Learning to Schedule in Multiclass Many Server Queues

Yueyang Zhong
The University of Chicago Booth School of Business

John Birge
The University of Chicago Booth School of Business

Amy Ward
The University of Chicago Booth School of Business

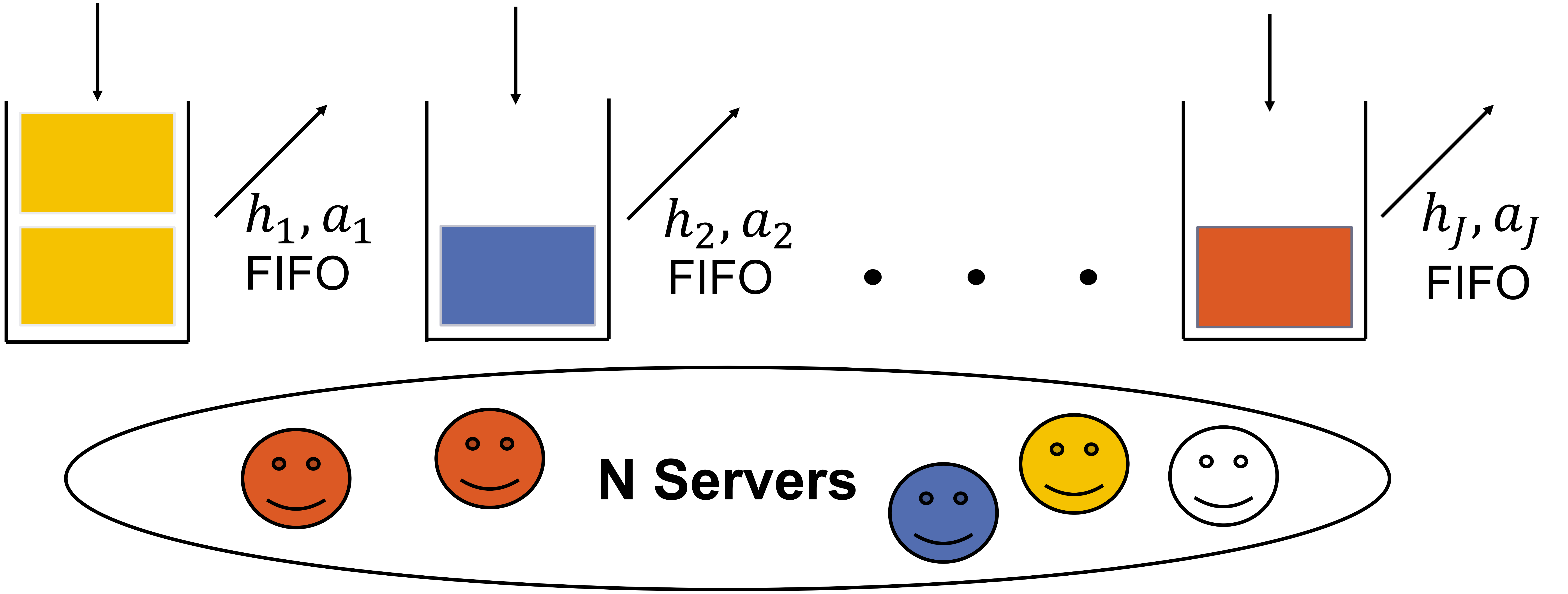
July 21-22, 2022

CHICAGO BOOTH
The University of Chicago Booth School of Business

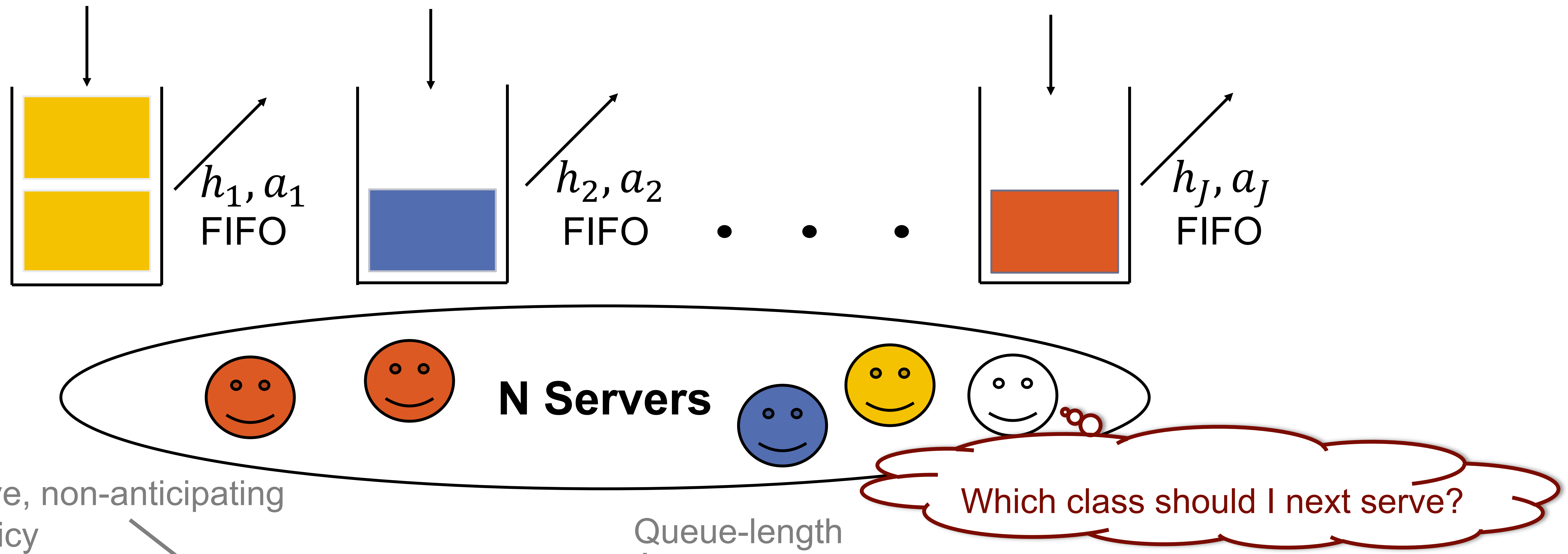
A Canonical Model For Service Systems

Service systems with different classes of customers who have limited patience:

