Analysis of Scheduling Policies Under Correlated Job Sizes

VARUN GUPTA
Carnegie Mellon University

With:

Michelle Burroughs
(CMU)

Mor Harchol-Balter
(CMU)

- Mean response time
- Tail of response time
- Buffer overflow prob.
- Prioritized service
- …

Scheduling = which job to serve to optimize performance

A typical paper:

We analyze [the metric] under [the scheduling policy]. We assume that the job sizes are i.i.d. and find that [the scheduling policy is good].

Our Message:

size-independent

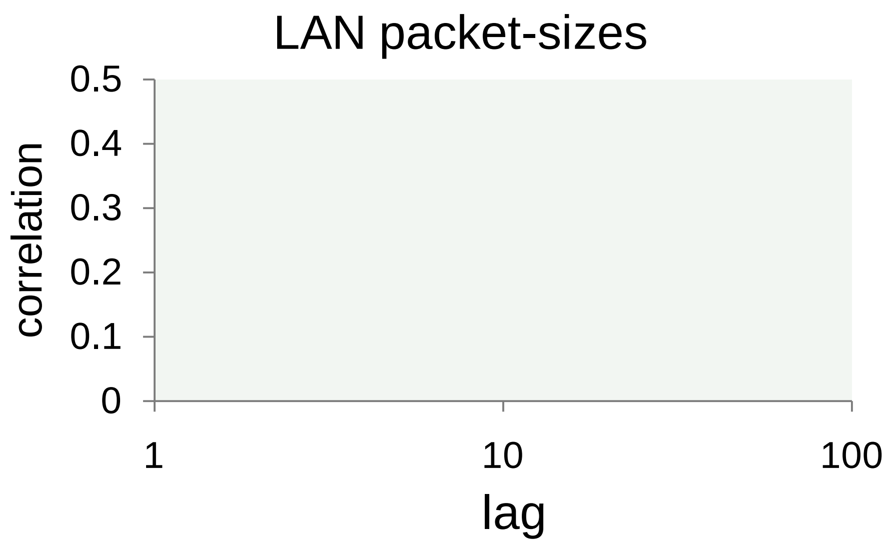
Correlation in sizes hugely impacts "qualitative properties" of size-independent policies

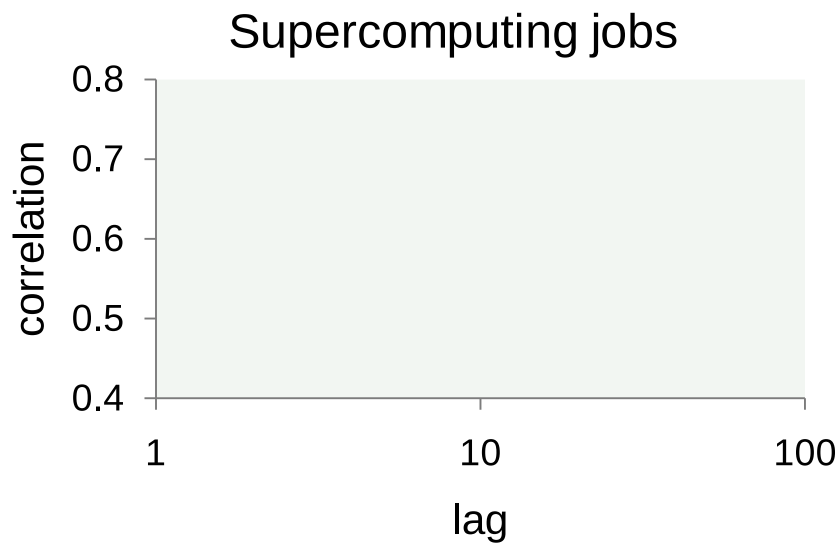
Optimality

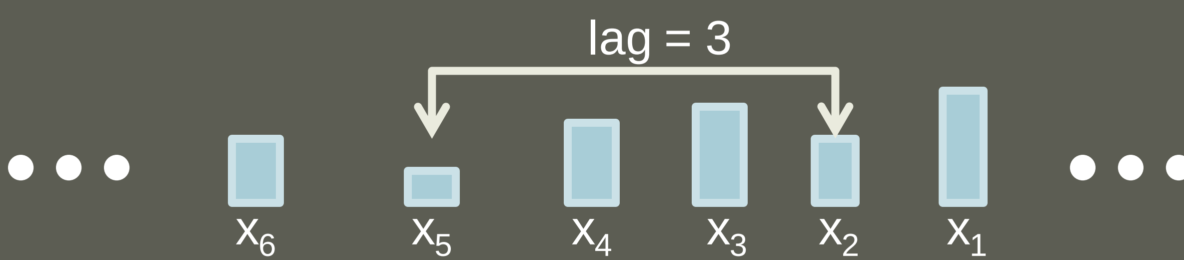
Effect of job size distribution

Correlation in sizes is important!

