Two Extensions of Kingman's GI/G/1 Bound

Florin Ciucu
University of Warwick

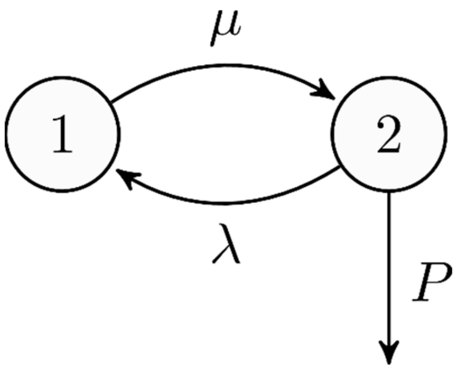
Felix Poloczek

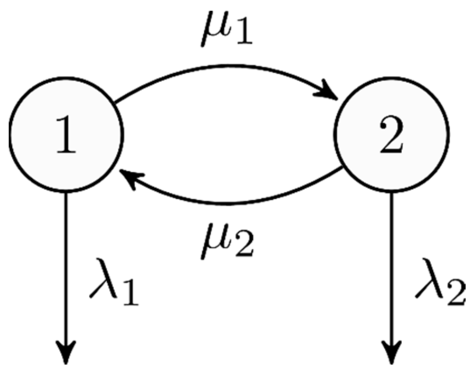
Outline

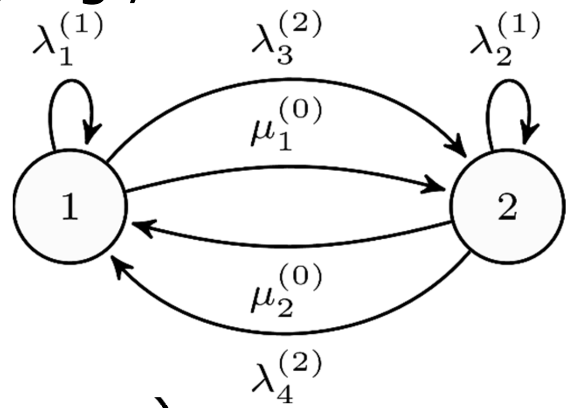
- Kingman's bound (GI/G/1 queue - renewal input)
- new sufficient condition for martingales
- extension 1: the ∑GI/G/1 queue (non-renewal / non-stationary)
- extension 2: queues with Markov input (non-renewal)

Goal: One Queue - One Method

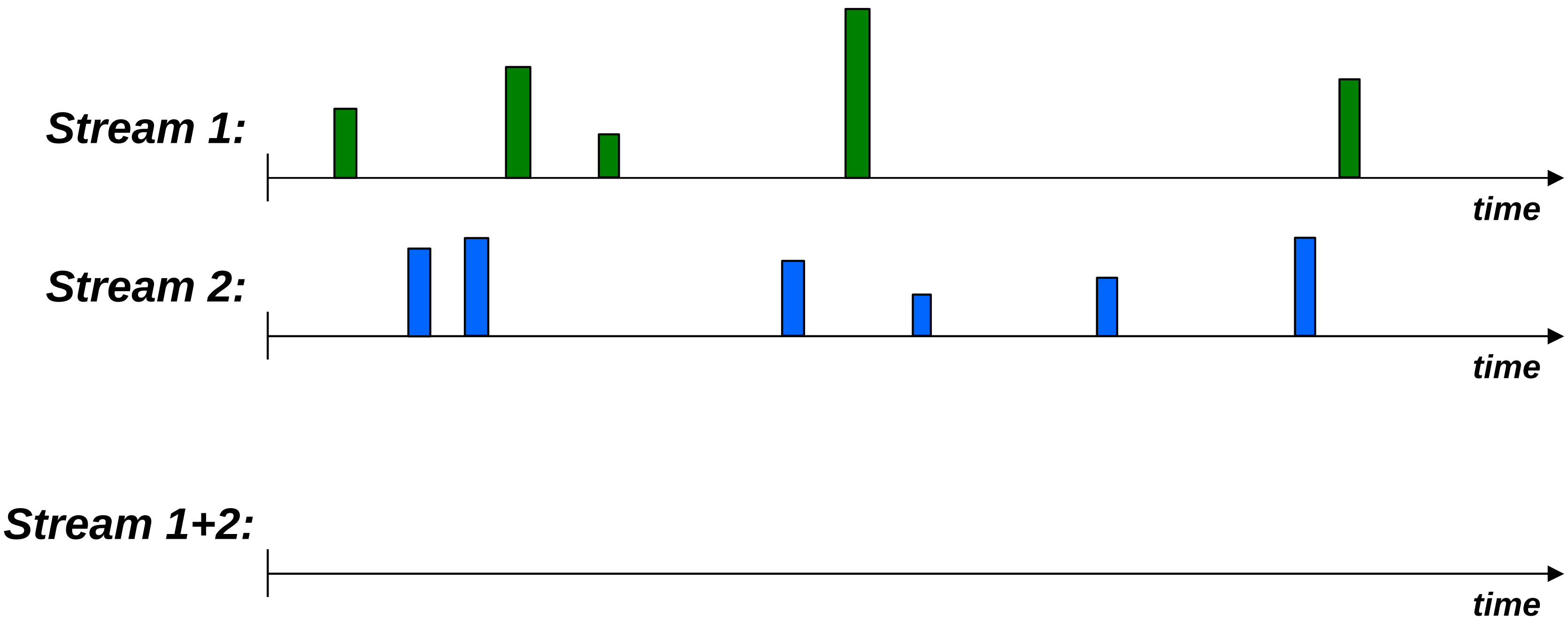
- A "unified" method for non-renewal arrivals, e.g.,