- Multiclass many server queue with abandonment



A Canonical Model For Service Systems



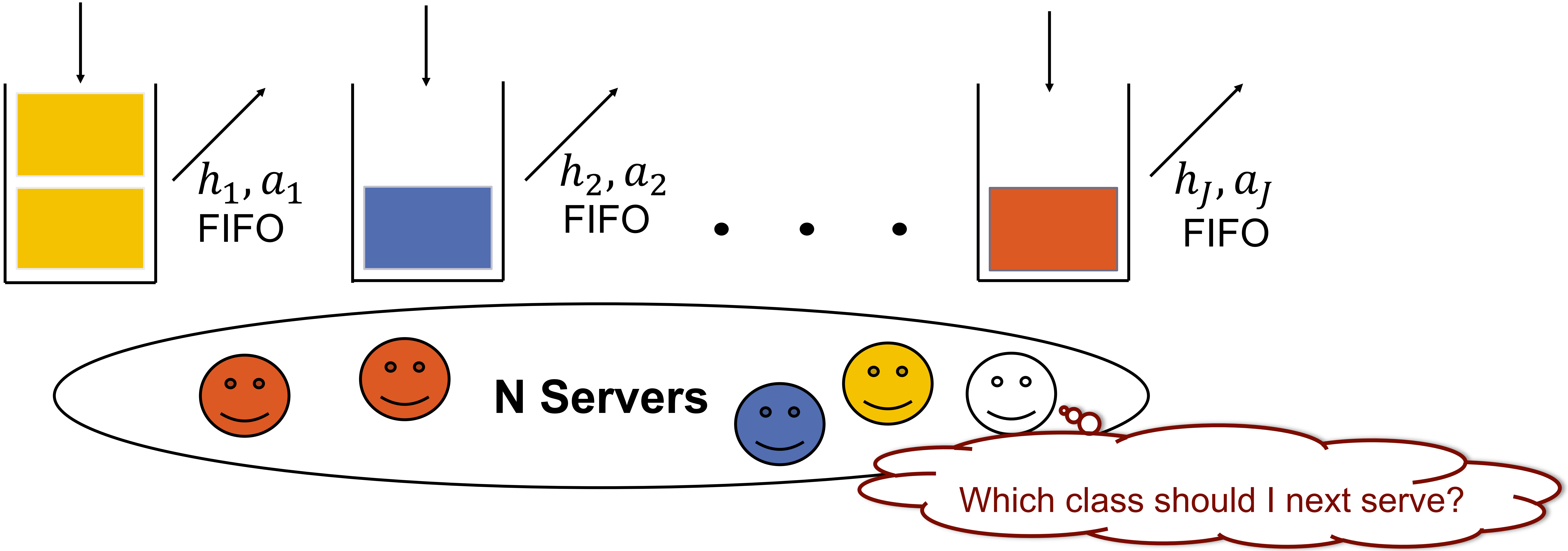
Non-preemptive, non-anticipating scheduling policy

Queue-length

Finite Horizon Cost: $C_T(\pi) \coloneqq \sum_{j=1}^{J} a_j R_j(T,\pi) + \int_0^T h_j Q_j(t,\pi)dt$

Cumulative reneging

A Canonical Model For Service Systems

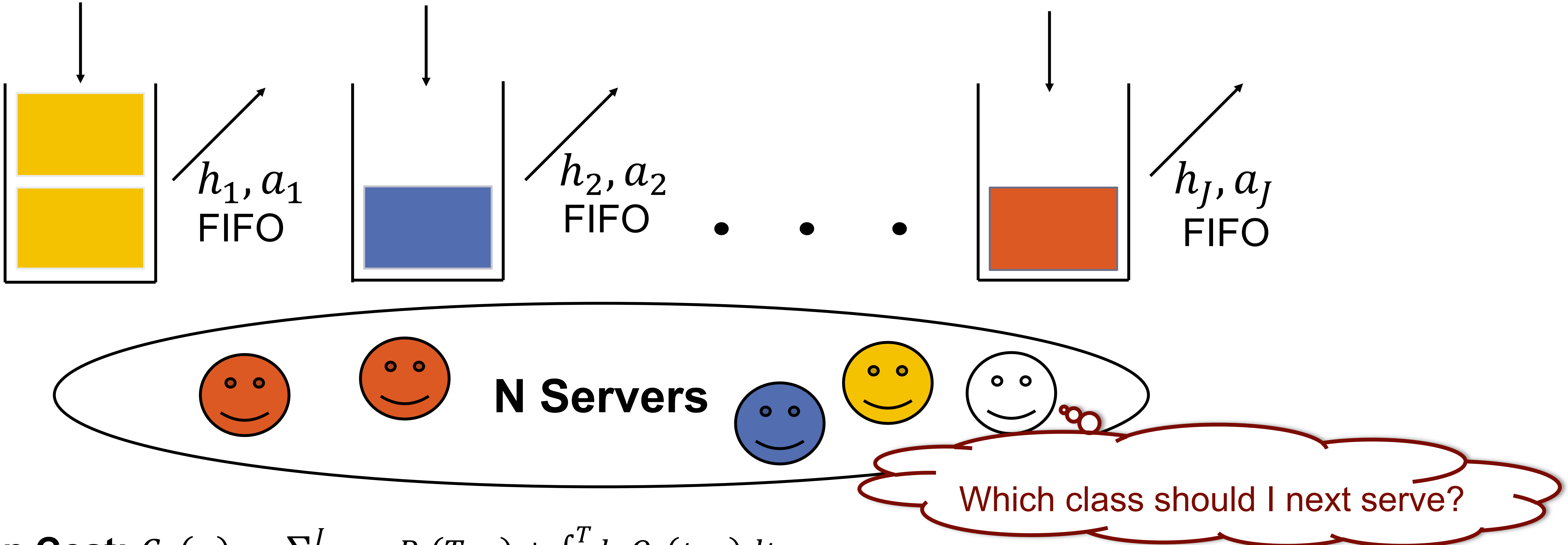


Finite Horizon Cost: $C_T(\pi) \coloneqq \sum_{j=1}^{J} a_j R_j(T,\pi) + \int_0^T h_j Q_j(t,\pi)dt$

Long-Run Average Cost: $C(\pi) \coloneqq \limsup_{T\to\infty} \mathbb{E}[C_T(\pi)]/T$

Known: a_j, h_j
Unknown: $\lambda_j(t)$, μ_j, θ_j

A Canonical Model For Service Systems



Finite Horizon Cost: $C_T(\pi) := \sum_{j=1}^{J} a_j R_j(T, \pi) + \int_0^T h_j Q_j(t, \pi) dt$

Long-Run Average Cost: $C(\pi) := \limsup_{T\to\infty} \mathbb{E}[C_T(\pi)]/T$

an optimal policy

Objective is to Minimize Regret, $\mathcal{R}(T, \pi) := \mathbb{E}[C_T(\pi)] - \mathbb{E}[C_T(\pi^\star)]$

CHICAGO BOOTH
The University of Chicago Booth School of Business

Selected Literature Review

Combining statistical learning and optimal control

- Inventory control

  [Kunnumkal and Topaloglu, 2008]

  [Huh and Rusmevichientong, 2014]
- Assortment optimization

  [Saure and Zeevi, 2013]
- Revenue management

  [Besbes and Zeevi, 2012]

  [den Boer and Zwart, 2015]

Not adapted to queueing settings

Learning the scheduling in queueing systems

[Krishnasamy et al., 2018]

[Krishnasamy et al., 2021]

[Stahlbuhk et al., 2021]

[Choudhury et al., 2021]

[Lee and Vojnovic, 2021]

Discrete time,
No abandonment

MAB

[Gittins, 1979]

[Lai and Robbins, 1985]

Parameter estimation in queues

[Walton and Xu, 2021] (survey)

[Asanjarina et al., 2021] (survey)

