Stein's Method for Steady-State Diffusion Approximations

Anton Braverman
and
Jim Dai

School of Operations Research and Information Engineering
Cornell University

May 25, 2015

Outline

Diffusion approximations

- Many-server queues:
  multi-dimensional piecewise Ornstein–Uhlenbeck (OU) processes.
- Networks of single server queues:
  multi-dimensional semimartingale reflecting Brownian motions (SRBMs).

Current status:

- There is a huge literature on stochastic process convergence.
- However, there is little work on rate of convergence for steady-state approximations.

Outline

Diffusion approximations

- Many-server queues:
  multi-dimensional piecewise Ornstein–Uhlenbeck (OU) processes.
- Networks of single server queues:
  multi-dimensional semimartingale reflecting Brownian motions (SRBMs).

Current status:

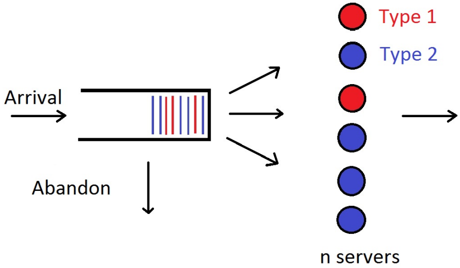
- There is a huge literature on stochastic process convergence.
- However, there is little work on rate of convergence for steady-state approximations.

Barbour, A.D., Stein's method for diffusion approximations, 1990.

Many-server queues

An $M/Ph/n + M$ System

We consider a sequence of $M/Ph/n + M$ queues indexed by the number of servers n.



- Arrival rate λ_n.
- Phase-type service times with mean service time $1/\mu$.
- A waiting customer abandons the queue when his waiting time exceeds his patience, which is exponentially distributed with rate α.

Phase-Type Random Variables

Definition (Neuts 1981)

A phase-type random variable corresponds to the hitting time of a continuous time Markov chain (CTMC) to an absorbing state. Inputs: (p, P, ν).

- For example, an H_2 (hyper-exponential) random variable S has the following representation:

$$S = \begin{cases} S_1 & \text{with probability } p_1, \\ S_2 & \text{with probability } p_2, \end{cases} \quad \text{and} \quad \begin{array}{l} p_1 + p_2 = 1, \\ S_i \sim \text{ Exponential}(\nu_i), \end{array}$$

$$\text{mean service time } = \frac{1}{\mu} = p_1 \frac{1}{\nu_1} + p_2 \frac{1}{\nu_2},$$

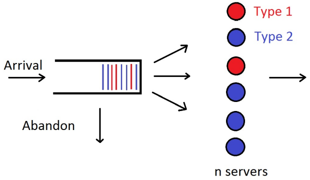
$$p = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \quad P = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \nu = \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}.$$

A Markov Chain Representation

The system can be modeled as a continuous time Markov chain (CTMC)

$$U^n = \Big\{ U^n(t) \in \big(\{1, 2, \ldots, d\}\big)^\infty, t \geq 0 \Big\}.$$

- An example of the state u is given by



$$u = (1, 2, 1, 2, 2, 2, |, 2, 1, 2, 2, 1, 1, 2, 2)$$

- Because of customer abandonment, the CTMC U^n is positive recurrent.

An $M/Ph/n+M$ System (cont.)

- Let $X_1^n(t)$ be the number of type 1 customers in system at time t.
- Let $X_2^n(t)$ be the number of type 2 customers in system at time t.
- Denote

$$X^n(\infty) = \Big(X_1^n(\infty), X_2^n(\infty)\Big)$$

  to be the random vector having the stationary distribution.
- The computation of the distribution of $X^n(\infty)$ can be expensive or unrealistic.

Fig. 1 of Dai-He (2013): an $M/H_2/500 + M$ System

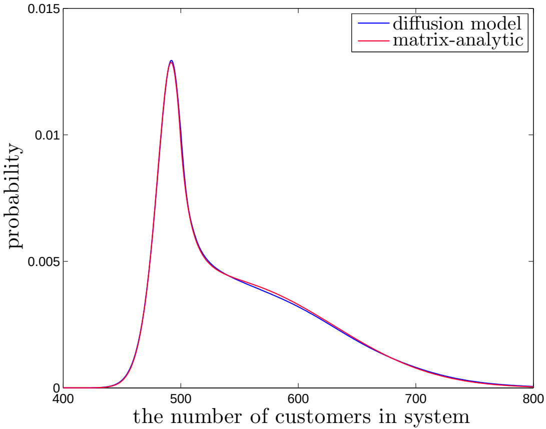


Figure : $\lambda = 522$, $p = (0.9351, 0.0649)$, $1/\nu = (0.1069, 13.89)$, mean patience time $= 2$.