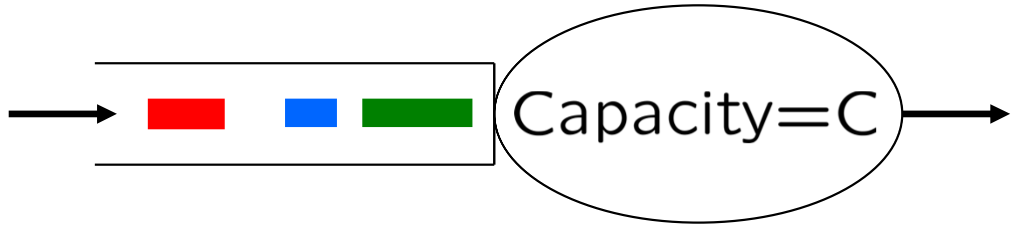


- … or non-stationary arrivals (e.g., ∑GI/G/1 queue)



An analogy

$$A(t) = \sum_{i=1}^{N(t)} X_i$$



- Lindley/Reich's equation

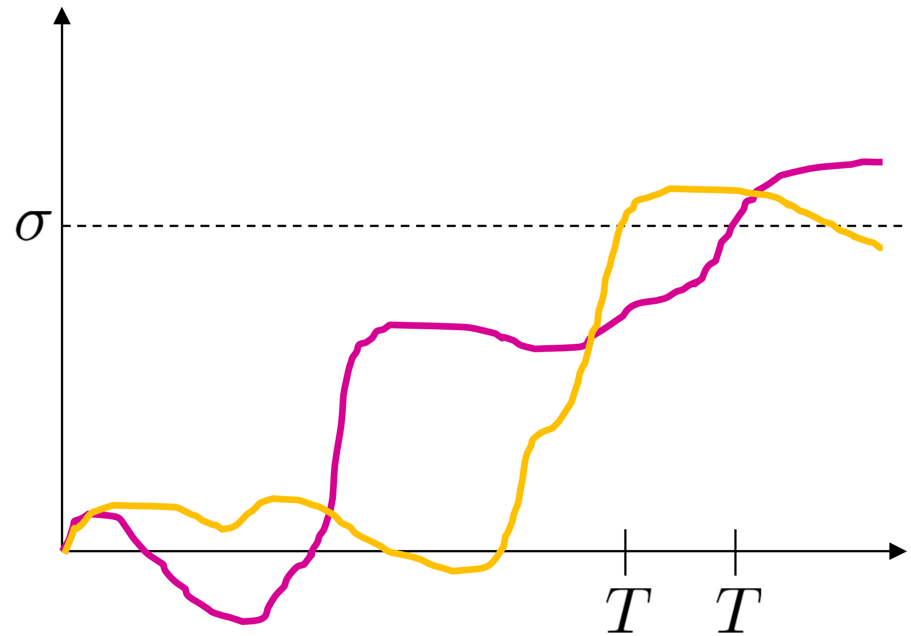
$$Q = \sup_{t \geq 0} \{A(t) - Ct\}$$

- define

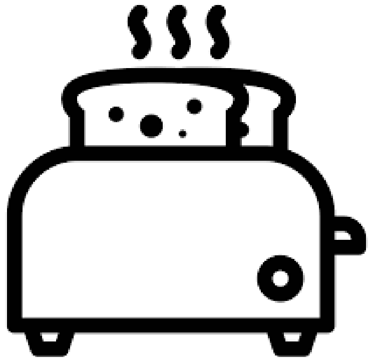
$$T := \inf \{t : A(t) - Ct \geq \sigma\}$$

- then

$$\mathbb{P}(Q \geq \sigma) = \mathbb{P}(T < \infty)$$

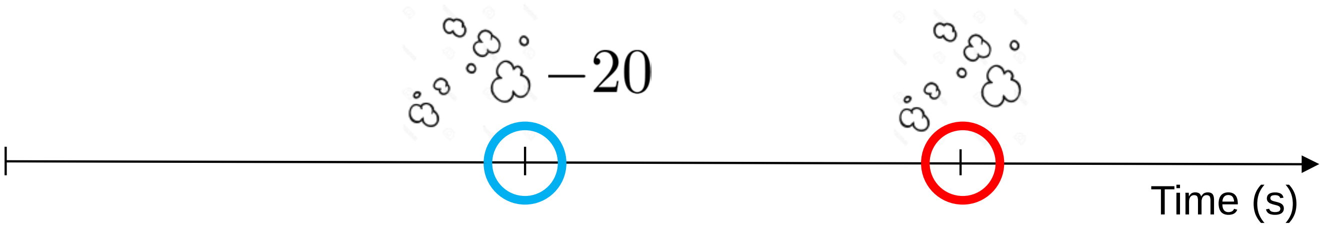


Perfect Toasting Time?



There's an art of knowing when.
Never try to guess.
Toast until it smokes and then
twenty seconds less.

(Piet Hein)



Stopping Time

- take r.v.'s $X_1, X_2, X_3, \ldots$
  - subscript is "time"
  - X_i encodes information (e.g., burning smells)

- A stopping time is a r.v. $N : \Omega \to \{1, 2, \ldots\} \cup \{\infty\}$ such that
  $\{N = n\}$ depends on $X_1, X_2, \ldots, X_n$ only

- first passage/hitting time

$$N = \min\{n \geq 1 \mid X_n \in A\}$$

  - time to buy/sell a stock; time "it smokes"

- $N = \infty$ w.p. > 0 (?) : an asymmetric random walk

$$X_n = \pm 1 \text{ w.p. } < 0.5$$

$$N = \min\{n \mid X_1 + X_2 + \cdots + X_n = 1\}$$

Stopping times are misleading ☹

- take iid r.v.'s $X_1, X_2, X_3, \ldots$
- by definition

$$E\left[X_n\right] = E\left[X_1\right]$$

- however, if N is a stopping time, then in general

$$E\left[X_N\right] \neq E\left[X_1\right]$$

- e.g., X_n are Bernoulli and $N := \min\{n \mid X_n = 1\}$

… but behave nicely for martingales

- **Def**: a sequence of r.v.'s $X_1, X_2, X_3, \ldots$ is a martingale if

$$E\left[\| X_n \| \right] < \infty$$

$$E\left[X_{n+1} \mid X_1, X_2, \ldots, X_n \right] = X_n$$

$$\Leftrightarrow E\left[X_{n+1} - X_n \mid X_1, X_2, \ldots, X_n \right] = 0$$

\- intuitive properties

  – it has "memory"
  – ensures a "fair game"