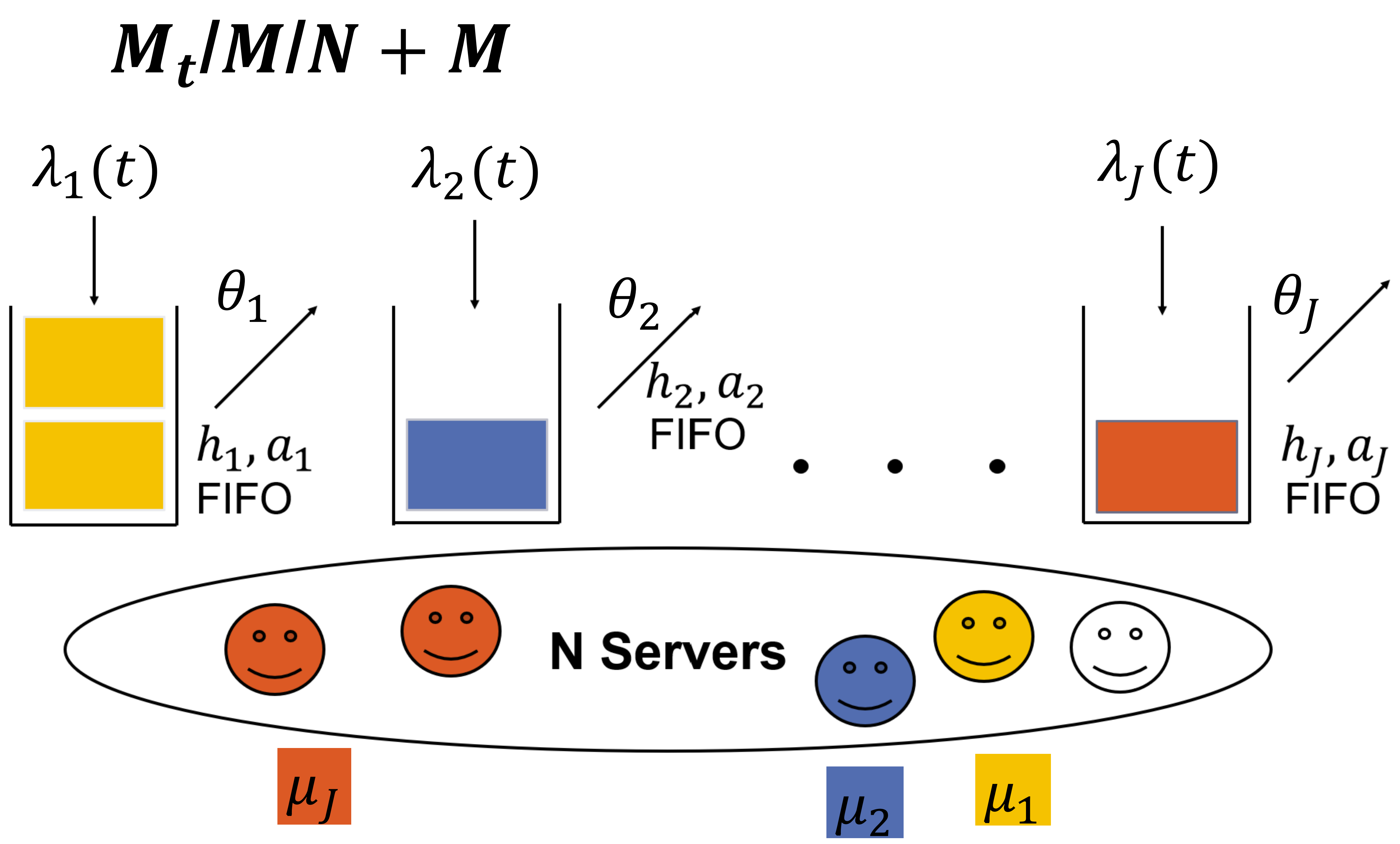
Time-varying queues

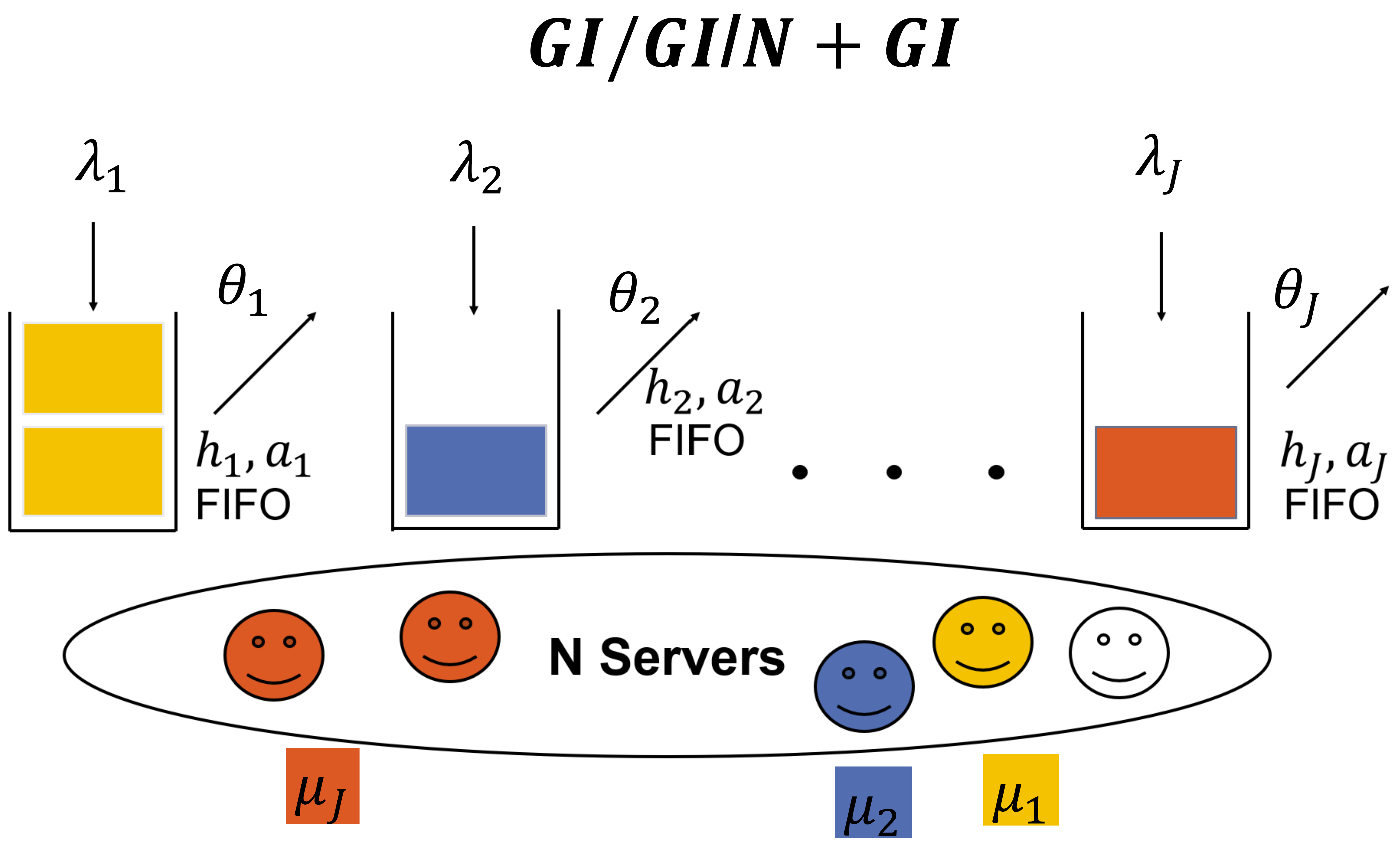
[Mandelbaum et al., 2002]

[Liu and Whitt, 2012]

[Zeifman, 1995]

An Optimal **Policy** π^*





Objective is to Minimize Regret, $\mathcal{R}(T,\pi) := \mathbb{E}[C_T(\pi)] - \mathbb{E}[C_T(\pi^\star)]$

$\pi^\star =$????????

Outline

- Regret Benchmark
- Algorithm & Regret Bounds
- Empirical Performance
- Current Work & Concluding Remarks

An Asymptotically Optimal Policy π^*

Simple and easy to learn policies!!!!