auto-correlation functions from traces





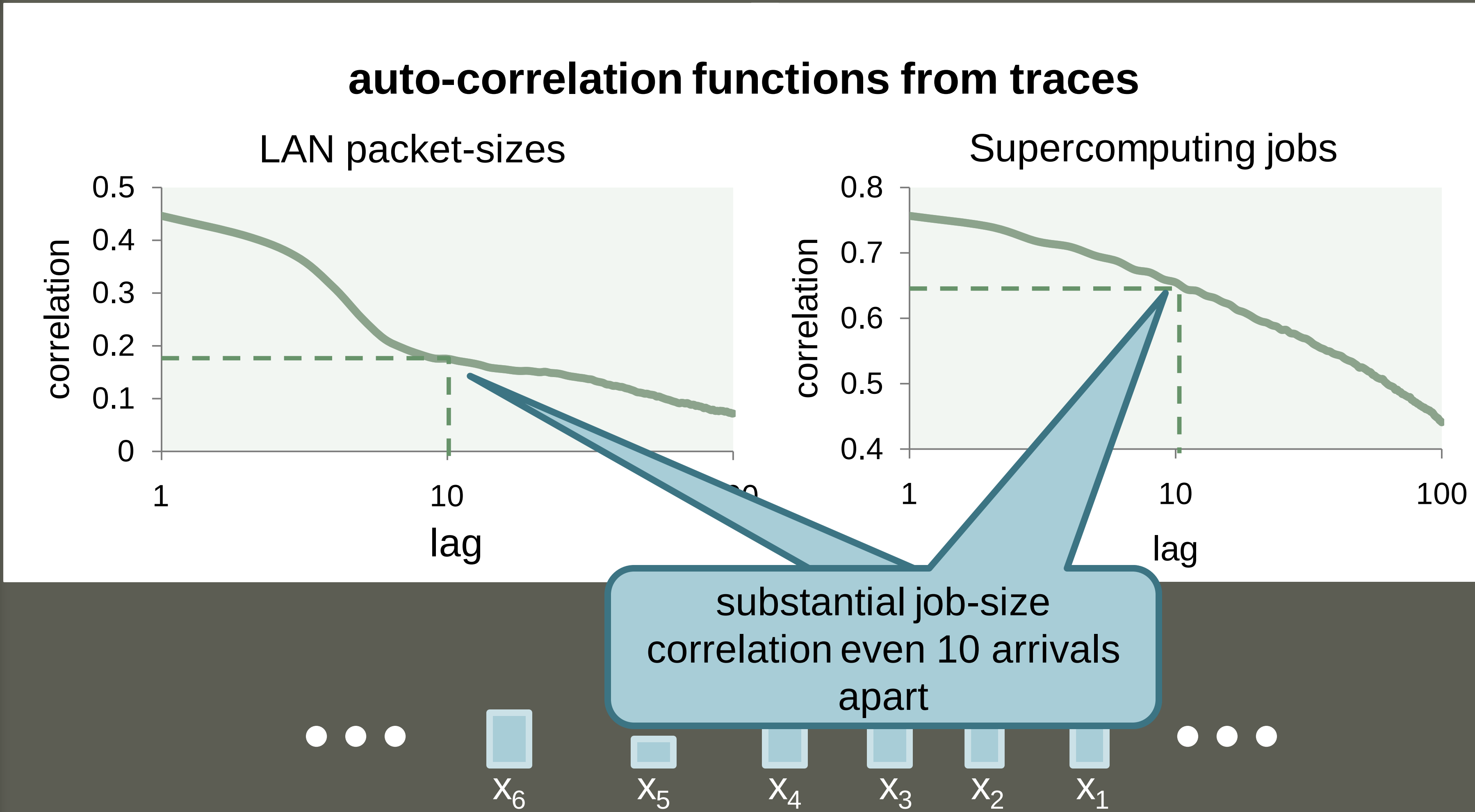


Recall:

correlation(X,Y) = 0 $\Rightarrow$ X,Y are *linearly independent*

correlation(X,Y) = 1 $\Rightarrow$ X = cY

Correlation in sizes is important!



Recall:

correlation(X,Y) = 0 $\Rightarrow$ X,Y are *linearly independent*

correlation(X,Y) = 1 $\Rightarrow$ X = cY

Most of existing scheduling theory

This talk

- **Analysis of common policies (mean response time)**
- **Impact of correlation on qualitative properties**

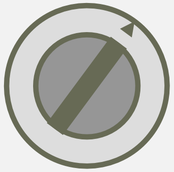
i.i.d. ⟷ **high correlation**

1. Scheduling theory refresher

3. Analysis of common scheduling policies

4. Simulation results

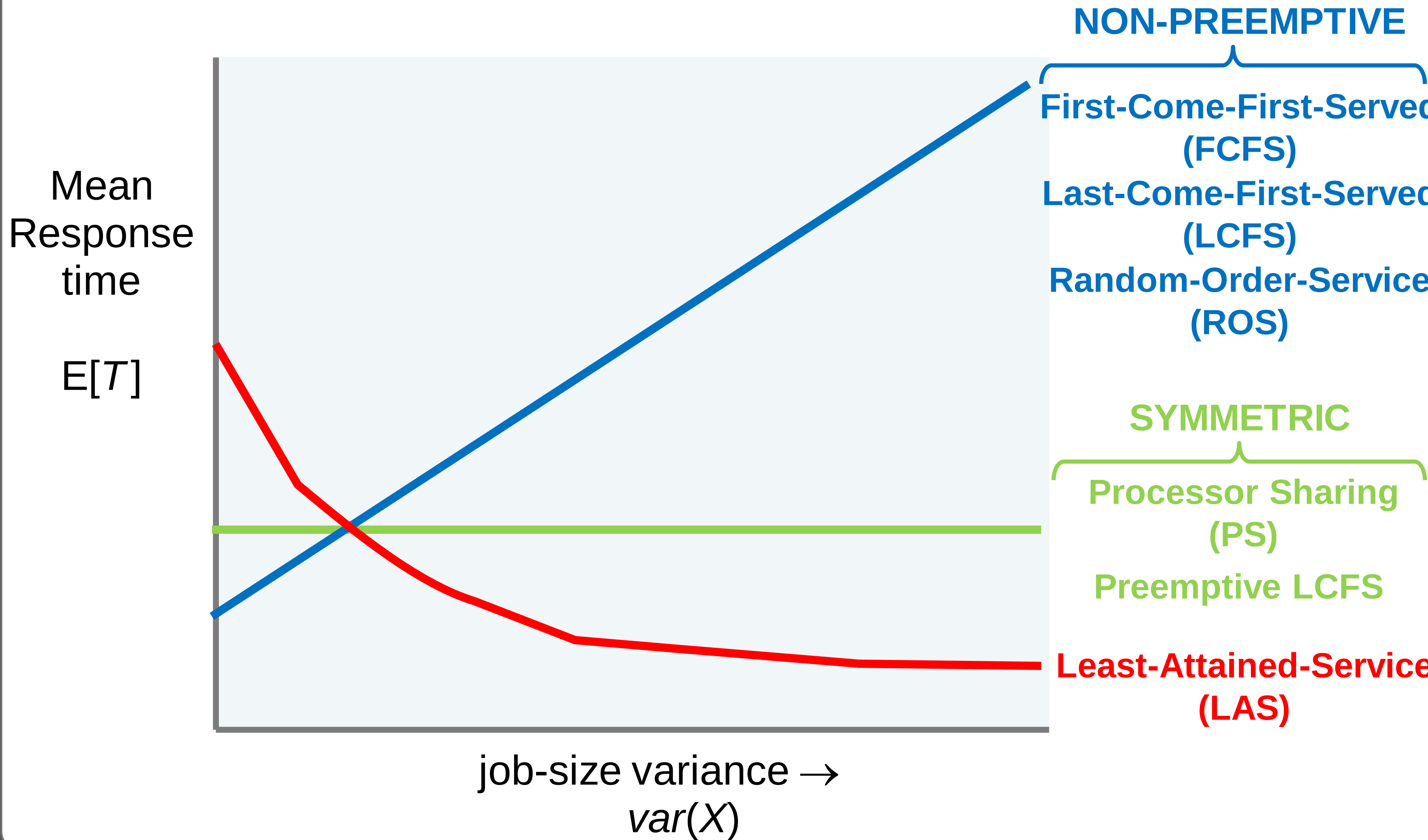
i.i.d. ⟷ **high correlation**



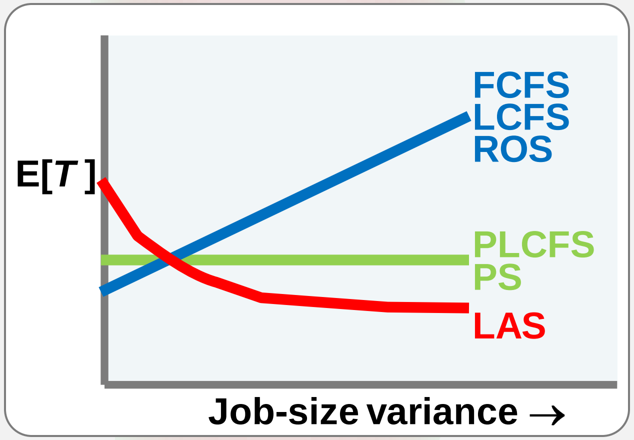
2. Analytically tractable correlation model

Scheduling refresher for *i.i.d.* job sizes

(and Poisson arrivals)



