

Control, Optimization, and Diffusion Limits for Queuing Systems

Anat Lev-Ari

PhD Thesis Seminar, July 17

Advisor: Prof. Rami Atar

Department of Electrical Engineering

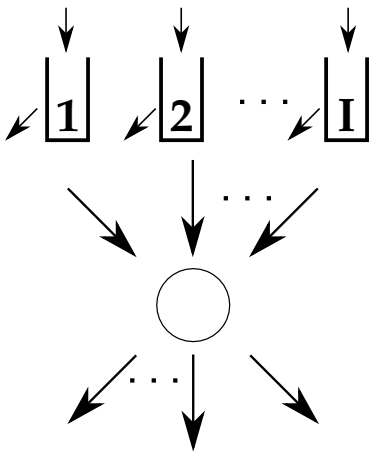


Technion
Israel Institute
of Technology

- 1 The Multiclass Single Server Queue with Reneging
 - The Model and Control problem
 - The Brownian Control Problem (BCP)
 - Solving the BCP
 - The Bellman Equation
 - Optimal Control Derivation
 - Main Result
- 2 The $G/G/1$ Queue with Retrials
 - The Model
 - First Result - Diffusion Limits
 - Second Result - Optimal Buffer Size
 - Simulation
- 3 Acknowledgments

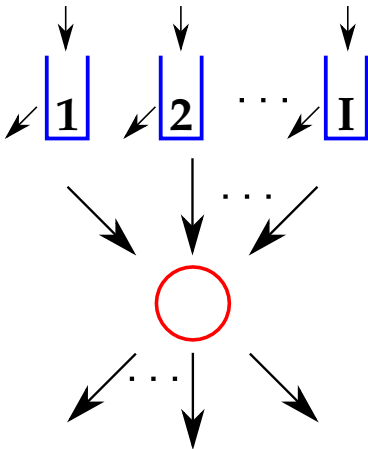
The Multiclass Single Server Queue with Reneging

The Model



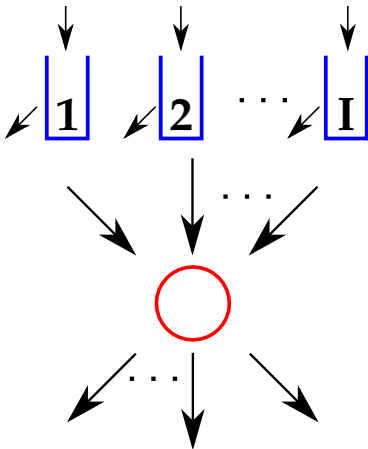
The Model

- / customer classes and
1 server (processor sharing).

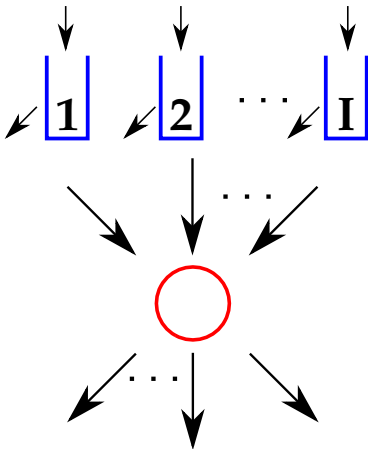


The Model

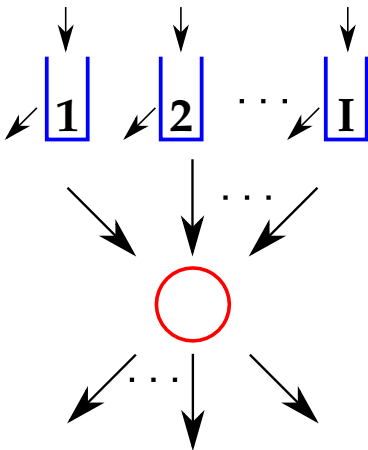
- / customer classes and 1 server (processor sharing).
- Arrivals - renewal processes.



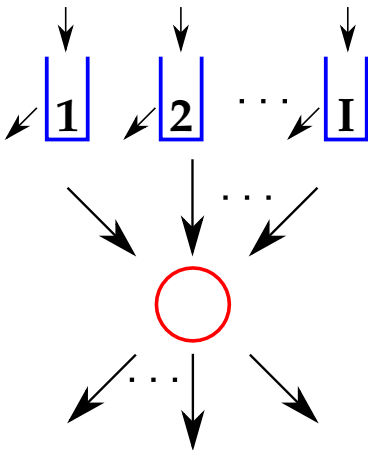
- / customer classes and 1 server (processor sharing).
- Arrivals - renewal processes.
- Service times - i.i.d.



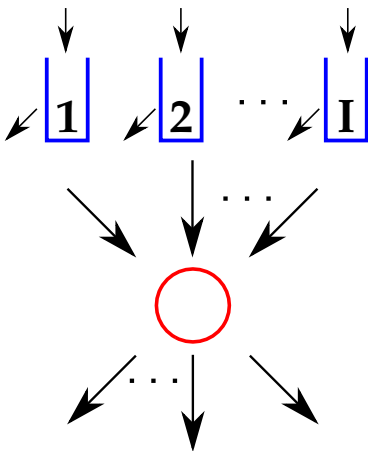
- I customer classes and 1 server (processor sharing).
- Arrivals - renewal processes.
- Service times - i.i.d.
- Abandonments - follow exponential clock.



- I customer classes and 1 server (processor sharing).
- Arrivals - renewal processes.
- Service times - i.i.d.
- Abandonments - follow exponential clock.
- Cost - holding cost + abandonments count.



- I customer classes and 1 server (processor sharing).
- Arrivals - renewal processes.
- Service times - i.i.d.
- Abandonments - follow exponential clock.
- Cost - holding cost + abandonments count.



Problem: Choose who to serve next to minimize the cost

- $X_i(t)$ = the number of customers in the i-th class at time t

$$= \underbrace{X_i(0)}_{\text{initial state}} + \underbrace{A_i(t)}_{\text{Arrivals}} - \underbrace{S_i\left(\int_0^t B_i(s)ds\right)}_{\text{Departures (service)}} - \underbrace{R_i(t)}_{\text{Reneging}}$$

- $X_i(t)$ = the number of customers in the i -th class at time t

$$= \underbrace{X_i(0)}_{\text{initial state}} + \underbrace{A_i(t)}_{\text{Arrivals}} - \underbrace{S_i\left(\int_0^t B_i(s)ds\right)}_{\text{Departures (service)}} - \underbrace{R_i(t)}_{\text{Reneging}}$$

- **Cost:**

$$J(B) = \mathbb{E}\left(\int_0^\infty e^{-\alpha t} b' dR(t)\right) = \alpha \mathbb{E}\left(\int_0^\infty e^{-\alpha t} c' X(t) dt\right)$$

where $\alpha > 0$, b_i reneging cost for class i , c_i holding cost for class i ($c_i = \theta_i b_i$, where θ_i is abandonment rate).

- $X_i(t)$ = the number of customers in the i -th class at time t

$$= \underbrace{X_i(0)}_{\text{initial state}} + \underbrace{A_i(t)}_{\text{Arrivals}} - \underbrace{S_i\left(\int_0^t B_i(s)ds\right)}_{\text{Departures (service)}} - \underbrace{R_i(t)}_{\text{Reneging}}$$

- **Cost:**

$$J(B) = \mathbb{E}\left(\int_0^\infty e^{-\alpha t} b' dR(t)\right) = \alpha \mathbb{E}\left(\int_0^\infty e^{-\alpha t} c' X(t) dt\right)$$

where $\alpha > 0$, b_i reneging cost for class i , c_i holding cost for class i ($c_i = \theta_i b_i$, where θ_i is abandonment rate).

- **Problem:** Find a control process $B(t)$ that minimizes $J(B)$.

- $X_i(t)$ = the number of customers in the i -th class at time t

$$= \underbrace{X_i(0)}_{\text{initial state}} + \underbrace{A_i(t)}_{\text{Arrivals}} - \underbrace{S_i\left(\int_0^t B_i(s)ds\right)}_{\text{Departures (service)}} - \underbrace{R_i(t)}_{\text{Reneging}}$$

- **Cost:**

$$J(B) = \mathbb{E}\left(\int_0^\infty e^{-\alpha t} b' dR(t)\right) = \alpha \mathbb{E}\left(\int_0^\infty e^{-\alpha t} c' X(t) dt\right)$$

where $\alpha > 0$, b_i reneging cost for class i , c_i holding cost for class i ($c_i = \theta_i b_i$, where θ_i is abandonment rate).

- **Problem:** Find a control process $B(t)$ that minimizes $J(B)$.
- **Not solvable analytically.**

- $X_i(t)$ = the number of customers in the i -th class at time t

$$= \underbrace{X_i(0)}_{\text{initial state}} + \underbrace{A_i(t)}_{\text{Arrivals}} - \underbrace{S_i\left(\int_0^t B_i(s)ds\right)}_{\text{Departures (service)}} - \underbrace{R_i(t)}_{\text{Reneging}}$$

- **Cost:**

$$J(B) = \mathbb{E}\left(\int_0^\infty e^{-\alpha t} b' dR(t)\right) = \alpha \mathbb{E}\left(\int_0^\infty e^{-\alpha t} c' X(t) dt\right)$$

where $\alpha > 0$, b_i reneging cost for class i , c_i holding cost for class i ($c_i = \theta_i b_i$, where θ_i is abandonment rate).