The Static Scheduling Problem:

Long-run average class j arrival rate

Decision variables:

$$\mathbb{B} := \left\{ b \in \mathbb{R}_+^J : b_j \le \frac{\Lambda_j}{\mu_j} \ \forall j \in [J], \sum_{j=1}^J b_j \le 1 \right\}$$

Long-run average fraction of server capacity devoted to class j

Objective ($GI/GI/N+GI$):
[Puha and Ward, 2019]

$$\inf_{b \in \mathbb{B}} \sum_{j=1}^J a_j(\lambda_j - b_j\mu_j) + h_j q_j(b_j) = \sum_{j=1}^J a_j\lambda_j - \sup_{b \in \mathbb{B}} \sum_{j=1}^J a_j \mu_j b_j$$

$= \frac{\lambda_j - b_j\mu_j}{\theta_j}$, if exponential abandonment

Objective ($M_t/M/N+M$):

$$\inf_{b \in \mathbb{B}} \sum_{j=1}^J a_j(\Lambda_j - b_j\mu_j) + h_j \frac{\Lambda_j - b_j\mu_j}{\theta_j} = \sum_{j=1}^J \Lambda_j \left(\frac{a_j\theta_j + h_j}{\theta_j} \right) - \sup_{b \in \mathbb{B}} \sum_{j=1}^J \frac{a_j\theta_j + h_j}{\theta_j} b_j \mu_j$$

Solution ~~($M_t/M/N+M$):~~

($GI/GI/N+GI$):

Assume $h_j = 0$.

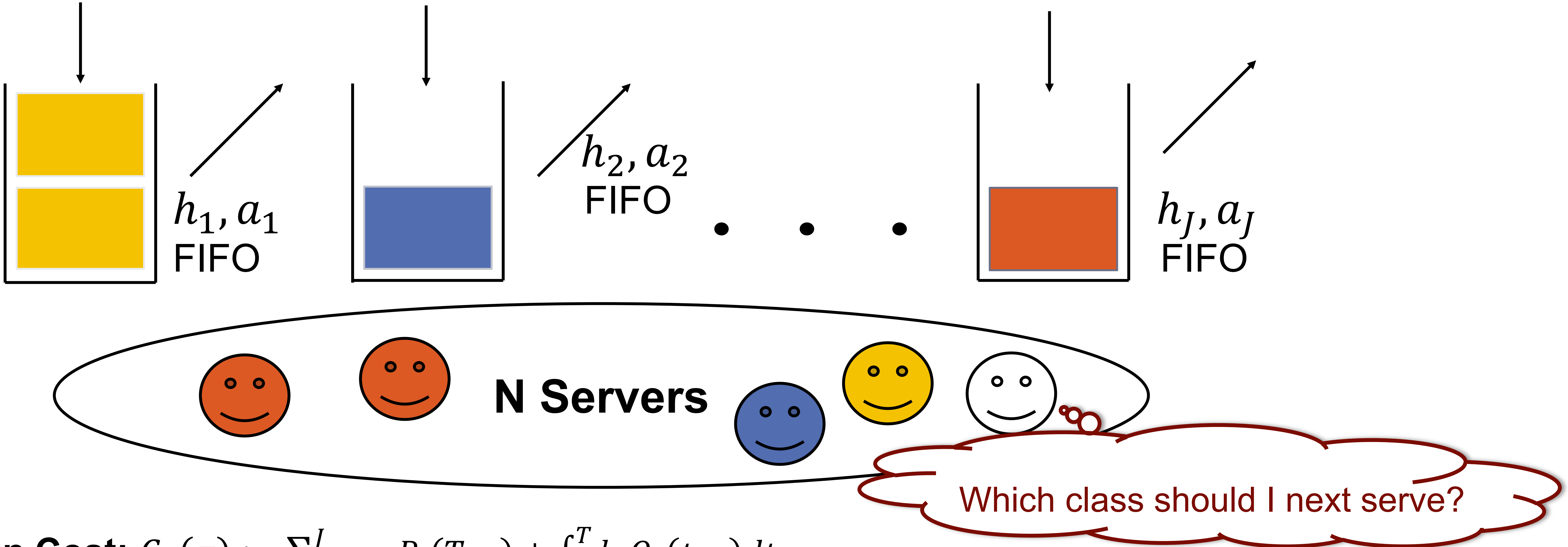
$c\mu/\theta$ static priority policy $\pi_{c\mu/\theta}$ (non-idling):

$$(a_1\theta_1 + h_1)\frac{\mu_1}{\theta_1} > (a_2\theta_2 + h_2)\frac{\mu_2}{\theta_2} > \cdots > (a_J\theta_J + h_J)\frac{\mu_J}{\theta_J}$$

Becomes $a_1\mu_1 > a_2\mu_2 > \cdots > a_J\mu_J$.

CHICAGO BOOTH
The University of Chicago Booth School of Business

A Canonical Model For Service Systems



Finite Horizon Cost: $C_T(\pi) \coloneqq \sum_{j=1}^{J} a_j R_j(T,\pi) + \int_0^T h_j Q_j(t,\pi)dt$

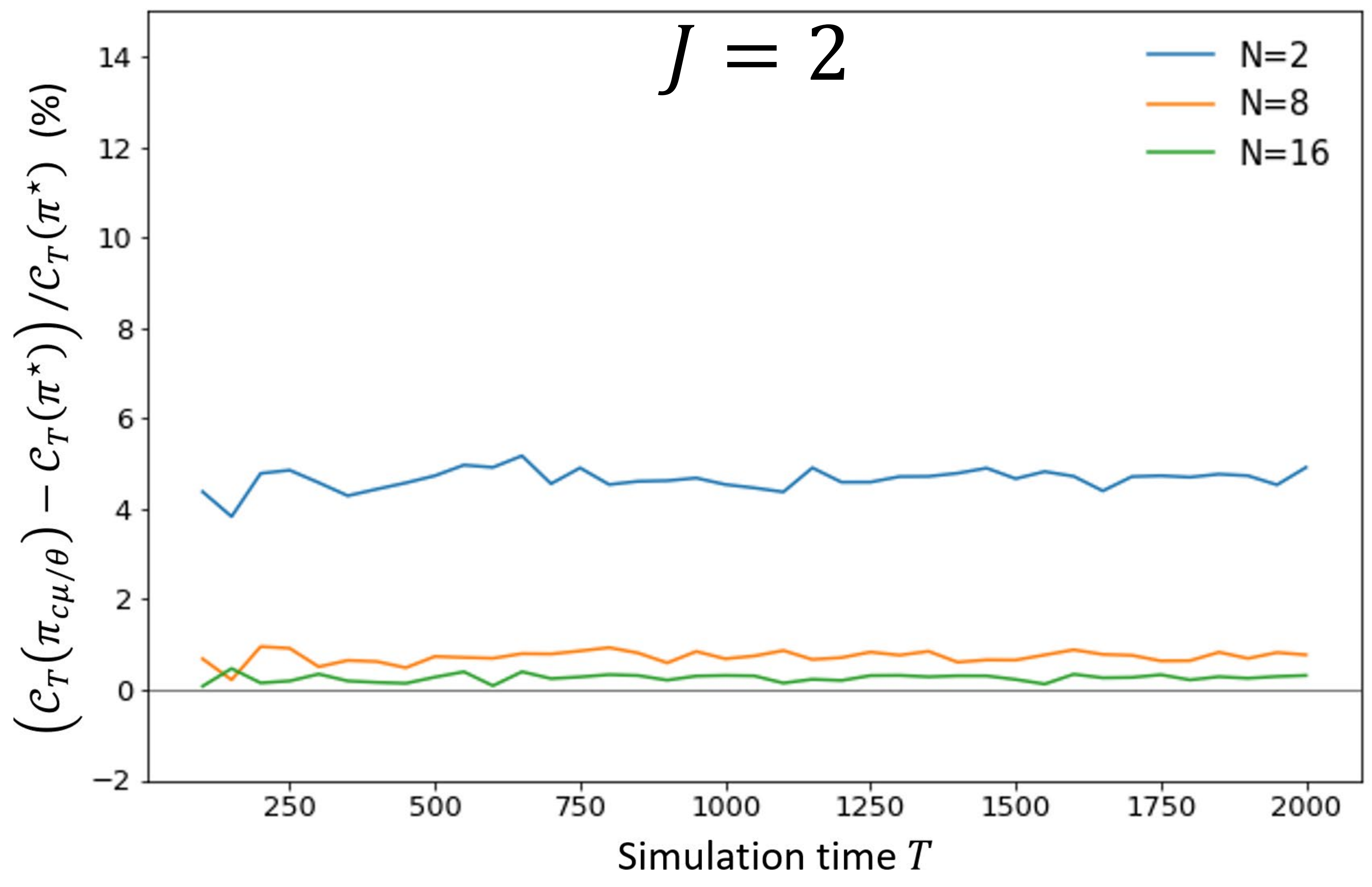
Long-Run Average Cost: $C(\pi) \coloneqq \limsup_{T\to\infty} \mathbb{E}[C_T(\pi)]/T$