Fig. 2 of Dai-He (2013): an $M/H_2/20 + M$ System

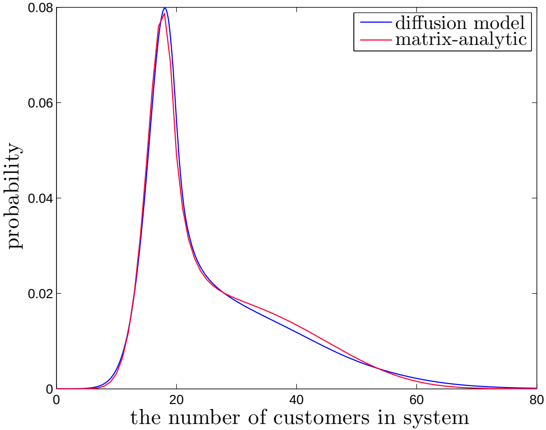


Figure : $\lambda = 22.2$, $p = (0.9351, 0.0649)$, $1/\nu = (0.1069, 13.89)$, mean patience time $= 2$.

Main Results

- Let $\beta \in \mathbb{R}$ be fixed.
- Assume the number of server n follows the square-root safety staffing rule:

$$n\mu = \lambda_n + \beta\sqrt{\lambda_n}, \tag{1}$$

  where λ_n is the arrival rate.
- The sequence of systems is in the Quality- and Efficiency-Driven (QED) regime, also known as the Halfin-Whitt (1981) regime.

Theorem 1

There exists a constant $C = C(\alpha, \beta, p, P, \nu)$ such that

$$\sup_{h \in \text{Lip}(1)} \left| \mathbb{E}h(\tilde{X}^n(\infty)) - \mathbb{E}h(Y(\infty)) \right| \le C\lambda_n^{-1/4}, \quad \forall n \ge 1, \tag{2}$$

where

$$\tilde{X}^n(t) = \frac{1}{\sqrt{\lambda_n}}\Big(X^n(t) - \gamma n\Big), \qquad \gamma_i = \frac{p_i/\nu_i}{p_1/\nu_1 + p_2/\nu_2}.$$

Main Results (cont.)

Theorem 2

For each integer $m > 0$, there exists a constant $C_m > 0$ such that if $h(x) : \mathbb{R}^d \longrightarrow \mathbb{R}$ satisfies

$$|h(x)| \leq |x|^m, \quad \text{for } x \in \mathbb{R}^d,$$

then

$$\left|\mathbb{E}h(\tilde{X}^n(\infty)) - \mathbb{E}h(Y(\infty))\right| \leq C_m \lambda_n^{-1/4}, \quad \forall n \geq 1.$$

Piece-Wise Ornstein-Uhlenbeck (OU) Process

A piece-wise OU process in $\mathbb{R}$ is a diffusion process satisfying

$$Y(t) = Y(0) + \sigma B(t) + \theta t + \alpha_1 \int_0^t Y(s)^- ds - \alpha_2 \int_0^t Y(s)^+ ds.$$

- $B = \{B(t), t \geq 0\}$ is the one-dimensional standard Brownian motion.
- When $\alpha_1 = \alpha_2 = \alpha$, Y becomes a $(\sigma^2, \theta, \alpha)$-OU process whose stationary distribution is normal
$$N\Big(\theta/\alpha, \sigma^2/(2\alpha)\Big).$$
- The generator is
$$Gf(x) = \frac{1}{2}\sigma^2 f''(x) + \theta f'(x) - \alpha x f'(x) \quad \text{for } f \in C^2(\mathbb{R}).$$

Piece-Wise Ornstein-Uhlenbeck (OU) Process (cont.)

