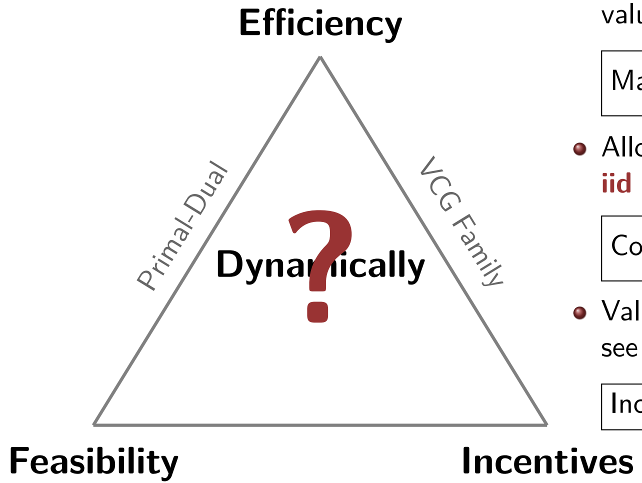
- T rounds, K agents, value $v_{t,i} \sim$ **unknown** $\mathcal{V}_i$

  Max value: $\sum_t v_{t,i_t}$

- Value $v_{t,i}$ private; only see **strategic report** $u_{t,i}$

  Incentivize $u_{t,i} \approx v_{t,i}$

Efficiency-Feasibility-Incentives Trilemma



- T rounds, K agents, value $v_{t,i} \sim$ **unknown** $\mathcal{V}_i$

$\boxed{\text{Max value: } \sum_t v_{t,i_t}}$

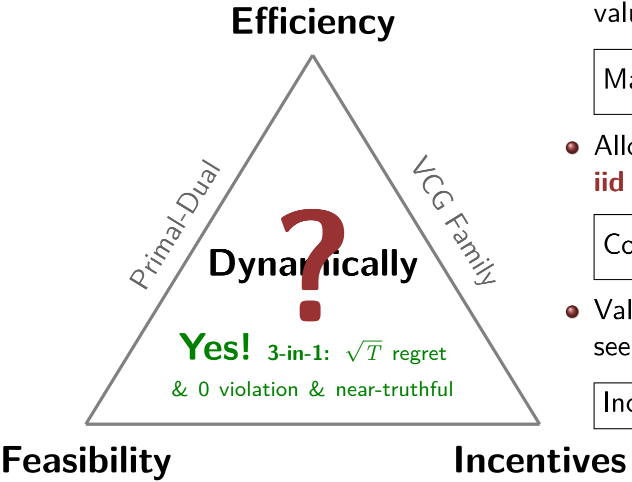
- Alloc cost $\boldsymbol{c}_{t,i}$ (d-dim), **iid** $\sim$ **unknown** $\mathcal{C}_i$

$\boxed{\text{Constr: } \sum_t \boldsymbol{c}_{t,i_t} \leq T\boldsymbol{\rho}}$

- Value $v_{t,i}$ private; only see **strategic report** $u_{t,i}$

$\boxed{\text{Incentivize } u_{t,i} \approx v_{t,i}}$

Efficiency-Feasibility-Incentives Trilemma



- T rounds, K agents, value $v_{t,i} \sim$ **unknown** $\mathcal{V}_i$

$$\boxed{\text{Max value: } \sum_t v_{t,i_t}}$$

- Alloc cost $\boldsymbol{c}_{t,i}$ (d-dim), **iid** $\sim$ **unknown** $\mathcal{C}_i$

$$\boxed{\text{Constr: } \sum_t \boldsymbol{c}_{t,i_t} \le T\boldsymbol{\rho}}$$

- Value $v_{t,i}$ private; only see **strategic report** $u_{t,i}$

$$\boxed{\text{Incentivize } u_{t,i} \approx v_{t,i}}$$

Formal Setup: Agents, Planner, Objectives

T rounds, K agents, 1 indivisible resource, d constraints.

Formal Setup: Agents, Planner, Objectives

T rounds, K agents, 1 indivisible resource, d constraints.

Type of agent $i \in [K]$ in round $t \in [T]$:

- **Private value** $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ (distribution $\mathcal{V}_i$ unknown)
- **Public cost** $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ (d-dim'l; distribution $\mathcal{C}_i$ unknown)

Formal Setup: Agents, Planner, Objectives

T rounds, K agents, 1 indivisible resource, d constraints.

Type of agent $i \in [K]$ in round $t \in [T]$:

- **Private value** $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ (distribution $\mathcal{V}_i$ unknown)
- **Public cost** $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ (d-dim'l; distribution $\mathcal{C}_i$ unknown)

Interaction protocol in round $t \in [T]$:

- Each agent $i \in [K]$ submits **arbitrary report** $u_{t,i}$
- Planner decides **allocation** i_t **& payment** p_{t,i_t} from $\boldsymbol{u}_t$ & $\boldsymbol{c}_t$

Formal Setup: Agents, Planner, Objectives

T rounds, K agents, 1 indivisible resource, d constraints.

Type of agent $i \in [K]$ in round $t \in [T]$:

- **Private value** $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ (distribution $\mathcal{V}_i$ unknown)
- **Public cost** $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ (d-dim'l; distribution $\mathcal{C}_i$ unknown)

Interaction protocol in round $t \in [T]$:

- Each agent $i \in [K]$ submits **arbitrary report** $u_{t,i}$
- Planner decides **allocation** i_t **& payment** p_{t,i_t} from $\boldsymbol{u}_t$ & $\boldsymbol{c}_t$

Agents' Objectives

Maximize γ-discounted utility:

$$\max \ \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t \mathbb{1}[i_t = i](v_{t,i} - p_{t,i})\right]$$

Formal Setup: Agents, Planner, Objectives

T rounds, K agents, 1 indivisible resource, d constraints.

Type of agent $i \in [K]$ in round $t \in [T]$:

- **Private value** $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ (distribution $\mathcal{V}_i$ unknown)
- **Public cost** $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ (d-dim'l; distribution $\mathcal{C}_i$ unknown)

Interaction protocol in round $t \in [T]$:

- Each agent $i \in [K]$ submits **arbitrary report** $u_{t,i}$
- Planner decides **allocation** i_t **& payment** p_{t,i_t} from $\boldsymbol{u}_t$ & $\boldsymbol{c}_t$

Agents' Objectives

Maximize γ-discounted utility:

$$\max \ \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t \mathbb{1}[i_t = i](v_{t,i} - p_{t,i})\right]$$

Planner's Objective

Maximize social welfare:

$$\max \ \mathbb{E}\left[\sum_{t=1}^{T} v_{t,i_t}\right] \text{ s.t. } \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{c}_{t,i_t} \le \boldsymbol{\rho}$$

Formal Setup: Agents, Planner, Objectives

T rounds, K agents, 1 indivisible resource, d constraints.

Type of agent $i \in [K]$ in round $t \in [T]$:

- **Private value** $v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i$ (distribution $\mathcal{V}_i$ unknown)
- **Public cost** $\boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$ (d-dim'l; distribution $\mathcal{C}_i$ unknown)

Interaction protocol in round $t \in [T]$:

- Each agent $i \in [K]$ submits **arbitrary report** $u_{t,i}$
- Planner decides **allocation** i_t **& payment** p_{t,i_t} from $\boldsymbol{u}_t$ & $\boldsymbol{c}_t$

Agents' Objectives

Maximize γ-discounted utility:

$$\max \ \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t \mathbb{1}[i_t = i](v_{t,i} - p_{t,i})\right]$$

Planner's Objective

Maximize social welfare:

$$\max \ \mathbb{E}\left[\sum_{t=1}^{T} v_{t,i_t}\right] \ \text{s.t.} \ \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{c}_{t,i_t} \le \boldsymbol{\rho}$$

Regret. $\mathbb{E}[\sum_t (v_{t,i_t^*} - v_{t,i_t})]$ with $\{i_t^*\}_{t\in[T]}$ constrained offline opt

Standard Tool: Primal-Dual Framework (when Truthful)

Primal-Dual Framework (when Truthful)

Primal (Good Allocations)

Dual (Track Constraints)

Standard Tool: Primal-Dual Framework (when Truthful)

Primal-Dual Framework (when Truthful)

Maintain dual variables $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_T \in \mathbb{R}^d_{\geq 0}$ as **implicit prices**:

Primal (Good Allocations)

Dual (Track Constraints)

Standard Tool: Primal-Dual Framework (when Truthful)

Primal-Dual Framework (when Truthful)

Maintain dual variables $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_T \in \mathbb{R}^d_{\geq 0}$ as **implicit prices**:

- **Round-t primal.** Make good "dual-adjusted" allocations

  (based on $\tilde{v}_{t,i} := v_{t,i} - \boldsymbol{\lambda}_t^\top \boldsymbol{c}_{t,i}$; $v_{t,i}$ = true value & $\boldsymbol{c}_{t,i}$ = true cost)

Primal (Good Allocations)
$i_t \leftarrow \text{argmax}_i (v_{t,i} - \boldsymbol{\lambda}_t^\top \boldsymbol{c}_{t,i})$

Dual $\boldsymbol{\lambda}_t \in \mathbb{R}^d_{\geq 0}$

Dual (Track Constraints)

Standard Tool: Primal-Dual Framework (when Truthful)

Primal-Dual Framework (when Truthful)

Maintain dual variables $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_T \in \mathbb{R}^d_{\geq 0}$ as **implicit prices**:

- **Round-t primal.** Make good "dual-adjusted" allocations (based on $\tilde{v}_{t,i} := v_{t,i} - \boldsymbol{\lambda}_t^\mathsf{T} \boldsymbol{c}_{t,i}$; $v_{t,i}$ = true value & $\boldsymbol{c}_{t,i}$ = true cost)
- **Round-t dual.** Update dual $\boldsymbol{\lambda}_{t+1}$ based on cost $\boldsymbol{c}_{t,i_t}$ (j-th budget over-spent $\Rightarrow \boldsymbol{\lambda}_{t+1,j}$ goes up)

Primal (Good Allocations)
$i_t \leftarrow \text{argmax}_i(v_{t,i} - \boldsymbol{\lambda}_t^\mathsf{T} \boldsymbol{c}_{t,i})$

Dual $\boldsymbol{\lambda}_t \in \mathbb{R}^d_{\geq 0}$

Allocation $i_t \in [K]$

Dual (Track Constraints)
$\boldsymbol{\lambda}_{t+1} \leftarrow \text{Upd}(\boldsymbol{\lambda}_t, \boldsymbol{c}_{t,i_t})$

Standard Tool: Primal-Dual Framework (when Truthful)

Primal-Dual Framework (when Truthful)

Maintain dual variables $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_T \in \mathbb{R}^d_{\geq 0}$ as **implicit prices**:

- **Round-t primal.** Make good "dual-adjusted" allocations (based on $\tilde{v}_{t,i} := v_{t,i} - \boldsymbol{\lambda}_t^\top \boldsymbol{c}_{t,i}$; $v_{t,i}$ = true value & $\boldsymbol{c}_{t,i}$ = true cost)
- **Round-t dual.** Update dual $\boldsymbol{\lambda}_{t+1}$ based on cost $\boldsymbol{c}_{t,i_t}$ (j-th budget over-spent $\Rightarrow \boldsymbol{\lambda}_{t+1,j}$ goes up)

Primal (Good Allocations)
$i_t \leftarrow \text{argmax}_i(v_{t,i} - \boldsymbol{\lambda}_t^\top \boldsymbol{c}_{t,i})$

Dual $\boldsymbol{\lambda}_t \in \mathbb{R}^d_{\geq 0}$

Allocation $i_t \in [K]$

Dual (Track Constraints)
$\boldsymbol{\lambda}_{t+1} \leftarrow \text{Upd}(\boldsymbol{\lambda}_t, \boldsymbol{c}_{t,i_t})$

What's Wrong when Agents are Strategic?

Immediate feedback loop is target of **strategic manipulations!**

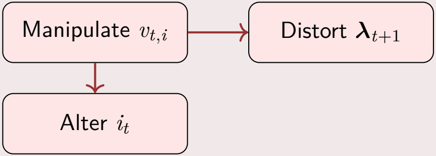
Standard Tool Fails: Vicious Cycle of Manipulation

What Happens With Strategic Agents?

$$\begin{tikzcd}
\text{Manipulate } v_{t,i} \arrow[d] \\
\text{Alter } i_t
\end{tikzcd}$$

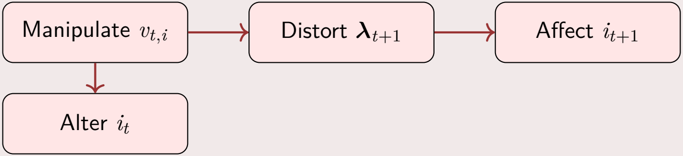
Standard Tool Fails: Vicious Cycle of Manipulation

What Happens With Strategic Agents?



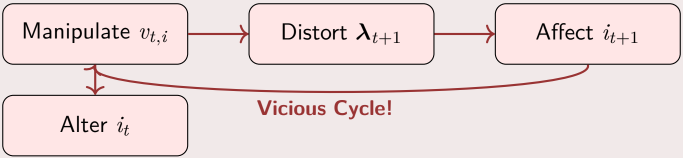
Standard Tool Fails: Vicious Cycle of Manipulation

What Happens With Strategic Agents?



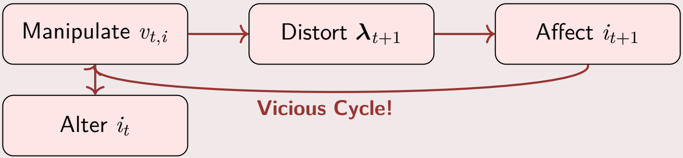
Standard Tool Fails: Vicious Cycle of Manipulation

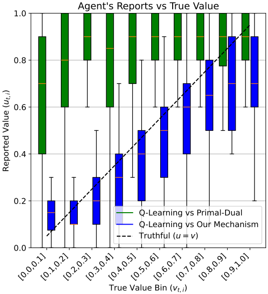
What Happens With Strategic Agents?

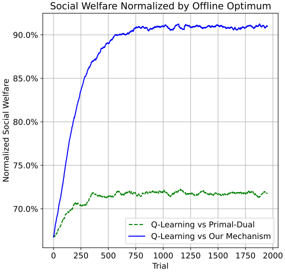


Standard Tool Fails: Vicious Cycle of Manipulation

What Happens With Strategic Agents?



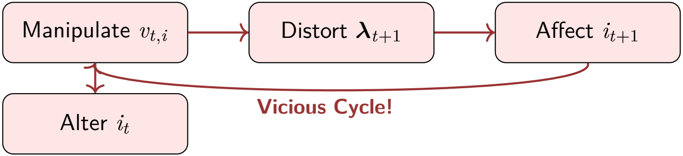




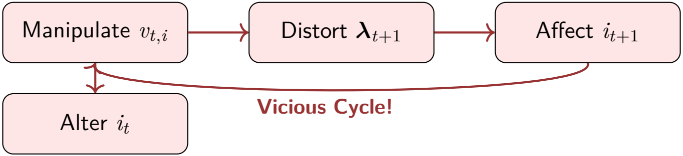
Today

1. Setup: A Trilemma in Online Mechanism Design
   - Efficiency-Incentive-Feasibility Trilemma
   - Formal Setup: Agents, Planner, Objectives
   - Vicious Cycle: Classical Primal-Dual Fails
2. Primal Side: Incentive-Aware Allocation
   - Pricing to Evict "Short-Term" Impact
   - Epoching to Mitigate "Long-Term" Impact
   - Exploration to Remove "Long-Term" Impact
3. Dual Side: Online Learning for Updates
   - Online Learning View of Dual Updates
   - Our Plan: Promise of Predictability
   - Beyond Optimistic FTRL: O-FTRL-FP

Pricing to Evict "Short-Term" Impact



Pricing to Evict "Short-Term" Impact

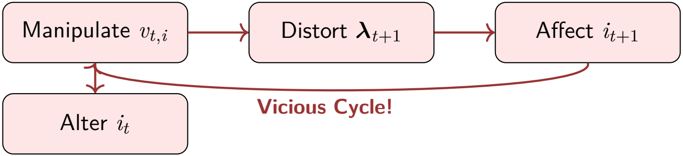


Ingredient 1: Dual-Adjusted Pricing

- VCG-like auction on "dual-adjusted reports", defined as:

$$\tilde{u}_{t,i} := u_{t,i} - \boldsymbol{\lambda}_t^\mathsf{T} \boldsymbol{c}_{t,i}, \quad \forall i \in [K].$$

Pricing to Evict "Short-Term" Impact



Ingredient 1: Dual-Adjusted Pricing

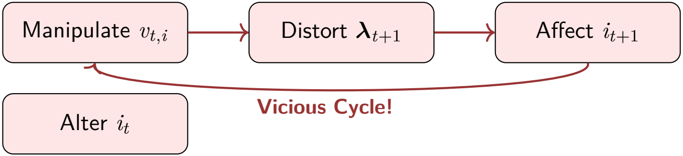
- VCG-like auction on "dual-adjusted reports", defined as:

$$\tilde{u}_{t,i} := u_{t,i} - \boldsymbol{\lambda}_t^\top \boldsymbol{c}_{t,i}, \quad \forall i \in [K].$$