1. Scheduling theory refresher

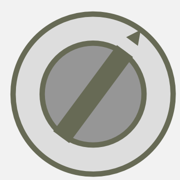


3. Analysis of common scheduling policies

4. Simulation results

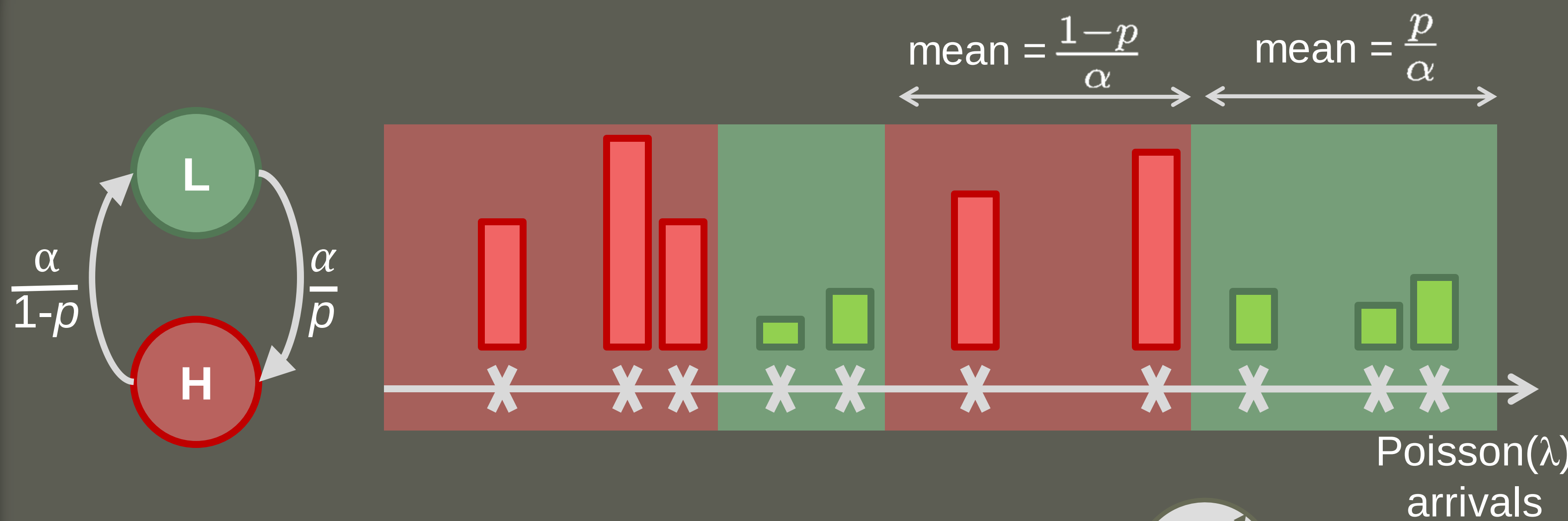
low correlation

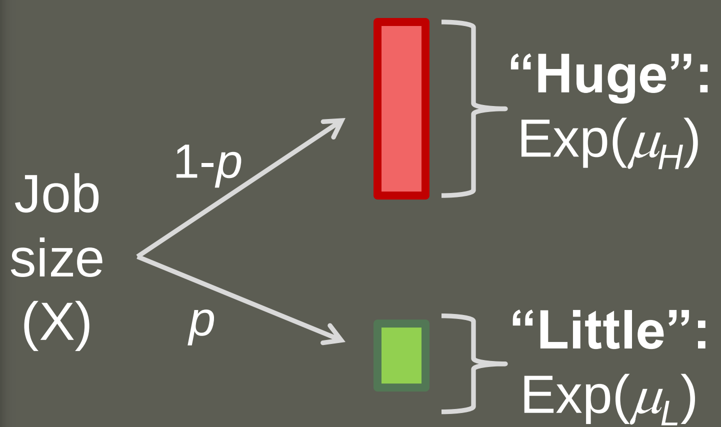
high correlation

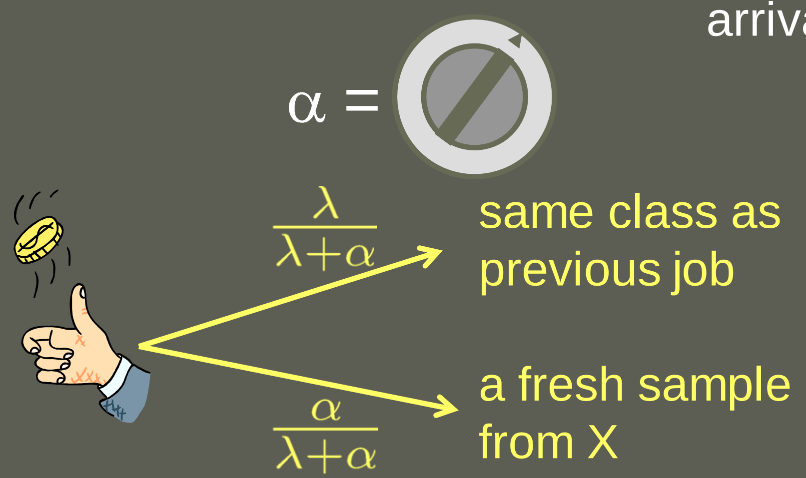


2. Analytically tractable correlation model

A Markov-Modulated correlation model

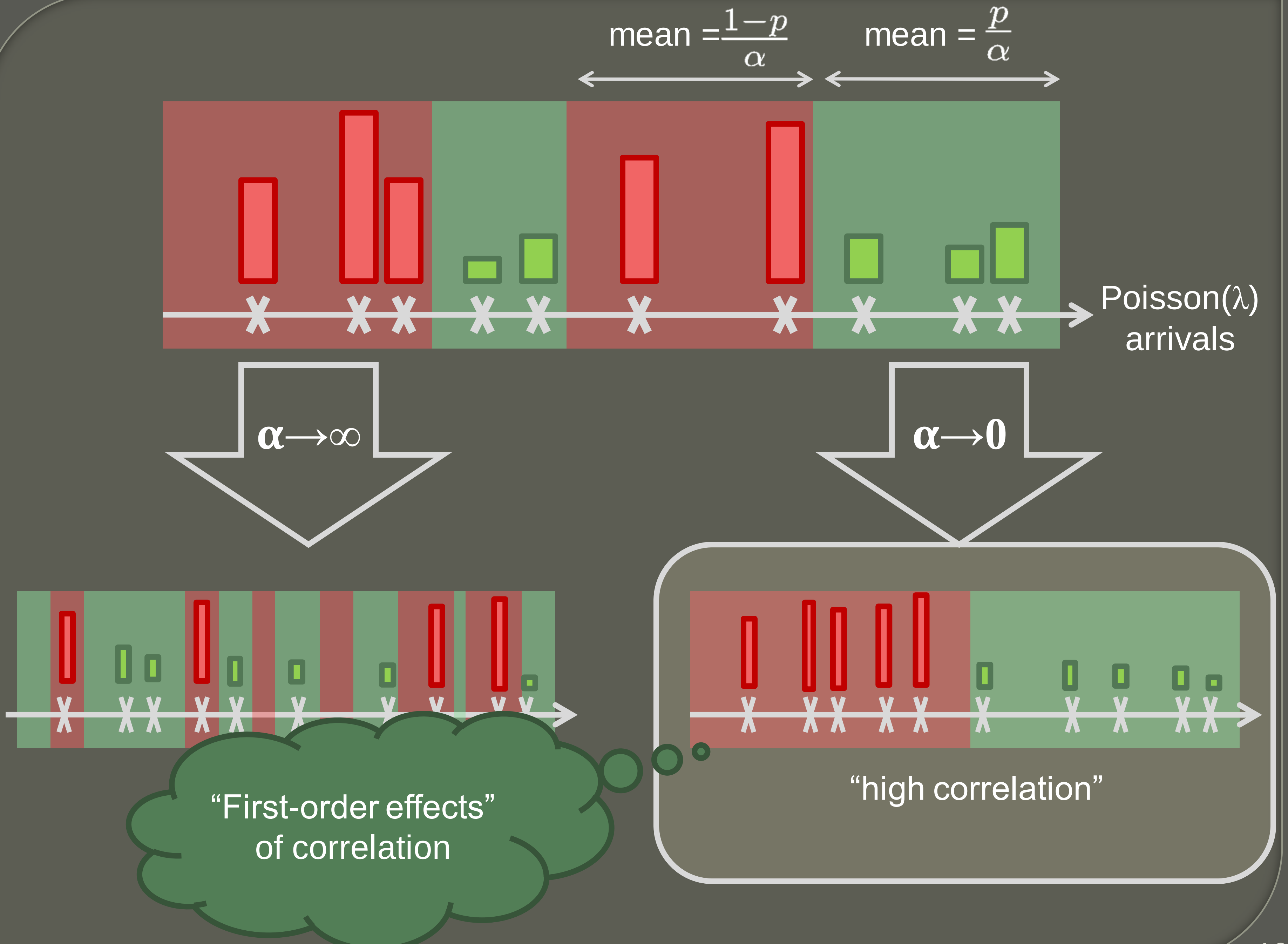




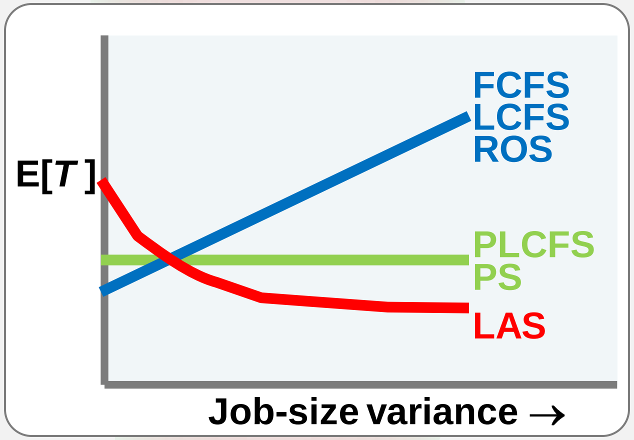


this talk: $\frac{\lambda}{\mu_H} > 1 > \frac{\lambda}{\mu_L}$

$$\text{lag } n \text{ correlation} = \frac{var(X)}{2E[X^2]}\left(\frac{\lambda}{\lambda+\alpha}\right)^n$$


