Random fluid limit of an overloaded polling model

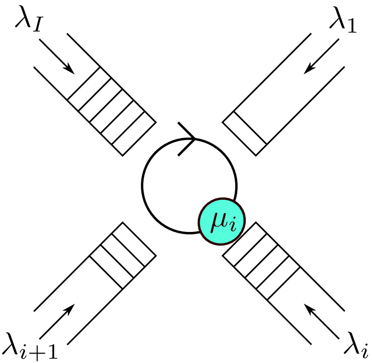
Maria Frolkova (CWI)
Sergey Foss (Heriot-Watt)
Bert Zwart (CWI, EURANDOM, VU, Georgia Tech)

Young European Queueing Theorists Workshop
EURANDOM 2012

Outline

Outline

“Classical" polling model



- cyclic routing $1, 2, \ldots, I, 1, 2, \ldots$
- independent infinite-buffer queues, each with
  - Poisson arrivals
  - i.i.d. service times
  - i.i.d. switchover times
- policies: gated, exhaustive and more

Multigated service

Per visit, *the server gates the queue*
at most *the gating index* number of times.

$X_1, \ldots, X_I$ — r.v.'s with values in $\mathbb{Z}_+ \cup \{\infty\}$

The service discipline in queue i is *X_i-gated*,
i.e. gating indices for visits of queue i are i.i.d. copies of X_i.

$$\text{Gating index} = \begin{cases} 1 & \Rightarrow & \text{gated} \\ \infty & \Rightarrow & \text{exhaustive} \\ \text{const} & \Rightarrow & \text{fair and efficient}^1 \end{cases} \quad \text{service}$$

[1] van Wijk, Adan, Boxma, Wierman 2012

Additional assumptions

- B_i — service time in queue i,
  $\mathbb{E}B_i \log B_i < \infty$ for all i
- zero switchover times
- *overload* : $\lambda_i/\mu_i < 1$ for all i and $\sum_{i=1}^{I} \lambda_i/\mu_i > 1$
- *empty initial state*

Outline

1. Model description
2. Fluid limit
3. Moment conditions
4. Summary

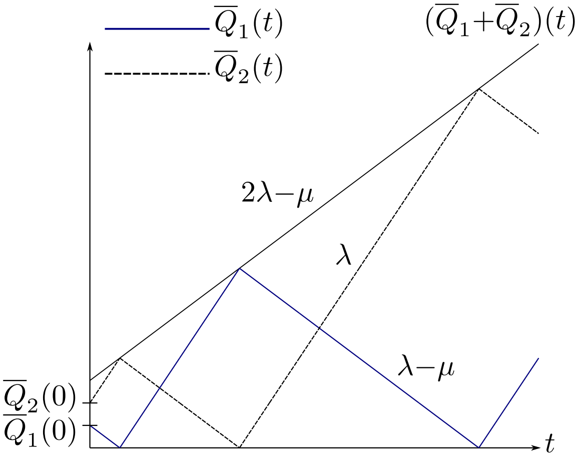
Example

$Q(\cdot) = (Q_1, \ldots, Q_I)(\cdot)$ — queue length process

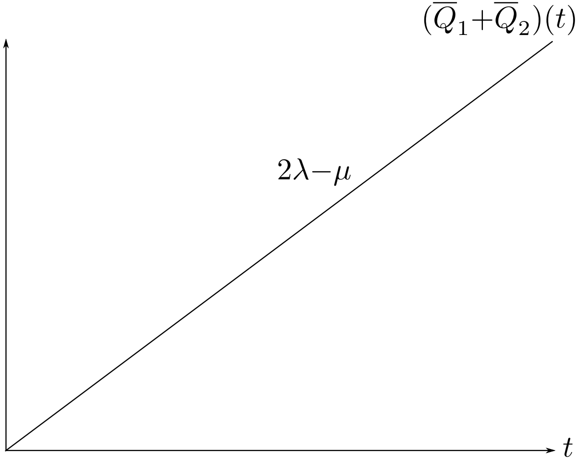
$x^{(n)} \to \infty$, then $\dfrac{Q(x^{(n)}\cdot)}{x^{(n)}} \to$ *fluid limit* $\overline{Q}(\cdot)$ (a.s.)

