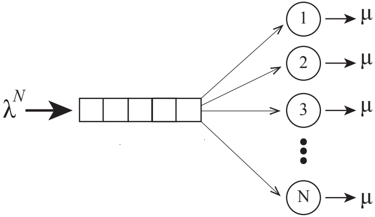
Asymptotic Coupling of an SPDE, with Applications to Many-Server Queues

Mohammadreza Aghajani
joint work with Kavita Ramanan

Brown University

March 2014

Many-Server Queues



Where do they arise?

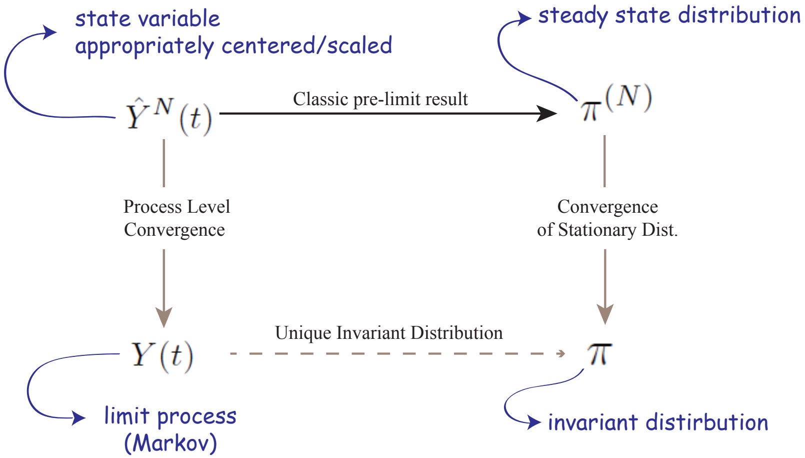
- Call Centers
- Health Care
- Data Centers

Relevant Steady state performance:

- $\mathbb{P}_{ss}\{\text{all } N \text{ servers are busy}\}$
- $\mathbb{P}_{ss}\{\text{waiting time} > t \text{ seconds}\}$

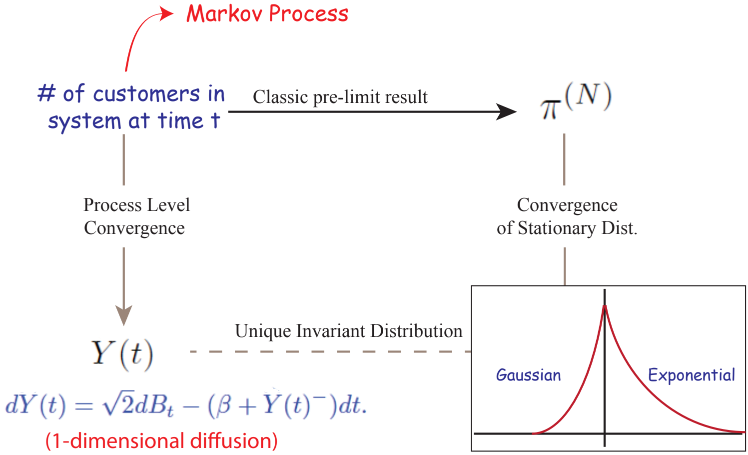
Asymptotic Analysis

Exact analysis for finite N is typically infeasible.



Exponential Service Distribution

Exponential service time case is studied in [Halfin-Whitt'81]



General Service Distribution

- Statistical data shows that service times are generally distributed (Lognormal, Pareto, etc. see e.g. [Brown et al. '05])

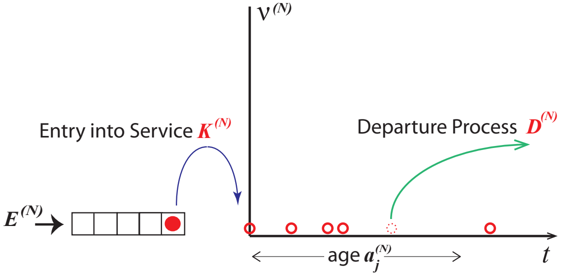
Goal: To extend the result to general service distribution

- **Challenges**
  - $\{X^{(N)} = \#$ of customers in system$\}$ is no longer a Markov Process
  - need to keep track of residual times or ages of customers in service to make the process Markovian
  - Dimension of any finite-dim. Markovian representation grows with N
  - Results using X^N obtained by [Puhalskii and Reed], [Reed], [Mandelbaum and Momcilovic], [Dai and He], etc.
  - However, few results on stationary distribution (beyond phase-type distributions)