- Let $Y = \{Y(t) \in \mathbb{R}^d, t \geq 0\}$ be the piece-wise OU process satisfying

$$\begin{aligned} Y(t) &= Y(0) - p\beta t - R\int_0^t \Big(Y(s) - p(e'Y(s))^+\Big) ds \\ &\quad -\alpha p \int_0^t (e'Y(s))^+ ds + \sqrt{\Sigma} B(t). \end{aligned}$$

- $B(t)$ is the standard d-dimensional Brownian motion,

$$\Sigma = \text{diag}(p) + \sum_{k=1}^{d} \gamma_k \nu_k H^k + (I - P^T)\text{diag}(\nu)\text{diag}(\gamma)(I - P),$$

$$H^k_{ii} = P_{ki}(1 - P_{ki}), \quad H^k_{ij} = -P_{ki}P_{kj} \quad \text{for } j \neq i.$$

- $e' = (1, \ldots, 1)$ and $R = (I - P')\text{diag}(\nu)$.
- The drift vector

$$b(x) = -\beta p - R(x - p(e'x)) - \alpha p(e'x)^+ \quad x \in \mathbb{R}^d.$$

Gurvich, Huang and Mandelbaum (2014, MOR)

- In an $M/M/n+M$ system (exponentially distributed service times), known as the Erlang-A model,

$$\sup_{x\in\mathbb{R}} \left| \mathbb{P}\left\{ \tilde{X}(\infty) \le x \right\} - \mathbb{P}\left\{ Y(\infty) \le x \right\} \right| \le C(\alpha,\mu)\frac{1}{\sqrt{\lambda}} \quad \text{for } \lambda \ge 1. \qquad (3)$$

- Universal for any $n \ge 1$ without assuming (1), but require a new centering at $x(\infty)$, where $x(\infty)$ satisfies

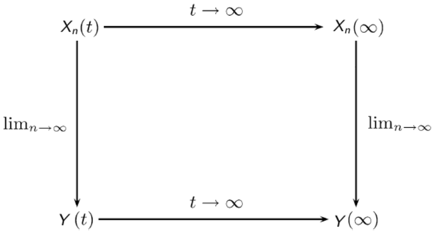
$$\lambda = n\mu + \alpha(x(\infty) - n)^+.$$

- Using an excursion-based approach.

Gurvich (2014, AAP)

- Establishes a framework, "reinventing" Stein's method in the context of steady-state diffusion approximations.
- Introduces generator coupling.
- Relies on the existence of a Lyapunov function $V(x)$ to establish uniform geometric ergodicity for the diffusion processes $\{Y^n, n \geq 1\}$.
- The CTMCs $\{X^n\}$ and the diffusion processes $\{Y^n\}$ have the same dimension; as a consequence, state space collapse (SSC) is not explored.

Standard method: Limit Interchange



- Prove process convergence $X^n(\cdot) \Rightarrow Y(\cdot)$; D-He-Tezcan (2010).
- Process convergence does not imply $X^n(\infty) \Rightarrow Y(\infty)$.
- Justify the limit interchange; D-Dieker-Gao (2014)

$$\lim_{n\to\infty} \lim_{t\to\infty} X^n(t) = \lim_{t\to\infty} \lim_{n\to\infty} X^n(t).$$

Limit Interchange Justifications

- Networks of single-server queues
  - Gamarnik & Zeevi (2006)
  - Budhiraja & Lee (2009)
  - Zhang & Zwart (2008)
  - Katsuda (2010, 2011)
  - Yao & Ye (2012)
  - Gurvich (MOR, 2014)
- Many-server systems
  - Tezcan (2008)
  - Gamarnik & Stolyar (2012)
  - D., Dieker & Gao (2014)
- No convergence rates

An outline of the proof for Theorem 1

Poisson equation

- Let

$$G_Y f(x) = \frac{1}{2}\sum_{i,j=1}^{d} \Sigma_{ij}\partial_{ij} f(x) + \sum_{i=1}^{d} \partial_i f(x) b_i(x) \quad \text{for } x \in \mathbb{R}^d,$$

be the generator of the piecewise OU process $Y = \{Y(t) \in \mathbb{R}^d, t \geq 0\}$.