~~an optimal policy~~

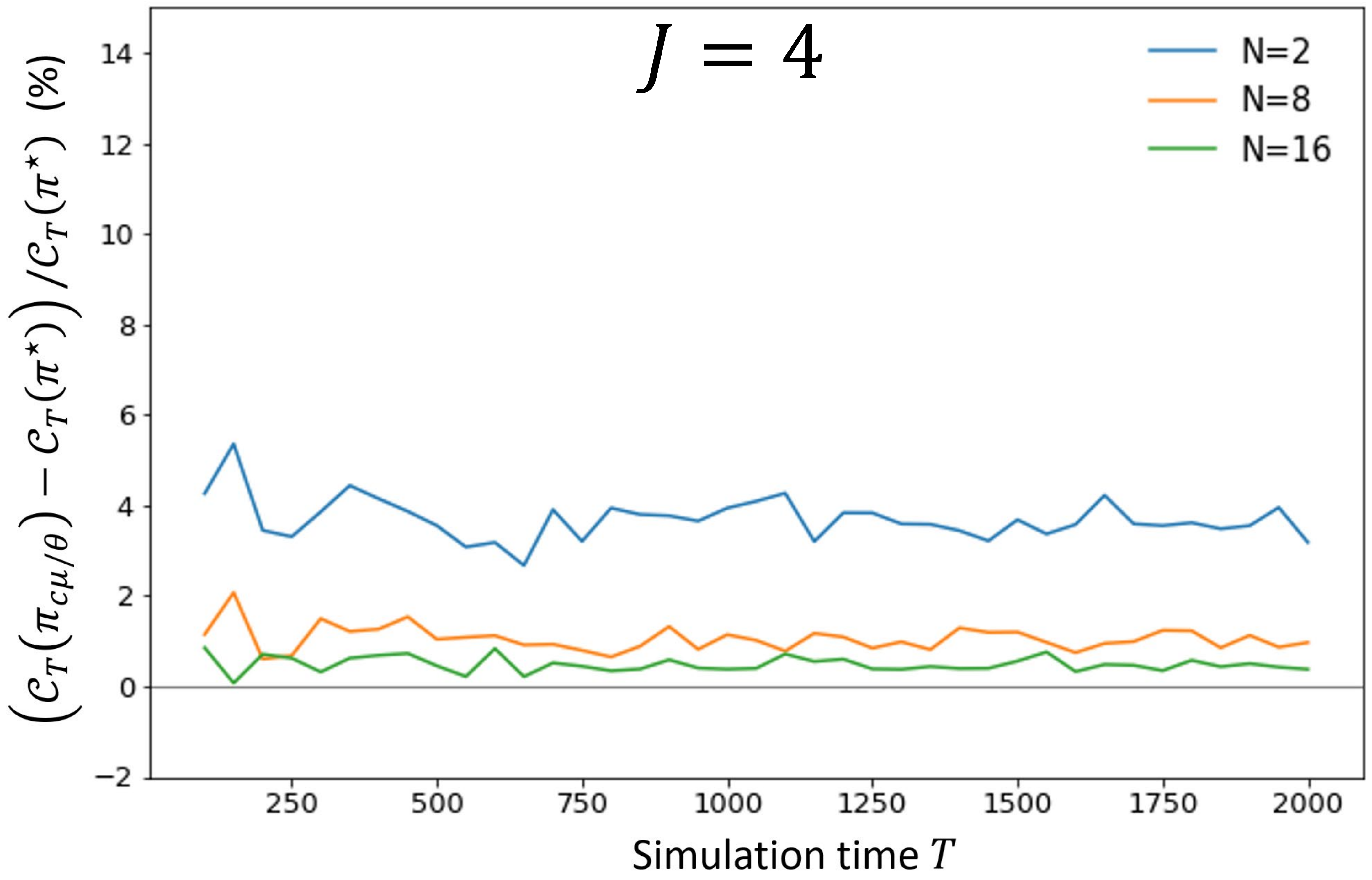
$\pi_{c\mu/\theta}$ the optimal $c\mu/\theta$ policy

Objective is to Minimize Regret, $\mathcal{R}(T,\pi) \coloneqq \mathbb{E}[C_T(\pi)] - \mathbb{E}[C_T\cancel{(\pi^\star)}]$

Performance of $c\mu/\theta$ in a Multiclass *M/M/N+M* Queue



$\lambda_1 = 2N, \lambda_2 = 4N, \mu_1 = \mu_2 = 4, \theta_1 = 1, \theta_2 = 4,$
$h_1 = 1, h_2 = 0.4, a_1 = 1, a_2 = 6$



$\lambda_1 = \lambda_3 = 2N, \lambda_2 = \lambda_4 = 4N, \mu_1 = \mu_2 = \mu_3 = \mu_4 = 4,$
$\theta_1 = \theta_3 = 1, \theta_2 = \theta_4 = 4,$
$h_1 = h_3 = 1, h_2 = h_4 = 0.4, a_1 = a_3 = 1, a_2 = a_4 = 6$

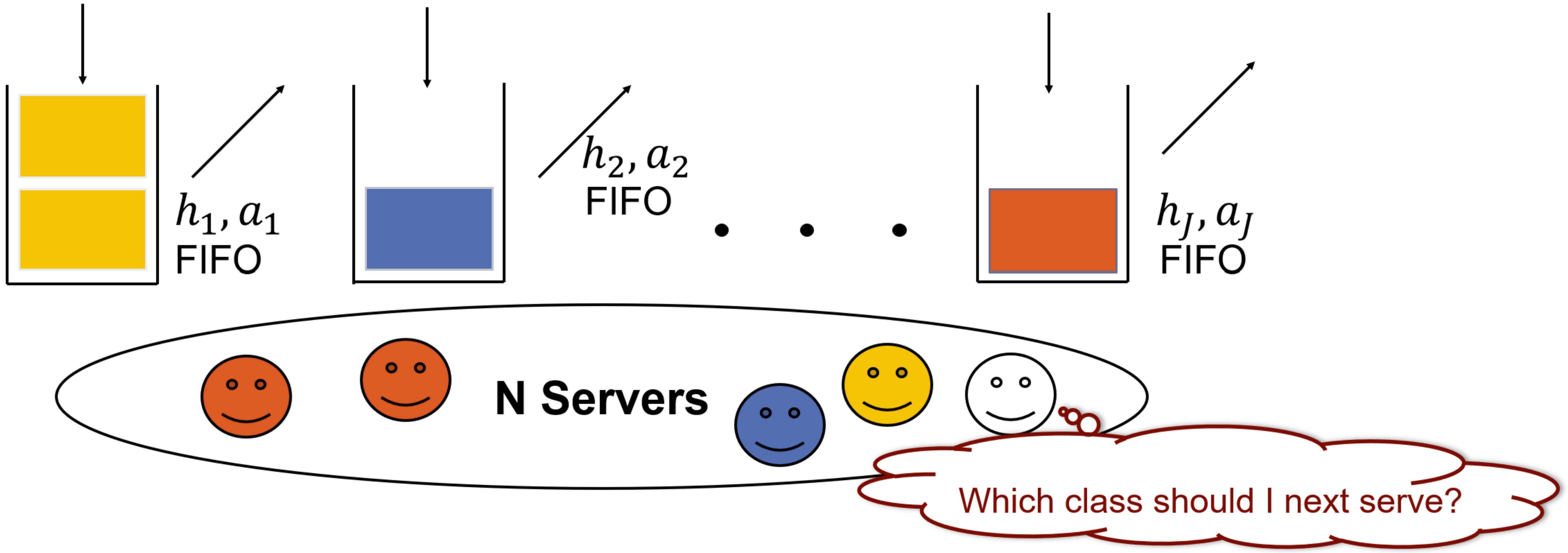
Proposition 3: For the $M_t/M/N + M$ queue, the $c\mu/\theta$ rule is asymptotically optimal, as $N \to \infty$, and then $t \to \infty$.