mean = $\frac{1-p}{\alpha}$
mean = $\frac{p}{\alpha}$
Poisson(λ) arrivals
α→∞
α→0
"First-order effects" of correlation
"high correlation"

1. Scheduling theory refresher



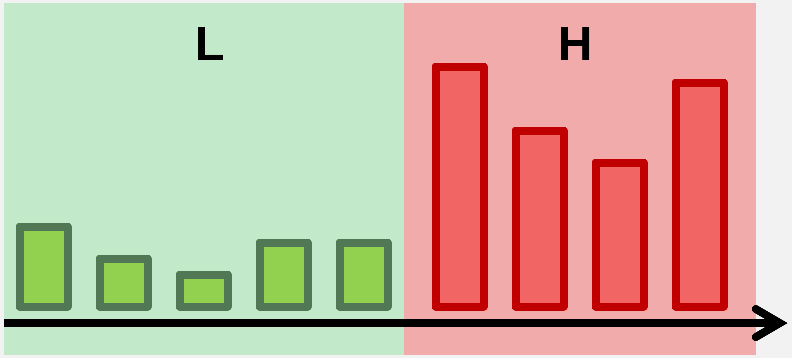
4. Simulation results

3. Analysis of common size-independent scheduling policies

i.i.d.

high correlation

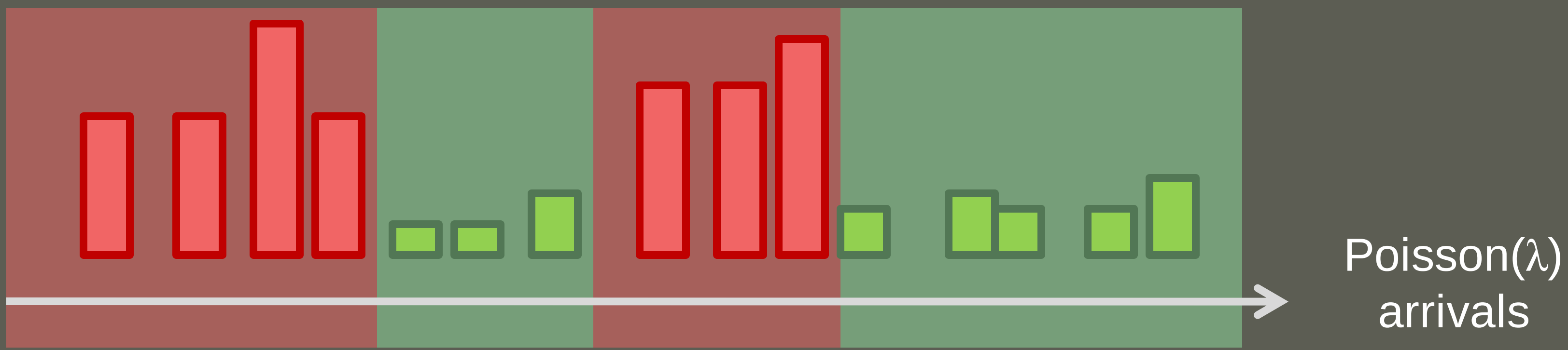
2. MMAP correlation model



Analysis Roadmap

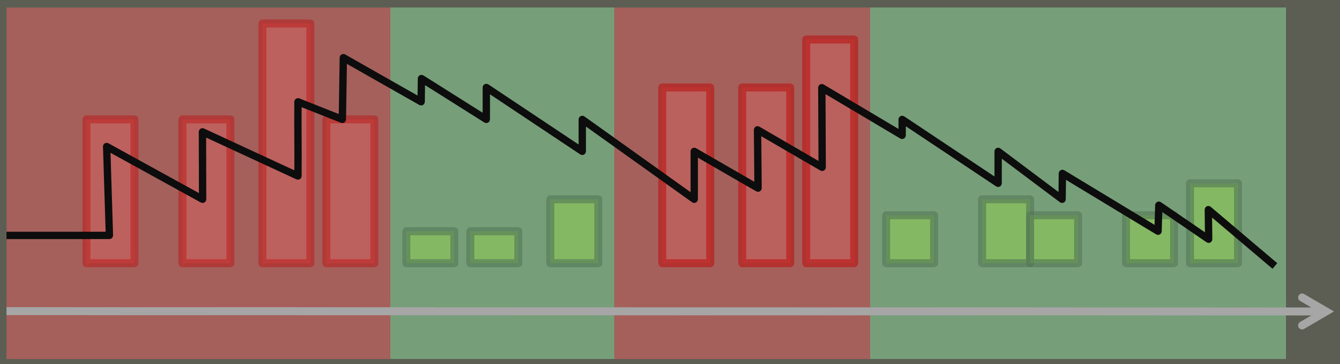
- Warm-up: Analysis of workload
  - FCFS
- A hypothetical OPT policy
- Preemptive LCFS
- Least-Attained-Service