- For $h : \mathbb{R}^d \to \mathbb{R}$, find a solution f_h to the Poisson equation

$$G_Y f_h(x) = h(x) - \mathbb{E}h(Y(\infty)). \tag{4}$$

- Then, the following Stein equation holds

$$\mathbb{E}\big[h(\tilde{X}^n(\infty))\big] - \mathbb{E}\big[h(Y(\infty))\big] = \mathbb{E}\big[G_Y f_h(\tilde{X}^n(\infty))\big]. \tag{5}$$

Basic Adjoint Relationship (BAR)

Lemma

Random vector $Y(\infty) \in \mathbb{R}^d$ *has the stationary distribution of the diffusion process* $Y = \{Y(t), t \geq 0\}$ *if and only if the following basic adjoint relationship (BAR) holds:*

$$\mathbb{E}\big[G_Y f(Y(\infty)\big] = 0 \quad \text{for all "good" } f \in C^1(\mathbb{R}^d). \tag{6}$$

- Echeverria (1982): Markov processes without boundary.
- Weiss (1981): Markov processes with boundaries.
- Harrison and Williams (1987), Dai and Kurtz (1994): semimartingale reflecting Brownian motions (SRBMs).
- Glynn and Zeevi (2008) provides conditions on f for (6) to hold for Markov chains.

Generator Coupling

- From the Stein equation,

$$\begin{aligned}\mathbb{E}\big[h(\tilde{X}^n(\infty))\big] - \mathbb{E}\big[h(Y(\infty))\big] &= \mathbb{E}\big[G_Y f_h(\tilde{X}^n(\infty))\big] \\ &= \mathbb{E}\big[G_Y f_h(\tilde{X}^n(\infty))\big] - \mathbb{E}\big[G_{U^n} f_h(\tilde{X}^n(\infty))\big] \\ &= \mathbb{E}\Big[G_Y f_h(\tilde{X}^n(\infty)) - G_{U^n} f_h(\tilde{X}^n(\infty))\Big].\end{aligned}$$

- Doing Taylor expansion on $G_{U^n} f_h(x)$ to bound

$$|G_Y f_h(x) - G_{U^n} f_h(x)| \qquad \text{for } x = \frac{i - \gamma n}{\sqrt{\lambda_n}} \text{with } i \in \mathbb{Z}_+^d.$$

Taylor Expansion: An Illustration for $M/M/n+M$

- Set $x = \frac{i-n}{\sqrt{\lambda_n}}$ for $i \in \mathbb{Z}_+$. The generator of the birth-death process $\tilde{X}^n$ is

$$\begin{aligned} G_{\tilde{X}^n} f(x) &= \lambda_n \Big(f\big(x + \frac{1}{\sqrt{\lambda_n}}\big) - f(x) \Big) \\ &\quad + \big(\mu(i \wedge n) + \alpha(i-n)^+\big) \Big(f\big(x - \frac{1}{\sqrt{\lambda_n}}\big) - f(x) \Big). \end{aligned}$$

- The generator of Y is

$$G_Y f(x) = \frac{\lambda_n - n\mu}{\sqrt{\lambda_n}} f'(x) + \left(\mu x^- - \alpha x^+\right) f'(x) + f''(x).$$

- Using Taylor expansion, we can write

$$\begin{aligned} G_X^n f_h(x) - G_Y f_h(x) &= \frac{f_h''(x)}{2\sqrt{\lambda_n}} \Big[\beta + \alpha x^+ - \mu x^- \Big] + \frac{1}{6} f_h'''(\xi) \frac{1}{\sqrt{\lambda_n}} \\ &\quad - \frac{1}{\sqrt{\lambda_n}} \Big(\frac{n\mu}{\lambda_n} + \frac{1}{\sqrt{\lambda_n}} \big(-\mu x^- + \alpha x^+\big) \Big) \frac{1}{6} f_h'''(\eta). \end{aligned}$$

Proof

Therefore, for any Lipschitz continuous h, one has

$$\left|\mathbb{E}h(\tilde{X}^n(\infty)) - \mathbb{E}h(Y(\infty))\right| = \left|\mathbb{E}G_X f_h(\tilde{X}^n(\infty)) - \mathbb{E}G_Y f_h(\tilde{X}^n(\infty))\right|$$
$$\leq \frac{||f_h''||}{2\sqrt{\lambda_n}}\Big(\beta + (\alpha+\mu)B\Big) + \frac{||f_h'''||}{6\sqrt{\lambda_n}}\Big(1 + \frac{n\mu}{\lambda_n} + \frac{1}{\sqrt{\lambda_n}}(\alpha+\mu)B\Big),$$

where $||g|| = \sup_x |g(x)|$ and $B \equiv \sup_n \mathbb{E}\left|\tilde{X}^n(\infty)\right| < \infty$.