- not everything is a martingale, e.g.,
  – an iid sequence (ignorance implies unfairness)
  – a Markov process; requires some "transform"

Optional Stopping Theorem (OST)

- immediate property of a martingale $X_1, X_2, X_3, \ldots$

$$E[X_n] = E[X_1]$$

- property preserved for stopping times, i.e.,

$$\boxed{E[X_N] = E[X_1]}$$

subject to

$$N \text{ is bounded}$$

counterexample

$$Y_n = \pm 1 \text{ w.p. } 0.5$$

$$N = \min\{n \mid Y_1 + Y_2 + \cdots + Y_n = 1\}$$

facts

$$X_n := Y_1 + Y_2 + \cdots + Y_n$$

$$1 = E[X_N] \neq E[X_1] = 0$$

Kingman's bound

- recall

$$Q = \max_{t \geq 0}\{A(t) - Ct\}$$

$$\mathbb{P}(Q \geq \sigma) = \mathbb{P}(T < \infty), \; T := \min\{t : A(t) - Ct \geq \sigma\}$$

- construct the martingale

$$X_t := e^{\theta(A(t) - Ct)}$$

- apply the OST with $n \to \infty$

$$\begin{aligned} 1 = \mathbb{E}[X_0] = \mathbb{E}[X_{T\wedge n}] &= \mathbb{E}[X_{T\wedge n} 1_{T\leq n}] + \mathbb{E}[X_{T\wedge n} 1_{T>n}] \\ &= \mathbb{E}[X_T 1_{T\leq n}] \\ &= \mathbb{E}\left[e^{\theta(A(T)-CT)} 1_{T\leq n}\right] \\ &\geq e^{\theta\sigma}\mathbb{E}[1_{T\leq n}] = e^{\theta\sigma}\mathbb{P}(T \leq n) \;. \end{aligned}$$

- hence

$$\mathbb{P}(Q \geq \sigma) \leq e^{-\theta\sigma}$$

Towards "One Queue – One Method" goal

Def. A bivariate process $(A(t), M_t)_t$ is a *Markov Additive Process* iff

1. the pair $(A(t), M_t)$ is a Markov process in $\mathbb{R}^2$,
2. $A(0) = 0$ and $A(t)$ is nondecreasing,
3. the (joint and conditional) distribution of

$$(A(s,t), M_t \mid A(s), M_s)$$

depends only on M_s.

- intuitive aspects:
  - M_t is a background/modulating Markov process
  - $A(t)$ is the additive component
    - » not necessarily Markov
    - » has *conditionally* independent increments

MAP Martingale

- few processes are both Markov and martingales, e.g.,
  - symmetric random walk
  - Brownian motion

Lemma. *For a Markov Additive Process* $(A(t), M_t)$, *functions* $h : \mathrm{rng}(M) \to \mathbb{R}^+$, $y \in \mathrm{rng}(M)$, *and* $C, \theta, s \geq 0$, *define*

$$\varphi_y(s) := \mathbb{E}\left[h(M_s) e^{\theta(A(s)-sC)} \,\middle|\, M_0 = y \right] .$$

If $\frac{d}{ds}\varphi_y(s)\big|_{s=0} = 0$ *for all* $y \in \mathrm{rng}(M) \subseteq \mathbb{R}^2$, *then the process*

$$h(M_t) e^{\theta(A(t)-tC)}$$

is a martingale.

- seemingly obscure result …

a more general and intuitive result

Lemma. *An (integrable) process X_t with continuous $\mathbb{E}[X_t]$ is a martingale iff*

$$\lim_{\Delta s \to 0} \frac{\mathbb{E}\left[X_{s+\Delta s} - X_s \mid \mathcal{F}_s\right]}{\Delta s} = 0 \ \forall s$$

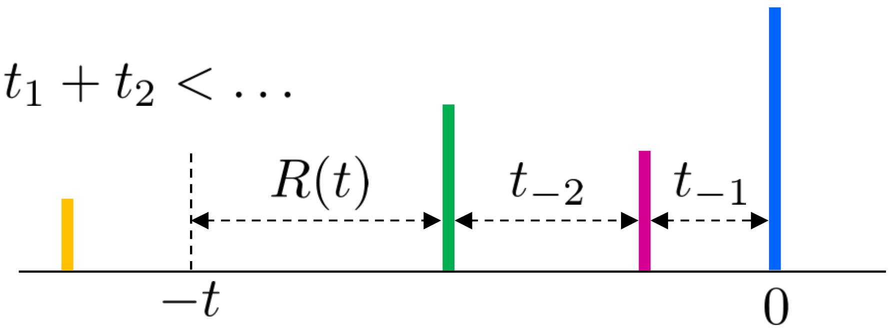
- compare condition to

$$\mathbb{E}\left[X_{n+1} - X_n \mid X_1, X_2, \ldots, X_n\right] = 0 \ \forall n$$

Extension 1: ΣGI/G/1 queue

- (aggregate) arrival process not stationary
- start with one GI/G/1
- arrival points

$$\cdots < -t_{-2} - t_{-1} < -t_{-1} < 0 < t_1 < t_1 + t_2 < \dots$$



- service times $(x_j)_{j\in\mathbb{Z}}$
- define the compound process

$$A(t) := \sum_{j=1}^{N(t)} x_{-j}, \text{ where } N(t) := \max\left\{ n \in \mathbb{N} \mid \sum_{j=1}^{n} t_{-j} \le t \right\}$$