- Winner $i_t := \operatorname{argmax}_i \tilde{u}_{t,i}$ pays 2nd highest price:

$$p_{t,i_t} := \boldsymbol{\lambda}_t^\top \boldsymbol{c}_{t,i_t} + \max_{j \neq i_t} \tilde{u}_{t,j}.$$

Pricing to Evict "Short-Term" Impact



Ingredient 1: Dual-Adjusted Pricing

- VCG-like auction on "dual-adjusted reports", defined as:

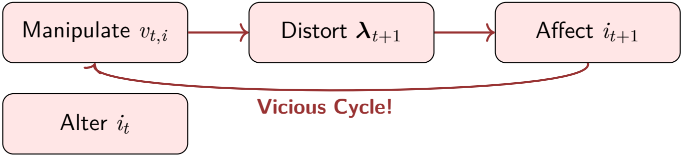
$$\tilde{u}_{t,i} := u_{t,i} - \boldsymbol{\lambda}_t^\mathsf{T} \boldsymbol{c}_{t,i}, \quad \forall i \in [K].$$

- Winner $i_t := \operatorname{argmax}_i \tilde{u}_{t,i}$ pays 2nd highest price:

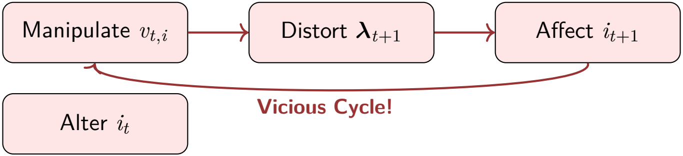
$$p_{t,i_t} := \boldsymbol{\lambda}_t^\mathsf{T} \boldsymbol{c}_{t,i_t} + \max_{j \neq i_t} \tilde{u}_{t,j}.$$

$\Longrightarrow$ `truth` **dominating** when maximizing round-t gain

Epoching to Mitigate "Long-Term" Impact



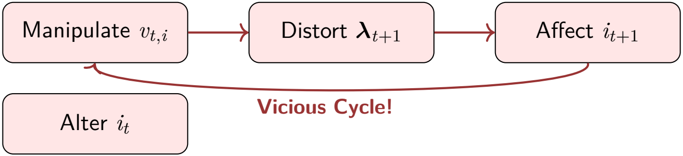
Epoching to Mitigate "Long-Term" Impact



Ingredient 2: Epoch-Based Lazy Updates

Break $[T]$ into *epochs* $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_L$. Fix dual $\boldsymbol{\lambda}_t$ within $t \in \mathcal{E}_\ell$.

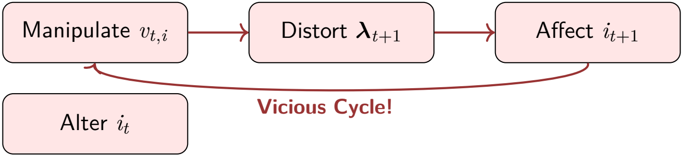
Epoching to Mitigate "Long-Term" Impact



Ingredient 2: Epoch-Based Lazy Updates

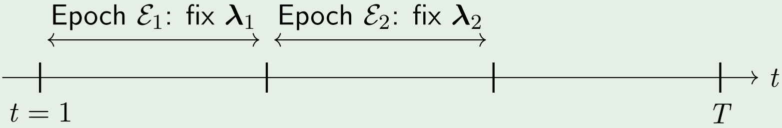
Break $[T]$ into *epochs* $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_L$. Fix dual $\boldsymbol{\lambda}_t$ within $t \in \mathcal{E}_\ell$.

Epoching to Mitigate "Long-Term" Impact

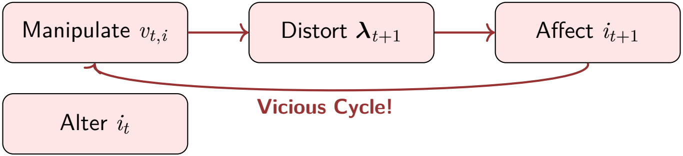


Ingredient 2: Epoch-Based Lazy Updates

Break $[T]$ into *epochs* $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_L$. Fix dual $\boldsymbol{\lambda}_t$ within $t \in \mathcal{E}_\ell$.

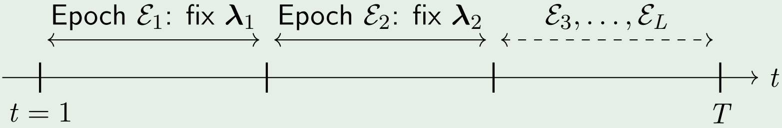


Epoching to Mitigate "Long-Term" Impact

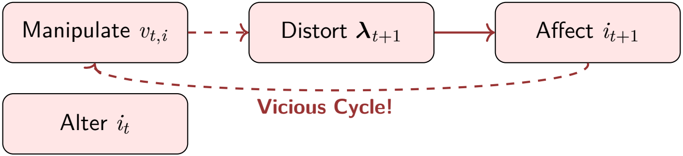


Ingredient 2: Epoch-Based Lazy Updates

Break $[T]$ into *epochs* $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_L$. Fix dual $\boldsymbol{\lambda}_t$ within $t \in \mathcal{E}_\ell$.

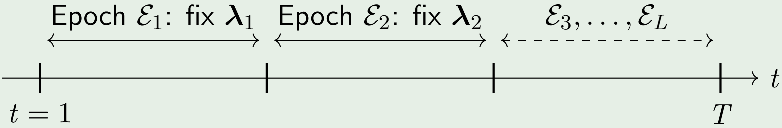


Epoching to Mitigate "Long-Term" Impact



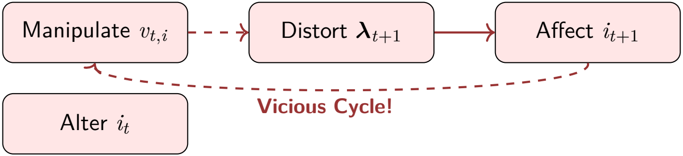
Ingredient 2: Epoch-Based Lazy Updates

Break $[T]$ into *epochs* $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_L$. Fix dual $\boldsymbol{\lambda}_t$ within $t \in \mathcal{E}_\ell$.

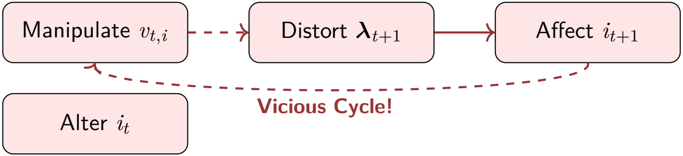


$\Longrightarrow$ **Make it hard** to affect dual updates

Exploration to Remove "Long-Term" Impact



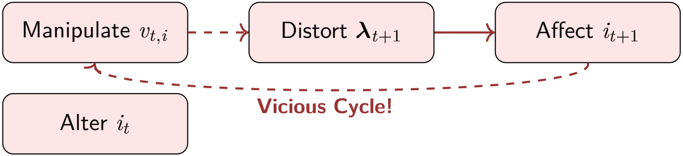
Exploration to Remove "Long-Term" Impact



Ingredient 3: Randomized Exploration Rounds

With low probability ϵ, in each round $t \in \mathcal{E}_\ell$, we do...

Exploration to Remove “Long-Term” Impact

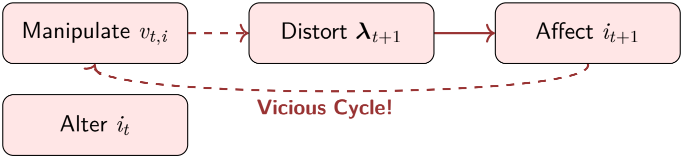


Ingredient 3: Randomized Exploration Rounds

With low probability ϵ, in each round $t \in \mathcal{E}_\ell$, we do...

- Random price $p \sim \text{Unif}[0,1]$ to random agent $i \sim \text{Unif}[K]$.

Exploration to Remove "Long-Term" Impact

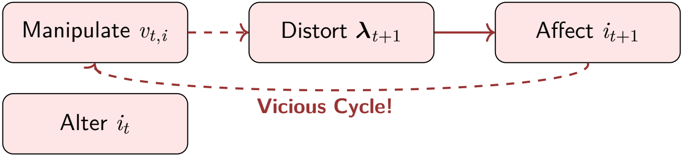


Ingredient 3: Randomized Exploration Rounds

With low probability ϵ, in each round $t \in \mathcal{E}_\ell$, we do...