Lemma (Gradient Bounds in One-Dimension)

There exists a constant $C = C(\alpha, \beta, \mu) > 0$ such that, for any h that is Lipschitz continuous, the solution f_h to Poisson equation

$$G_Y f_h(x) = h(x) - \mathbb{E}[h(Y(\infty))]$$

satisfies

$$||f_h'|| \leq C||h'||, \quad ||f_h''|| \leq C||h'||, \quad ||f_h''|| \leq C||h'||.$$

Multi-dimensional Gradient Bounds

Lemma (Gurvich (2015))

Suppose $|h(x)| \le |x|^{2m}$ for some $m > 0$, then the solution to Poisson equation

$$G_Y f_h(x) = h(x) - \mathbb{E}h(Y(\infty))$$

satisfies

$$|f(x)| \le C_m(1 + |x|^2)^m,$$

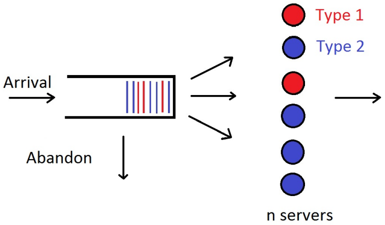
$$|Df(x)| \le C_m(1 + |x|^2)^m(1 + |x|),$$

$$\left|D^2 f(x)\right| \le C_m(1 + |x|^2)^m(1 + |x|)^2,$$

$$\sup_{|y-x|<1, y\neq x} \frac{\left|D^2 f(x) - D^2 f(y)\right|}{|x - y|} \le C_m(1 + |x|^2)^m(1 + |x|)^3.$$

Generator Coupling

- Recall that $f_h : \mathbb{R}^2 \to \mathbb{R}$ is a solution to the Poisson equation.
- $\mathbb{E}\big[G_{U^n} f_h(\tilde{X}^n(\infty))\big]$ is not well defined.
- With a general phase-type service distribution, system size process $\{(X_1^n(t), X_2^n(t)), t \geq 0\}$ is no longer a CTMC.
- U^n is a CTMC living on state space $\mathcal{U} = \{1, 2\}^\infty$. Its generator G_{U^n} acts on functions $F : \mathcal{U} \to \mathbb{R}$.
- BAR gives $\mathbb{E}\big[G_{U^n} A f_h(U^n(\infty))\big] = 0$.



Applying Gradient Bounds

- Recall $\delta = \frac{1}{\sqrt{\lambda_n}}$.
- For $u = (1, 2, 1, 2, 2, 2, |, 2, 1, 2, 2, 1, 1, 2, 2)$, $(z_1, z_2, q_1, q_2) = (2, 4, 3, 5)$ and

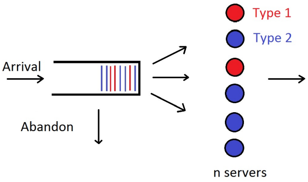
$$\begin{aligned} G_U A f_h(u) = {} & \lambda p_1 f_h(x_1 + \delta, x_2) + \lambda p_2 f_h(x_1, x_2 + \delta) \\ & + \alpha q_1 f_h(x_1 - \delta, x_2) + \alpha q_2 f_h(x_1, x_2 - \delta) \\ & + z_1 \nu_1 f_h(x_1, x_2) + z_1 \nu_1 f_h(x_1, x_2 - \delta) \\ & - (\lambda + \alpha q + z_1 \nu_1 + z_2 \nu_2) f_h(x_1, x_2). \end{aligned}$$

Lemma

There exists a constant $C(m) = C(m, \beta, \alpha, p, \nu, P)$ such that for any $u \in \mathcal{U}$,

$$\begin{aligned} |G_U A f_h(u) - G_Y f_h(x)| \quad \leq \quad & C(m)(1 + |x|^2)^m (1 + |x|) \left| \delta q - p(e^T x)^+ \right| \\ & + \delta\, C(m)(1 + |x|^2)^m (1 + |x|)^4. \end{aligned}$$

State Space Collapse in $M/Ph/n+M$ Case



- The total queue size $\left(X_1^n(\infty) + X_2^n(\infty) - n\right)^+ = (e'X^n(\infty) - n)^+$.

Lemma (State-Space Collapse)

There exists $C(m) > 0$ such that $\forall n \geq 1$,

$$\mathbb{E}\left|\delta\left(Q_i^n(\infty) - p_i(e'X^n(\infty) - n)^+\right)\right|^{2m} \leq C(m)\delta^m \mathbb{E}[(e'\tilde{X}^n(\infty))^+]^m \quad \text{for } i = 1, 2.$$

Is the Convergence Rate in Theorem 1 Sharp?