Warm-up : fluid analysis of workload



$$\left(\frac{\lambda}{\mu_H} > 1 > \frac{\lambda}{\mu_L}\right)$$

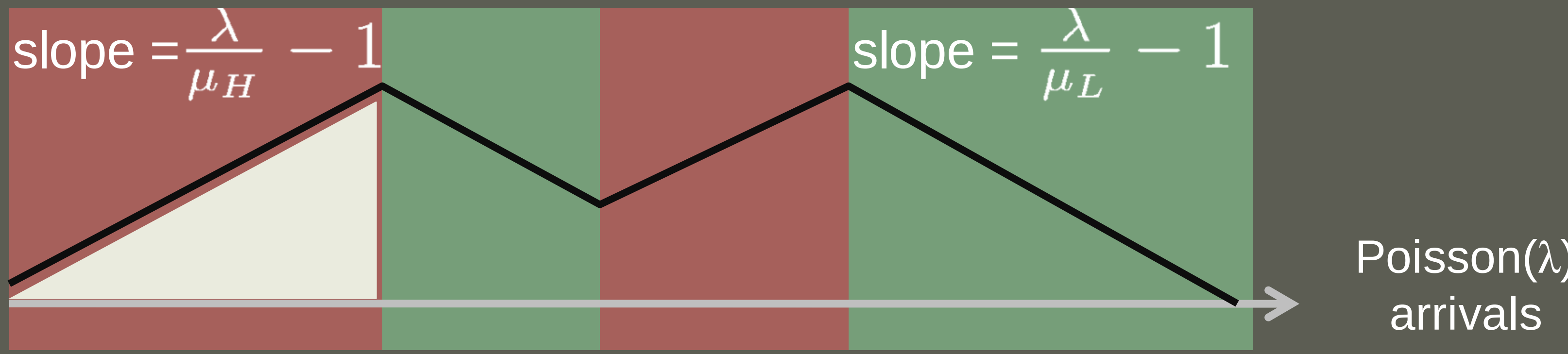
Warm-up : fluid analysis of workload



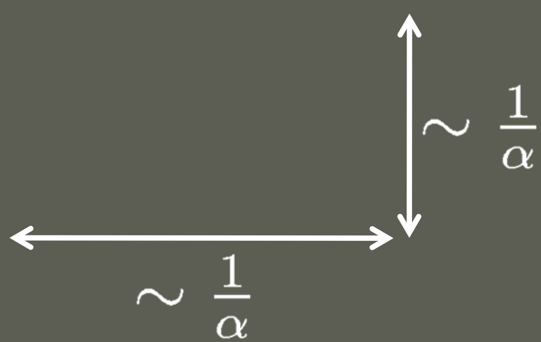
as $\alpha \to 0$

$$\left(\frac{\lambda}{\mu_H} > 1 > \frac{\lambda}{\mu_L} \right)$$

Warm-up : fluid analysis of workload



as $\alpha \to 0$



$$w = f(\lambda, \mu_L, \mu_H, p)$$

For all policies, workload:

$$\mathrm{E}[W] \sim \frac{w}{\alpha}$$

$\Rightarrow$

For policy π :

$$\mathrm{E}[T^{\pi}] \sim \frac{K^{\pi}}{\alpha}$$

For FCFS (by PASTA): $\mathrm{E}[T^{\mathrm{FCFS}}] \sim \frac{w}{\alpha}$

Analysis Roadmap

- Warm-up: Workload analysis
  - FCFS

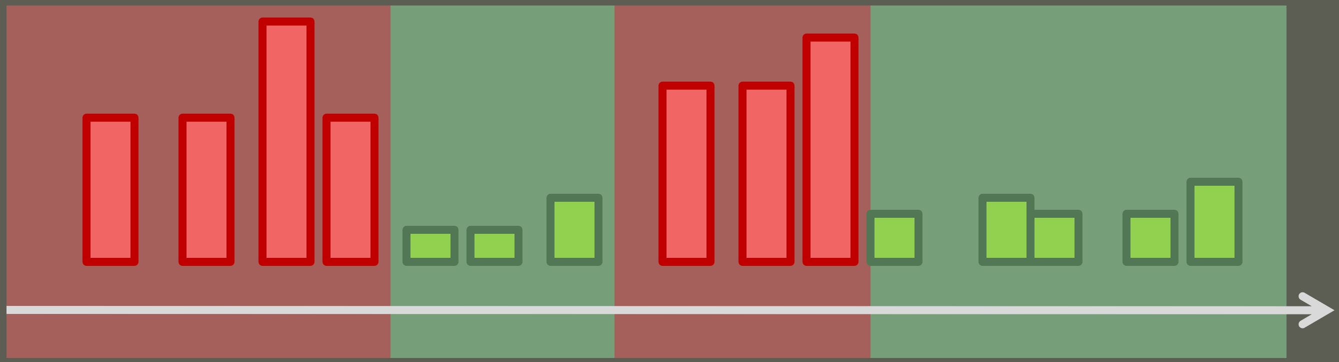
$$\mathrm{E}[T^{\pi}] \sim \frac{K^{\pi}}{\alpha}$$

$$\mathrm{E}[T^{\mathrm{FCFS}}] \sim \mathrm{E}[W] \sim \frac{w}{\alpha}$$

- A hypothetical OPT policy
- Preemptive LCFS
- Least-Attained-Service

A hypothetical OPT policy

OPT: preemptive priority to (little) L jobs $\Rightarrow$ Lower bound on the best “predictive” policy



For L jobs: $E[T_L^{OPT}] = O(1)$

$E[T_L^{OPT}]$ is bounded as $\alpha \to 0$

$\Rightarrow$ Workload almost entirely from H jobs

$\Rightarrow \quad \mathrm{E}[\#\text{ of jobs}] \sim \mu_H \cdot \mathrm{E}[W]$

$\Rightarrow \quad \mathrm{E}[T^{\mathrm{OPT}}] \sim \mu_H \cdot \mathrm{E}[W]/\lambda$

Analysis Roadmap

- Warm-up: Workload analysis
  - FCFS

$$\mathrm{E}[T^{\pi}] \sim \frac{K^{\pi}}{\alpha}$$

$$\mathrm{E}[T^{\mathrm{FCFS}}] \sim \mathrm{E}[W] \sim \frac{w}{\alpha}$$

- A hypothetical OPT policy

$$\mathrm{E}[T_L^{\mathrm{OPT}}] = O(1)$$

$$\mathrm{E}[T^{\mathrm{OPT}}] \sim \frac{\mu_H}{\lambda} \cdot \frac{w}{\alpha}$$