- Random price $p \sim \text{Unif}[0, 1]$ to random agent $i \sim \text{Unif}[K]$.
- Allocate at price p if $u_{t,i} \geq p$, void the item otherwise:

Exploration to Remove "Long-Term" Impact

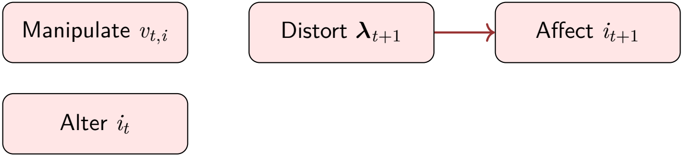


Ingredient 3: Randomized Exploration Rounds

With low probability ϵ, in each round $t \in \mathcal{E}_\ell$, we do...

- Random price $p \sim$ Unif$[0, 1]$ to random agent $i \sim$ Unif$[K]$.
- Allocate at price p if $u_{t,i} \geq p$, void the item otherwise:
  1. If $u_{t,i} > v_{t,i}$: pay higher when $v_{t,i} < p < u_{t,i}$
  2. If $u_{t,i} < v_{t,i}$: lose profit when $u_{t,i} < p < v_{t,i}$

Exploration to Remove "Long-Term" Impact



Ingredient 3: Randomized Exploration Rounds

With low probability ϵ, in each round $t \in \mathcal{E}_\ell$, we do...

- Random price $p \sim \text{Unif}[0,1]$ to random agent $i \sim \text{Unif}[K]$.
- Allocate at price p if $u_{t,i} \ge p$, void the item otherwise:
  1. If $u_{t,i} > v_{t,i}$: pay higher when $v_{t,i} < p < u_{t,i}$
  2. If $u_{t,i} < v_{t,i}$: lose profit when $u_{t,i} < p < v_{t,i}$

$\Longrightarrow$ **Lying causes harm** $\ge \epsilon \frac{(u-v)^2}{2K}$ (outweighing benefits)

Primal Side Summary

1: **for** epoch $\ell = 1, 2, \dots, L$ **do**
2: Receive dual variable $\boldsymbol{\lambda}_\ell$ ▷ *Epoching*
3: **for** round t in epoch $\mathcal{E}_\ell$ **do**

3 Key Ingredients

1. Epoch-based lazy updates

Primal Side Summary

for epoch $\ell = 1, 2, \ldots, L$ **do**
  Receive dual variable $\boldsymbol{\lambda}_\ell$ ▷ *Epoching*
  for round t in epoch $\mathcal{E}_\ell$ **do**
    if w.p. $1 - \frac{1}{|\mathcal{E}_\ell|}$ **then** ▷ *Pricing*
      $i_t \leftarrow \text{argmax}_i (u_{t,i} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i})$, $p_{t,i_t} \leftarrow \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i_t} + \max_{j \neq i_t} (u_{t,j} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,j})$

3 Key Ingredients

1. Epoch-based lazy updates
2. Dual-adjusted pricing (i_t, p_t)

Primal Side Summary

for epoch $\ell = 1, 2, \ldots, L$ **do**
  Receive dual variable $\boldsymbol{\lambda}_\ell$ ▷ *Epoching*
  for round t in epoch $\mathcal{E}_\ell$ **do**
    if w.p. $1 - \frac{1}{|\mathcal{E}_\ell|}$ **then** ▷ *Pricing*
      $i_t \leftarrow \text{argmax}_i(u_{t,i} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i})$, $p_{t,i_t} \leftarrow \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i_t} + \max_{j \neq i_t}(u_{t,j} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,j})$
    else ▷ *Exploration*
      Let $i \sim \text{Unif}[K]$ and $p \sim \text{Unif}[0,1]$: If $u_{t,i} \geq p$, set $i_t \leftarrow i$ & $p_t \leftarrow p$; otherwise, no allocation $i_t \leftarrow 0$.

3 Key Ingredients

1. Epoch-based lazy updates
2. Dual-adjusted pricing (i_t, p_t)
3. Randomized exploration

Primal Side Summary

for epoch $\ell = 1, 2, \ldots, L$ **do**
  Receive dual variable $\boldsymbol{\lambda}_\ell$ ▷ *Epoching*
  for round t in epoch $\mathcal{E}_\ell$ **do**
    if w.p. $1 - \frac{1}{|\mathcal{E}_\ell|}$ **then** ▷ *Pricing*
      $i_t \leftarrow \text{argmax}_i(u_{t,i} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i})$, $p_{t,i_t} \leftarrow \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i_t} + \max_{j \neq i_t}(u_{t,j} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,j})$
    else ▷ *Exploration*
      Let $i \sim \text{Unif}[K]$ and $p \sim \text{Unif}[0,1]$: If $u_{t,i} \geq p$, set $i_t \leftarrow i$ & $p_t \leftarrow p$; otherwise, no allocation $i_t \leftarrow 0$.
    Some unimportant stuff here

3 Key Ingredients

1. Epoch-based lazy updates
2. Dual-adjusted pricing (i_t, p_t)
3. Randomized exploration

Theorem (Informal): This Framework ensures Near-Truthfulness

For each epoch $\mathcal{E}_\ell$, there exists a PBE where agents are **nearly truthful** ($\widetilde{\mathcal{O}}(1)$ misreports per epoch)

Primal Side Summary

for epoch $\ell = 1, 2, \ldots, L$ **do**
  Receive dual variable $\boldsymbol{\lambda}_\ell$ ▷ *Epoching*
  for round t in epoch $\mathcal{E}_\ell$ **do**
    if w.p. $1 - \frac{1}{|\mathcal{E}_\ell|}$ **then** ▷ *Pricing*
      $i_t \leftarrow \text{argmax}_i(u_{t,i} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i})$, $p_{t,i_t} \leftarrow \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i_t} + \max_{j \neq i_t}(u_{t,j} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,j})$
    else ▷ *Exploration*
      Let $i \sim \text{Unif}[K]$ and $p \sim \text{Unif}[0,1]$: If $u_{t,i} \geq p$, set $i_t \leftarrow i$ & $p_t \leftarrow p$; otherwise, no allocation $i_t \leftarrow 0$.
    Some unimportant stuff here

3 Key Ingredients

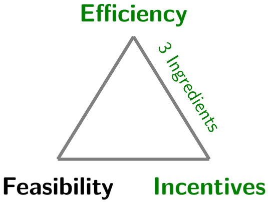
1. Epoch-based lazy updates
2. Dual-adjusted pricing (i_t, p_t)
3. Randomized exploration

Theorem (Informal): This Framework ensures Near-Truthfulness

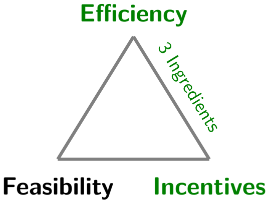
For each epoch $\mathcal{E}_\ell$, there exists a PBE where agents are **nearly truthful** ($\widetilde{\mathcal{O}}(1)$ misreports per epoch) $\Longrightarrow$ **Primal regret** $\widetilde{\mathcal{O}}(L)$

Today

Dual Updates

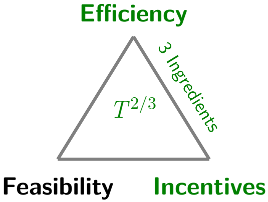


Dual Updates



Tune duals $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots$ **before** knowing costs

Dual Updates: Online Learning gives $\widetilde{\mathcal{O}}(T^{2/3})$

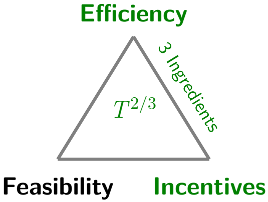


Tune duals $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots$ **before** knowing costs

Theorem 1: Sublinear Regret ✓

3 ingredients (primal) + GD / FTRL (dual)
$\Longrightarrow \widetilde{\mathcal{O}}(T^{2/3})$ **regret** (“no-regret” guarantee)

Dual Updates: Online Learning gives $\widetilde{\mathcal{O}}(T^{2/3})$



Tune duals $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots$ **before** knowing costs

Theorem 1: Sublinear Regret ✓

3 ingredients (primal) + GD / FTRL (dual)
$\Longrightarrow \widetilde{\mathcal{O}}(T^{2/3})$ **regret** ("no-regret" guarantee)

Can we do better than $\Theta(T^{2/3})$?