λ_n	$\mathbb{E}\left[\tilde{X}^n(\infty)^2\right]$	$\mathbb{E}\left[Y(\infty)^2\right]$	Diff. (δ_n)	δ_n/δ_{2n}
50	1.3811	1.4221	0.0410	
100	1.3929	1.4221	0.0292	$1.4045 = \sqrt{1.9725}$
200	1.4013	1.4221	0.0208	$1.4069 = \sqrt{1.9793}$

Table : $M/H_2/n+M$ queue with parameters $p = (67.41\%, 32.59\%)$, $\nu = (0.6741, 0.3259)$ and $\beta = 0$.

- If $\lambda_n^{-1/4}$ is a sharp convergence rate, expect that as λ_n doubles, error decreases by a factor of $2^{1/4} = 1.1892$.
- If $\lambda_n^{-1/2}$ is a sharp convergence rate, expect that as λ_n doubles, error decreases by a factor of $2^{1/2} = 1.4142$.
- Error appears to decrease at a rate of $\lambda_n^{-1/2}$.

Networks of single-server queues and open problems

A $G/G/1$ Queue

Consider a single-server queue operating under first-come-first-serve discipline.

- $A, A_1, A_2, \ldots$ i.i.d. inter-arrival times with mean $1/\lambda = 1$.
- $S, S_1, S_2, \ldots$ i.i.d. service times with mean m.
- Traffic intensity $\rho = \lambda m = m$.

Lindley recursion for waiting times:

- Recursive formula for W_n – the nth customer's waiting time in queue:

$$W_{n+1} = (W_n + S_n - A_{n+1})^+, \qquad x^+ := \max(x, 0).$$

- A_n, S_n – inter-arrival and service time of nth customer, respectively.

Steady-State Behavior in Heavy Traffic

- Steady-state customer waiting time $W(\infty)$.
- As $\rho = m \uparrow 1$, $W(\infty) \to \infty$.
- The scaled version $\widetilde{W} = (1-\rho)W(\infty)$ does not blow up.
- $$\widetilde{W}^* \stackrel{d}{=} (\widetilde{W} + (1-\rho)X)^+,$$

  where

  $$\widetilde{W}^* \stackrel{d}{=} \widetilde{W}, \quad X \perp \widetilde{W}, \quad X \stackrel{d}{=} S - A, \quad \mathbb{E}X = m - \frac{1}{\lambda} = \rho - 1.$$
- Define

  $$G_{\widetilde{W}} f(w) := \mathbb{E}\Big[f\big((w + (1-\rho)X)^+\big)\Big] - f(w), \quad w \geq 0.$$

Basic Adjoint Relationship (BAR)

For all 'nice' functions f, we have BAR

$$\mathbb{E}\Big[G_{\widetilde{W}} f(\widetilde{W})\Big] = \mathbb{E}\Big[f\Big((\widetilde{W} + (1-\rho)X)^+\Big) - f(\widetilde{W})\Big] = 0,$$

where $\widetilde{W}$ and X are independent.

- Suppose $f \in C^3(\mathbb{R})$, use Taylor expansion:

$$\begin{aligned}
&\mathbb{E}\Big[f\Big((\widetilde{W} + (1-\rho)X)^+\Big) - f(\widetilde{W})\Big] \\
=&\mathbb{E}\Big[f\big(\widetilde{W} + (1-\rho)X\big) - f(\widetilde{W}) + \Big(f(0) - f(\widetilde{W} + (1-\rho)X)\Big)1_{\{\widetilde{W}+(1-\rho)X\le 0\}}\Big] \\
=&\mathbb{E}\Big[f'(\widetilde{W})(1-\rho)\mathbb{E}X + \frac{1}{2}f''(\widetilde{W})(1-\rho)^2\mathbb{E}X^2 - f'(0)(1-\rho)\mathbb{E}X\Big] \\
&+\mathbb{E}\Big[\frac{1}{6}(1-\rho)^3 f'''(\xi)\mathbb{E}X^3 - \frac{1}{2}(\widetilde{W} + (1-\rho)X)^2 f''(\eta)1_{\{\widetilde{W}+(1-\rho)X\le 0\}}\Big],
\end{aligned}$$

where we have used

$$\mathbb{E}\Big[(\widetilde{W} + (1-\rho)X)1_{\{\widetilde{W}+(1-\rho)X\le 0\}}\Big] = (1-\rho)\mathbb{E}X.$$