A way out: Common State Space (infinite-dimensional)

A Measure-Valued Representation

A measure-valued state descriptor [Kaspi-Ramanan'11]: $(X^{(N)}, \nu^{(N)})$



- $E^{(N)}$ represents the cumulative external arrivals
- $a_j^{(N)}$ represents age of the jth customer to enter service
- $\nu^{(N)}$ keeps track of the ages of all the customers in service

$$\nu_t^{(N)} = \sum_j \delta_{a_j^{(N)}(t)}$$

A Central Limit Theorem

An important functional:

$$Z_t^{(N)}(r) = \sum_{j \text{ in service}} \frac{1 - G(a_j(t) + r)}{1 - G(a_j(t))}, \ r \in (0, \infty).$$

Take the state variable $Y_t^{(N)} = (X_t^{(N)}, Z_t^{(N)}) \in \mathbb{R} \times \mathbb{H}^1(0, \infty)$

Lemma (Aghajani and 'R'13, Corollary of Kaspi and 'R'13)

Under Assumptions on G, if $\widehat{y}_0^{(N)} \Rightarrow y_0$ in $\mathbb{R} \times \mathbb{H}^1(0, \infty)$, then for every $t \geq 0$,

$$\widehat{Y}_t^{(N)} \Rightarrow Y_t \qquad \text{in } \mathbb{R} \times \mathbb{H}^1(0, \infty),$$

for some continuous Markov process $\{Y_t = (X_t, Z_t); t \geq 0\}$.

- To make Y have all the desired properties, the choice of space is subtle.

Main focus: Limit Process

When $z_0 \in \mathbb{H}^1(0,\infty) \cap \mathbb{C}^1[0,\infty)$, the limit process is the unique solution of the coupled Ito process/SPDE:

$$\begin{cases} dX_t = dW_t + dB_t - \beta dt + Z_t'(0)dt, \\ dZ_t(r) = \left[Z_t'(r) - \bar{G}(r)Z_t'(0)\right] dt - dM_t(r) + \bar{G}(r)dZ_t(0), \end{cases}$$

with boundary condition $Z_t(0) = -X_t^- = X_t \wedge 0$.

- M is a martingale measure, correlated to the Brownian motion W.
- W and B are independent Brownian motions.

Remarks:

- Unusual SPDE – boundary conditions enter the drift term.
- The process lies in the infinite-dimensional space $\mathbb{R} \times \mathbb{H}^1(0,\infty)$.
- For general initial condition z_0, Y is a solution to a slightly more complicated SPDE.

We are the interested in the invariant distribution of the Markov process Y.

Invariant Distribution: Existence

Theorem (Aghajani and 'R'13)

Under assumptions on G, the limit process has an invariant distribution.

- Main tool: Krylov-Bogoliubov ['37] Theorem. Need to show the tightness for family of occupation measures

$$\mu(\Gamma) = \frac{1}{T}\int_0^T 1_\Gamma\left(Y_t\right) dt$$

- Starting from a particular initial condition introduced in [Gamarnik and Goldberg '11], we use the bounds they obtained on $\widehat{X}^{(N)}$.

Invariant Distribution: Uniqueness

- Key challenge: State Space $\mathcal{Y} \doteq \mathbb{R} \times \mathbb{H}^1$ is infinite dimensional
  - Traditional recurrence methods are not easily applicable.
  - In some cases, traditional methods fail [Hairer et. al.'11].
- We invoke the asymptotic coupling method (Hairer, Mattingly, Sheutzow, Bakhtin, et al.)

Theorem (Hairer et al.'11, continuous-time version)

Assume for a Markov kernel $\{\mathcal{P}_t\}$ on a polish state space $\mathcal{Y}$, there exists a measurable set $A \subseteq \mathcal{Y}$ with the following two properties:

(I) $\mu(A) > 0$ *for any invariant probability measure* μ *of* $\mathcal{P}_t$.