Analysis Assumptions

Assumption 1 (Arrival Rates):

- $\lambda_j(t)$ is locally integrable on $(0, \infty)$
- $0 < \lambda_j^L \leq \lambda_j(t) \leq \lambda_j^U$,

Finite constants



Assumption 2 (Index Separation):

For any $i \neq j$,

$$\left| \frac{(a_i\theta_i + h_i)\mu_i}{\theta_i} - \frac{(a_j\theta_j + h_j)\mu_j}{\theta_j} \right| \geq \delta$$

Finite positive constant

Objective is to Minimize Regret

$$\mathcal{R}(T, \pi) := \mathbb{E}[C_T(\pi)] - \mathbb{E}[C_T(\pi_{c\mu/\theta})]$$

Outline

- Regret Benchmark
- Algorithm & Regret Bounds
- Empirical Performance
- Current Work & Concluding Remarks

Regret Lower Bound

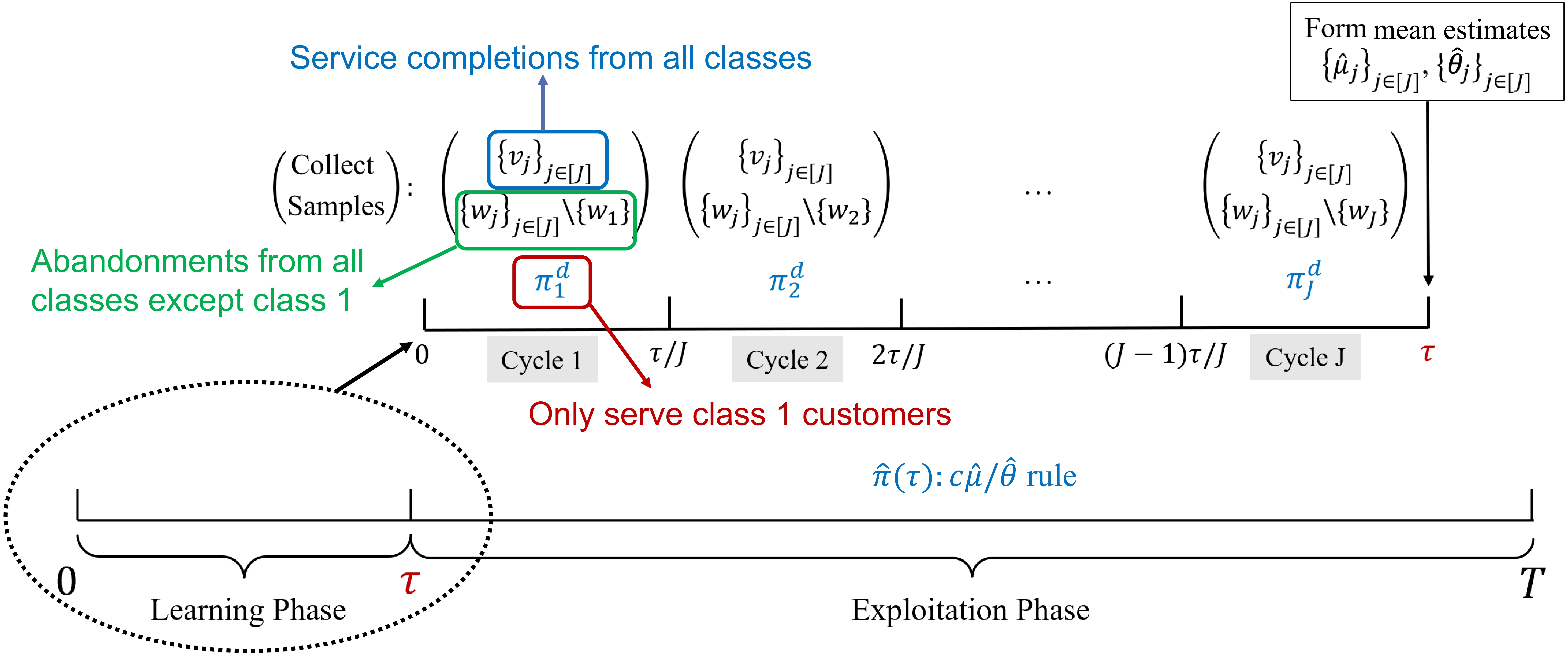
Theorem (Lower Bound on Regret): For any *consistent* non-anticipating and non-preemptive (and possibly randomized and idling) priority scheduling policy φ, there exists a finite constant $c > 0$ such that

$$\mathcal{R}(T, \varphi) \geq c \cdot \log T$$

- Consistency assumption: any "sub-optimal" scheduling policy will be applied for no more than $\mathcal{O}(n^a)$ cumulative fraction of time by the n-th arrival, so the "optimal" scheduling policy (i.e., the benchmark policy) can be eventually identified.

- Multi-Armed Bandit: Lai and Robbins (1985)

Algorithm: Learn-Then-Schedule $\pi_{LTS}(\tau)$



Algorithm Performance

Theorem (Upper Bound on Regret): When $\tau = \mathcal{O}(\log T)$, there exists a finite constant $C > 0$ such that

$$\mathcal{R}(T, \pi_{LTS}(\tau)) \leq C \cdot \log T$$

Theorem (Lower Bound on Regret): $\mathcal{R}(T, \varphi) \geq c \cdot \log T$

Corollary: When $\tau = \mathcal{O}(\log T)$, $\mathcal{R}(T, \pi_{LTS}(\tau)) = \Theta(\log T)$

Regret Analysis Framework

Regret incurred between time τ and time T.

$$\mathcal{R}(T, \pi_{LTS}(\tau)) = \mathcal{R}^{Learning}(T, \pi_{LTS}(\tau)) + \mathcal{R}^{Exploitation}(T, \pi_{LTS}(\tau))$$

Regret incurred between time 0 and τ.