Consider a symmetric model, 2 queues, exhaustive service, $1/2 < \lambda/\mu < 1$

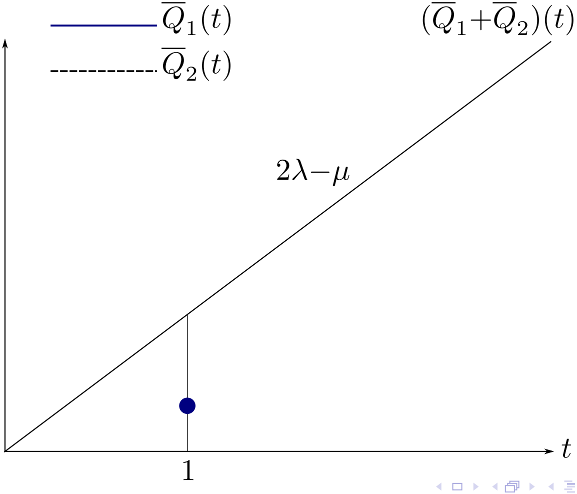
Example, starting non-empty



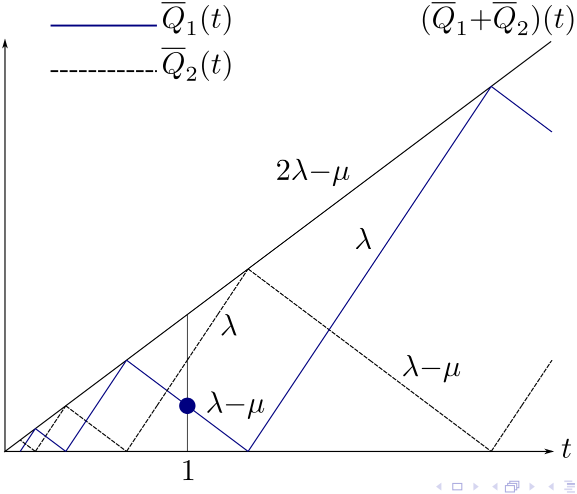
Example, starting empty



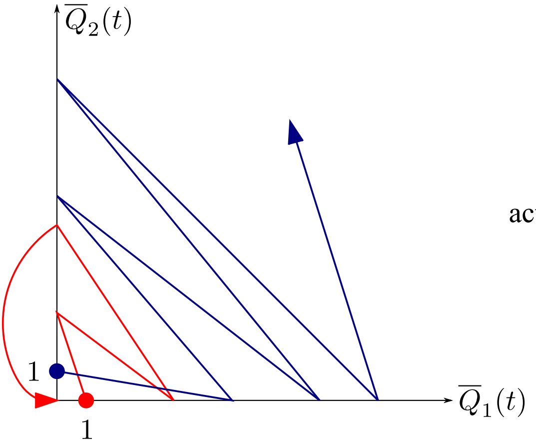
Example, starting empty



Example, starting empty



Why is the fluid limit random?

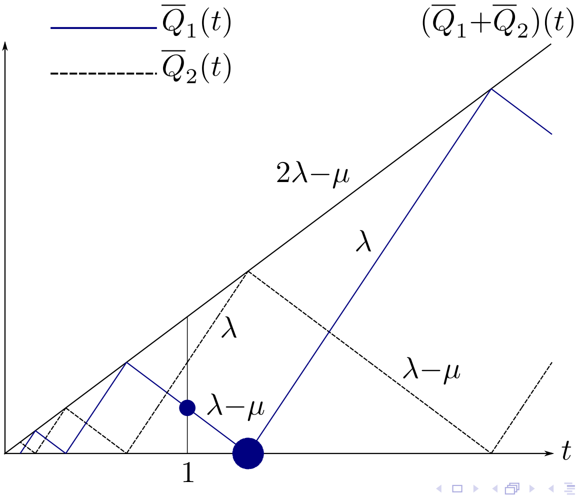


overload

+

actual initial state ●

Example, starting empty



Embedded MTBP

Property

For all $i = 1, \ldots, I$, the customers found in queue i at a polling instant get replaced during the course of the visit by i.i.d. populations of customers jointly in all queues.

