Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

EXACT GRADIENT SIMULATION FOR STOCHASTIC FLUID NETWORKS IN STEADY STATE

Xinyun Chen

Department of Applied Mathematics and Statistics
Stony Brook University
Stony Brook, NY 11794, USA

ABSTRACT

In this paper, we develop a new simulation algorithm that generates unbiased gradient estimators for the steady-state workload of a stochastic fluid network, with respect to the throughput rate of each server. Our algorithm is based on the perfect sampling algorithm developed in Blanchet and Chen (2014), and the infinitesimal perturbation analysis (IPA) method. We illustrate the performance of our algorithm with two multidimensional examples, including its formal application in the case of multidimensional reflected Brownian motion.

1 INTRODUCTION

Stochastic fluid network is a class of queueing model that is good for telecommunication networks and data processing systems where the main source of uncertainty in the network is the external inputs (see Kella and Whitt 1996). One important performance measure for a stochastic fluid network is the workload process $Y(t)$, which is a vector consisting of the amount of outstanding workload at each server. Performance evaluation of such systems involves computing the steady-state expectation $E[Y^*]$. In many application settings, for instance, the service capacity management problem as discussed in Dieker, Ghosh, and Squillante (2014) , it is of interest to compute the sensitivity of $E[Y^*]$ with respect to some model parameters. However, analytic characterization for the steady-state distribution of Y^* and the steady-state gradient remains an open problem. As an alternative approach, Blanchet and Chen (2014) provides a Monte Carlo algorithm that generates unbiased samples from the steady-state distribution of Y^*. Based on their perfect sampling algorithm, we develop a Monte Carlo algorithm to generate unbiased estimators for the steady-state workload gradient with respect to the throughput rates, or equivalently, the service rates of the servers.

Our algorithm is a combination of the classic Infinitesimal Perturbation Analysis (IPA) and the Coupling from the Past (CFTP) techniques. In detail, we apply IPA to define a stochastic process V coupled with $Y(\cdot)$ such that its steady-state expectation $E[V^*]$ exists and equals the gradient. Then we show that the coalescence time for Y, as defined and simulated in the CFTP perfect sampling algorithm, is also a coalescence time for the process V. As a result, it is possible to generate exact samples from the joint steady-state distribution (Y^*,V^*) and compute unbiased estimators for the gradient.

In the end, we show that V^* can be computed deterministically from the output sample path $\{Y^*(t) : -\tau \leq t \leq 0\}$ of the perfect sampling algorithm Blanchet and Chen (2014). The computation basically involves a fixed number of matrix operation at each time step and does not requires generating extra random numbers. Therefore, the gradient simulation algorithm inherits from the perfect simulation algorithm Blanchet and Chen (2014) a polynomial complexity in term of the number of servers in the network.

A benchmark algorithm for unbiased steady-state gradient simulation is the regenerative algorithm (see Glynn 2006 and the references therein). In our setting, the stochastic fluid network clearly has a regenerative structure. However, it is difficult to specify a proper likelihood ratio, which is a key step in the regenerative algorithm, as a function of the throughput rate μ. In particular, we show in Section 4.1 that in a network consisting of tandem servers, the sample path spaces of the workload process, corresponding to different throughput rates μ and $\mu+(0,\Delta,0,..,0)'$, have empty intersection for any $\Delta \neq 0$. As a consequence, the likelihood ratio can not be defined as a finite function of the throughput rate.

The rest of the paper is organized as following: in Section 2, we state our model assumptions, and give a brief review on the IPA method; in Section 3, we give all the details of our gradient simulation algorithm; in Section 4, we report the numerical results of our algorithm. In particular, in Section 4.2, we apply our algorithm to discrete approximations of reflected Brownian motions (RBM). The numerical results indicate that the gradient estimator of the discrete approximation process is a good approximation for the steady-state gradient of the RBM. Of course, rigorous analysis is in need in future work to validate the convergence result.

Throughout the paper we shall use boldface ($\mathbf{Y}$) for vectors in $\mathbb{R}^d$, and superscript i (Y^i) for the the i-th component of a vector.