Online Learning View of Dual Updates

Online Learning Formulation

1. **Setup.** L-round with actions $\boldsymbol{\lambda}_\ell \in \boldsymbol{\Lambda} \subseteq \mathbb{R}^d_{\geq 0}$ (a bounded set)

Online Learning View of Dual Updates

Online Learning Formulation

1. **Setup.** L-round with actions $\boldsymbol{\lambda}_\ell \in \boldsymbol{\Lambda} \subseteq \mathbb{R}^d_{\geq 0}$ (a bounded set)
2. **Loss Function.** After choosing dual $\boldsymbol{\lambda}_\ell$ in epoch ℓ:

$$F_\ell(\boldsymbol{\lambda}) := \sum_{t \in \mathcal{E}_\ell} (\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\mathsf{T} \boldsymbol{\lambda}, \quad \forall \boldsymbol{\lambda} \in \boldsymbol{\Lambda},$$

where i_t is the *allocation based on $\boldsymbol{\lambda}_\ell$* $\approx \operatorname{argmax}_i (v_{t,i} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i})$

Online Learning View of Dual Updates

Online Learning Formulation

1. **Setup.** L-round with actions $\boldsymbol{\lambda}_\ell \in \boldsymbol{\Lambda} \subseteq \mathbb{R}^d_{\geq 0}$ (a bounded set)
2. **Loss Function.** After choosing dual $\boldsymbol{\lambda}_\ell$ in epoch ℓ:

$$F_\ell(\boldsymbol{\lambda}) := \sum_{t \in \mathcal{E}_\ell} (\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\mathsf{T} \boldsymbol{\lambda}, \quad \forall \boldsymbol{\lambda} \in \boldsymbol{\Lambda},$$

   where i_t is the *allocation based on $\boldsymbol{\lambda}_\ell$* $\approx \text{argmax}_i (v_{t,i} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i})$
3. **Objective.** Minimize online learning regret (derivation omitted)

$$\mathfrak{R}^{\boldsymbol{\lambda}} := \mathbb{E}\left[\max_{\boldsymbol{\lambda}^* \in \boldsymbol{\Lambda}} \sum_{\ell=1}^{L} (F_\ell(\boldsymbol{\lambda}_\ell) - F_\ell(\boldsymbol{\lambda}^*))\right]$$

Price of Incentives: An $\Omega(T^{2/3})$ Barrier

Dilemma of Lazy Updates

- Lazy updates **good for primal** (less incentives & ablities)

Price of Incentives: An $\Omega(T^{2/3})$ Barrier

Dilemma of Lazy Updates

- Lazy updates **good for primal** (less incentives & ablities)
- Lazy updates **bad for dual**

Price of Incentives: An $\Omega(T^{2/3})$ Barrier

Dilemma of Lazy Updates

- Lazy updates **good for primal** (less incentives & ablities)
- Lazy updates **bad for dual**

No Lazy $\quad \boldsymbol{c}_{t-1,i_{t-1}} \rightarrow \boldsymbol{\lambda}_t \rightarrow \boldsymbol{c}_{t,i_t} \rightarrow \boldsymbol{\lambda}_{t+1}$

Price of Incentives: An $\Omega(T^{2/3})$ Barrier

Dilemma of Lazy Updates

- Lazy updates **good for primal** (less incentives & ablities)
- Lazy updates **bad for dual**

Start of Epoch ℓ

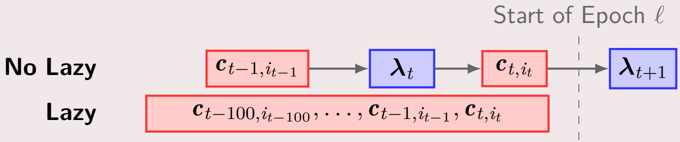
No Lazy $\quad \boldsymbol{c}_{t-1,i_{t-1}} \rightarrow \boldsymbol{\lambda}_t \rightarrow \boldsymbol{c}_{t,i_t} \rightarrow \boldsymbol{\lambda}_{t+1}$

Lazy

Price of Incentives: An $\Omega(T^{2/3})$ Barrier

Dilemma of Lazy Updates

- Lazy updates **good for primal** (less incentives & ablities)
- Lazy updates **bad for dual**



Price of Incentives: An $\Omega(T^{2/3})$ Barrier

Dilemma of Lazy Updates

- Lazy updates **good for primal** (less incentives & ablities)
- Lazy updates **bad for dual**

Start of Epoch ℓ

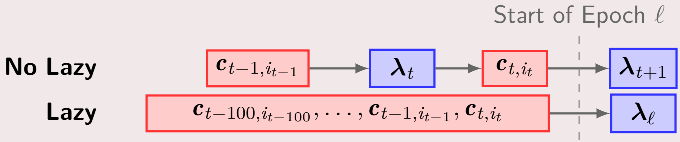
No Lazy $\boxed{\boldsymbol{c}_{t-1,i_{t-1}}} \rightarrow \boxed{\boldsymbol{\lambda}_t} \rightarrow \boxed{\boldsymbol{c}_{t,i_t}} \rightarrow \boxed{\boldsymbol{\lambda}_{t+1}}$

Lazy $\boxed{\boldsymbol{c}_{t-100,i_{t-100}}, \ldots, \boldsymbol{c}_{t-1,i_{t-1}}, \boldsymbol{c}_{t,i_t}} \rightarrow \boxed{\boldsymbol{\lambda}_\ell}$

Price of Incentives: An $\Omega(T^{2/3})$ Barrier

Dilemma of Lazy Updates

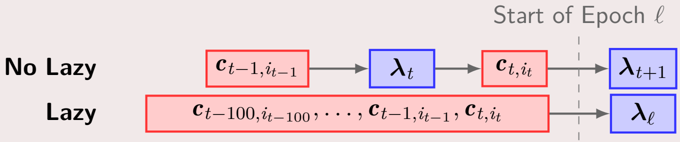
- Lazy updates **good for primal** (less incentives & ablities)
- Lazy updates **bad for dual** (cannot react to $\boldsymbol{c}_{t,i_t}$ promptly)



Price of Incentives: An $\Omega(T^{2/3})$ Barrier

Dilemma of Lazy Updates

- Lazy updates **good for primal** (less incentives & ablities)
- Lazy updates **bad for dual** (cannot react to $\boldsymbol{c}_{t,i_t}$ promptly)

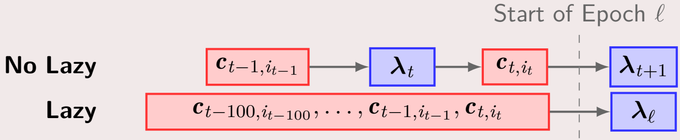


Theorem. "Low-switching online learning" has $\Omega(T/\sqrt{L})$ **regret**

Price of Incentives: An $\Omega(T^{2/3})$ Barrier

Dilemma of Lazy Updates

- Lazy updates **good for primal** (less incentives & ablities)
- Lazy updates **bad for dual** (cannot react to $\boldsymbol{c}_{t,i_t}$ promptly)



Theorem. "Low-switching online learning" has $\Omega(T/\sqrt{L})$ **regret**

Hardness (?) $\Omega(L)$ primal $+$ $\Omega(T/\sqrt{L})$ dual $=$ $\Omega(T^{2/3})$ **overall**

Our Plan: Promise of Predictability

Observation. $\Omega(T/\sqrt{L})$ is for fully adversarial loss functions

Our Plan: Promise of Predictability

Observation. $\Omega(T/\sqrt{L})$ is for fully adversarial loss functions

Key Insight: (Near-)Truthful Reports $\Longrightarrow$ (Near-)Predictability

Our Plan: Promise of Predictability

Observation. $\Omega(T/\sqrt{L})$ is for fully adversarial loss functions

Key Insight: (Near-)Truthful Reports $\Longrightarrow$ (Near-)Predictability

1. **Truthful $\Longrightarrow$ i.i.d. future allocations** in current epoch $\mathcal{E}_\ell$:

$$i_t = \operatorname*{argmax}_{i \in [K]} (v_{t,i} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i}), \quad \text{where } v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i, \boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$$

Our Plan: Promise of Predictability

Observation. $\Omega(T/\sqrt{L})$ is for fully adversarial loss functions

Key Insight: (Near-)Truthful Reports $\Longrightarrow$ (Near-)Predictability

1. **Truthful $\Longrightarrow$ i.i.d. future allocations** in current epoch $\mathcal{E}_\ell$:

$$i_t = \operatorname*{argmax}_{i\in[K]}(v_{t,i} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i}), \quad \text{where } v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i, \boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$$