- $N(t)$ is inhomogeneous Poisson process with random rate $\lambda(R(t))$

$$R(t) := t - \sum_{j=1}^{N(t)} t_{-j}, \ \lambda(s) := \lim_{\Delta s \to 0} \mathbb{P}\left(s < t_1 \le s + \Delta s \mid s < t_1\right) = \frac{f(s)}{1 - F(s)}$$

GI/G/1 Martingale (time domain)

Lemma: Let θ satisfying $E\left[e^{-\theta t_1}\right] E\left[e^{\theta x_1}\right] = 1$ and

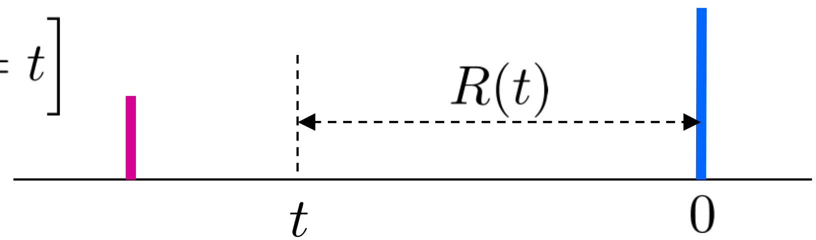
$$h(t) := \frac{1 - E\left[e^{\theta x_1}\right] \int_0^t e^{-\theta s} f(s) ds}{e^{-\theta t}\left(1 - F(t)\right)} .$$

Then the process

$$h(R(t)) e^{\theta(A(t)-t)}$$

is a martingale.

$$\varphi_{t,t}(s) := \mathbb{E}\left[h(M_{t+s}) e^{\theta(A(t,t+s)-Cs)} \;\middle|\; R(t) = t\right]$$



Proof: The martingale condition:

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t}\left[\lambda(t)\Delta t h(0)\mathbb{E}[e^{\theta x_1}] e^{-\theta \Delta t} + (1 - \lambda(t)\Delta t) h(t + \Delta t) e^{-\theta \Delta t} - h(t)\right] = 0$$

yields the ODE

$$h'(t) = h(t)\left(\lambda(t) + \theta\right) - \lambda(t) h(0) \mathbb{E}[e^{\theta x_1}]$$

GI/G/1 Martingales: time and space domains

Lemma: Let θ satisfying $E\left[e^{-\theta t_1}\right] E\left[e^{\theta x_1}\right] = 1$. Then

$$X_t := h(R(t))e^{\theta(A(t)-t)} \qquad X_n := e^{\theta(x_1+\cdots+x_n-t_1-\cdots-t_n)}$$

are martingales.

- both yield the same GI/G/1 bounds

$$Q = \sup_{t\geq 0}\{A(t)-t\} \qquad Q = \max_{n\geq 0}\{\sum x_i - \sum t_i\}$$

- only former works for the ∑GI/G/1 queue

$$Q = \sup_{t\geq 0}\{\sum A_k(t)-t\} \qquad Q = \max_{n\geq 0}\{\sum_k(\sum x_{k,i} - \sum t_{k,i})\}$$

  - martingales are closed under multiplication
  - need the same θ,i.e., $h_k(R_k(t))e^{\theta(A_k(t)-w_k t)}$

Example 1: ∑Weibull/G/1

Take $\mathbb{P}(t_{1,1} \leq t) = 1 - e^{-t^2}$. A bound on the waiting time for each class is

$$\mathbb{P}(W \geq \sigma) \leq K(\theta)^{N-1} e^{-\theta N \sigma} \ ,$$

where

$$K(\theta) := E\left[e^{\theta N x_{1,1}}\right] e^{\frac{\theta^2}{4}} erfc\left(\frac{\theta}{2}\right)$$

and θ satisfies $E\left[e^{-\theta t_1}\right] E\left[e^{\theta N x_{1,1}}\right] = 1$.

Example 2: ∑Erlang-k/G/1

A bound on the waiting time for each class is

$$\mathbb{P}(W \geq \sigma) \leq K(\theta)^{N-1} e^{-\theta N \sigma} \ ,$$

where

$$K(\theta) := \frac{\lambda}{k} \frac{E\left[e^{\theta N x_{1,1}}\right] - 1}{\theta}$$

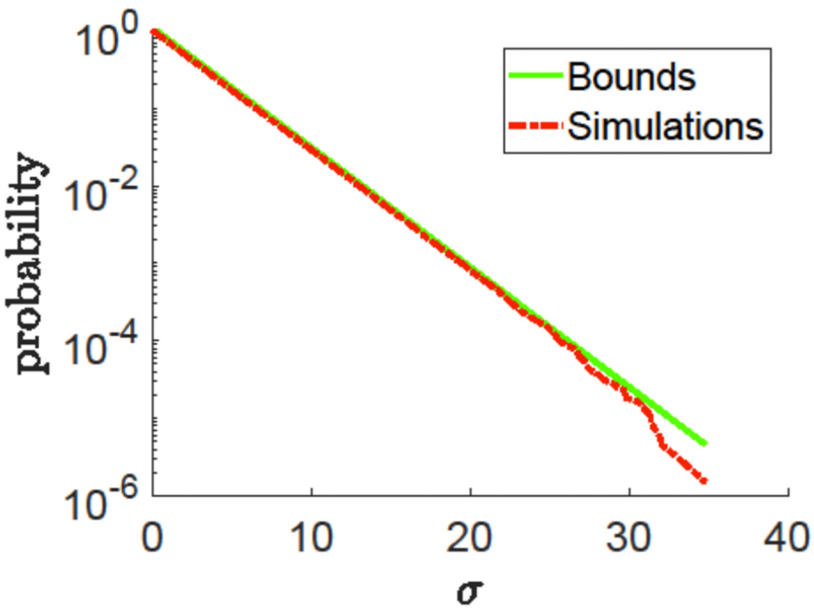
and θ satisfies $\left(1 + \frac{\theta}{\lambda}\right)^{-k} E\left[e^{\theta N x_{1,1}}\right] = 1$.