Proposition 1: $\mathcal{R}^{Learning}(T, \pi_{LTS}(\tau)) \leq \left(\sum_{j=1}^{J} \left(\frac{h_j}{\theta_j} + a_j \right) \lambda_j^U \right) \cdot \tau$

Proposition 2: There exist $C_1, C_2, l_1 > 0$ such that

$$\mathcal{R}^{Exploitation}(T, \pi_{LTS}(\tau)) \leq C_1 + C_2(T - \tau)e^{-\ell_1 \tau}.$$

Regret Analysis Framework

$$\mathcal{R}^{Exploitation}(T, \pi_{LTS}(\tau))$$

$$= \mathbb{P}\{\hat{\pi}(\tau) = \pi_{c\mu/\theta}\} \cdot \mathcal{R}^{Exploitation}\left(T, \pi_{LTS}(\tau) | \hat{\pi}(\tau) = \pi_{c\mu/\theta}\right)$$

$$+ \mathbb{P}\{\hat{\pi}(\tau) \neq \pi_{c\mu/\theta}\} \cdot \mathcal{R}^{Exploitation}\left(T, \pi_{LTS}(\tau) | \hat{\pi}(\tau) = \pi_{c\mu/\theta}\right)$$

$$\leq \boxed{\mathcal{R}^{Exploitation}\left(T, \pi_{LTS}(\tau) | \hat{\pi}(\tau) = \pi_{c\mu/\theta}\right)} + \boxed{\mathbb{P}\{\hat{\pi}(\tau) \neq \pi_{c\mu/\theta}\}} \cdot \boxed{\mathcal{R}^{Exploitation}\left(T, \pi_{LTS}(\tau) | \hat{\pi}(\tau) \neq \pi_{c\mu/\theta}\right)}$$

$$\mathbb{E}\left[\left|Q_j^1(t) - Q_j^2(t)\right|^m \mid \left(Q_j^1(0), Q_j^2(0)\right)\right]$$

$$\leq \Psi \cdot e^{-d_j \cdot t} \cdot \left|Q_j^1(0) - Q_j^2(0)\right|^m$$

Exponential loss of memory

(Weakly ergodicity of $M_t/M/N{+}M$)

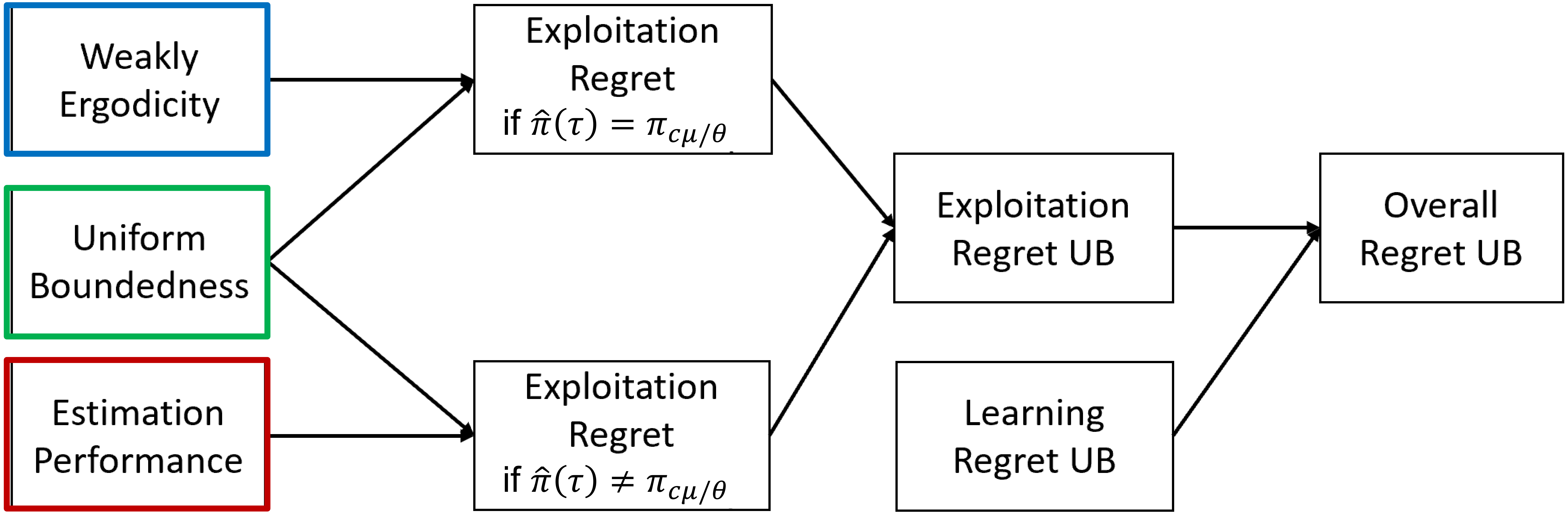
$$\mathbb{P}\{\hat{\pi}(\tau) \neq \pi_{c\mu/\theta}\} \leq 6J \cdot e^{-\ell_1 \cdot \tau}$$