Poisson Equation and Gradient Bounds

- Consider Poisson equation

$$G_Z f_h(w) := \frac{1}{2}\sigma^2 f_h''(w) - \theta f_h'(w) + \theta f_h'(0) = h(w) - \mathbb{E}h(Z),$$

where

$$\sigma^2 = (1-\rho)^2 \mathbb{E}X^2, \quad \theta = -(1-\rho)\mathbb{E}X > 0$$

and Z is an exponential random variable with mean $\sigma^2/2\theta$.

- A solution satisfying $f_h'(0) = 0$ also satisfies

$$\|f_h''\| \le \frac{\|h'\|}{\theta} \quad \text{and} \quad \|f_h'''\| \le \frac{4}{\sigma^2}\|h'\|.$$

$G/G/1$ Waiting Time Approximation

Using Stein equation

$$\begin{aligned}\mathbb{E}h(\widetilde{W}) - \mathbb{E}h(Z) &= \mathbb{E}\Big[G_Z f_h(\widetilde{W})\Big] - \mathbb{E}\Big[G_W f_h(\widetilde{W})\Big] \\ &= (1-\rho)^3 \mathbb{E}\Big[\frac{1}{6} f'''(\xi)\Big]\mathbb{E}X^3 \\ &\quad - \mathbb{E}\Big[\frac{1}{2}(\widetilde{W} + (1-\rho)X)^2 f''(\eta) 1_{\{\widetilde{W}+(1-\rho)X \le 0\}}\Big],\end{aligned}$$

we obtain:

Lemma

Assume $\mathbb{E}X^3 < \infty$. Then,

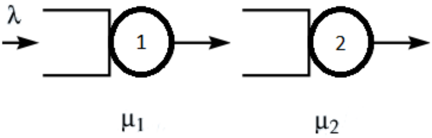
$$d_{\mathcal{W}}\Big(\widetilde{W}, Z\Big) \le C\sqrt{(1-\rho)}.$$

Furthermore, if $\mathbb{E}X^m < \infty$ for all $m \ge 1$, then for any $\epsilon > 0$, there exists a constant C_ϵ such that

$$d_{\mathcal{W}}\Big(\widetilde{W}, Z\Big) \le C_\epsilon (1-\rho)^{1-\epsilon}.$$

Multidimensional SRBMs

Consider the $M/M/1 \to \cdot/M/1$ tandem system, we are interested in the queue lengths.



- Assume $\lambda = 1$. Heavy traffic: $\mu_i = \mu_i^{(n)}$ and $\lambda - \mu_i^{(n)} = -\beta_i/\sqrt{n} < 0$.
- The approximating diffusion process is a two-dimensional semimartingale reflecting Brownian motion (SRBM)

$$Z = \{(Z_1(t), Z_2(t)) \in \mathbb{R}_+^2, t \geq 0\}.$$

- See Williams (1995) for a review of SRBMs.

PDE in an orthant with oblique boundary derivatives

Open Problem

Consider the operator

$$\mathcal{A}_n f(x) = \frac{1}{2}\sum_{i,j=1}^{2} \Sigma_{ij} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{2} \nu_i \frac{\partial f(x)}{\partial x_i} + \sum_{i=1}^{2} \beta_i \langle R^{(i)}, \nabla f(x)|_{x_i=0} \rangle,$$

where

$$\nu = \frac{1}{n}\begin{pmatrix} -\beta_1 \\ \beta_1 - \beta_2 \end{pmatrix}, \quad \Sigma = \frac{1}{n}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad R = \frac{1}{n}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

and $R^{(i)}$ is the ith column of R. If $h : \mathbb{R}_+^2 \to \mathbb{R}$ is a Lipschitz-1 function, **under what conditions on** $\langle R^{(i)}, \nabla f(x)|_{x_i=0} \rangle$, does the solution to the PDE

$$\mathcal{A}_n f_h(x) = h(x) - \mathbb{E}h(Z_n(\infty))$$

satisfy

$$\|D^2 f_h\| \le C_1 n \quad \text{and} \quad \|D^3 f_h\| \le C_2 n.$$

References I

Braverman, A. and Dai, J. G. (2015).
Stein's method for steady-state diffusion approximations.
Preprint, URL http://people.orie.cornell.edu/jdai/publications/preprints.html.