2 PROBLEM SETTING AND PRELIMINARIES

2.1 Stochastic Fluid Network

Consider a stochastic fluid network (SFN) consisting of d servers. Jobs arrive to the servers according to some renewal process $N(\cdot)$ with i.i.d. inter-arrival times $\{A_n\}$. In particular, we assume A_n be the inter-arrival time between the n-th and $(n+1)$-th job. Suppose the n-th job brings W_n^i units of workload to server i. Let $\mathbf{W_n}=(W_n^1,...,W_n^d)'$ be the vector of workload and we assume $\{\mathbf{W}_n\}$ forms a sequence of i.i.d. random vectors in $\mathbb{R}^d$. The sequences $\{\mathbf{W_n}\}$ and $\{A_n\}$ are independent. Let r_i be the constant service rate of server i. Let P_{ij} be the proportion of workload that goes to server j after departure server i so that $\sum_j P_{ij} \leq 1$. Then the $d\times d$ matrix $R=(I-P)'$ ($R_{ij}=\delta_i^j-P_{ji}$, $\delta_i^j=1$ if $i=j$, and $=0$ otherwise) is called the reflection matrix. The throughput rate of the network is $\mu=R\mathbf{r}$.

The workload $\mathbf{Y}_n=(Y_n^1,...,Y_n^d)$ is defined as the amount of workload observed by the n-th job upon its arrival. The value of $\mathbf{Y}_{n+1}$ is totally determined by $(\mathbf{Y}_n,\mathbf{W}_n,A_n)$ and (R,μ) via the Skorokhod map. In detail, define a process $\mathbf{x}(t)=\mathbf{Y}_n+\mathbf{W}_n-\mu t$ and its corresponding reflected process

$$\mathbf{y}(t)=\Gamma(\mathbf{x})(t)=\mathbf{x}(t)+R\mathbf{l}(t),$$

where Γ is the Skorokhod map and $\mathbf{l}(t)$ is the minimal nondecreasing nonnegative cádlág process to keep $\mathbf{y}(t)\geq 0$. For rigorous definition of the Skorokhod map, see for instance Harrison and Reiman (1981). Then, $\mathbf{Y}_{n+1}=\mathbf{y}(A_n)$ is determined by $(\mathbf{Y}_n,\mathbf{W}_n,A_n)$ and (R,μ). As result, we can write a Markov recursion for the sequence $\{\mathbf{Y}_n\}$:

$$\mathbf{Y_{n+1}}=\phi(\mathbf{Y_n},\mu,\mathbf{W}_n,A_n,R). \tag{1}$$

In Harrison and Reiman (1981), it is proved that the Skorokhod map, Γ, is Lipschitz continuous in the uniform norm. As a direct consequence, the recursion function ϕ is Lipschitz continuous in μ and $\mathbf{y}=\mathbf{Y}_n$. If the network is stable, its steady-steady distribution is of course sensitive to the throughput rate μ. In many application setting, it is of interest to compute the gradient $\partial E[\mathbf{Y}^*]/\partial\mu$. Note that both $E[\mathbf{Y}^*]$ and μ are d-dimensional vectors and therefore the gradient is a d-by-d matrix. Our goal is to develop a Monte Carlo algorithm that generates i.i.d. sample of random matrix V^* such that $E[V^*]=\partial E[\mathbf{Y}^*]/\partial\mu$, especially, $E[V_{ij}^*]=\partial E[Y^{*,i}]/\partial\mu^j$. For the algorithm to work, we shall impose the following assumptions on the stochastic fluid network:

Assumption 1 $R^{-1}(E[\mathbf{W}_1]-\mu E[A_1])<0$.

Assumption 2 The inter-arrival times A_1 has an infinite support and a continuous distribution function.

Assumption 3 There exists $\theta>0$, $\theta\in\mathbb{R}^d$ such that $E[\exp(\theta'\mathbf{W}_1)]<\infty$.

Remark: Assumption 1 ensures the existence of stationary distribution for the workload process and Assumption 2 for the gradient process, which we shall define in a moment. Assumption 3 ensures the perfect sampling algorithm Blanchet and Chen (2014) is applicable.