- **Problem:** Find a control process $B(t)$ that minimizes $J(B)$.
- **Not solvable analytically.**
- **Approach:** Use diffusion approximation.

c_i - holding cost of class i , μ_i - service rate, θ_i - abandonment rate

c_i - holding cost of class i , μ_i - service rate, θ_i - abandonment rate

- $c\mu$ -rule: Prioritize according to the $c_i\mu_i$ index

c_i - holding cost of class i , μ_i - service rate, θ_i - abandonment rate

- $c\mu$ -rule: Prioritize according to the $c_i\mu_i$ index
⇒ optimal without abandonments (Smith '56)

c_i - holding cost of class i , μ_i - service rate, θ_i - abandonment rate

- $c\mu$ -rule: Prioritize according to the $c_i\mu_i$ index
 - \Rightarrow optimal without abandonments (Smith '56)
 - \Rightarrow asymptotically optimal without abandonments, including an extended (nonlinear) holding cost (special case of van Mieghem '95)

c_i - holding cost of class i , μ_i - service rate, θ_i - abandonment rate

- $c\mu$ -rule: Prioritize according to the $c_i\mu_i$ index
 - \Rightarrow optimal without abandonments (Smith '56)
 - \Rightarrow asymptotically optimal without abandonments, including an extended (nonlinear) holding cost (special case of van Mieghem '95)
- $c\mu/\theta$ -rule: Prioritize according to the $c_i\mu_i/\theta_i$ index

c_i - holding cost of class i , μ_i - service rate, θ_i - abandonment rate

- $c\mu$ -rule: Prioritize according to the $c_i\mu_i$ index
 - \Rightarrow optimal without abandonments (Smith '56)
 - \Rightarrow asymptotically optimal without abandonments, including an extended (nonlinear) holding cost (special case of van Mieghem '95)
- $c\mu/\theta$ -rule: Prioritize according to the $c_i\mu_i/\theta_i$ index
 - \Rightarrow asymptotically optimal with abandonments under the fluid scale (Atar Giat Shimkin 11')

c_i - holding cost of class i , μ_i - service rate, θ_i - abandonment rate

- $c\mu$ -rule: Prioritize according to the $c_i\mu_i$ index
 - \Rightarrow optimal without abandonments (Smith '56)
 - \Rightarrow asymptotically optimal without abandonments, including an extended (nonlinear) holding cost (special case of van Mieghem '95)
- $c\mu/\theta$ -rule: Prioritize according to the $c_i\mu_i/\theta_i$ index
 - \Rightarrow asymptotically optimal with abandonments under the fluid scale (Atar Giat Shimkin 11')

These results DO NOT apply in Diffusion scale.

- ① Samim Ghamami and Amy R. Ward. Dynamic scheduling of a two-server parallel server system with complete resource pooling and reneging in heavy traffic: asymptotic optimality of a two-threshold policy.
Math. Oper. Res., 38(4):761–824, 2013
- ② Jeunghyun Kim and Amy R. Ward. Dynamic scheduling of a $GI/GI/1 + GI$ queue with multiple customer classes.
Queueing Syst., 75(2-4):339–384, 2013
- ③ Barış Ata and Mustafa H. Tongarлак. On scheduling a multiclass queue with abandonments under general delay costs.
Queueing Syst., 74(1):65–104, 2013
- ④ Melanie Rubino and Baris Ata. Dynamic control of a make-to-order, parallel-server system with cancellations.
Operations Research, 57(1):94–108, 2009

- 1 **Define BCP:** consider a sequence of systems, generated from the original system under the diffusion scale, to arrive at a Brownian Control Problem (BCP).

- 1 **Define BCP**: consider a sequence of systems, generated from the original system under the diffusion scale, to arrive at a Brownian Control Problem (BCP).
- 2 **Solve BCP**: compute the corresponding Bellman equation and deduce the control from it.
(The Bellman equation characterizes the value function)

- ① **Define BCP**: consider a sequence of systems, generated from the original system under the diffusion scale, to arrive at a Brownian Control Problem (BCP).
- ② **Solve BCP**: compute the corresponding Bellman equation and deduce the control from it.
(The Bellman equation characterizes the value function)
- ③ **AO**: prove this control is Asymptotically Optimal.

Step 1 - Define BCP

- $X_t = x + \underbrace{W_t}_{\text{BM}} - \int_0^t \underbrace{\Theta}_{\text{diag}(\theta)} X_s ds + \underbrace{Y_t}_{\text{Control}} \in \mathbb{R}_+^I$

- $$X_t = x + \underbrace{W_t}_{\text{BM}} - \int_0^t \underbrace{\Theta}_{\text{diag}(\theta)} X_s ds + \underbrace{Y_t}_{\text{Control}} \in \mathbb{R}_+^I$$

- **Heavy traffic condition:** $\sum_{i=1}^I \lambda_i / \mu_i = 1$

λ, μ are first order approximations for the arrival rate and service rate

- $$X_t = x + \underbrace{W_t}_{\text{BM}} - \int_0^t \underbrace{\Theta}_{\text{diag}(\theta)} X_s ds + \underbrace{Y_t}_{\text{Control}} \in \mathbb{R}_+^I$$

- Heavy traffic condition:** $\sum_{i=1}^I \lambda_i / \mu_i = 1$
 λ, μ are first order approximations for the arrival rate and service rate

- Goal:** To minimize the cost function

$$J(x, Y) = \mathbb{E} \left(\int_0^\infty e^{-\alpha t} \underbrace{c'}_{\text{holding cost}} X_t dt \right).$$

for Y being admissible

Definition

admissible control system with initial condition $x \in \mathbb{R}_+^I$ is a tuple $(\bar{\Sigma}, W_t, Y_t, X_t)$, where $\bar{\Sigma} = (\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P})$ is a filtered probability space, $\{W_t\}$ is an I -dimensional $\bar{\mathcal{F}}_t$ -adapted (y, σ) -BM, the processes $\{X_t\}$ and $\{Y_t\}$ have sample paths in $\mathbb{D}_{\mathbb{R}^I}(\mathbb{R}_+)$ and are $\bar{\mathcal{F}}_t$ -adapted, and the following hold:

- i. For all $t, s \geq 0$, $W_{t+s} - W_t$ is independent of $\bar{\mathcal{F}}_t$ under \bar{P} ,
- ii. X_t satisfies $X_t \in \mathbb{R}_+^I$ for all t \bar{P} -a.s.,
- iii. The process $m'Y$, where $m = (1/\mu_1, \dots, 1/\mu_I)$, is non-negative and non-decreasing.

Definition

admissible control system with initial condition $x \in \mathbb{R}_+^I$ is a tuple $(\bar{\Sigma}, W_t, Y_t, X_t)$, where $\bar{\Sigma} = (\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P})$ is a filtered probability space, $\{W_t\}$ is an I -dimensional $\bar{\mathcal{F}}_t$ -adapted (y, σ) -BM, the processes $\{X_t\}$ and $\{Y_t\}$ have sample paths in $\mathbb{D}_{\mathbb{R}^I}(\mathbb{R}_+)$ and are $\bar{\mathcal{F}}_t$ -adapted, and the following hold:

- i. For all $t, s \geq 0$, $W_{t+s} - W_t$ is independent of $\bar{\mathcal{F}}_t$ under \bar{P} ,
- ii. X_t satisfies $X_t \in \mathbb{R}_+^I$ for all t \bar{P} -a.s.,
- iii. The process $m'Y$, where $m = (1/\mu_1, \dots, 1/\mu_I)$, is non-negative and non-decreasing.

Step 1 - Define BCP

Admissible Control

Definition

admissible control system with initial condition $x \in \mathbb{R}_+^I$ is a tuple $(\bar{\Sigma}, W_t, Y_t, X_t)$, where $\bar{\Sigma} = (\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{P})$ is a filtered probability space, $\{W_t\}$ is an I -dimensional $\bar{\mathcal{F}}_t$ -adapted (y, σ) -BM, the processes $\{X_t\}$ and $\{Y_t\}$ have sample paths in $\mathbb{D}_{\mathbb{R}^I}(\mathbb{R}_+)$ and are $\bar{\mathcal{F}}_t$ -adapted, and the following hold:

- i. For all $t, s \geq 0$, $W_{t+s} - W_t$ is independent of $\bar{\mathcal{F}}_t$ under \bar{P} ,
- ii. X_t satisfies $X_t \in \mathbb{R}_+^I$ for all t \bar{P} -a.s.,
- iii. The process $m'Y$, where $m = (1/\mu_1, \dots, 1/\mu_I)$, is non-negative and non-decreasing.

- $\mathcal{A}(x)$ the set of all admissible control systems with initial condition $x \in \mathbb{R}_+^I$.
- **The value function**

$$V(x) = \inf_{\mathcal{A}(x)} J(x, Y), \quad x \in \mathbb{R}_+^I.$$

Step 2- Solve BCP

The Bellman Equation

l -th dimensional equation.

Step 2- Solve BCP

The Bellman Equation

I -th dimensional equation.

- **Problem:** hard to solve.

Step 2- Solve BCP

The Bellman Equation

I -th dimensional equation.

- **Problem:** hard to solve.
- **Solution:** Reduction to 1-dimension - Reduced Brownian Control Problem (RBCP).
Often occurs and is called State Space Collapse property (SSC)

Step 2- Solve BCP

The Bellman Equation

I -th dimensional equation.

- **Problem:** hard to solve.
- **Solution:** Reduction to 1-dimension - Reduced Brownian Control Problem (RBCP).
Often occurs and is called State Space Collapse property (SSC)

The two problems are equivalent

Step 2- Solve BCP

The RBCP

- Projection to the workload vector $m = (1/\mu_1, \dots, 1/\mu_l)$.