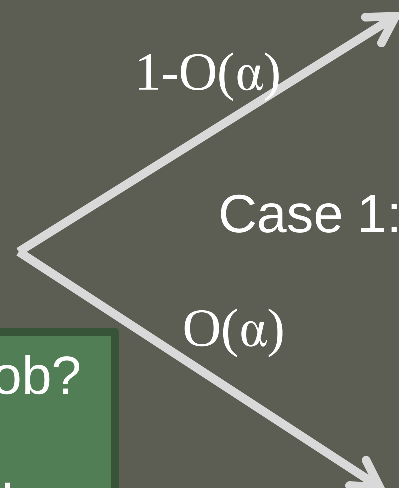
- Preemptive LCFS
- Least-Attained-Service

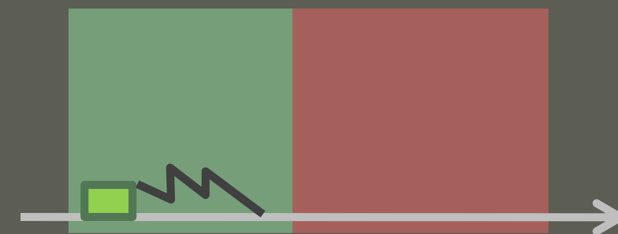
Preemptive Last-Come-First-Served

THEOREM: $E[T_L^{PLCFS}] = O(1)$ **COR:** $E[T^{PLCFS}] \sim E[T^{OPT}]$

Q: Response time of the L job?

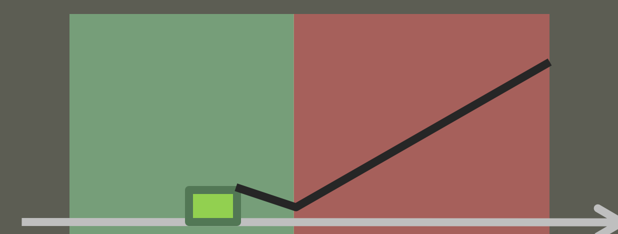
A: Busy period generated by the L job





Case 1: busy period ends in the L state

$E[T_L] = O(1)$



Case 2: busy period overflows into H state

$E[T_L] = O(1/\alpha)$

Analysis Roadmap

- Warm-up: Workload analysis
  - FCFS

$$\mathrm{E}[T^{\pi}] \sim \frac{K^{\pi}}{\alpha}$$

$$\mathrm{E}[T^{\mathrm{FCFS}}] \sim \mathrm{E}[W] \sim \frac{w}{\alpha}$$

- A hypothetical OPT policy

$$\mathrm{E}[T_L^{\mathrm{OPT}}] = O(1)$$

$$\mathrm{E}[T^{\mathrm{OPT}}] \sim \frac{\mu_H}{\lambda} \cdot \frac{w}{\alpha}$$

- Preemptive LCFS

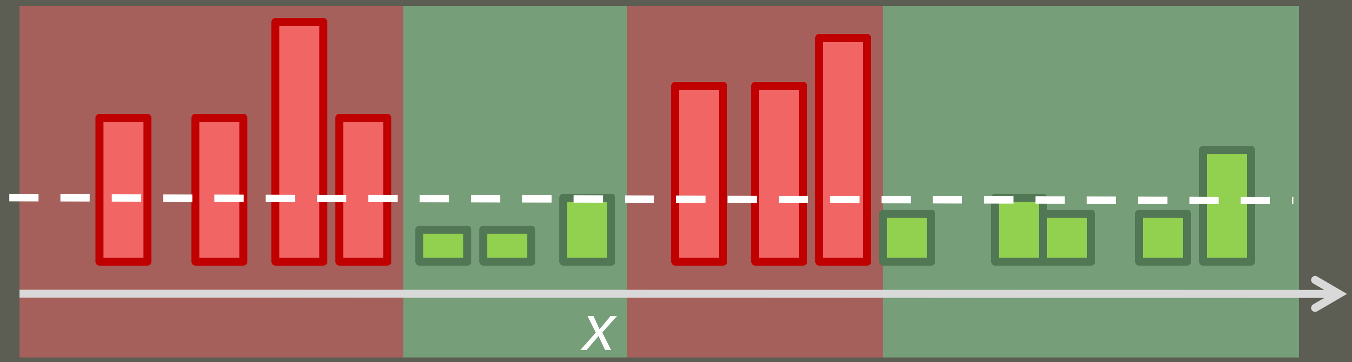
$$\mathrm{E}[T_L^{\mathrm{PLCFS}}] = O(1)$$

PLCFS is asymptotically OPT

- Least-Attained-Service

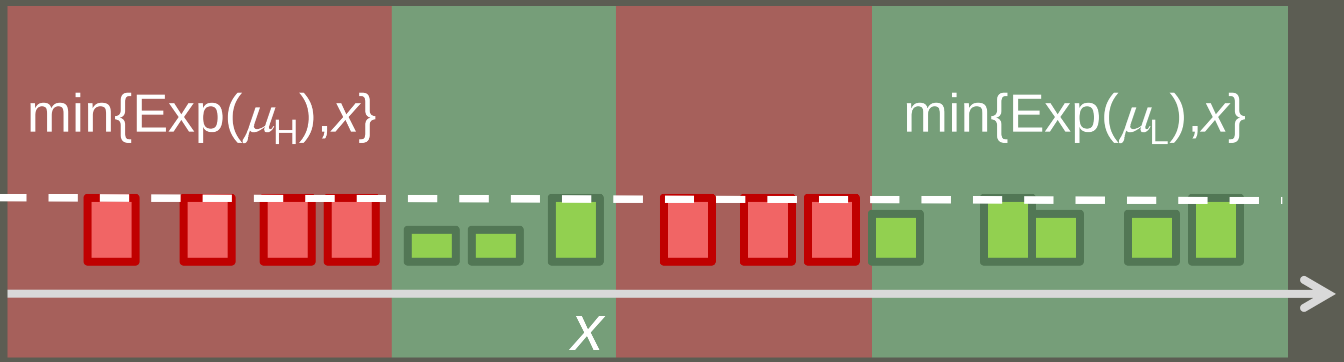
Least-Attained-Service

THEOREM: $\mathrm{E}[T_L^{\mathrm{LAS}}] \sim \frac{\kappa}{\alpha}$ **COR.:** LAS strictly suboptimal



Least-Attained-Service

THEOREM: $\mathrm{E}[T_L^{\mathrm{LAS}}] \sim \frac{\kappa}{\alpha}$ **COR.:** LAS strictly suboptimal



Response time of tagged job of size x = time until empty under modified job sizes

Consider tagged L job of size $x > x^* = \frac{1}{\mu_H}\log\left(\frac{\mu_H}{\lambda - \mu_H}\right)$

⇒ tagged job sees overloaded H states
⇒ Response time of tagged job = $\Omega(1/\alpha)$

Analysis Roadmap

- Warm-up: Workload analysis
  - FCFS

$$\mathrm{E}[T^{\pi}] \sim \frac{K^{\pi}}{\alpha}$$

$$\mathrm{E}[T^{\mathrm{FCFS}}] \sim \mathrm{E}[W] \sim \frac{w}{\alpha}$$

- A hypothetical OPT policy

$$\mathrm{E}[T_L^{\mathrm{OPT}}] = O(1)$$

$$\mathrm{E}[T^{\mathrm{OPT}}] \sim \frac{\mu_H}{\lambda} \cdot \frac{w}{\alpha}$$