2.2 Infinitesimal Perturbation Analysis

Infinitesimal perturbation analysis (IPA) is a gradient simulation approach based on derivatives of the random variables. In our case, the waiting time sequence $\{\mathbf{Y_n}\}$ satisfies a recursion of the form:

$$\mathbf{Y}_{n+1} = \phi(\mathbf{Y}_n, \mu). \tag{2}$$

Here we omit the variables $\mathbf{W}_n$, R and A_n in the function ϕ not only for simplicity, but also because in the IPA algorithm only compute the sensitivity of $\phi(\mathbf{y}, \mu)$ with respect to the initial position $\mathbf{y}$ and the throughput rate μ only, as we shall see in a moment.

Given Assumption 2, we can prove (in Proposition 1) that $\phi(\mathbf{y}, \mu)$ is continuously differentiable in a neighborhood of $(\mathbf{y}, \mu)$ with probability 1. Formally defining $V_n = \frac{d\mathbf{Y}_n}{d\mu} \in \mathbb{R}^{d\times d}$ and differentiating (2), we get

$$V_{n+1} = \frac{d\phi}{d\mathbf{y}}(\mathbf{Y}_n, \mu)V_n + \frac{d\phi}{d\mu}(\mathbf{Y}_n, \mu).$$

Now we substitute $\{\mathbf{Y}_n\}$ with a stationary sequence $\{\mathbf{Y}_n^*\}$ and define a stochastic process V_n in $\mathbb{R}^{d\times d}$ according to

$$V_{n+1} = C_n V_n + B_n, \text{ where } C_n = \frac{d\phi}{d\mathbf{y}}(\mathbf{Y}_n^*, \mu) \text{ and } B_n = \frac{d\phi}{d\mu}(\mathbf{Y}_n^*(\mu), \mu) \in \mathbb{R}^{d\times d}. \tag{3}$$

According to Glasserman (1992), when the recursion function ϕ is Lipschitz continuous and

$$P\left(\frac{d\phi}{d\mathbf{y}}(\mathbf{Y}_n^*, \mu) = 0\right) > 0, \tag{4}$$

the gradient process $\{V_n\}$ is stable. In particular, the process $\{V_n\}$ admits a unique stationary distribution V^* and converges weakly to its stationary distribution, i.e. $V_n \Rightarrow V^*$. Besides, V^* is an unbiased estimator for $\frac{dE[\mathbf{Y}^*]}{d\mu}$ such that $E[V^*] = \frac{dE[\mathbf{Y}^*]}{d\mu}$.

3 EXACT GRADIENT SIMULATION ALGORITHM

In this section, we fit the IPA algorithm into a backwards framework and combine it with coupling from the past (CFTP) technique to develop an unbiased Monte Carlo algorithm that generates exact samples from the stationary distribution of V^*.

3.1 Coalescence Time and Perfect Sampling

In Blanchet and Chen (2014), the authors developed a CFTP algorithm for SFN with renewal arrival process and general workload distributions. The key steps of the algorithm is to construct a coupled dominating process, simulate the stationary dominating process backwards in time until it hits the origin, and then simulate the workload process forward starting from the origin. The first time, going backwards to the past, when the dominating process hits the origin is called the coalescence time and let's denote it by τ. Given Assumption 1 and 2, $E[\tau] < \infty$. Let $n(\tau)$ be number of job arrivals from τ to time 0. Then, the output of the CFTP algorithm is a stationary sequence of $\{\mathbf{Y}_k^*\}$ for $k = -n(\tau), -n(\tau)+1, ..., 0$ with the corresponding sequence of inter-arrival times $\{A_k\}$ and workload vectors $\{\mathbf{W}_k\}$. As we shall prove in Section 3.2, $B_{k-1} = C_{k-1} = 0$, whenever $\mathbf{Y}_k^* = 0$. In Blanchet and Chen (2014), it is proved that under Assumption 1 and 2, $P(\mathbf{Y}_k^* = 0) > 0$. Therefore, (4) holds and the gradient sequence $\{V_k\}$ is stable. Let