Step 2- Solve BCP

The RBCP

- Projection to the workload vector $m = (1/\mu_1, \dots, 1/\mu_I)$.
- **Workload process**

$$\tilde{X}_t = m'X(t) = \tilde{x} + \tilde{W}_t - \int_0^t \theta' U_s \tilde{X}_s ds + \tilde{Y}_t.$$

- Projection to the workload vector $m = (1/\mu_1, \dots, 1/\mu_I)$.
- **Workload process**

$$\tilde{X}_t = m'X(t) = \tilde{x} + \tilde{W}_t - \int_0^t \theta' U_s \tilde{X}_s ds + \tilde{Y}_t.$$

- **Control:** the pair (U, \tilde{Y}) .

- Projection to the workload vector $m = (1/\mu_1, \dots, 1/\mu_I)$.
- **Workload process**

$$\tilde{X}_t = m'X(t) = \tilde{x} + \tilde{W}_t - \int_0^t \theta' U_s \tilde{X}_s ds + \tilde{Y}_t.$$

- **Control:** the pair (U, \tilde{Y}) .
 - $\tilde{Y} = m'Y$

- Projection to the workload vector $m = (1/\mu_1, \dots, 1/\mu_I)$.
- **Workload process**

$$\tilde{X}_t = m'X(t) = \tilde{x} + \tilde{W}_t - \int_0^t \theta' U_s \tilde{X}_s ds + \tilde{Y}_t.$$

- **Control:** the pair (U, \tilde{Y}) .
 - $\tilde{Y} = m'Y$
 - U is an I-dimensional process corresponds to the division of the workload between the classes $\rightarrow U_i$ is the fraction of workload dedicated to class i .

- Projection to the workload vector $m = (1/\mu_1, \dots, 1/\mu_I)$.
- **Workload process**

$$\tilde{X}_t = m'X(t) = \tilde{x} + \tilde{W}_t - \int_0^t \theta' U_s \tilde{X}_s ds + \tilde{Y}_t.$$

- **Control:** the pair (U, \tilde{Y}) .
 - $\tilde{Y} = m'Y$
 - U is an I-dimensional process corresponds to the division of the workload between the classes $\rightarrow U_i$ is the fraction of workload dedicated to class i .
- **Cost:**

$$\tilde{J}(\tilde{x}, (U, \tilde{Y})) = \mathbb{E}\left(\int_0^\infty e^{-\alpha t} q' U_t \tilde{X}_t dt\right).$$

Definition

An *admissible control system* with initial condition $\tilde{x} \in \mathbb{R}_+$ is a tuple $(\tilde{\Sigma}, \tilde{W}, U, \tilde{Y}, \tilde{X})$, where $\tilde{\Sigma} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ is a filtered probability space, $\{\tilde{W}_t\}$ is a 1-dimensional $\tilde{\mathcal{F}}_t$ -adapted $(\tilde{y}, \tilde{\sigma})$ -BM, and the processes $\{U_t\}$, $\{\tilde{Y}_t\}$, $\{\tilde{X}_t\}$ have sample paths in $\mathbb{D}_{\mathcal{S}_1}(\mathbb{R}_+)$, $\mathbb{D}_{\mathbb{R}'}(\mathbb{R}_+)$ and $\mathbb{D}_{\mathbb{R}}(\mathbb{R}_+)$, resp., are $\tilde{\mathcal{F}}_t$ -adapted and the following hold:

- i. For all $t, s \geq 0$, $\tilde{W}_{t+s} - \tilde{W}_t$ is independent of $\tilde{\mathcal{F}}_t$ under \tilde{P} ,
- ii. \tilde{X}_t satisfies $\tilde{X}_t \geq 0$ for all t \tilde{P} -a.s.,
- iii. The process \tilde{Y} is non-negative and non-decreasing.

Definition

An *admissible control system* with initial condition $\tilde{x} \in \mathbb{R}_+$ is a tuple $(\tilde{\Sigma}, \tilde{W}, U, \tilde{Y}, \tilde{X})$, where $\tilde{\Sigma} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ is a filtered probability space, $\{\tilde{W}_t\}$ is a 1-dimensional $\tilde{\mathcal{F}}_t$ -adapted $(\tilde{y}, \tilde{\sigma})$ -BM, and the processes $\{U_t\}$, $\{\tilde{Y}_t\}$, $\{\tilde{X}_t\}$ have sample paths in $\mathbb{D}_{\mathcal{S}_1}(\mathbb{R}_+)$, $\mathbb{D}_{\mathbb{R}'}(\mathbb{R}_+)$ and $\mathbb{D}_{\mathbb{R}}(\mathbb{R}_+)$, resp., are $\tilde{\mathcal{F}}_t$ -adapted and the following hold:

- i. For all $t, s \geq 0$, $\tilde{W}_{t+s} - \tilde{W}_t$ is independent of $\tilde{\mathcal{F}}_t$ under \tilde{P} ,
- ii. \tilde{X}_t satisfies $\tilde{X}_t \geq 0$ for all t \tilde{P} -a.s.,
- iii. The process \tilde{Y} is non-negative and non-decreasing.

Definition

An *admissible control system* with initial condition $\tilde{x} \in \mathbb{R}_+$ is a tuple $(\tilde{\Sigma}, \tilde{W}, U, \tilde{Y}, \tilde{X})$, where $\tilde{\Sigma} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{P})$ is a filtered probability space, $\{\tilde{W}_t\}$ is a 1-dimensional $\tilde{\mathcal{F}}_t$ -adapted $(\tilde{y}, \tilde{\sigma})$ -BM, and the processes $\{U_t\}$, $\{\tilde{Y}_t\}$, $\{\tilde{X}_t\}$ have sample paths in $\mathbb{D}_{S^1}(\mathbb{R}_+)$, $\mathbb{D}_{\mathbb{R}'}(\mathbb{R}_+)$ and $\mathbb{D}_{\mathbb{R}}(\mathbb{R}_+)$, resp., are $\tilde{\mathcal{F}}_t$ -adapted and the following hold:

- i. For all $t, s \geq 0$, $\tilde{W}_{t+s} - \tilde{W}_t$ is independent of $\tilde{\mathcal{F}}_t$ under \tilde{P} ,
- ii. \tilde{X}_t satisfies $\tilde{X}_t \geq 0$ for all t \tilde{P} -a.s.,
- iii. The process \tilde{Y} is non-negative and non-decreasing.

- $\tilde{\mathcal{A}}(\tilde{x})$ the set of all the admissible controls for the initial condition \tilde{x} .
- **Value function:**

$$\tilde{V}(\tilde{x}) = \inf_{(U, \tilde{Y}) \in \tilde{\mathcal{A}}(\tilde{x})} \tilde{J}(\tilde{x}, (U, \tilde{Y})), \quad \tilde{x} \in \mathbb{R}_+.$$

Step 2- Solve BCP

The Bellman Equation

$$\left\{ \begin{array}{ll} -\frac{\tilde{\sigma}^2}{2} \frac{d^2 v}{dx^2} - \tilde{y} \frac{dv}{dx} + x F\left(\frac{dv}{dx}\right) + \alpha v = 0, & 0 < x < \infty, \\ \frac{dv}{dx}(0) = 0, \quad |v(x)| \leq C(1+x)^C, & x \in [0, \infty). \end{array} \right.$$

Step 2- Solve BCP

The Bellman Equation

$$\begin{cases} -\frac{\tilde{\sigma}^2}{2} \frac{d^2 v}{dx^2} - \tilde{y} \frac{dv}{dx} + x F\left(\frac{dv}{dx}\right) + \alpha v = 0, & 0 < x < \infty, \\ \frac{dv}{dx}(0) = 0, \quad |v(x)| \leq C(1+x)^C, & x \in [0, \infty). \end{cases}$$

where

- $F(y) = \max_u g(u, y)$, u corresponds to the original control.

Step 2- Solve BCP

The Bellman Equation

$$\begin{cases} -\frac{\tilde{\sigma}^2}{2} \frac{d^2 v}{dx^2} - \tilde{y} \frac{dv}{dx} + x F\left(\frac{dv}{dx}\right) + \alpha v = 0, & 0 < x < \infty, \\ \frac{dv}{dx}(0) = 0, \quad |v(x)| \leq C(1+x)^C, & x \in [0, \infty). \end{cases}$$

where

- $F(y) = \max_u g(u, y)$, u corresponds to the original control.
- \tilde{y} , $\tilde{\sigma}$ - system parameters.

Step 2- Solve BCP

The Bellman Equation

$$\begin{cases} -\frac{\tilde{\sigma}^2}{2} \frac{d^2 v}{dx^2} - \tilde{y} \frac{dv}{dx} + x F\left(\frac{dv}{dx}\right) + \alpha v = 0, & 0 < x < \infty, \\ \frac{dv}{dx}(0) = 0, \quad |v(x)| \leq C(1+x)^C, & x \in [0, \infty). \end{cases}$$

where

- $F(y) = \max_u g(u, y)$, u corresponds to the original control.
- \tilde{y} , $\tilde{\sigma}$ - system parameters.
- The Bellman equation admits a unique classical solution - the value function \tilde{V} .

Step 2- Solve BCP

The Bellman Equation

$$\begin{cases} -\frac{\tilde{\sigma}^2}{2} \frac{d^2 v}{dx^2} - \tilde{y} \frac{dv}{dx} + x F\left(\frac{dv}{dx}\right) + \alpha v = 0, & 0 < x < \infty, \\ \frac{dv}{dx}(0) = 0, \quad |v(x)| \leq C(1+x)^C, & x \in [0, \infty). \end{cases}$$

where

- $F(y) = \max_u g(u, y)$, u corresponds to the original control.
- \tilde{y} , $\tilde{\sigma}$ - system parameters.
- The Bellman equation admits a unique classical solution - the value function \tilde{V} .
- The optimal control is deduced from finding the maximum in $F(y) = \max_u g(u, y)$.