Example 3: ∑Weibull + Erlang-k/G/1

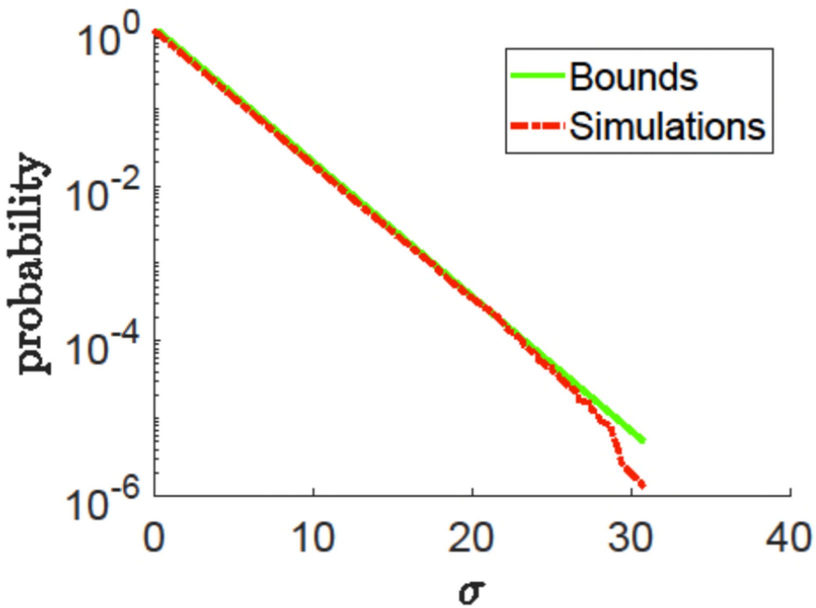
A bound on the waiting-time of a Weibull class is

$$\mathbb{P}(W \geq \sigma) \leq K_W(\theta)^{N_1 - 1} K_E(\theta)^{N_2} e^{-\theta\sigma}$$

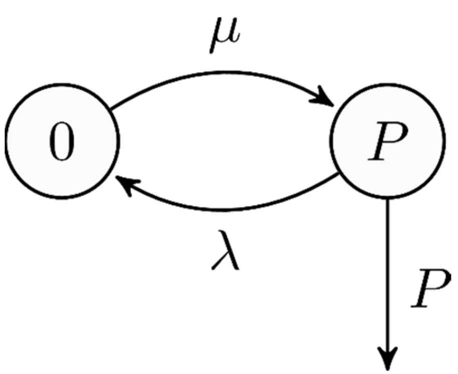
$N_1 = 1,\ N_2 = 4$



$N_1 = 4,\ N_2 = 1$



Extension 2. Markov Fluid (MF)



$$A(t) = \int_0^t M_s ds \ .$$

Let

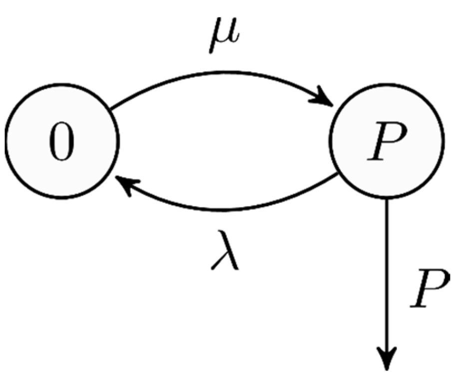
$$\theta = \frac{\lambda}{P - C} - \frac{\mu}{C} \ , \quad h(P) = \frac{\theta C + \mu}{\mu} \ , \quad \text{and} \quad h(0) = 1 \ .$$

Then the process

$$h(M_t) e^{\theta(A(t) - tC)}$$

is a martingale.

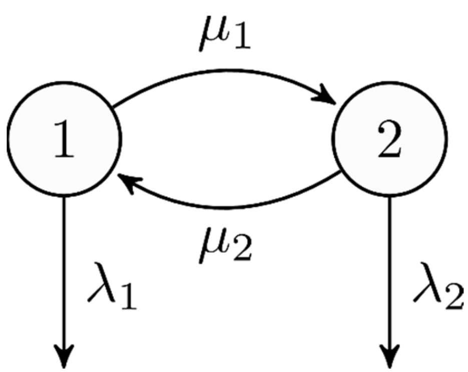
Proof



$$\varphi_y(s) := \mathbb{E}\left[h(M_s)e^{\theta(A(s)-Cs)} \;\middle|\; M_0 = y\right]$$

$$\begin{aligned}
&\frac{d}{ds}\varphi_P(s)\Big|_{s=0} \\
&\quad = \lim_{\Delta s \to 0} \frac{1}{\Delta s}\mathbb{E}\Big[h(M_{\Delta s})e^{\theta(A(\Delta s)-C\Delta s)} - h(P)\Big| M_0 = P\Big] \\
&\quad = \lim_{\Delta s \to 0} \frac{1}{\Delta s}\Big(\lambda \Delta s e^{-\theta C \Delta s} + (1-\lambda\Delta s)h(P)e^{\theta\Delta s(P-C)} - h(P)\Big) \\
&\quad = \lambda - \lambda h(P) + h(P)\theta(P-C) = 0
\end{aligned}$$

Markov Modulated Poisson Process (MMPP)



For $\theta > 0$, let T_θ denote the following 2×2-matrix:

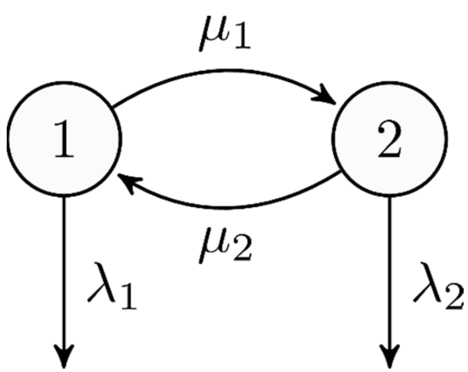
$$T_\theta := \begin{pmatrix} \lambda_1 e^\theta - \mu_1 - \lambda_1 & \mu_1 \\ \mu_2 & \lambda_2 e^\theta - \mu_2 - \lambda_2 \end{pmatrix} .$$

Further, let $\lambda(\theta)$ denote its spectral radius. Pick $\theta > 0$ such that $\lambda(\theta) = \theta C$, and let $h = (h_1, h_2)$ denote an eigenvector corresponding to T_θ and $\lambda(\theta)$. Then the process

$$h(M_t)e^{\theta(A(t)-tC)}$$

is a martingale.

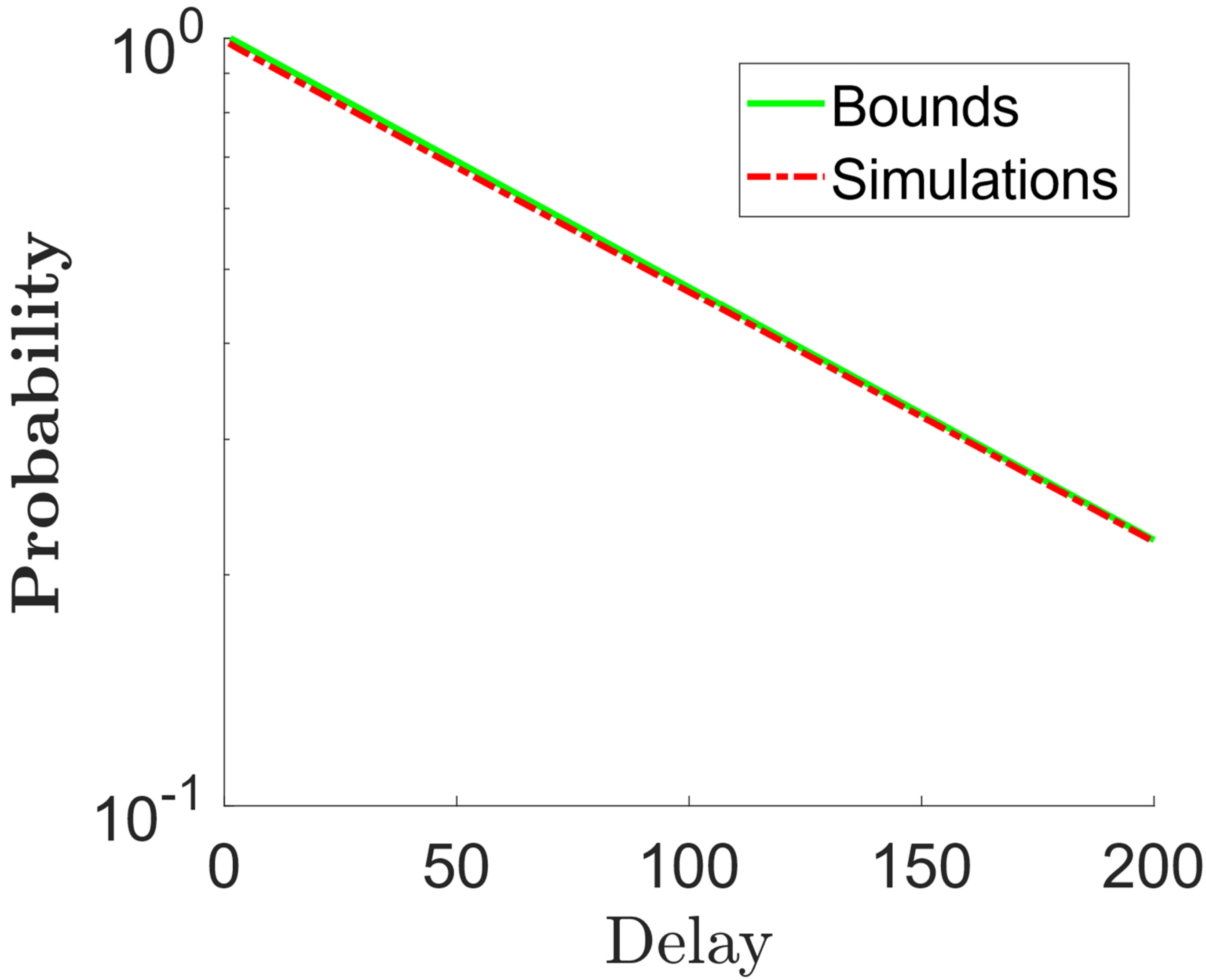
Proof



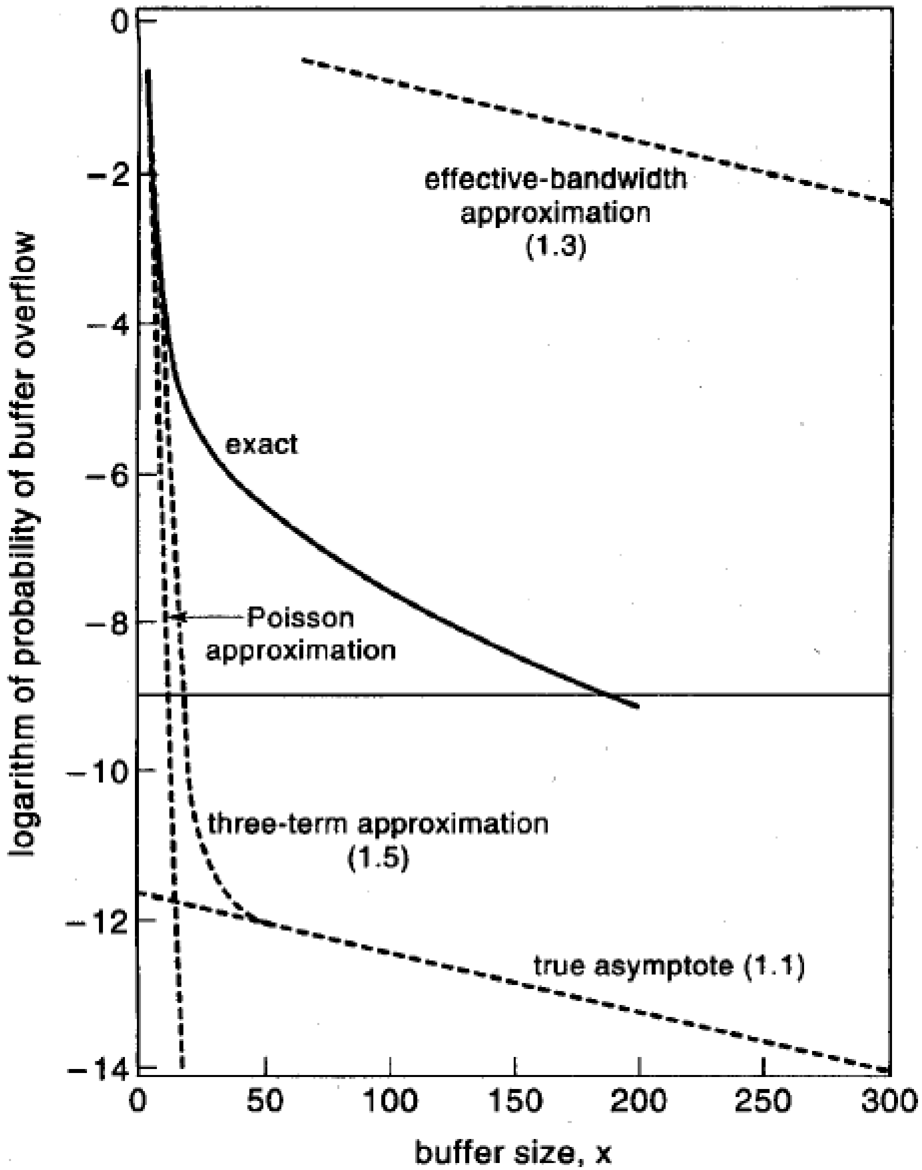
$$\varphi_y(s) := \mathbb{E}\left[h(M_s)e^{\theta(A(s)-Cs)} \;\middle|\; M_0 = y\right]$$

$$\begin{aligned}
&\frac{d}{ds}\varphi_1(s)\Big|_{s=0} \\
&\quad = \lim_{\Delta s \to 0} \frac{1}{\Delta s}\mathbb{E}\Big[h(M_{\Delta s})e^{\theta(A(\Delta s)-C\Delta s)} - h_1 \Big| M_0 = 1\Big] \\
&\quad = \lim_{\Delta s \to 0} \frac{1}{\Delta s}\Big((1-\mu_1\Delta s)\,\lambda_1\Delta s\, h_1\, e^{\theta(1-C\Delta s)} + (1-\mu_1\Delta s)\,(1-\lambda_1\Delta s)\, h_1\, e^{-\theta C\Delta s} \\
&\qquad\qquad + \mu_1\Delta s\,(1-\lambda_1\Delta s)\, h_2\, e^{-\theta C\Delta s} + o(\Delta s) - h_1\Big) = 0
\end{aligned}$$

Bounds vs Simulations



A plot from the '90s



60 MMPP flows

$$(1.3)\ \mathbb{P}(Q > x) \approx e^{-\eta x}$$

$$(1.1)\ \approx \alpha e^{-\eta x}$$