2. **Truthful $\Longrightarrow$ reliable historical reports** to make predictions:

$$\mathcal{V}_i, \mathcal{C}_i \text{ can be restored from previous } \{v_{\tau,i}, \boldsymbol{c}_{\tau,i}\}_{\tau<t}$$

Our Plan: Promise of Predictability

Observation. $\Omega(T/\sqrt{L})$ is for fully adversarial loss functions

Key Insight: (Near-)Truthful Reports $\Longrightarrow$ (Near-)Predictability

1. **Truthful $\Longrightarrow$ i.i.d. future allocations** in current epoch $\mathcal{E}_\ell$:

$$i_t = \operatorname*{argmax}_{i \in [K]} (v_{t,i} - \boldsymbol{\lambda}_\ell^\mathsf{T} \boldsymbol{c}_{t,i}), \quad \text{where } v_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{V}_i, \boldsymbol{c}_{t,i} \overset{\text{i.i.d.}}{\sim} \mathcal{C}_i$$

2. **Truthful $\Longrightarrow$ reliable historical reports** to make predictions:

$$\mathcal{V}_i, \mathcal{C}_i \text{ can be restored from previous } \{v_{\tau,i}, \boldsymbol{c}_{\tau,i}\}_{\tau < t}$$

Truthfulness not only boosts efficiency, but also helps dual!

Full View: Primal-Dual in Strategic Environments

Primal Side

Dual Side

Full View: Primal-Dual in Strategic Environments

Primal Side

Incentive-aware $\{(i_t, p_{t,i_t})\}_{t \in \mathcal{E}_\ell}$ for a fixed dual price $\boldsymbol{\lambda}_\ell$

$\boldsymbol{\lambda}_\ell \in \mathbb{R}^d_{\geq 0}$

Dual Side

Allocation & Pricing

Given $\boldsymbol{\lambda}_\ell \in \mathbb{R}^d_{\geq 0}$ from dual:

- Allocate to an i_t maximizing $\tilde{v}_{t,i} \Longrightarrow$ **Efficiency**
- Incentivize truthfulness via $p_{t,i_t} \Longrightarrow$ **Incentives**

Full View: Primal-Dual in Strategic Environments

Primal Side
Incentive-aware $\{(i_t, p_{t,i_t})\}_{t\in\mathcal{E}_\ell}$
for a fixed dual price $\boldsymbol{\lambda}_\ell$

$\boldsymbol{\lambda}_\ell \in \mathbb{R}^d_{\geq 0}$

$i_t \in [K]$

Dual Side
Learn next $\boldsymbol{\lambda}_{\ell+1}$
utilizing predictability

Allocation & Pricing

Given $\boldsymbol{\lambda}_\ell \in \mathbb{R}^d_{\geq 0}$ from dual:

- Allocate to an i_t maximizing $\tilde{v}_{t,i} \Longrightarrow$ **Efficiency**
- Incentivize truthfulness via $p_{t,i_t} \Longrightarrow$ **Incentives**

Learning via Predictability

Given $\{i_t\}_{t\in\mathcal{E}_\ell}$ from primal:

- Learn a better $\boldsymbol{\lambda}_{\ell+1}$ for $\mathcal{E}_{\ell+1}$ $\Longrightarrow$ **Feasibility**
- Bypass $\Omega(T^{2/3})$ barrier via predictability $\Longrightarrow$ **Efficiency**

Exploiting Predictability with Optimism: O-FTRL

Optimistic FTRL Framework (O-FTRL)

O-FTRL considers *predicted loss* $\hat{F}_\ell \colon \boldsymbol{\Lambda} \to \mathbb{R}$ before deciding $\boldsymbol{\lambda}_\ell$
(where loss function $F_\ell(\boldsymbol{\lambda}) := \sum_{t \in \mathcal{E}_\ell} f_t(\boldsymbol{\lambda}) = \sum_{t \in \mathcal{E}_\ell} (\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\mathsf{T} \boldsymbol{\lambda}$):

Exploiting Predictability with Optimism: O-FTRL

Optimistic FTRL Framework (O-FTRL)

O-FTRL considers *predicted loss* $\hat{F}_\ell \colon \mathbf{\Lambda} \to \mathbb{R}$ before deciding $\boldsymbol{\lambda}_\ell$ (where loss function $F_\ell(\boldsymbol{\lambda}) := \sum_{t\in\mathcal{E}_\ell} f_t(\boldsymbol{\lambda}) = \sum_{t\in\mathcal{E}_\ell} (\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\mathsf{T}\boldsymbol{\lambda}$):

$$\boldsymbol{\lambda}_\ell = \operatorname*{argmin}_{\boldsymbol{\lambda}\in\Lambda} \left(\sum_{\ell'<\ell} F_{\ell'}(\boldsymbol{\lambda}) + \hat{F}_\ell(\boldsymbol{\lambda}) + \frac{1}{\eta_\ell}\Psi(\boldsymbol{\lambda})\right), \quad \forall \ell \in [L].$$

Exploiting Predictability with Optimism: O-FTRL

Start of Epoch ℓ

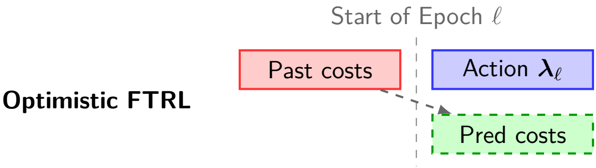
Lazy FTRL | Past costs → Action $\boldsymbol{\lambda}_\ell$ | $\Omega(T/\sqrt{L})$ **regret**

Optimistic FTRL Framework (O-FTRL)

O-FTRL considers *predicted loss* $\hat{F}_\ell \colon \boldsymbol{\Lambda} \to \mathbb{R}$ before deciding $\boldsymbol{\lambda}_\ell$ (where loss function $F_\ell(\boldsymbol{\lambda}) := \sum_{t\in\mathcal{E}_\ell} f_t(\boldsymbol{\lambda}) = \sum_{t\in\mathcal{E}_\ell}(\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\mathsf{T}\boldsymbol{\lambda}$):

$$\boldsymbol{\lambda}_\ell = \operatorname*{argmin}_{\boldsymbol{\lambda}\in\Lambda}\left(\sum_{\ell'<\ell} F_{\ell'}(\boldsymbol{\lambda}) + \hat{F}_\ell(\boldsymbol{\lambda}) + \frac{1}{\eta_\ell}\Psi(\boldsymbol{\lambda})\right), \quad \forall \ell \in [L].$$

Exploiting Predictability with Optimism: O-FTRL



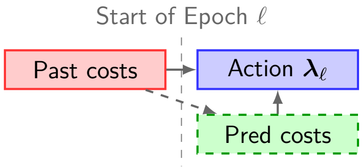
Optimistic FTRL Framework (O-FTRL)

O-FTRL considers *predicted loss* $\hat{F}_\ell \colon \boldsymbol{\Lambda} \to \mathbb{R}$ before deciding $\boldsymbol{\lambda}_\ell$ (where loss function $F_\ell(\boldsymbol{\lambda}) := \sum_{t \in \mathcal{E}_\ell} f_t(\boldsymbol{\lambda}) = \sum_{t \in \mathcal{E}_\ell} (\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\top \boldsymbol{\lambda}$):

$$\boldsymbol{\lambda}_\ell = \operatorname*{argmin}_{\boldsymbol{\lambda} \in \Lambda} \left(\sum_{\ell' < \ell} F_{\ell'}(\boldsymbol{\lambda}) + \hat{F}_\ell(\boldsymbol{\lambda}) + \frac{1}{\eta_\ell} \Psi(\boldsymbol{\lambda}) \right), \quad \forall \ell \in [L].$$

Exploiting Predictability with Optimism: O-FTRL

Optimistic FTRL



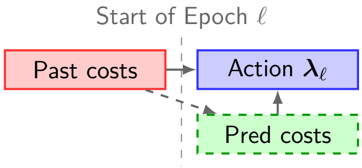
Optimistic FTRL Framework (O-FTRL)

O-FTRL considers *predicted loss* $\hat{F}_\ell \colon \boldsymbol{\Lambda} \to \mathbb{R}$ before deciding $\boldsymbol{\lambda}_\ell$ (where loss function $F_\ell(\boldsymbol{\lambda}) := \sum_{t \in \mathcal{E}_\ell} f_t(\boldsymbol{\lambda}) = \sum_{t \in \mathcal{E}_\ell} (\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\mathsf{T} \boldsymbol{\lambda}$):

$$\boldsymbol{\lambda}_\ell = \operatorname*{argmin}_{\boldsymbol{\lambda} \in \Lambda} \left(\sum_{\ell' < \ell} F_{\ell'}(\boldsymbol{\lambda}) + \hat{F}_\ell(\boldsymbol{\lambda}) + \frac{1}{\eta_\ell} \Psi(\boldsymbol{\lambda}) \right), \quad \forall \ell \in [L].$$