Step 2- Solve BCP

The Bellman Equation

$$\begin{cases} -\frac{\tilde{\sigma}^2}{2} \frac{d^2 v}{dx^2} - \tilde{y} \frac{dv}{dx} + x F\left(\frac{dv}{dx}\right) + \alpha v = 0, & 0 < x < \infty, \\ \frac{dv}{dx}(0) = 0, \quad |v(x)| \leq C(1+x)^C, & x \in [0, \infty). \end{cases}$$

where

- $F(y) = \max_u g(u, y)$, u corresponds to the original control.
- \tilde{y} , $\tilde{\sigma}$ - system parameters.
- The Bellman equation admits a unique classical solution - the value function \tilde{V} .
- The optimal control is deduced from finding the maximum in $F(y) = \max_u g(u, y)$.
- Can be solved numerically (equation in one variable).

Step 2- Solve BCP

Optimal Control Derivation

- $g(u, y) = \theta \cdot uy - q \cdot u$

Step 2- Solve BCP

Optimal Control Derivation

- $g(u, y) = \theta \cdot uy - q \cdot u$
- $u \in \{a \in \mathbb{R}_+^I \mid \sum_i a_i = 1\}$

- $g(u, y) = \theta \cdot uy - q \cdot u$
- $u \in \{a \in \mathbb{R}_+^I \mid \sum_i a_i = 1\}$

\Rightarrow The maximal solution is one of the extreme points:

$u \in \{e_1, \dots, e_I\}$ - the standard basis

- $g(u, y) = \theta \cdot uy - q \cdot u$
- $u \in \{a \in \mathbb{R}_+^I \mid \sum_i a_i = 1\}$

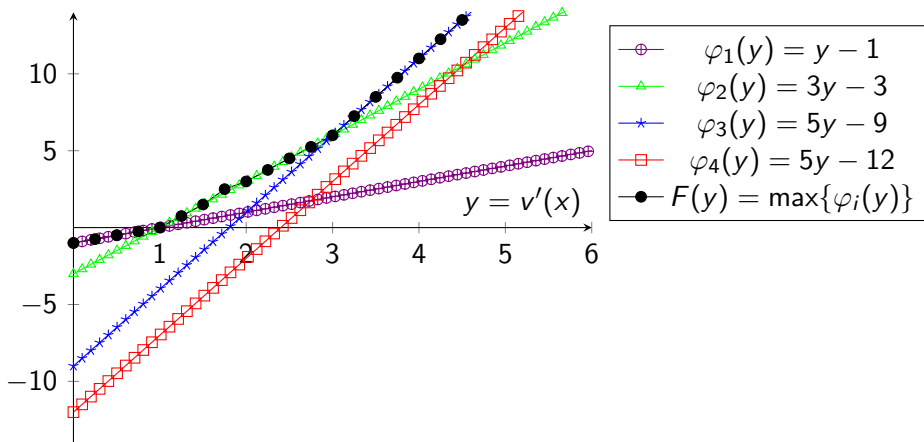
\Rightarrow The maximal solution is one of the extreme points:

$u \in \{e_1, \dots, e_I\}$ - the standard basis

$$\Rightarrow F(y) = \max_u g(u, y) = \max_i \varphi_i(y), \quad \varphi_i(y) = \theta_i y - c_i \mu_i, \quad y \in \mathbb{R}_+$$

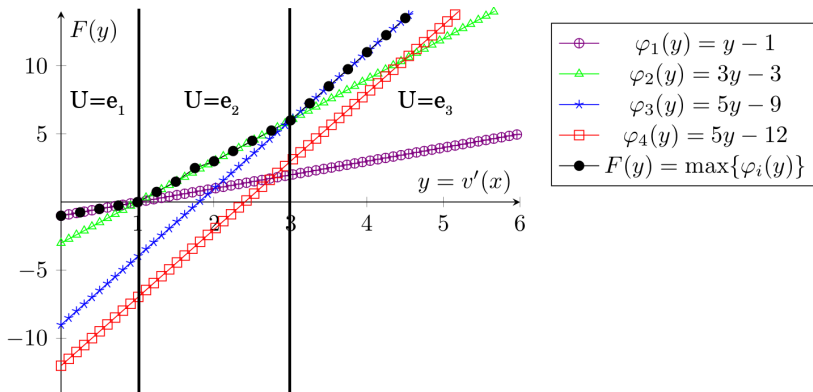
Step 2- Solve BCP

Optimal Control Derivation - Example



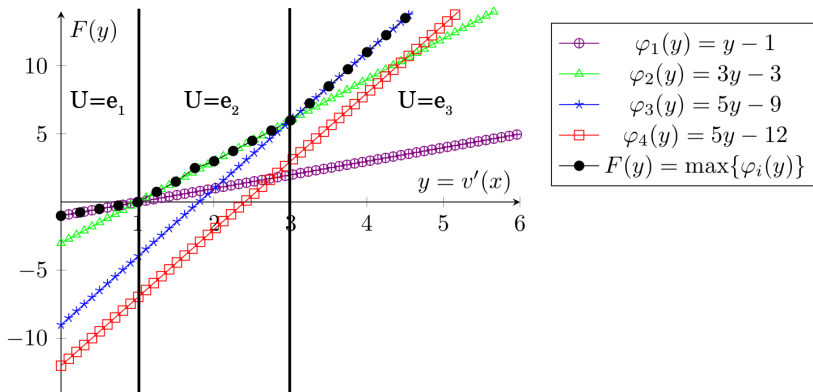
Step 2- Solve BCP

Optimal Control Derivation - Example



Step 2- Solve BCP

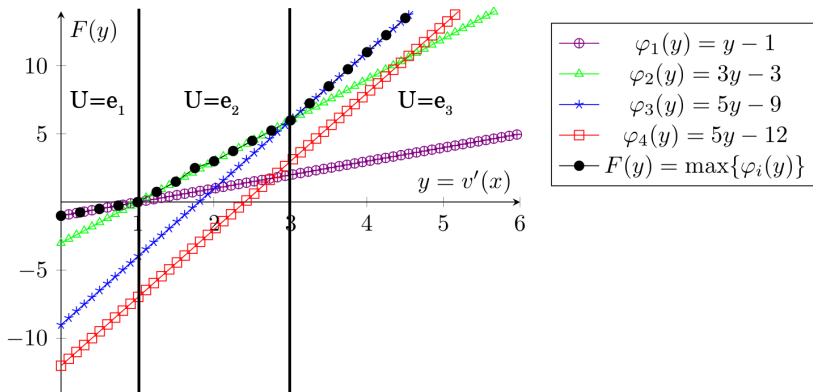
Optimal Control Derivation - Example



- 3 different intervals for v' : $[0, 1)$, $[1, 3)$, $[3, \infty)$

Step 2- Solve BCP

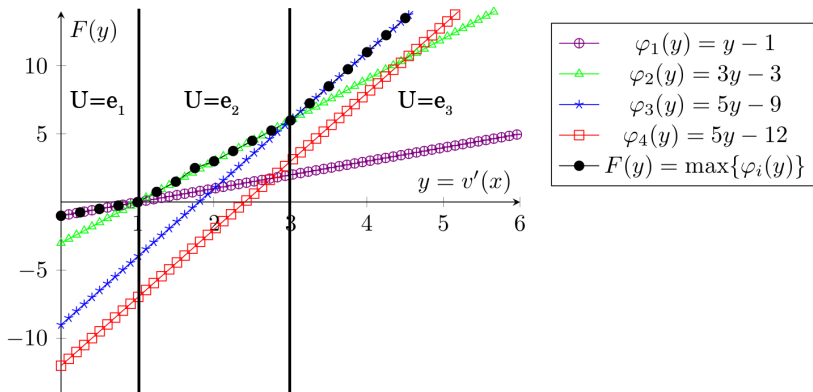
Optimal Control Derivation - Example



- 3 different intervals for v' : $[0, 1)$, $[1, 3)$, $[3, \infty)$
- On each interval the control process satisfy $U = e_i$

Step 2- Solve BCP

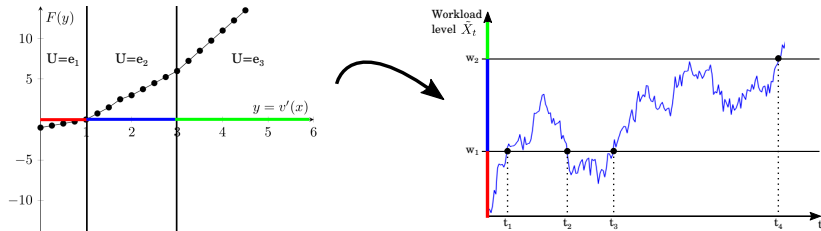
Optimal Control Derivation - Example



- 3 different intervals for v' : $[0, 1)$, $[1, 3)$, $[3, \infty)$
- On each interval the control process satisfy $U = e_i$
→ all the workload remains in the i -th class.