$$(1.5)\ \approx \alpha_1 e^{-\eta_1 x} + \alpha_2 e^{-\eta_2 x} + \alpha_3 e^{-\eta_3 x}$$

$$(!)\ \mathbb{P}(Q > x) \approx \beta e^{-N\gamma} e^{-\eta x}$$

G. Choudhury, D. Lucantoni, and W. Whitt, Squeezing the Most out of ATM, IEEE Transactions on Communications, 1996

Markovian Arrival Process (MArP)

A Markovian Arrival Process is defined via a pair (D_0, D_1) of $n \times n$-matrices such that:

$$d_{i,j} := D_0(i,j) \geq 0 \ , i \neq j \ , \quad d'_{i,j} := D_1(i,j) \geq 0 \ ,$$
$$d_{i,i} := D_0(i,i) = -\sum_{i \neq j} d_{i,j} - \sum_{j} d'_{i,j} \ .$$

The background process M_t is a Markov process with generator $D_0 + D_1$ and steady-state distribution π. If a transition of M_t is triggered by an element of D_1, a packet is generated and $A(t)$ increases by 1 (*active transitions*); transitions triggered by D_0 do not increase $A(t)$ (*hidden transitions*):

$$\mathbb{P}\left(A(t, t + \Delta t) = 0, M_{t+\Delta t} = j \mid M_t = i\right) = D_0(i,j)\Delta t + o(\Delta t) \ ,$$

and

$$\mathbb{P}\left(A(t, t + \Delta t) = 1, M_{t+\Delta t} = j \mid M_t = i\right) = D_1(i,j)\Delta t + o(\Delta t) \ .$$

MArP Martingale

For $\theta > 0$, let $\lambda(\theta)$ denote the spectral radius of the matrix

$$D_0 + e^{\theta} D_1 \ ,$$

If $\lambda(\theta) = \theta C$, and h is a corresponding eigenvector, then the process

$$h(M_t)e^{\theta(A(t)-tC)}$$

is a martingale.