- Preemptive LCFS

$$\mathrm{E}[T_L^{\mathrm{PLCFS}}] = O(1)$$

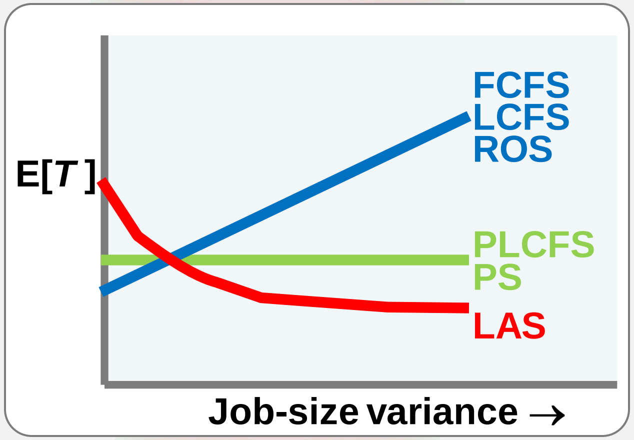
PLCFS is asymptotically OPT

opposite of i.i.d. case

- Least-Attained-Service

LAS is strictly suboptimal

1. Scheduling theory refresher



4. Simulation results

3. Analysis of common scheduling policies

Fluid Workload

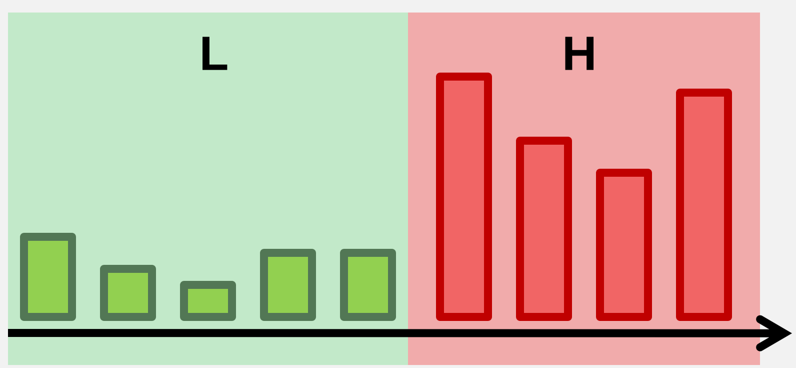
Busy period analysis

ODEs

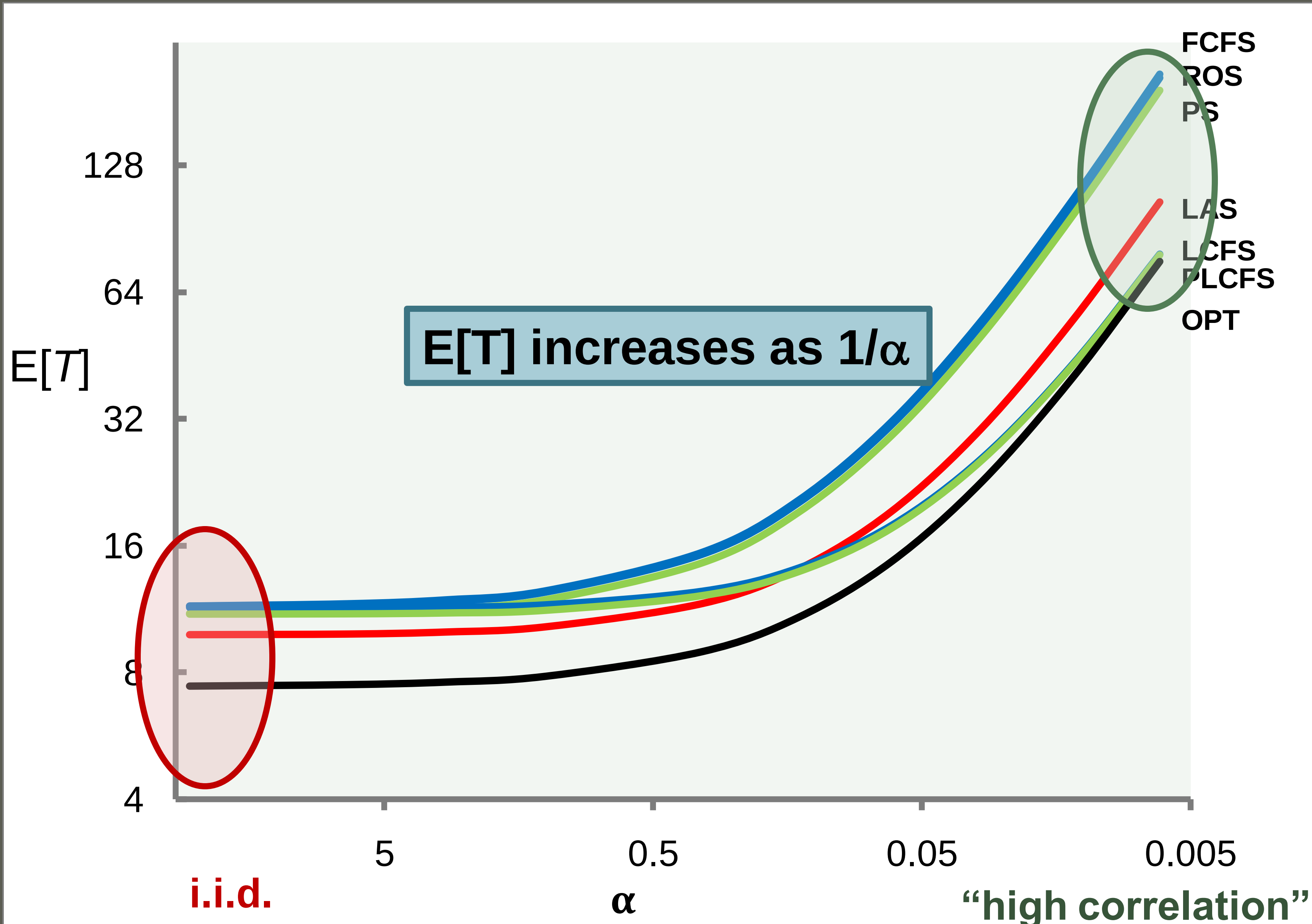
i.i.d.