Exploiting Predictability with Optimism: O-FTRL

Optimistic FTRL



$\widetilde{\mathcal{O}}(T/L)$ **regret**

Optimistic FTRL Framework (O-FTRL)

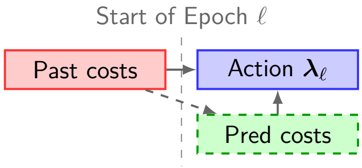
O-FTRL considers *predicted loss* $\hat{F}_\ell \colon \boldsymbol{\Lambda} \to \mathbb{R}$ before deciding $\boldsymbol{\lambda}_\ell$
(where loss function $F_\ell(\boldsymbol{\lambda}) := \sum_{t\in\mathcal{E}_\ell} f_t(\boldsymbol{\lambda}) = \sum_{t\in\mathcal{E}_\ell} (\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\mathsf{T}\boldsymbol{\lambda}$):

$$\boldsymbol{\lambda}_\ell = \operatorname*{argmin}_{\boldsymbol{\lambda}\in\Lambda} \left(\sum_{\ell'<\ell} F_{\ell'}(\boldsymbol{\lambda}) + \hat{F}_\ell(\boldsymbol{\lambda}) + \frac{1}{\eta_\ell}\Psi(\boldsymbol{\lambda}) \right), \quad \forall \ell \in [L].$$

If $\hat{F}_\ell$'s are accurate, regret improves from $\widetilde{\mathcal{O}}(T/\sqrt{L})$ to $\widetilde{\mathcal{O}}(T/L)$.

Technical Challenge: Circular Dependency

Optimistic FTRL

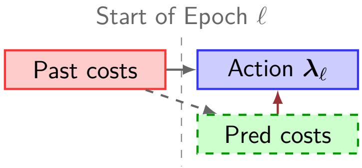


$\widetilde{\mathcal{O}}(T/L)$ **regret?**

Technical Challenge: Circular Dependency between $\boldsymbol{\lambda}_\ell$ and $\hat{F}_\ell$

Technical Challenge: Circular Dependency

Optimistic FTRL



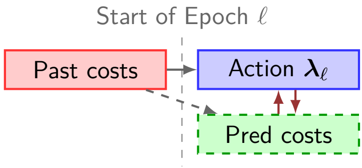
$\widetilde{\mathcal{O}}(T/L)$ **regret?**

Technical Challenge: Circular Dependency between $\boldsymbol{\lambda}_\ell$ and $\hat{F}_\ell$

- Yield $\boldsymbol{\lambda}_\ell$ **as-if true costs = pred costs**

Technical Challenge: Circular Dependency

Optimistic FTRL



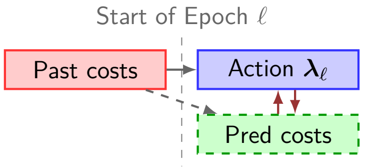
$\widetilde{\mathcal{O}}(T/L)$ **regret?**

Technical Challenge: Circular Dependency between $\boldsymbol{\lambda}_\ell$ and $\hat{F}_\ell$

- Yield $\boldsymbol{\lambda}_\ell$ **as-if true costs = pred costs**
- Yield pred costs **as-if true dual = $\boldsymbol{\lambda}_\ell$**

Technical Challenge: Circular Dependency

Optimistic FTRL



$\widetilde{\mathcal{O}}(T/L)$ **regret?**

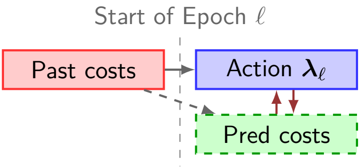
Technical Challenge: Circular Dependency between $\boldsymbol{\lambda}_\ell$ and $\hat{F}_\ell$

- Yield $\boldsymbol{\lambda}_\ell$ **as-if true costs = pred costs**
- Yield pred costs **as-if true dual = $\boldsymbol{\lambda}_\ell$**

Action $\boldsymbol{\lambda}_\ell$

Technical Challenge: Circular Dependency

Optimistic FTRL



$\widetilde{\mathcal{O}}(T/L)$ **regret?**

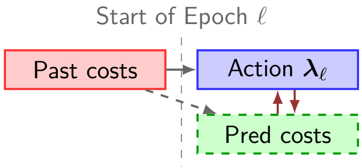
Technical Challenge: Circular Dependency between $\boldsymbol{\lambda}_\ell$ and $\hat{F}_\ell$

- Yield $\boldsymbol{\lambda}_\ell$ **as-if true costs = pred costs**
- Yield pred costs **as-if true dual = $\boldsymbol{\lambda}_\ell$**

Action $\boldsymbol{\lambda}_\ell \Longrightarrow$ Allocations $\{i_t\}_{t \in \mathcal{E}_\ell}$ $(i_t \approx \operatorname{argmax}_i (v_{t,i} - \boldsymbol{\lambda}_\ell^\top \boldsymbol{c}_{t,i}))$

Technical Challenge: Circular Dependency

Optimistic FTRL



$\widetilde{\mathcal{O}}(T/L)$ **regret?**

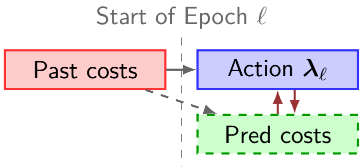
Technical Challenge: Circular Dependency between $\boldsymbol{\lambda}_\ell$ and $\hat{F}_\ell$

- Yield $\boldsymbol{\lambda}_\ell$ **as-if true costs = pred costs**
- Yield pred costs **as-if true dual = $\boldsymbol{\lambda}_\ell$**

Action $\boldsymbol{\lambda}_\ell \Longrightarrow$ Allocations $\{i_t\}_{t\in\mathcal{E}_\ell}$ $(i_t \approx \text{argmax}_i(v_{t,i} - \boldsymbol{\lambda}_\ell^\top \boldsymbol{c}_{t,i})) \Longrightarrow$ Loss $F_\ell(\boldsymbol{\lambda})$ $(F_\ell(\boldsymbol{\lambda}) = \sum_{t\in\mathcal{E}_\ell}(\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\top \boldsymbol{\lambda})$

Technical Challenge: Circular Dependency

Optimistic FTRL



$\widetilde{\mathcal{O}}(T/L)$ **regret?**

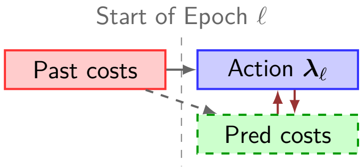
Technical Challenge: Circular Dependency between $\boldsymbol{\lambda}_\ell$ and $\hat{F}_\ell$

- Yield $\boldsymbol{\lambda}_\ell$ **as-if true costs = pred costs**
- Yield pred costs **as-if true dual = $\boldsymbol{\lambda}_\ell$**

Action $\boldsymbol{\lambda}_\ell \Longrightarrow$ Allocations $\{i_t\}_{t\in\mathcal{E}_\ell}$ $(i_t \approx \operatorname{argmax}_i(v_{t,i} - \boldsymbol{\lambda}_\ell^\top \boldsymbol{c}_{t,i})) \Longrightarrow$ Loss $F_\ell(\boldsymbol{\lambda})$ $(F_\ell(\boldsymbol{\lambda}) = \sum_{t\in\mathcal{E}_\ell}(\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\top \boldsymbol{\lambda}) \Longrightarrow$ Prediction $\hat{F}_\ell(\boldsymbol{\lambda})$ (using historical data to estimate Law(i_t))

Technical Challenge: Circular Dependency

Optimistic FTRL



$\widetilde{\mathcal{O}}(T/L)$ **regret?**

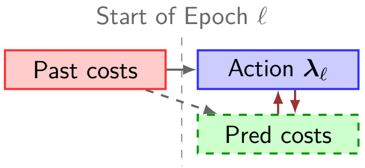
Technical Challenge: Circular Dependency between $\boldsymbol{\lambda}_\ell$ and $\hat{F}_\ell$

- Yield $\boldsymbol{\lambda}_\ell$ **as-if true costs = pred costs**
- Yield pred costs **as-if true dual = $\boldsymbol{\lambda}_\ell$**

Action $\boldsymbol{\lambda}_\ell \Longrightarrow$ Allocations $\{i_t\}_{t\in\mathcal{E}_\ell}$ $(i_t \approx \operatorname{argmax}_i(v_{t,i} - \boldsymbol{\lambda}_\ell^\top \boldsymbol{c}_{t,i})) \Longrightarrow$
Loss $F_\ell(\boldsymbol{\lambda})$ $(F_\ell(\boldsymbol{\lambda}) = \sum_{t\in\mathcal{E}_\ell}(\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\top \boldsymbol{\lambda}) \Longrightarrow$ Prediction $\hat{F}_\ell(\boldsymbol{\lambda})$
(using historical data to estimate Law(i_t)) $\Longrightarrow$ Action $\boldsymbol{\lambda}_\ell$ (O-FTRL)

Technical Challenge: Circular Dependency

Optimistic FTRL



Circular issue!