Multiplexing MArPs

In the situation with two MArPs, for $\theta > 0$, let $\lambda(\theta)$ and $\lambda'(\theta)$ denote the spectral radii of the matrices

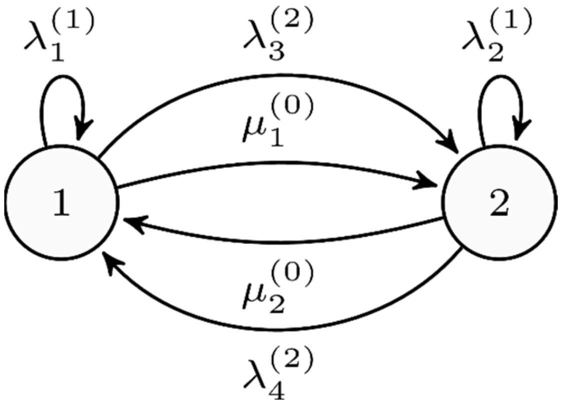
$$D_0 + e^{\theta} D_1 \text{ and } D_0' + e^{\theta} D_1' \ ,$$

respectively. If $\lambda(\theta) + \lambda'(\theta) = \theta C$ and h a corresponding eigenvector, then the process

$$h(M_t) e^{\theta\left(A(t) + A'(t) - tC\right)}$$

is a martingale.
(!) No blow-up of numerical complexity.

Generalized MArP (GMArP)

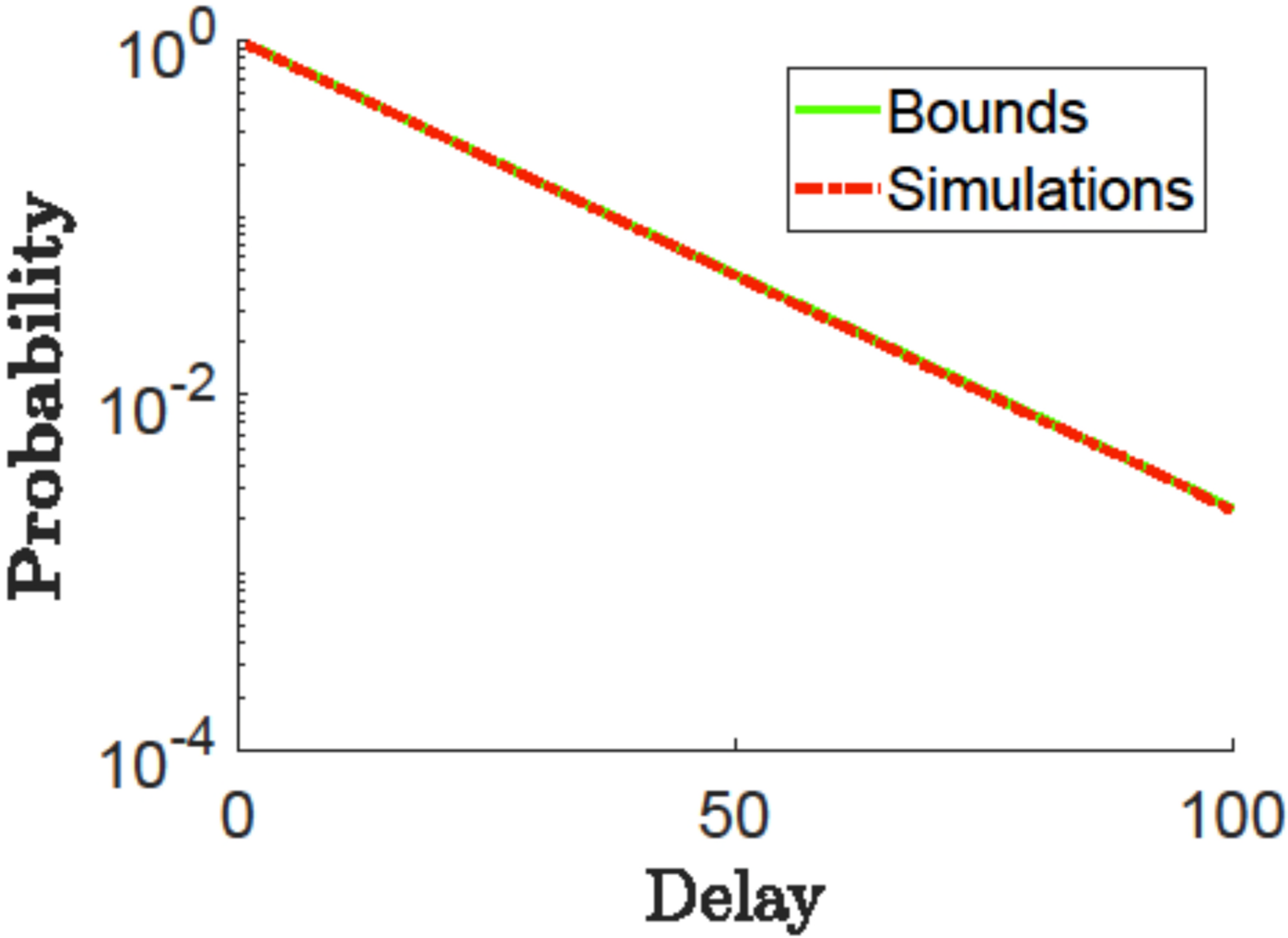


$$D_0 = \begin{bmatrix} -\lambda_1 - \lambda_3 - \mu_1 & \mu_1 \\ \mu_2 & -\lambda_2 - \lambda_4 - \mu_2 \end{bmatrix}$$

$$D_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$D_2 = \begin{bmatrix} 0 & \lambda_3 \\ \lambda_4 & 0 \end{bmatrix}$$

Bounds vs Simulations



Conclusions

- Kingman's bound: inspect the queue at a stopping time and extract information through a martingale

- a sufficient condition for martingale constructions from MAPs (inhomogeneous + uncountable state space)

- two extensions of Kingman's bounds to queues with non-renewal / non-stationary input

- 3rd extension (finite buffer queueing systems): inspect the queue at two stopping times