high correlation

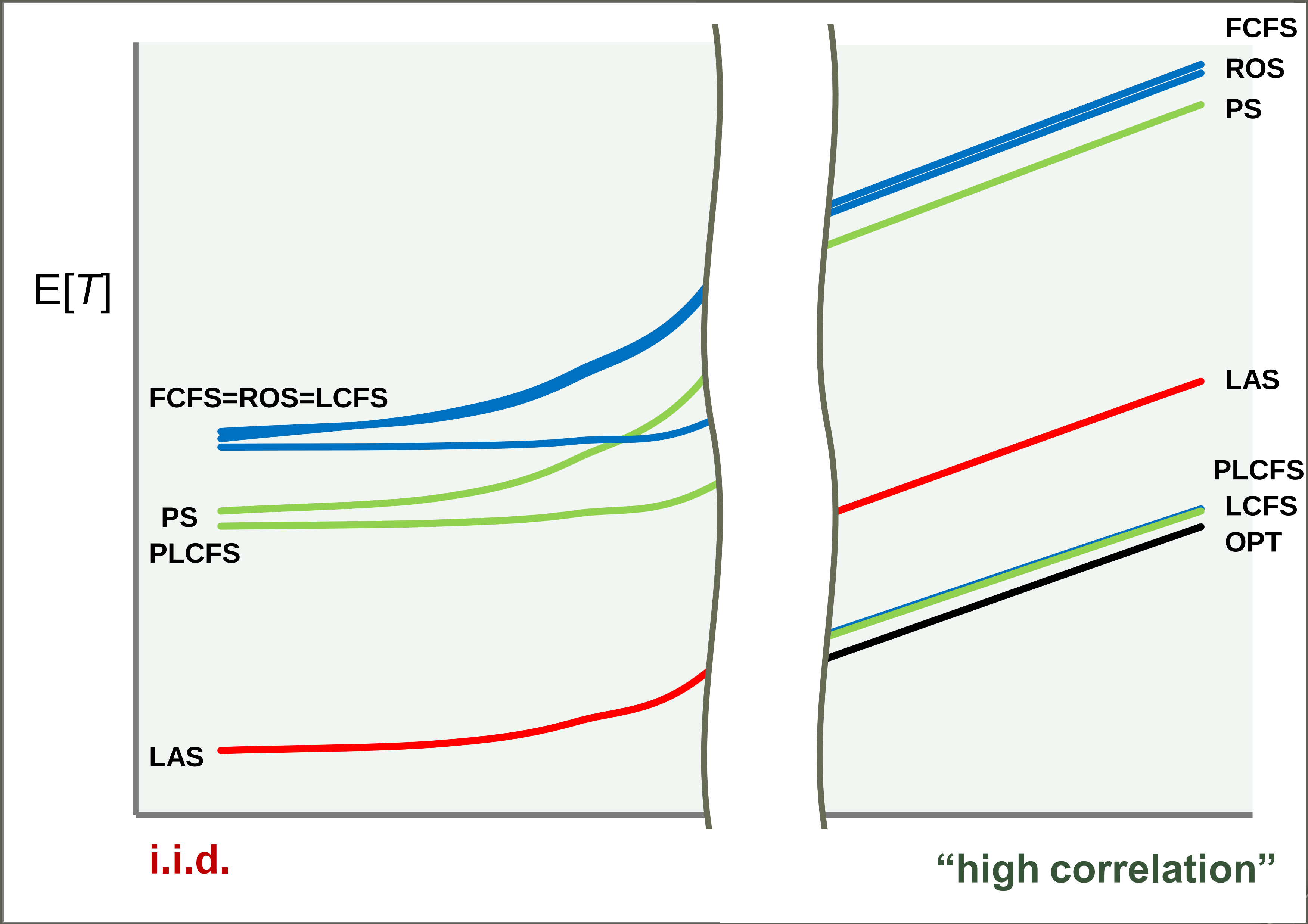
2. MMAP correlation model



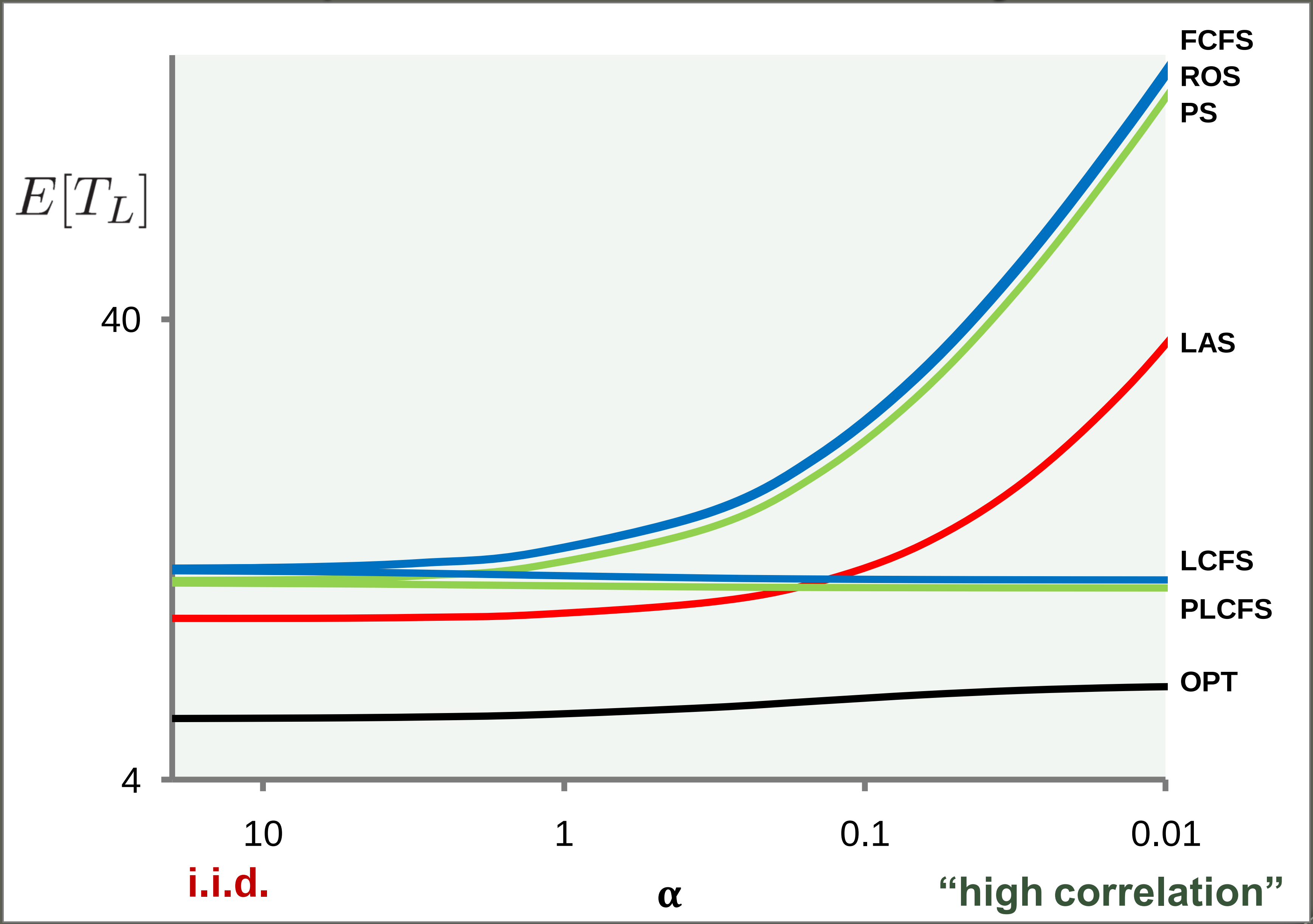
Mean Response time



Mean Response time



Response time of "Little" jobs



Simulation Results

- Order of E[T] changes:
  (i.i.d.) FCFS=ROS=LCFS>PS=P-LCFS>LAS
  $\Downarrow$
  (corr.) FCFS>ROS>PS>LAS>LCFS=P-LCFS
- LCFS, PLCFS overtake LAS very slowly
- LAS or PLCFS are near-optimal for all α

Conclusions

- Qualitative properties of scheduling policies vastly different under *i.i.d.* and correlated job sizes

- Challenges in modeling correlation
  - performance not just a function of auto-correlation function

- Simple size/correlation-oblivious policies