$t_1^{(n)}, t_2^{(n)}, \ldots, t_I^{(n)}$, $n \geq 0$ — successive polling instants

$Q(t_1^{(n)})$, $n \geq 0$ — MTBP with immigration in the empty state[2]

[2] Resing 1993

Embedded MTBP

- $\rho > 1$ — biggest eigenvalue of the mean offspring matrix
- $v \in (0, \infty)^I$ — eigenvector for ρ

Theorem (Kesten-Stigum)

$$\frac{Q(t_1^{(n)})}{\rho^n} \to \zeta v \text{ a.s., where } \zeta > 0$$

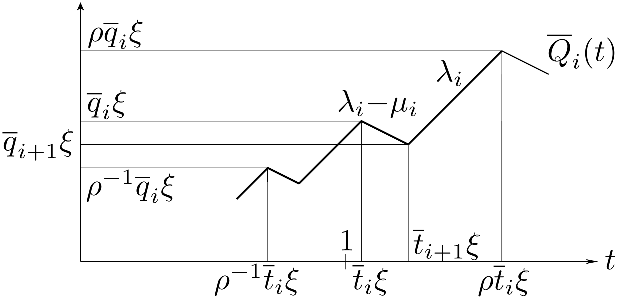
$$\frac{Q(\rho^n \cdot)}{\rho^n} \to \textit{fluid limit } \overline{Q}(\cdot) \quad \text{a.s.}$$

Fluid limit

$$(t_1^{\eta_n}, \ldots, t_I^{\eta_n}) = \min\{(t_1^{(n)}, \ldots, t_I^{(n)}) \colon t_1^{(n)} \geq \rho^n\}$$

$$(t_1^{\eta_n}, \ldots, t_I^{\eta_n})/\rho^n \to (\bar{t}_1, \ldots, \bar{t}_I)\xi \text{ a.s.}$$

$$(Q_1(t_1^{\eta_n}), \ldots, Q_I(t_I^{\eta_n}))/\rho^n \to (\bar{q}_1, \ldots, \bar{q}_I)\xi \text{ a.s.}$$



Fluid limit

Kesten-Stigum theorem ($Q(t_1^{(n)})/\rho^n \to \zeta v$) + LLN $\Rightarrow$

$$\bar{t}_1 = \frac{\sum_i v_i/\mu_i}{\sum_i \lambda_i/\mu_i - 1}$$

$$t_1^{(n)} + \sum_i \sum_{E_i(t_1^{(n)})-Q_i(t_1^{(n)})}^{E_i(t_1^{(n)})} B_i^{(k)} = \sum_i \sum_{k=1}^{E_i(t_1^{(n)})} B_i^{(k)}$$

Fluid limit

Kesten-Stigum theorem $(Q(t_1^{(n)})/\rho^n \to \zeta v)$ + LLN $\Rightarrow$

for all i,

$$\overline{q}_i = v_1 + \lambda_i(\overline{t}_i - \overline{t}_1) \qquad Q_i(t_i^{(n)}) = Q_i(t_1^{(n)}) + E_i(t_i^{(n)}) - E_i(t_1^{(n)})$$

$$\overline{t}_{i+1} = \overline{t}_i + \overline{q}_i\gamma_i \qquad t_{i+1}^{(n)} = t_i^{(n)} + \sum_{k=1}^{Q_i(t_i^{(n)})} V_i^{(n,k)}$$

$$\gamma_i := \mathbb{E}(t_{i+1}^{(n)} - t_i^{(n)} | Q_i(t_i^{(n)}) = 1) = \frac{1 - \mathbb{E}(\lambda_i/\mu_i)^{X_i}}{\mu_i - \lambda_i}$$