$\{V_k^*\}$ be the stationary sequence of the gradient process $\{V_k\}$ coupled with the stationary sequence $\{\mathbf{Y}_k^*\}$ that is simulated by the CFTP algorithm. Then, as $\mathbf{Y}_{-n(\tau)}^* = 0$ indicates $B_{-n(\tau)-1} = C_{-n(\tau)-1} = 0$,

$$V_0^* = \sum_{k=-1}^{-\infty} B_k \prod_{i=k+1}^{-1} C_i = \sum_{k=-1}^{-n(\tau)} B_k \prod_{i=k+1}^{-1} C_i,$$

which becomes a finite summation. As we shall prove in Section 3.2, the matrix B_k and C_k has an explicit expression in terms of $(\mathbf{Y}_k^*, A_k, \mathbf{W}_k)$ and (μ, R) for all $k \in \{-n(\tau), ..., -1\}$. As a result, V_0^* can be computed from the output sequence $\{\mathbf{Y}_k : k = -n(\tau), ..., 0\}$ of the CFTP algorithm and is an unbiased estimation for $\frac{dE[\mathbf{Y}^*]}{d\mu}$.

Now we are ready to give the algorithm to compute V_0^*:

Algorithm 1: Perfect Gradient Simulation

1. Input: A stationary sample path $\{(\mathbf{Y}_k^*, A_k, \mathbf{W}_k) : k = -n(\tau), -n(\tau)+1, ..., 0\}$ with $\mathbf{Y}_{-n(\tau)}^* = 0$.
2. Initially set $V = 0$
3. For $k = -n(\tau) : -1$, update $V = V \cdot C_k + B_k$. Here $C_k = \frac{d\phi}{d\mathbf{y}}(\mathbf{Y}_k^*, \mu)$ and $B_k = \frac{d\phi}{d\mu}(\mathbf{Y}_k^*, \mu)$ are computed according to (6).
4. Output V which follows the stationary distribution of V^*.

3.2 Computation of the Gradient Recursion

The following results give explicit expressions for (B_k, C_k) in terms of $(\mathbf{Y}_k, A_k, \mathbf{W}_k)$ and (μ, R). Moreover, they also show that $C_k = 0$ whenever $Y_{k+1} = 0$. First, let's introduce some notation. Recall that the recursion function ϕ is defined by the Skorokhod map Γ which has an expression $\Gamma(\mathbf{x})(t) = \mathbf{x}(t) + R\mathbf{l}(t)$. Let's define $s_i = \inf\{s : l^i(s) > 0\}$. The following result is crucial to simplify the computation of the gradient matrix (B_k, C_k)

Lemma 1 Let $\mathbf{x}(t) = \mathbf{Y}_k + \mathbf{W}_k - \mu t$, which is a linear function, and suppose $\Gamma(\mathbf{x})(t) = \mathbf{x}(t) + R\mathbf{l}(t)$. Then, the process $\mathbf{l}(t)$ is piecewise linear and $l^i(t)$ is strictly increasing on $[s_i, \infty)$.

Proof. As is proved in Harrison and Reiman (1981), the complimentary process $\mathbf{l}(t)$ is the unique solution to a fixed point problem

$$l^i(t) = [\sup_{s\in[0,t]} (-x^i(s) - \sum_{j\neq i} R_{ij} l(s))] \vee 0. \tag{5}$$

In our case, the function $\mathbf{x}(t)$ is a linear function and we can solve the fixed point equation (5) explicitly. In particular, the solution $\mathbf{l}(t)$ is a piecewise linear function on $[0, A_k]$. Let $\{t_0 = 0 < t_1 < t_2 < ... < t_m = t_{m+1} =\leq A_k\}$ such that $\mathbf{l}(t)$ is linear on $[t_r, t_{r+1})$. Let $\mathbf{l}_r$ be the linear coefficient of $l(t)$ on $[t_, t_{r+1})$, then the sequence of $\{(t_r, \mathbf{l}_r)\}$ is given by the following recursion:

$$l_r^i = \begin{cases} 0 & \text{if } y^i(t_r) > 0, \\ (\mu^i - \sum_{j\neq i} R_{ij} l_r^j) \vee 0 & \text{if } y^i(t_r) = 0, \end{cases}$$

and

$$t_{r+1} = t_r + \min_{y^j(t_r)>0, \mu_r^j>0} \left(\frac{y^j(t_r)}{\mu_r^j}, A_k \right) \text{ where } \mu_r = \mu + R\mathbf{l}_r.$$

Recall that $s_i = \inf\{s : l^i(s) > 0\}$ and let $I_r = \{i : 1 \leq i \leq d, l_r^i > 0\}$. Given the above recursion, we can check the following facts by induction

1. $y^i(t) = 0$ for all $t > s_i$.
2. I_r is increasing in r.