Dai, J. G. and He, S. (2013).
Many-server queues with customer abandonment: Numerical analysis of their diffusion model.
Stochastic Systems, **3** 96–146.
URL http://dx.doi.org/10.1214/11-SSY029.

Dieker, A. and Gao, X. (2013).
Positive recurrence of piecewise Ornstein–Uhlenbeck processes and common quadratic Lyapunov functions.
The Annals of Applied Probability, **23** 1291–1317.
URL http://dx.doi.org/10.1214/12-AAP870.

References II

GAMARNIK, D. and ZEEVI, A. (2006).
Validity of heavy traffic steady-state approximation in generalized Jackson networks.
Ann. Appl. Probab., **16** 56–90.

GLYNN, P. W. and ZEEVI, A. (2008).
Bounding stationary expectations of Markov processes.
In *Markov processes and related topics: a Festschrift for Thomas G. Kurtz*, vol. 4 of *Inst. Math. Stat. Collect.* Inst. Math. Statist., Beachwood, OH, 195–214.
URL http://dx.doi.org/10.1214/074921708000000381.

GURVICH, I. (2014).
Diffusion models and steady-state approximations for exponentially ergodic Markovian queues.
The Annals of Applied Probability, **24** 2527–2559.
URL http://dx.doi.org/10.1214/13-AAP984.

References III

GURVICH, I., HUANG, J. and MANDELBAUM, A. (2014).
Excursion-based universal approximations for the Erlang-A queue in steady-state.
Mathematics of Operations Research, **39** 325–373.
URL http://dx.doi.org/10.1287/moor.2013.0606.

REIMAN, M. I. (1984).
Open queueing networks in heavy traffic.
Mathematics of Operations Research, **9** 441–458.

WILLIAMS, R. J. (1995).
Semimartingale reflecting Brownian motions in the orthant.
In *Stochastic Networks* (F. P. Kelly and R. J. Williams, eds.), vol. 71 of *The IMA Volumes in Mathematics and its Applications*. Springer, New York, 125–137.

Gradient Bounds for Elliptic PDEs

- Based on Gurvich (2015).
- Consider the elliptic differential operator

$$Lf(x) = \sum_{1 \leq i,j \leq d} a_{ij} D_{ij} f(x) + \sum_{1 \leq i \leq d} b_i(x) D_i f(x).$$

- The matrix A defined by $A_{ij} = a_{ij}$ is positive definite.
- $b(x) = (b_1(x), ..., b_d(x))$ satisfies the Lipschitz condition

$$|b(x) - b(y)| \leq c_b \, |x - y| \, .$$

Schauder Interior Estimates

- For $x \in \mathbb{R}^d$, let $B_x = \{y \in \mathbb{R}^d : |y - x| \le \frac{1}{1+|x|}\}$.

Lemma (Gilbarg & Trudinger (2001))

Let $f(x)$ be a solution to the PDE

$$Lf(x) = h(x).$$

There exists a constant C depending only on A and c_b, such that

$$|Df(x)| + \left|D^2 f(x)\right| + \sup_{y,z \in B_x, y \neq z} \frac{\left|D^2 f(z) - D^2 f(y)\right|}{|z-y|}$$

$$\le C\left(\sup_{y \in B_x} |f(y)| + \sup_{y \in B_x} |h(y)| + \sup_{y,z \in B_x, y \neq z} \frac{|h(z) - h(y)|}{|z-y|}\right)(1+|x|)^3.$$

Lyapunov Functions

- If the elliptic operator L is the generator of some diffusion process $Y = \{Y(t), t \geq 0\}$, then the solution to

$$G_Y f(x) = h(x) - \mathbb{E}h(Y(\infty)) =: \tilde{h}(x)$$

satisfies

$$f(x) = \int_0^\infty \mathbb{E}_x \tilde{h}(Y(t))dt.$$

- Suppose we know that

$$|\mathbb{E}_x h(Y(t)) - \mathbb{E}h(Y(\infty))| \leq V(x)e^{-\eta t}, \quad \eta > 0.$$

- Then

$$|f(x)| \leq \int_0^\infty \left|\mathbb{E}_x \tilde{h}(Y(t))\right| dt \leq CV(x).$$