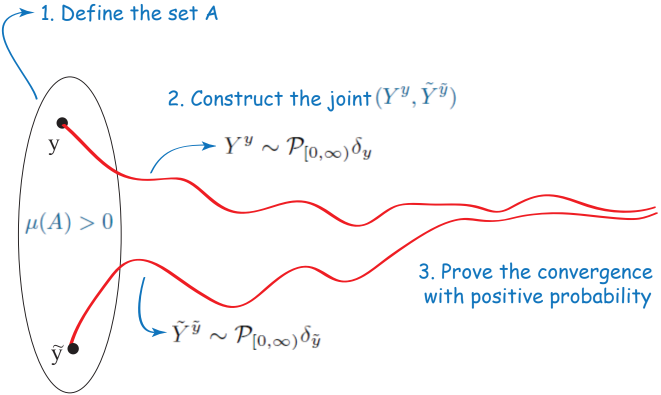
(II) *For every* $y, \tilde{y} \in A$*, there exists a measurable map* $\Gamma_{y,\tilde{y}} : A \times A \to \tilde{\mathcal{M}}(\mathcal{P}_{[0,\infty)}\delta_y, \mathcal{P}_{[0,\infty)}\delta_{\tilde{y}})$*, such that* $\Gamma_{y,\tilde{y}}(\mathcal{D}) > 0$.

Then $\{\mathcal{P}_t\}$ *has at most one invariant probability measure.*

- $\mathcal{P}_{[0,\infty)}\delta_y \in \mathcal{M}(\mathcal{Y}^{[0,\infty)})$: distribution of a Markov process with kernel $\mathcal{P}$ starting at y.
- $\tilde{\mathcal{M}}(\mu_1, \mu_2) = \{\gamma \in \mathcal{M}(\mathcal{Y}^{[0,\infty)} \times \mathcal{Y}^{[0,\infty)}), \gamma_i \ll \mu_i, i = 1, 2.\}$
- $\mathcal{D} = \{(x, y) \in \mathcal{Y}^{[0,\infty)} \times \mathcal{Y}^{[0,\infty)}; \lim_{t\to\infty} d(x_t, y_t) = 0\}$

Invariant Distribution: Uniqueness

To utilize the the asymptotic coupling theorem, we have to:



Then $\Gamma_{y,\tilde{y}} = \text{Law}(Y^y, \tilde{Y}^{\tilde{y}})$ is a legitimate asymptotic coupling.

Invariant Distribution: Uniqueness

Theorem (Aghajani and 'R'13)

Under assumptions on G, the limit process has at most one invariant distribution.

Proof idea. Let $y = (x_0, z_0)$ and $\tilde{y} = (\tilde{x}_0, \tilde{y}_0)$. As defined before,

$$\begin{cases} dX_t = dW_t + dB_t - \beta dt + Z_t'(0)dt, \\ dZ_t(r) = \left[Z_t'(r) - \bar{G}(r)Z_t'(0)\right] dt - dM_t(r) + \bar{G}(r)dZ_t(0), \end{cases}$$

Now define

$$\begin{cases} d\tilde{X}_t = dW_t + d\tilde{B}_t - \beta dt + \tilde{Z}_t'(0)dt, \\ d\tilde{Z}_t(r) = \left[\tilde{Z}_t'(r) - \bar{G}(r)\tilde{Z}_t'(0)\right] dt - dM_t(r) + \bar{G}(r)d\tilde{Z}_t(0), \end{cases}$$

where

$$\boxed{\tilde{B}_t = B_t + \int_0^t \left(\Delta Z_s'(0) - \lambda \Delta X_s\right) ds.}$$

Invariant Distribution: Uniqueness

Define $A = \{(x, z) \in \mathcal{Y}; x \geq 0\}$. For every invariant dist. μ of $\mathcal{P}$, $\mu(A) > 0$.

Lemma

When $y, \tilde{y} \in A$, we have $\Delta Z'_\cdot(0) \in \mathbb{L}^2$

- **Asymptotic Convergence:**
  - $\Delta X_t = \Delta x_0 e^{-\lambda t} \quad \Rightarrow \quad \Delta X_t \to 0.$

$$\Delta Z_t(r) = \Delta z_0(t + r) + \bar{G}(r)\Delta X_t^- + \int_0^t \Delta X_s^- g(t + r - s)ds - \int_0^t \Delta Z'_s(0)\bar{G}(t + r - s)ds. \qquad (1)$$

  - Using Lemma (2), $\Delta Z_t \to 0$ in $\mathbb{H}^1(0, \infty)$.
- **Distribution of $\tilde{Y}$:**
  - By Lemma 2, the drift $\Delta Z'_\cdot(0) - \lambda \Delta X_\cdot$ satisfies the Novikov condition.
  - By Girsanov's Theorem, the distribution of $\tilde{B}$ is equivalent to a Brownian motion, and therefore, $\tilde{Y} \sim \mathcal{P}_{\lfloor \infty \rfloor}\delta_{\tilde{y}}$.