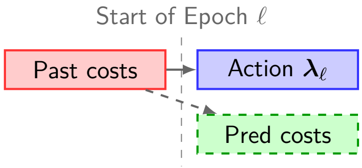
Technical Challenge: Circular Dependency between $\boldsymbol{\lambda}_\ell$ and $\hat{F}_\ell$

- Yield $\boldsymbol{\lambda}_\ell$ **as-if true costs = pred costs**
- Yield pred costs **as-if true dual = $\boldsymbol{\lambda}_\ell$**

Action $\boldsymbol{\lambda}_\ell \Longrightarrow$ Allocations $\{i_t\}_{t \in \mathcal{E}_\ell}$ ($i_t \approx \operatorname{argmax}_i(v_{t,i} - \boldsymbol{\lambda}_\ell^\top \boldsymbol{c}_{t,i})$) $\Longrightarrow$ Loss $F_\ell(\boldsymbol{\lambda})$ ($F_\ell(\boldsymbol{\lambda}) = \sum_{t \in \mathcal{E}_\ell}(\boldsymbol{\rho} - \boldsymbol{c}_{t,i_t})^\top \boldsymbol{\lambda}$) $\Longrightarrow$ Prediction $\hat{F}_\ell(\boldsymbol{\lambda})$ (using historical data to estimate Law(i_t)) $\Longrightarrow$ Action $\boldsymbol{\lambda}_\ell$ (O-FTRL)

Our Solution: O-FTRL with Fixed Points (O-FTRL-FP)

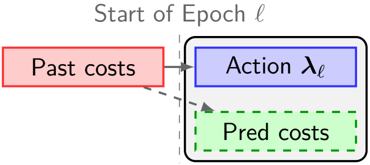
O-FTRL-FP



Reformulating Update as Fixed-Point Problem

Our Solution: O-FTRL with Fixed Points (O-FTRL-FP)

O-FTRL-FP

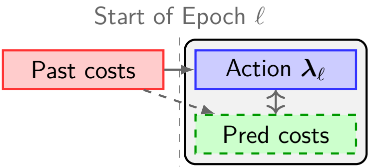


Reformulating Update as Fixed-Point Problem

Let $\hat{F}_\ell(\boldsymbol{\lambda}; \boldsymbol{\lambda}_\ell)$ = predicted loss function $F_\ell(\boldsymbol{\lambda})$ **when action is $\boldsymbol{\lambda}_\ell$**

Our Solution: O-FTRL with Fixed Points (O-FTRL-FP)

O-FTRL-FP



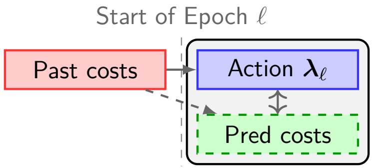
Reformulating Update as Fixed-Point Problem

Let $\hat{F}_\ell(\boldsymbol{\lambda}; \boldsymbol{\lambda}_\ell)$ = predicted loss function $F_\ell(\boldsymbol{\lambda})$ **when action is $\boldsymbol{\lambda}_\ell$**
Decide action $\boldsymbol{\lambda}_\ell$ as the fixed point solution to:

$$\boldsymbol{\lambda}_\ell = \operatorname*{argmin}_{\boldsymbol{\lambda} \in \Lambda} \left(\sum_{s<\ell} F_s(\boldsymbol{\lambda}) + \hat{F}_\ell(\boldsymbol{\lambda}; \boldsymbol{\lambda}_\ell) + \frac{1}{\eta_\ell} \Psi(\boldsymbol{\lambda}) \right). \quad (*)$$

Our Solution: O-FTRL with Fixed Points (O-FTRL-FP)

O-FTRL-FP



$\widetilde{\mathcal{O}}(T/L)$ **regret**

Reformulating Update as Fixed-Point Problem

Let $\hat{F}_\ell(\boldsymbol{\lambda}; \boldsymbol{\lambda}_\ell) =$ predicted loss function $F_\ell(\boldsymbol{\lambda})$ **when action is $\boldsymbol{\lambda}_\ell$**
Decide action $\boldsymbol{\lambda}_\ell$ as the fixed point solution to:

$$\boldsymbol{\lambda}_\ell = \operatorname*{argmin}_{\boldsymbol{\lambda} \in \Lambda} \left(\sum_{s<\ell} F_s(\boldsymbol{\lambda}) + \hat{F}_\ell(\boldsymbol{\lambda}; \boldsymbol{\lambda}_\ell) + \frac{1}{\eta_\ell} \Psi(\boldsymbol{\lambda}) \right). \qquad (*)$$

Lemma. In our case, Eq. (*) has **approximate fixed points**

Dual Side Summary

Goal. Deciding duals despite only L switches & strategic agents.

Challenge: Lazy Updates $\Longrightarrow \Omega(T^{2/3})$ Barrier

$\Longrightarrow$ Primal framework + FTRL can only get $\widetilde{\mathcal{O}}(T^{2/3})$ **regret**

Dual Side Summary

Goal. Deciding duals despite only L switches & strategic agents.

Challenge: Lazy Updates $\Longrightarrow \Omega(T^{2/3})$ Barrier

$\Longrightarrow$ Primal framework + FTRL can only get $\widetilde{\mathcal{O}}(T^{2/3})$ **regret**

Solution: Near-Truthfulness $\Longrightarrow$ Predictability

- Near-*i.i.d.* future + reliable historical data = **Predictability**

Dual Side Summary

Goal. Deciding duals despite only L switches & strategic agents.

Challenge: Lazy Updates $\Longrightarrow \Omega(T^{2/3})$ Barrier

$\Longrightarrow$ Primal framework + FTRL can only get $\widetilde{\mathcal{O}}(T^{2/3})$ **regret**

Solution: Near-Truthfulness $\Longrightarrow$ Predictability

- Near-*i.i.d.* future + reliable historical data = **Predictability**
- Fixed-point oracles to resolve circular dependencies between action $\boldsymbol{\lambda}_\ell$ & prediction $\hat{F}_\ell(\boldsymbol{\lambda}; \boldsymbol{\lambda}_\ell)$ = **O-FTRL-FP framework**

Dual Side Summary

Goal. Deciding duals despite only L switches & strategic agents.

Challenge: Lazy Updates $\Longrightarrow \Omega(T^{2/3})$ Barrier

$\Longrightarrow$ Primal framework + FTRL can only get $\widetilde{\mathcal{O}}(T^{2/3})$ **regret**

Solution: Near-Truthfulness $\Longrightarrow$ Predictability

- Near-*i.i.d.* future + reliable historical data = **Predictability**
- Fixed-point oracles to resolve circular dependencies between action $\boldsymbol{\lambda}_\ell$ & prediction $\hat{F}_\ell(\boldsymbol{\lambda}; \boldsymbol{\lambda}_\ell)$ = **O-FTRL-FP framework**

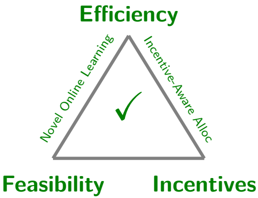
$\Longrightarrow$ Primal framework + our O-FTRL-FP = $\widetilde{\mathcal{O}}(\sqrt{T})$ **regret**

Main Results & Takeaway

Main Result

1st dynamic mechanism resolving **trilemma**:

- **Efficiency.** Optimal $\widetilde{\mathcal{O}}(\sqrt{T})$ regret
- **Feasibility.** Zero constraint violation
- **Incentives.** Robust to strategic agents



Key Techniques

- **Primal Side: Incentive-Aware Allocation.** Novel mixture of dual-adjusted pricing + epoching + random exploration
- **Dual Side: Online Learning for Updates.** Truthfulness ⇒ Predictability + novel O-FTRL-FP for circular dependencies

Questions are more than welcomed!

✉ yandai20@mit.edu; 🌐 https://yandaichn.github.io/