Fluid limit

Kesten-Stigum theorem $(Q(t_1^{(n)})/\rho^n \to \zeta v) \Rightarrow$

$$\xi = \rho^{\{\log_\rho(\bar{t}_1\zeta)\}}/\bar{t}_1, \quad \bar{t}_1\xi \in [1, \rho)$$

Fluid limit

Kesten-Stigum theorem $(Q(t_1^{(n)})/\rho^n \to \zeta v) \Rightarrow$

$$\xi = \rho^{\{\log_\rho(\bar{t}_1\zeta)\}}/\bar{t}_1, \quad \bar{t}_1\xi \in [1, \rho)$$

for $x \in [1, \rho)$, $\quad \mathbb{P}\{\bar{t}_1\xi \geq x\} =$

$$\frac{1}{1 - q_G} \sum_{\substack{k \in \mathbb{Z}^I, \\ |k| \geq 1}} G(k)$$

$$\times \sum_{\substack{l \leq k, \\ |l| \geq 1}} \binom{k}{l} ((1, \ldots, 1) - (q_1, \ldots, q_I))^l (q_1, \ldots, q_I)^{k-l}$$

$$\times \mathbb{P}\{\{\log_\rho(\bar{t}_1 \sum_{i=1}^{I} \sum_{j=1}^{l_i} \xi_i^{(j)})\} \geq \log_\rho x\},$$

Fluid limit

where

- q_i — extinction probabilities (no immigration)
- ξ_i — Kesten-Stigum limits (no immigration)
- $\xi_i^{(j)}$, $j \geq 1$ — i.i.d. copies of $\xi_i | \xi_i > 0$
- $G(k)$, $k \in \mathbb{Z}_+^I$ — immigration distribution
- q_G — probability for $Q(t_1^{(n)})$, $n \geq 0$, to return to zero

Outline

Moment conditions

$Q(t_1^{(n)})$ — MTBP, $L_{i,j}$ — queue j offspring of a type i customer

Kesten-Stigum theorem: $\mathbb{E}L_{i,j} \log L_{i,j}$ must be finite

Consider $M/G/1$ queue, $\lambda < \mu$, service time B, busy period τ.

$$\mathbb{E}f(B) < \infty \Rightarrow \mathbb{E}f(\tau) < \infty, \quad f(x) := (x \log x)^+$$

$\tau(k)$ — at most k gating stages, $\tau(k) \to \tau$ a.s.
$\tau(k+1) \stackrel{d}{=} B + \sum_{j=1}^{E(B)} \tau^{(j)}(k)$

Moment conditions

$$\begin{aligned}
&\mathbb{E}f(\tau(k)) \le \mathbb{E}f(\tau(k+1))\\
\le&(1/2)\mathbb{E}f(2B) + (1/2)\mathbb{E}f(2\sum_{j=1}^{E(B)} \tau^{(j)}(k))\\
\le&(1/2)\mathbb{E}f(2B) + (1/2)\mathbb{E}f(2E(B)\tau(k))\\
\le&(1/2)\mathbb{E}f(2B) + \underbrace{(1/2)\mathbb{E}(2E(B))}_{=\lambda/\mu}\mathbb{E}f(\tau(k)) + \underbrace{(1/2)\mathbb{E}\tau(k)}_{\le 1/(2(\mu-\lambda))}\underbrace{\mathbb{E}f(2E(B))}_{<\infty}
\end{aligned}$$

convexity, $f(xy) \le xf(y) + yf(x)$,
$\limsup f(cx)/f(x) < \infty$, $\lim \log f(x)/x = 0$

$$(1-\lambda/\mu)\underbrace{\mathbb{E}f(\tau(k))}_{\to \mathbb{E}f(\tau)} \le \mathbb{E}f(2B)/2 + \mathbb{E}f(2E(B))/(2(\mu-\lambda))$$

Outline

Summary

- precise description of the random fluid limit
- moments of the busy period in $M/G/1$ queue