Step 2- Solve BCP

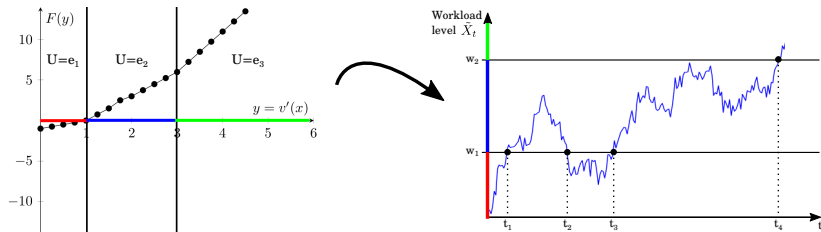
Optimal Control Derivation - Example



- v is nondecreasing and convex

Step 2- Solve BCP

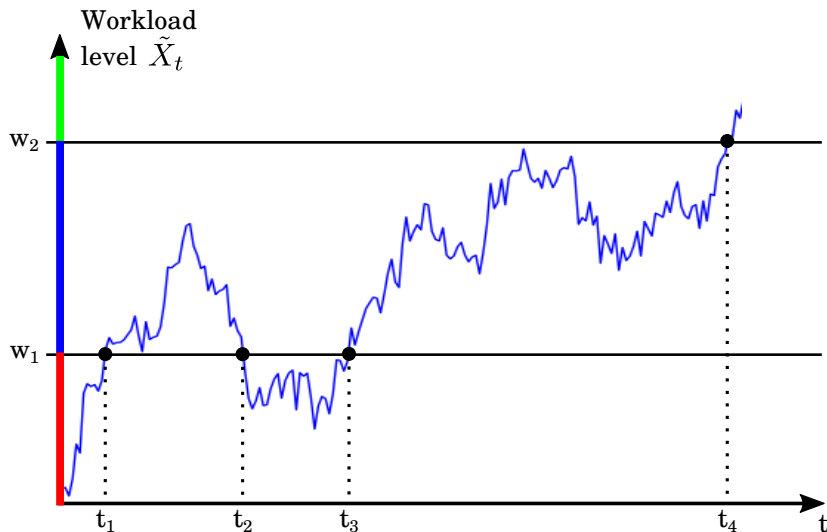
Optimal Control Derivation - Example



- v is nondecreasing and convex
- The intervals for v' correspond to interval for \tilde{X} - $[0, w_1), [w_1, w_2), [w_2, \infty)$

Step 2- Solve BCP

Optimal Control Derivation - Example



Ex. Below w_1 ($[0, t_1)$, $[t_2, t_3), \dots$): $U = e_1 \rightarrow$ least priority to class 1.

Step 2- Solve BCP

The Optimal Control $B^{n,*}$ - Definition

\mathcal{K} is a minimal subset of \mathcal{I} such that $\max_{k \in \mathcal{K}} \varphi_k(y) = \max_{i \in \mathcal{I}} \varphi_i(y)$

Step 2- Solve BCP

The Optimal Control $B^{n,*}$ - Definition

\mathcal{K} is a minimal subset of \mathcal{I} such that $\max_{k \in \mathcal{K}} \varphi_k(y) = \max_{i \in \mathcal{I}} \varphi_i(y)$

Assumptions:

① $c_1 \mu_1 \leq c_2 \mu_2 \leq \dots \leq c_K \mu_K.$

Step 2- Solve BCP

The Optimal Control $B^{n,*}$ - Definition

\mathcal{K} is a minimal subset of \mathcal{I} such that $\max_{k \in \mathcal{K}} \varphi_k(y) = \max_{i \in \mathcal{I}} \varphi_i(y)$

Assumptions:

- 1 $c_1 \mu_1 \leq c_2 \mu_2 \leq \dots \leq c_K \mu_K.$
- 2 $\{L_k\}_{k \in \mathcal{K}}$ a partition of $[0, \infty).$

Step 2- Solve BCP

The Optimal Control $B^{n,*}$ - Definition

\mathcal{K} is a minimal subset of \mathcal{I} such that $\max_{k \in \mathcal{K}} \varphi_k(y) = \max_{i \in \mathcal{I}} \varphi_i(y)$

Assumptions:

- 1 $c_1 \mu_1 \leq c_2 \mu_2 \leq \dots \leq c_K \mu_K$.
- 2 $\{L_k\}_{k \in \mathcal{K}}$ a partition of $[0, \infty)$.

The optimal policy:

Step 2- Solve BCP

The Optimal Control $B^{n,*}$ - Definition

\mathcal{K} is a minimal subset of \mathcal{I} such that $\max_{k \in \mathcal{K}} \varphi_k(y) = \max_{i \in \mathcal{I}} \varphi_i(y)$

Assumptions:

- ① $c_1 \mu_1 \leq c_2 \mu_2 \leq \dots \leq c_K \mu_K$.
- ② $\{L_k\}_{k \in \mathcal{K}}$ a partition of $[0, \infty)$.

The optimal policy:

- ① Non-interruptible, non-idling policy that allows no processor sharing.

Step 2- Solve BCP

The Optimal Control $B^{n,*}$ - Definition

\mathcal{K} is a minimal subset of \mathcal{I} such that $\max_{k \in \mathcal{K}} \varphi_k(y) = \max_{i \in \mathcal{I}} \varphi_i(y)$

Assumptions:

- 1 $c_1\mu_1 \leq c_2\mu_2 \leq \dots \leq c_K\mu_K$.
- 2 $\{L_k\}_{k \in \mathcal{K}}$ a partition of $[0, \infty)$.

The optimal policy:

- 1 Non-interruptible, non-idling policy that allows no processor sharing.
- 2 At time t , the server becomes available and some of the buffers are non-empty, the current value of the scaled workload \tilde{X}_t^n is computed and the **least priority class** k is set to be the unique i for which $\tilde{X}_t^n \in L_i$.

Step 2- Solve BCP

The Optimal Control $B^{n,*}$ - Definition

\mathcal{K} is a minimal subset of \mathcal{I} such that $\max_{k \in \mathcal{K}} \varphi_k(y) = \max_{i \in \mathcal{I}} \varphi_i(y)$

Assumptions:

- 1 $c_1\mu_1 \leq c_2\mu_2 \leq \dots \leq c_K\mu_K$.
- 2 $\{L_k\}_{k \in \mathcal{K}}$ a partition of $[0, \infty)$.

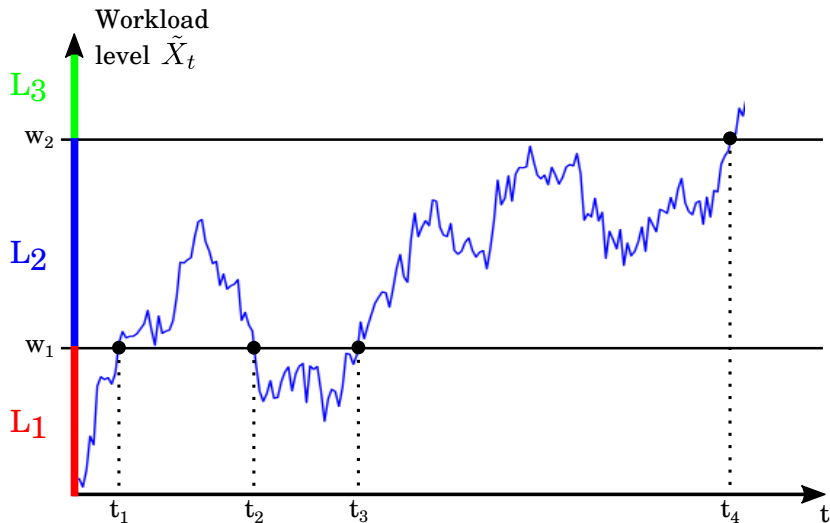
The optimal policy:

- 1 Non-interruptible, non-idling policy that allows no processor sharing.
- 2 At time t , the server becomes available and some of the buffers are non-empty, the current value of the scaled workload \tilde{X}_t^n is computed and the **least priority class** k is set to be the unique i for which $\tilde{X}_t^n \in L_i$.
- 3 The customer to be served is then picked according to the ordering

$$(k+1, k+2, \dots, l, 1, 2, \dots, k).$$

Step 2- Solve BCP

The Optimal Control - Example



$$\mathcal{I} = \{1, 2, 3, 4\}, \mathcal{K} = \{1, 2, 3\}, c_1\mu_1 \leq c_2\mu_2 \leq c_3\mu_3$$

Step 2- Solve BCP

The Optimal Control

- Dynamic index rule. **Not static index rule as $c\mu$ or $c\mu/\theta$.**

Step 2- Solve BCP

The Optimal Control

- Dynamic index rule. **Not static index rule as $c\mu$ or $c\mu/\theta$.**
- When workload is **low** the server gives least priority to class with lowest $c\mu$ index.

Step 2- Solve BCP

The Optimal Control

- Dynamic index rule. **Not static index rule as $c\mu$ or $c\mu/\theta$.**
- When workload is **low** the server gives least priority to class with lowest $c\mu$ index.
- When workload is **high** the server gives least priority to class with highest θ index.

Theorem

Let v denote the solution of the Bellman equation. Then

i. The limit value is determined by the function v . In particular,

$$\lim_{n \rightarrow \infty} \hat{V}^n = v(m'x_0).$$

ii. The control process $B^{n,*}$ described above is AO, that is,

$$\lim_{n \rightarrow \infty} \hat{J}^n(B^{n,*}) = v(m'x_0).$$

Step 3 - AO

Main Result - Proof Steps

Two parts:

Two parts:

① **Lower bound:**

For any sequence of admissible controls

$$\liminf_{n \rightarrow \infty} \hat{J}^n(B^n) \geq v(m'x_0).$$

Two parts:

① **Lower bound:**

For any sequence of admissible controls

$$\liminf_{n \rightarrow \infty} \hat{J}^n(B^n) \geq v(m'x_0).$$

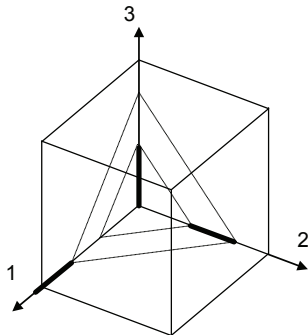
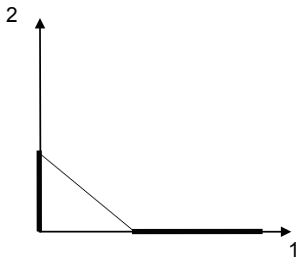
② **Upper bound:**

Under the control process $B^{n,*}$ described earlier

$$\limsup_{n \rightarrow \infty} \hat{J}^n(B^{n,*}) \leq v(m'x_0).$$

Step 3 - AO

Why The Process X is Discussions?