Summary and Conclusion

Key Challenge: Choosing the right space for Z

Space	Markov Property	SPDE Charac.	Uniqueness of Stat. Dist.
$\mathbb{C}[0,\infty)$	Yes	No	Unknown
$\mathbb{C}^1[0,\infty)$	Yes	Yes	Unknown
$\mathbb{L}_2(0,\infty)$	Unknown	No	Yes
$\mathbb{H}^1(0,\infty)$	Yes	Yes	Yes

Summary and Conclusions:

- Introduced a more tractable SPDE framework for the study of diffusion limits of many-server queues
- Use of the asymptotic coupling method (as opposed to Lyapunov function methods) to establishing stability properties of queueing networks: more suitable for establishing uniqueness of stationary distributions of infinite-dimensional processes
- Strengthened the Gamarnik-Goldberg tightness result to convergence of the X-marginal

Open Problem

The main open problem:

to characterize the unique invariant distribution of Y

A possible approach: to use the generator of the process Y.

Lemma

For all $F : \mathbb{R} \times \mathbb{H}^1(0,\infty) \cap \mathbb{C}^1[0,\infty) \to \mathbb{R}$ *with a representation* $F(x,z) = f(x, z^{(n)})$ *for some* $n \geq 1$, $r_1, ..., r_n \in [0,\infty)$, $z^{(n)} = (z(r_1), ..., z(r_n))$ *and* $f \in \mathbb{C}^1(\mathbb{R}^{n+1})$,

$$\mathcal{L}F(x,z) = f_x(x,z^{(n)})\{z'(0) - \beta\} + \sum_{j=1}^{n} f_j(x,z^{(n)})\{z'(r_j) - \bar{G}(r_j)m(x,z)\}$$

$$+ f_{xx}(x,z^{(n)}) + \sum_{j=1}^{n} b_j(x) f_{xj}(x,z^{(n)}) + \frac{1}{2}\sum_{i,j=1}^{n} c_{i,j}(x) f_{i,j}(x,z^{(n)}).$$

where b_i *and* $c_{i,j}$*'s are piecewise constant functions with discontinuity only at* $x = 0$, *and*

$$m(x,z) = \begin{cases} \beta & \text{if } x \leq 0 \\ z'(0) & \text{if } x > 0 \end{cases}.$$

Selected References

- Halfin, S. and Whitt, W. (1981). *Heavy-traffic limit theorems for queues with many servers*, Oper. Res. 29, 567-588.
- Hairer, M., Mattingly, J. C. and Scheutzow, M. (2011). *Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations.* Probab. Theory Related Fields 149, 223-259.
- Puhalskii, A.A. and Reiman, M.I (2000) *The Multiclass $GI/PH/N$ Queue in the Halfin-Whitt Regime.* Adv. Appl. Probab. 32, no. 2, pp. 564–595.
- Gamarnik, D. and Momcilovic, P. (2008) *Steady-state analysis of a multi-server queue in the Halfin-Whitt regime*, Advances in Applied Probability, 40, 548577.
- Reed, J. (2009) *The $G/GI/N$ queue in the Halfin-Whitt regime*, Ann. Appl. Probab. 19 No. 6, 2211-2269.
- Puhalskii, A.A. and Reed, J. (2010) *On many servers queues in heavy traffic*, Ann. Appl. Probab. 20 No.1, 129-195.
- Kang, W. and Ramanan, K. (2010) *Fluid limits of many-server queues with reneging* Ann. Appl. Probab. 20:2204-2260
- Kaspi, H. and Ramanan, K. (2013) *Law of large numbers limits for many-server queues*, Ann. Appl. Probab. 21 No. 1, 33-114..
- Kaspi, H. and Ramanan, K. (2013) *SPDE limits of many-server queues*, Ann. Appl. Probab. 23 No. 1, 145-229.