Estimation performance

$$\mathbb{E}[Q_j(t, \pi)] \leq \frac{\lambda_j^U}{\mu_j} + \mathbb{E}[Q_j(0)]$$

Uniform boundedness

Regret Analysis Framework

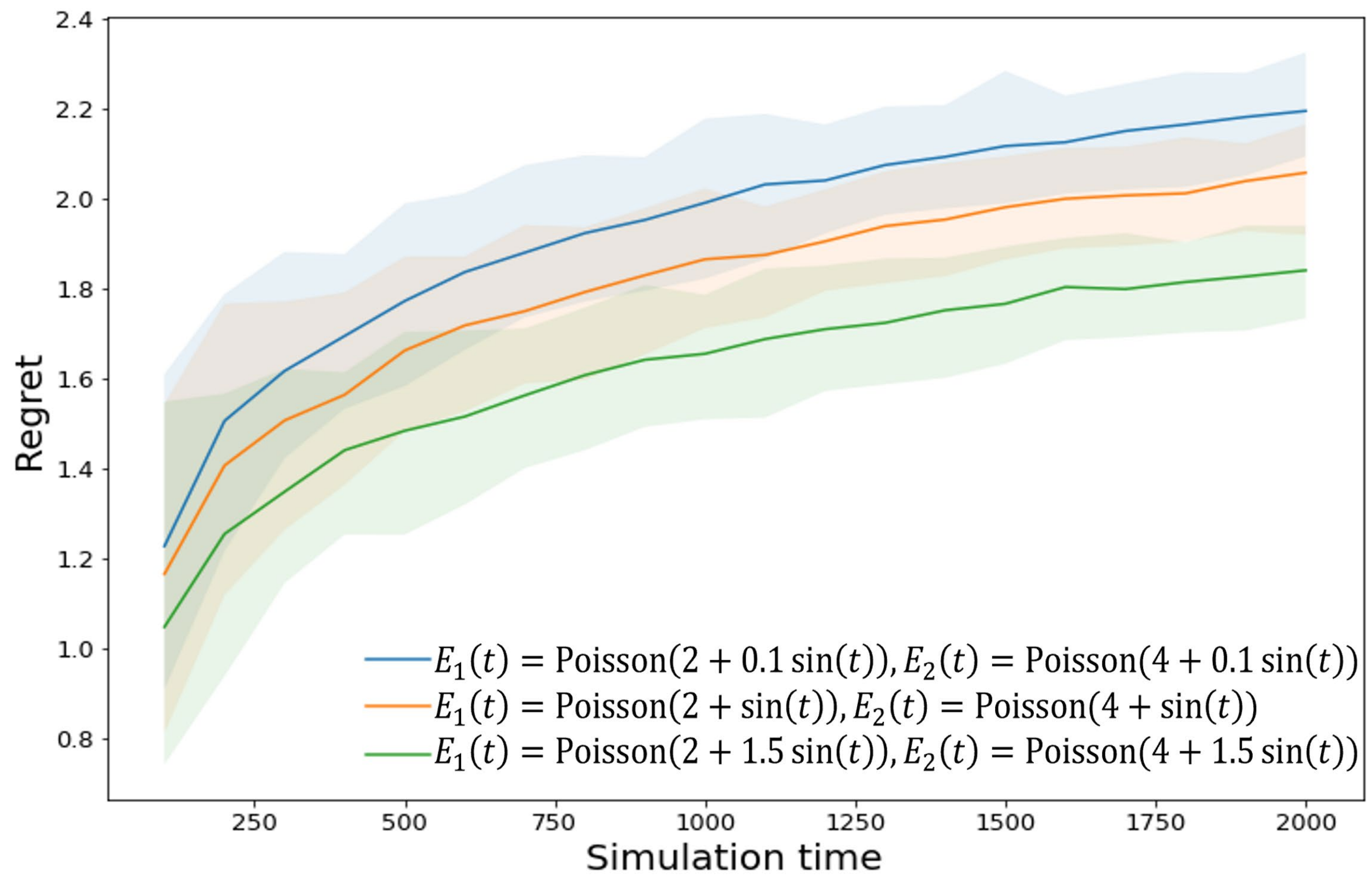


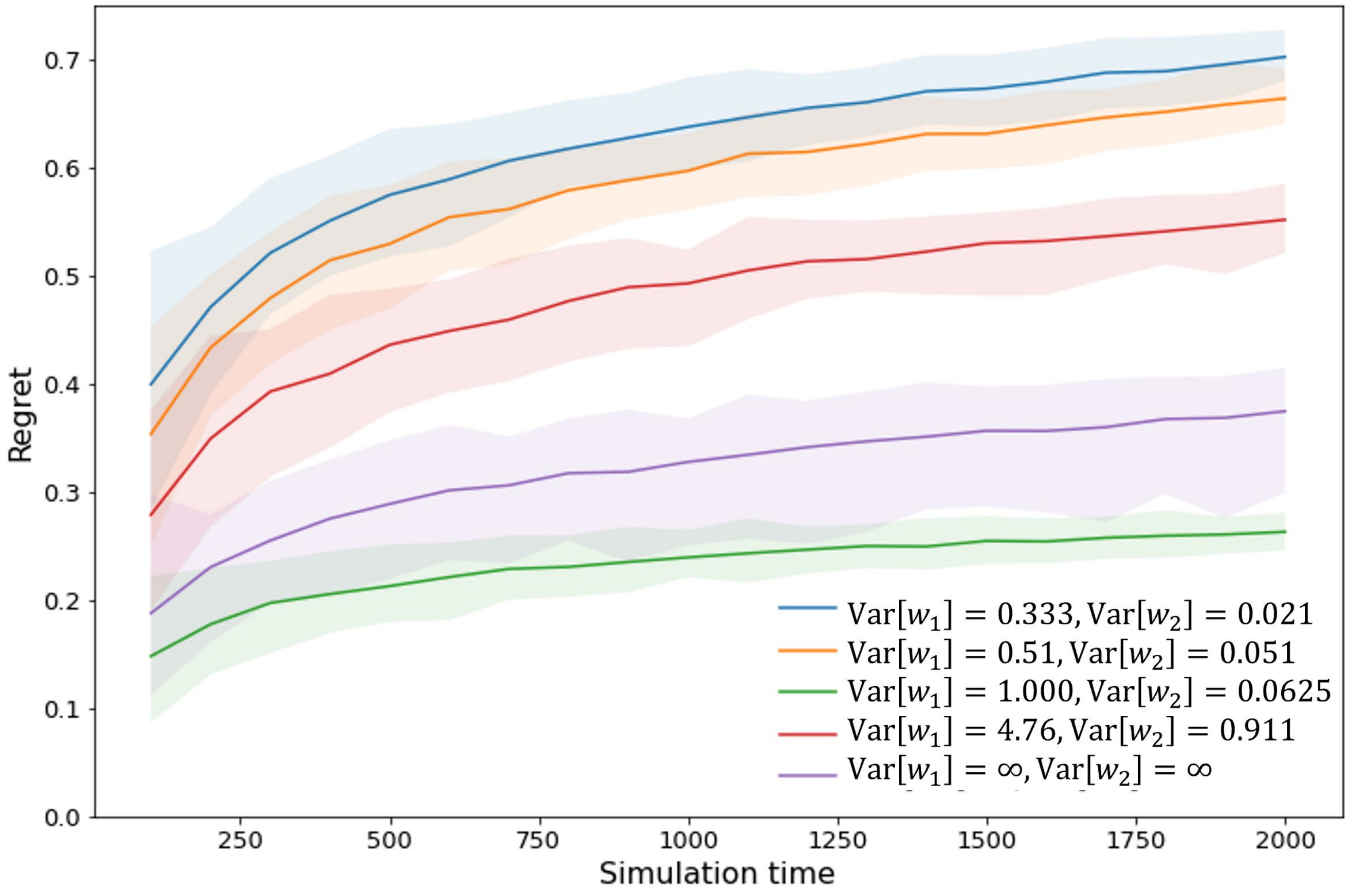
Outline

- Regret Benchmark
- Algorithm & Regret Bounds
- Empirical Performance
- Current Work & Concluding Remarks

Numerical Study: Two Classes ($N = 8$)

Class	Inter-arrival time	Service time	Patience time	Holding cost h_j	Abandon cost a_j	Index $\left(a_j\theta_j + h_j\right)\frac{\mu_j}{\theta_j}$
1	~~Poisson(2)~~	Exp(4)	~~Exp(1)~~	1	1	8.0
2	~~Poisson(4)~~	Exp(4)	~~Exp(4)~~	0.4	6	7.6





Outline

- Regret Benchmark
- Algorithm & Regret Bounds
- Empirical Performance
- Current Work & Concluding Remarks

Current Work: Instance Independent Regret

- Minimax regret:

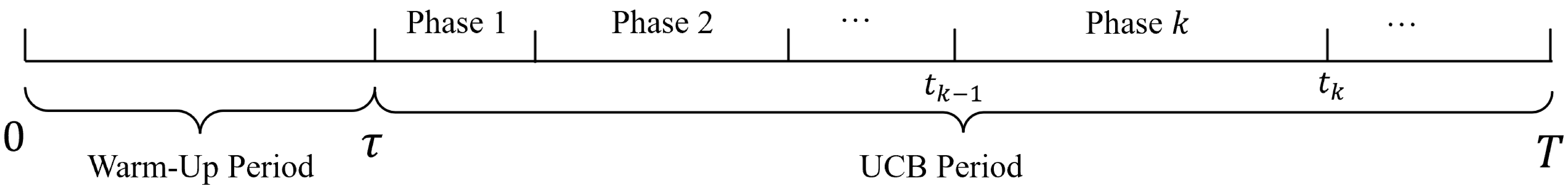
$$Reg(T,\pi) \coloneqq \sup_{\eta\in H} \mathcal{R}(T,\pi;\eta)$$

Parameter space

Theorem (LTS is Not Enough): $Reg(T,\pi_{LTS}(\tau)) = \Theta(T^{2/3})$ when $\tau = \mathcal{O}(T^{2/3})$.

Theorem (Lower Bound): $Reg(T,\pi) \geq \Omega(\sqrt{T})$.

Proposed algorithm: Phased-UCB with forced exploration $\pi_{PUCBFE}(\tau)$



Theorem (Upper Bound): $Reg\left(T,\pi_{PUCBFE}(\tau)\right) \leq \mathcal{O}(\sqrt{T})$, when $\tau = \mathcal{O}(\log T)$.

Concluding Remarks

Summary

- We solve a learning variant of a canonical scheduling problem in a multiclass many server queue with abandonment (specifically, the $M_t/M/N+M$ and the $GI/GI/N+GI$ systems), when model parameters are a priori unknown.
- We propose online algorithms to achieve optimal instance-dependent and minimax regrets.

Follow-On Work

- Extend the results to $G/GI/N+GI$ queue with holding costs and $G_t/GI/N+GI$ queue.
- Bound the suboptimality gap of $\pi_{c\mu/\theta}$ from the true optimal
- Multiple classes of servers

Vision for Learning Problems in Stochastic Systems

- *Exploit asymptotic analysis to define an easier learning problem*

Paper on SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4090021

Thanks!
Q&A

yzhong0@chicagobooth.edu