Step 3 - AO

Main Result - Proof of the Upper Bound

- **Problem:** The process X is discontinuous with $K - 1$ discontinuity points w_1, \dots, w_{K-1} .

Step 3 - AO

Main Result - Proof of the Upper Bound

- **Problem:** The process X is discontinuous with $K - 1$ discontinuity points w_1, \dots, w_{K-1} .
- **Solution:**

Step 3 - AO

Main Result - Proof of the Upper Bound

- **Problem:** The process X is discontinuous with $K - 1$ discontinuity points w_1, \dots, w_{K-1} .
- **Solution:**
 - 1 Show that in the interior of each interval $[w_i, w_{i+1})$ where the policy is fixed and gives least priority to class i , all the other classes (not the i -th class) empties. State Space Collapse property (SSC).

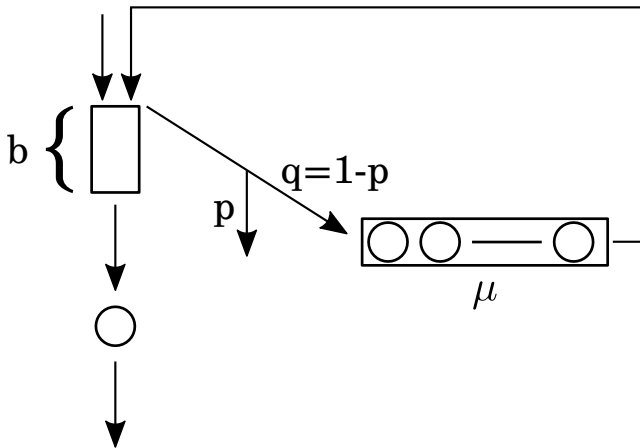
Step 3 - AO

Main Result - Proof of the Upper Bound

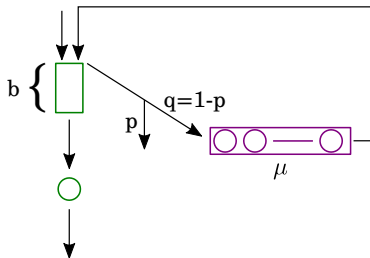
- **Problem:** The process X is discontinuous with $K - 1$ discontinuity points w_1, \dots, w_{K-1} .
- **Solution:**
 - 1 Show that in the interior of each interval $[w_i, w_{i+1})$ where the policy is fixed and gives least priority to class i , all the other classes (not the i -th class) empties. State Space Collapse property (SSC).
 - 2 Show that the time spent around the discontinuity points w_i is small \rightarrow can be neglected.

The $G/G/1$ Queue with Retrials

The Model



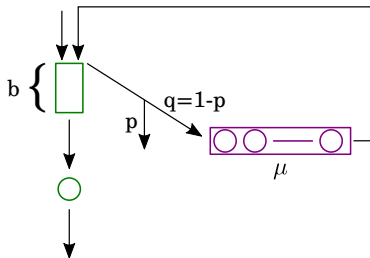
2 **stations**: **main** and **retrial**.



2 **stations:** **main** and **retrial**.

- **Main station:**

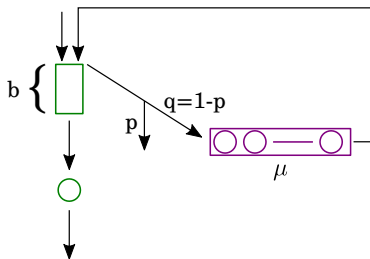
- $G/G/1$ queue, finite buffer size b .



2 **stations:** **main** and **retrial**.

- **Main station:**

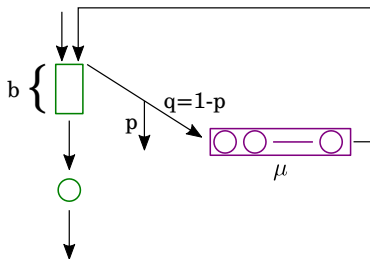
- $G/G/1$ queue, finite buffer size b .
- Reject incoming customers when buffer full.



2 **stations:** **main** and **retrial**.

- **Main station:**

- $G/G/1$ queue, finite buffer size b .
- Reject incoming customers when buffer full.
- Each rejected customer move to the retrial station w.p. q .



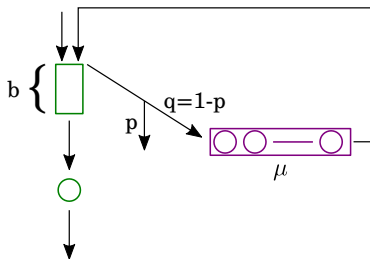
2 stations: main and retrial.

- Main station:

- $G/G/1$ queue, finite buffer size b .
- Reject incoming customers when buffer full.
- Each rejected customer move to the retrial station w.p. q .

- Retrial station:

- Infinite number of exponential servers.



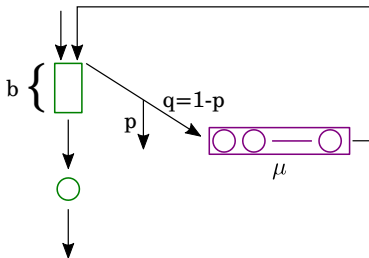
2 stations: main and retrial.

- Main station:

- $G/G/1$ queue, finite buffer size b .
- Reject incoming customers when buffer full.
- Each rejected customer move to the retrial station w.p. q .

- Retrial station:

- Infinite number of exponential servers.
- Service rate μ .



The Model

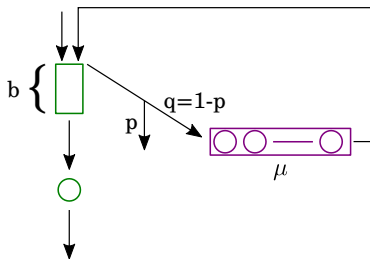
2 **stations:** **main** and **retrial**.

- **Main station:**

- $G/G/1$ queue, finite buffer size b .
- Reject incoming customers when buffer full.
- Each rejected customer move to the retrial station w.p. q .

- **Retrial station:**

- Infinite number of exponential servers.
- Service rate μ .



- **First Problem:** Diffusion limits.

- **Second Problem:** Optimizing b with cost

$$\mathbb{E} \int_0^\infty e^{-\alpha t} \left(c_1 \underbrace{X(t)}_{\text{system state}} dt + c_2 d \underbrace{C(t)}_{\text{rejections count}} \right)$$

- ① Avi Mandelbaum, William A Massey, and Martin I Reiman. [Strong approximations for markovian service networks.](#)
Queueing Systems, 30(1):149–201, 1998
- ② GI Falin and JR Artalejo. [Approximations for multiserver queues with balking/retrial discipline.](#)
OR Spectrum, 17(4):239–244, 1995

- As before, use diffusion scale to arrive to the diffusion model.

First Problem - The Diffusion Limits

The Diffusion Model

- As before, use diffusion scale to arrive to the diffusion model.
- The diffusion model consists of:

- As before, use diffusion scale to arrive to the diffusion model.
- The diffusion model consists of:
 - $X(t)$ - number of customers in the main station (at time t).

- As before, use diffusion scale to arrive to the diffusion model.
- The diffusion model consists of:
 - $X(t)$ - number of customers in the main station (at time t).
 - $R(t)$ - number of customers in the retrial station (at time t).

- As before, use diffusion scale to arrive to the diffusion model.
- The diffusion model consists of:
 - $X(t)$ - number of customers in the main station (at time t).
 - $R(t)$ - number of customers in the retrial station (at time t).
 - $C(t)$ - rejections count, $L(t)$ - idle time. Keep the process X in $[0, b]$.

- As before, use diffusion scale to arrive to the diffusion model.
- The diffusion model consists of:
 - $X(t)$ - number of customers in the main station (at time t).
 - $R(t)$ - number of customers in the retrial station (at time t).
 - $C(t)$ - rejections count, $L(t)$ - idle time. Keep the process X in $[0, b]$.
- The diffusion model is characterize using the Skorokhod Map on $[0, b]$.

Definition

The Skorokhod Problem on $[0, b]$ (Tanaka '79)

Given $\psi \in \mathcal{D}[0, \infty)$, find $(\phi, \eta_l, \eta_u) \in (\mathcal{D}[0, \infty))^3$, such that

- ① $\phi(t) = \psi(t) + \eta_l(t) - \eta_u(t), \quad \phi(t) \in [0, b] \quad \text{for all } t.$
- ② η_l, η_u are non negative and non decreasing and one has

$$\int_{[0, \infty)} \mathbb{I}_{\{\phi(t) > 0\}} d\eta_l(t) = 0, \quad \int_{[0, \infty)} \mathbb{I}_{\{\phi(t) < b\}} d\eta_u(t) = 0.$$

Definition

The Skorokhod Problem on $[0, b]$ (Tanaka '79)

Given $\psi \in \mathcal{D}[0, \infty)$, find $(\phi, \eta_l, \eta_u) \in (\mathcal{D}[0, \infty))^3$, such that

- ① $\phi(t) = \psi(t) + \eta_l(t) - \eta_u(t), \quad \phi(t) \in [0, b] \quad \text{for all } t.$
- ② η_l, η_u are non negative and non decreasing and one has

$$\int_{[0, \infty)} \mathbb{I}_{\{\phi(t) > 0\}} d\eta_l(t) = 0, \quad \int_{[0, \infty)} \mathbb{I}_{\{\phi(t) < b\}} d\eta_u(t) = 0.$$

- Has a unique solution.

Definition

The Skorokhod Problem on $[0, b]$ (Tanaka '79)

Given $\psi \in \mathcal{D}[0, \infty)$, find $(\phi, \eta_l, \eta_u) \in (\mathcal{D}[0, \infty))^3$, such that

- ① $\phi(t) = \psi(t) + \eta_l(t) - \eta_u(t), \quad \phi(t) \in [0, b] \quad \text{for all } t.$
- ② η_l, η_u are non negative and non decreasing and one has

$$\int_{[0, \infty)} \mathbb{I}_{\{\phi(t) > 0\}} d\eta_l(t) = 0, \quad \int_{[0, \infty)} \mathbb{I}_{\{\phi(t) < b\}} d\eta_u(t) = 0.$$

- Has a unique solution.
- The solution is called the *Skorokhod map* and is denoted by $\Gamma_{0,b}$.

Definition

The Skorokhod Problem on $[0, b]$ (Tanaka '79)

Given $\psi \in \mathcal{D}[0, \infty)$, find $(\phi, \eta_l, \eta_u) \in (\mathcal{D}[0, \infty))^3$, such that

- ① $\phi(t) = \psi(t) + \eta_l(t) - \eta_u(t), \quad \phi(t) \in [0, b] \quad \text{for all } t.$
- ② η_l, η_u are non negative and non decreasing and one has

$$\int_{[0, \infty)} \mathbb{I}_{\{\phi(t) > 0\}} d\eta_l(t) = 0, \quad \int_{[0, \infty)} \mathbb{I}_{\{\phi(t) < b\}} d\eta_u(t) = 0.$$

- Has a unique solution.
- The solution is called the *Skorokhod map* and is denoted by $\Gamma_{0,b}$.
- Thus $(\phi, \eta_l, \eta_u) = \Gamma_{0,b}(\psi)$.

First Problem - The Diffusion Limits

The Diffusion Model Definition

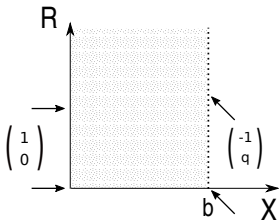
$$\begin{cases} (X, L, C) = \Gamma_{0,b}(Z) \\ Z(t) = x + \underbrace{W(t)}_{\text{BM}} + \int_0^t \mu R(s) ds \\ R(t) = r + qC(t) - \int_0^t \mu R(s) ds. \end{cases}$$

First Problem - The Diffusion Limits

The Diffusion Model Definition

$$\begin{cases} (X, L, C) = \Gamma_{0,b}(Z) \\ Z(t) = x + \underbrace{W(t)}_{\text{BM}} + \int_0^t \mu R(s) ds \\ R(t) = r + qC(t) - \int_0^t \mu R(s) ds. \end{cases}$$

Reflection vector field:

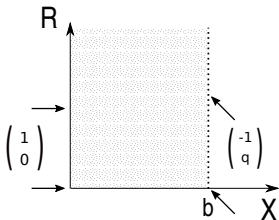


First Problem - The Diffusion Limits

The Diffusion Model Definition

$$\begin{cases} (X, L, C) = \Gamma_{0,b}(Z) \\ Z(t) = x + \underbrace{W(t)}_{\text{BM}} + \int_0^t \mu R(s) ds \\ R(t) = r + qC(t) - \int_0^t \mu R(s) ds. \end{cases}$$

Reflection vector field:



Two dimensional reflected diffusion process which is degenerate in the second diffusion coefficient

Theorem

As $n \rightarrow \infty$,

$$(\hat{X}^n, \hat{R}^n, \hat{L}^n, \hat{C}^n) \Rightarrow (X, R, L, C).$$

- **Cost:**

$$\mathbb{E}\left(\int_0^\infty e^{-\alpha t}(c_1 X(t)dt + c_2 dC(t))\right)$$

Second Result - The Diffusion Optimization Problem

Definition

- **Cost:**

$$\mathbb{E}\left(\int_0^\infty e^{-\alpha t}(c_1 X(t)dt + c_2 dC(t))\right)$$

- **Not solvable as before:**

- **Cost:**

$$\mathbb{E}\left(\int_0^\infty e^{-\alpha t}(c_1 X(t)dt + c_2 dC(t))\right)$$

- **Not solvable as before:**

- ① Not solvable explicitly - the Bellman equation is in two dimensions (correspond to (X, R))

- **Cost:**

$$\mathbb{E}\left(\int_0^\infty e^{-\alpha t}(c_1 X(t)dt + c_2 dC(t))\right)$$

- **Not solvable as before:**

- ① Not solvable explicitly - the Bellman equation is in two dimensions (correspond to (X, R))
- ② Not applicable - the solution require knowledge of the number of customers in retrial (not observable).

- **Cost:**

$$\mathbb{E}\left(\int_0^\infty e^{-\alpha t}(c_1 X(t)dt + c_2 dC(t))\right)$$

- **Not solvable as before:**

- ① Not solvable explicitly - the Bellman equation is in two dimensions (correspond to (X, R))
- ② Not applicable - the solution require knowledge of the number of customers in retrial (not observable).

- **New Goal:** Optimize over buffer size b (using the diffusion model).

- **Cost:**

$$\mathbb{E}\left(\int_0^\infty e^{-\alpha t}(c_1 X(t)dt + c_2 dC(t))\right)$$

- **Not solvable as before:**

- ① Not solvable explicitly - the Bellman equation is in two dimensions (correspond to (X, R))
- ② Not applicable - the solution require knowledge of the number of customers in retrial (not observable).

- **New Goal:** Optimize over buffer size b (using the diffusion model).

- **Value:**

$$V(x) = \inf_b \mathbb{E}\left(\int_0^\infty e^{-\alpha t}(c_1 X(t)dt + c_2 dC(t))\right)$$

Second Result - The Diffusion Optimization Problem

Solution

- Reduction to a 1-dimensional problem (SSC) for:

Second Result - The Diffusion Optimization Problem

Solution

- Reduction to a 1-dimensional problem (SSC) for:
 - ① $\mu \rightarrow 0$

Second Result - The Diffusion Optimization Problem

Solution

- Reduction to a 1-dimensional problem (SSC) for:
 - ① $\mu \rightarrow 0$
 - ② $\mu \rightarrow \infty$

Second Result - The Diffusion Optimization Problem

Solution

- Reduction to a 1-dimensional problem (SSC) for:
 - ① $\mu \rightarrow 0$
 - ② $\mu \rightarrow \infty$
- For general $\mu \in (0, \infty)$ - simulation.

Second Result - The Diffusion Optimization Problem

The Harrison-Taksar Free Boundary Problem

- Harrison Taksar '83.

Second Result - The Diffusion Optimization Problem

The Harrison-Taksar Free Boundary Problem

- Harrison Taksar '83.
- 1-dimensional problem.

Second Result - The Diffusion Optimization Problem

The Harrison-Taksar Free Boundary Problem

- Harrison Taksar '83.
- 1-dimensional problem.
- One server, infinite queue. No retrials.

Second Result - The Diffusion Optimization Problem

The Harrison-Taksar Free Boundary Problem

- Harrison Taksar '83.
- 1-dimensional problem.
- One server, infinite queue. No retrials.
- Controls the buffer size to minimize the same cost.

Second Result - The Diffusion Optimization Problem

The Harrison-Taksar Free Boundary Problem

$$\begin{cases} (\tilde{X}, \tilde{L}, \tilde{C}) = \Gamma_{0,b}(x + W), & \text{where } W \text{ is a BM} \\ J_{\text{HT}}^b[x; c_1, c_2] = \mathbb{E} \int_0^\infty e^{-\alpha t} [c_1 \tilde{X}_t + c_2 \tilde{C}_t] dt, \end{cases}$$

Second Result - The Diffusion Optimization Problem

The Harrison-Taksar Free Boundary Problem

$$\begin{cases} (\tilde{X}, \tilde{L}, \tilde{C}) = \Gamma_{0,b}(x + W), & \text{where } W \text{ is a BM} \\ J_{\text{HT}}^b[x; c_1, c_2] = \mathbb{E} \int_0^\infty e^{-\alpha t} [c_1 \tilde{X}_t + c_2 \tilde{C}_t] dt, \end{cases}$$

- Has a unique optimal buffer size: $b_{\text{HT}}[c_1, c_2]$.

Second Result - The Diffusion Optimization Problem

The Harrison-Taksar Free Boundary Problem

$$\begin{cases} (\tilde{X}, \tilde{L}, \tilde{C}) = \Gamma_{0,b}(x + W), & \text{where } W \text{ is a BM} \\ J_{\text{HT}}^b[x; c_1, c_2] = \mathbb{E} \int_0^\infty e^{-\alpha t} [c_1 \tilde{X}_t + c_2 \tilde{C}_t] dt, \end{cases}$$

- Has a unique optimal buffer size: $b_{\text{HT}}[c_1, c_2]$.
- The value function

$$v[x; c_1, c_2] = \inf_{b \in (0, \infty)} J_{\text{HT}}^b[x; c_1, c_2].$$

is the unique solution to a 1-dimensional Bellman equation:

$$\begin{cases} \frac{1}{2} \sigma^2 f'' + \hat{y} f' - \alpha f + c_1 = 0, & \text{in } (0, b_0), \\ f'(0) = 0, & f'(b_0) = \frac{c_2}{\alpha}. \end{cases}$$

(unknowns: f and b_0)

Second Result - The Diffusion Optimization Problem

The Case $\mu \rightarrow 0$

Theorem

One has

$$\liminf_{\mu \rightarrow 0} \liminf_{n \rightarrow \infty} V^{n,\mu} = \limsup_{\mu \rightarrow 0} \limsup_{n \rightarrow \infty} V^{n,\mu} = v[x; c_1, c_2],$$

and $b_{\text{HT}}[c_1, c_2]$ is an AO scaled buffer size, namely, with $b = b_{\text{HT}}[c_1, c_2]$,

$$\liminf_{\mu \rightarrow 0} \liminf_{n \rightarrow \infty} J^{n,\mu,b} = \limsup_{\mu \rightarrow 0} \limsup_{n \rightarrow \infty} J^{n,\mu,b} = v[x; c_1, c_2].$$

Second Result - The Diffusion Optimization Problem

The Case $\mu \rightarrow \infty$

Theorem

One has

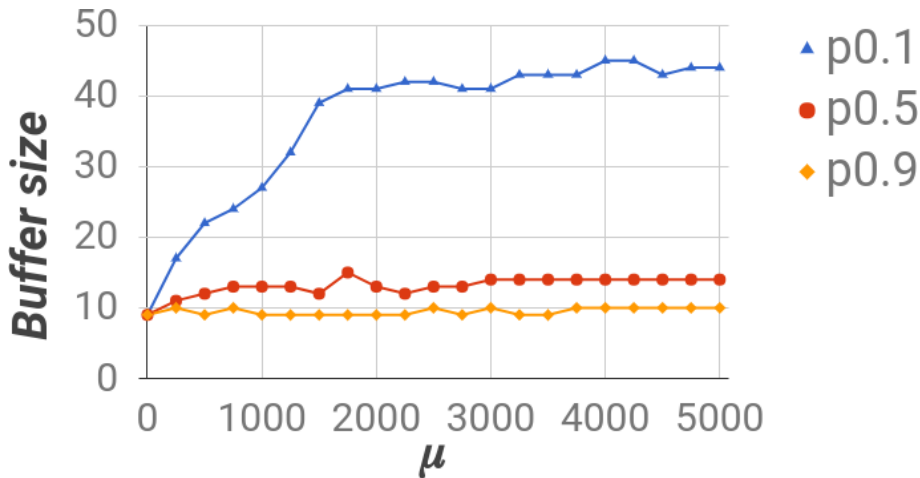
$$\liminf_{\mu \rightarrow \infty} \liminf_{n \rightarrow \infty} V^{n,\mu} = \limsup_{\mu \rightarrow \infty} \limsup_{n \rightarrow \infty} V^{n,\mu} = v[x; c_1, \frac{c_2}{p}].$$

Moreover, $b_{\text{HT}}[c_1, \frac{c_2}{p}]$ is an AO scaled buffer size, in the sense that, with $b = b_{\text{HT}}[c_1, \frac{c_2}{p}]$,

$$\liminf_{\mu \rightarrow \infty} \liminf_{n \rightarrow \infty} J^{n,\mu,b} = \limsup_{\mu \rightarrow \infty} \limsup_{n \rightarrow \infty} J^{n,\mu,b} = v[x; c_1, \frac{c_2}{p}].$$

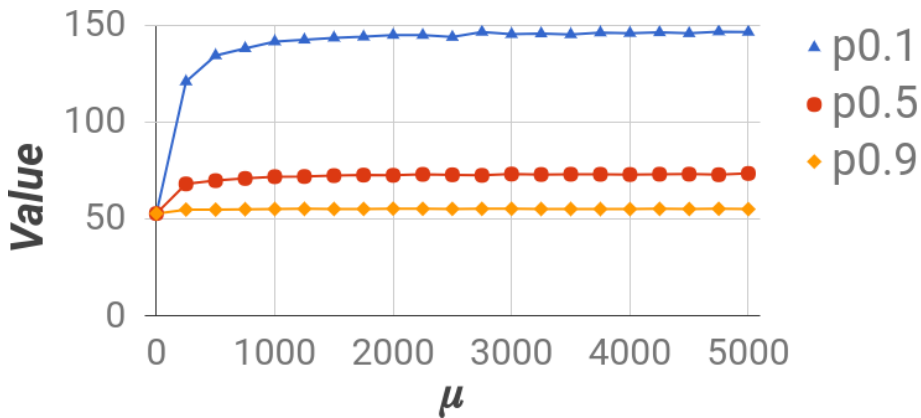
Simulation

Buffer Size, $\mu \in [0, 5000]$



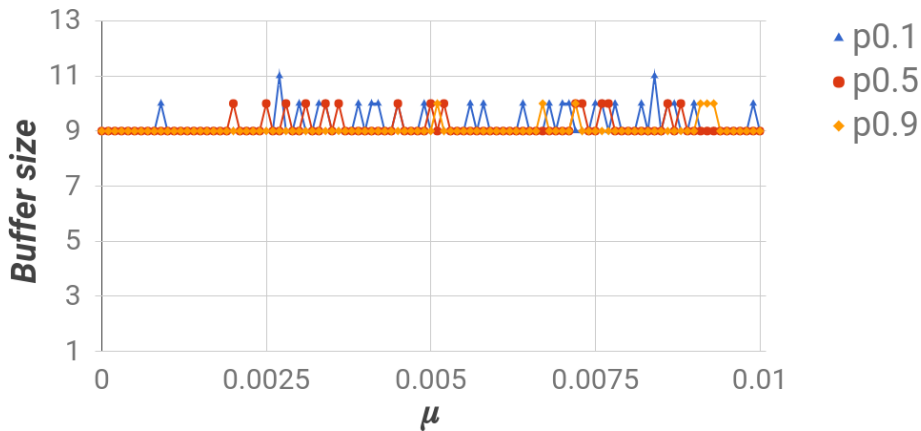
Simulation

Value Function, $\mu \in [0, 5000]$



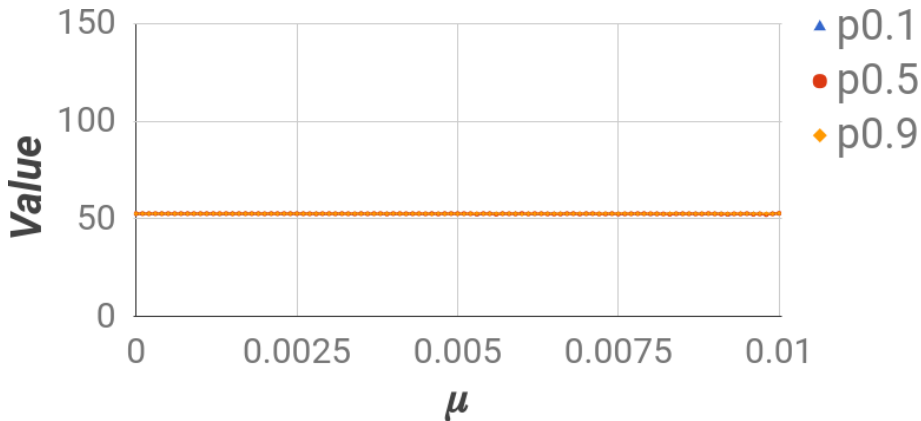
Simulation

Buffer Size, $\mu \in [0, 0.01]$



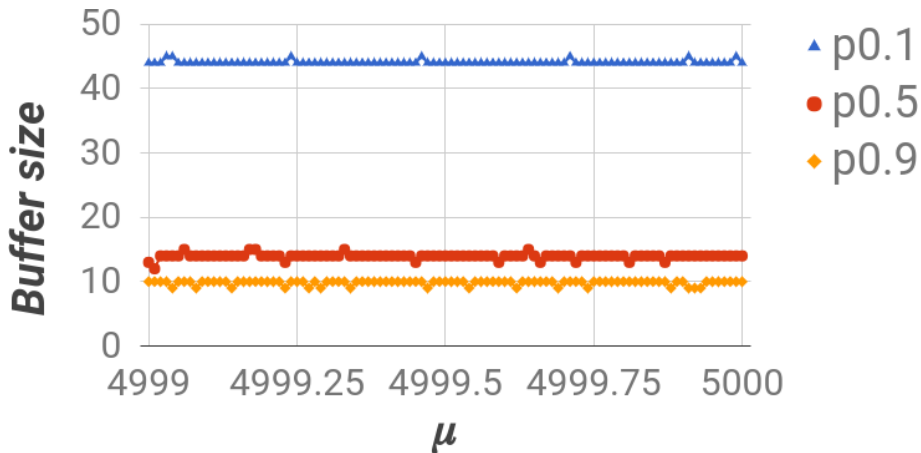
Simulation

Value of $\mu \in [0, 0.01]$



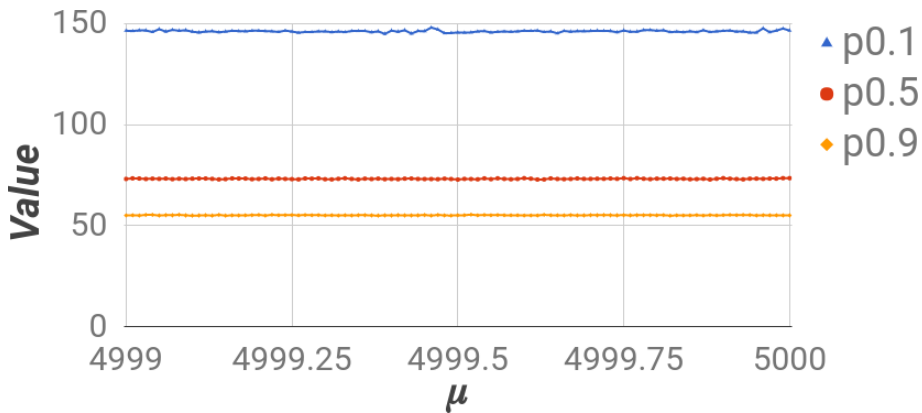
Simulation

Buffer Size, $\mu \in [4999, 5000]$



Simulation

Value Function, $\mu \in [4999, 5000]$



- **Advisor:** Prof. Rami Atar
- Probability group members
- **Administration:** Orly Babad-Tamir and Danit Cohen
- Andrew and Erna Viterbi Faculty of Electrical Engineering

Acknowledgments



Thank You!