3. Since $l_r^j \geq 0$ for all j and $R_{ij} < 0$ for all $j \neq i$, l_r^i is non-decreasing in r.

As we have showed, $l(t)$ is piecewise linear, and therefore there exists $\Delta t > 0$ such that $dl^i(t) > 0$ on $[s_i, s_i + \Delta t]$. Then, the claim that $l^i(t)$ is strictly increasing for $t > s^i$ follows immediately after fact 3. □

Given Lemma 1, we are ready to derive an explicit expression for (B_k, C_k) in terms of $(\mathbf{Y}_k, A_k, \mathbf{W}_k)$ and (μ, R).

Proposition 1 With probability 1, $C_k = E + RD$, where E is the $d \times d$ identity matrix and the matrix D can be solved from the linear system

$$D_{ij} = \begin{cases} 0 & \text{if } i \notin I, \\ -E_{ij} - \sum_{r \neq i} R_{ir} D_{rj} & \text{if } i \in I, \end{cases} \tag{6}$$

where

$$I = \{1 \leq i \leq d : [R^{-1}(\mathbf{Y}_{k+1} - \mathbf{Y}_k - \mathbf{W}_k + \mu A_k)]^i > 0\}$$

The matrix $B_k = A_k \cdot C_k$

Remark: Given Assumption 2 that the inter-arrival time A_k has continuous distribution function, with probability 1, the value set function I is constant in a neighborhood of $(\mathbf{y}, \mu)$, where $Y_k = \mathbf{y}$ is the initial value and μ is the throughput rate. Then, since the linear system (6) admits a unique solution for fixed set I, we can conclude that $C_k = \frac{\partial \phi}{\partial \mathbf{y}}(\mathbf{y}, \mu)$ and $B_k = \frac{\partial \phi}{\partial \mu}(\mathbf{y}, \mu)$ is continuous in a neighborhood of (Y_k, μ). In other words, ϕ is continuously differentiable in a neighborhood of (Y_k, μ) with probability 1 and this justified the gradient recursion (3).

Proof. We shall apply the results on direction derivative of Skorokhod map in Mandelbaum and Ramanan (2010) to compute the gradient matrix (B_k, C_k). According to Theorem 1.1. in Mandelbaum and Ramanan (2010), $C_k = C(A_k)$ where $C(t)$ is a matrix-value function such that

$$C(t) = E + RD(t),$$

where $D(t)$ is the unique solution to the system of equations

$$D_{ij}(t) = \begin{cases} 0 & \text{if } t < t_l^i, \\ \sup_{s \in \Phi^i(t)} [-E_{ij} - \sum_{r \neq i} R_{ir} D_{rj}(s)] \vee 0 & \text{if } t \in [t_l^i, t_u^i], \\ \sup_{s \in \Phi^i(t)} [-E_{ij} - \sum_{r \neq i} R_{ir} D_{rj}(s)] & \text{if } t > t_u^i, \end{cases}$$

where $t_l^i = \inf\{t \geq 0 : y^i(t) = 0\}$, $t_u^i = \inf\{t \geq 0 : l^i(t > 0)\}$ and $\Phi^i(t) = \{s \in [0,t] : y^i(s) = 0 \text{ and } l^i(s) = l^i(t)\}$.

Since A_k is a continuous random variable, $P(A_k = t_l^i, \text{ for some } 1 \leq i \leq d) = 0$ and hence $\{i : A_k \in [t_l^i, t_u^i]\}$ is empty with probability 1. By definition, $t_u^i = s_i$. Let's define $\tilde{I} = \{i : A_k > t_u^i\}$. Following the result in Lemma 1, we can conclude that $\Phi^i(t) = \{t\}$ for all $t \geq s_i$. As a result, $\Phi^i(A_k) = \{A_k\}$ for all $i \in \tilde{I}$. In summary, we can simplify the system of equations and get

$$D_{ij}(A_k) = \begin{cases} 0 & \text{if } i \notin \tilde{I}, \\ -E_{ij} - \sum_{r \neq i} R_{ir} D_{rj}(A_k) & \text{if } i \in \tilde{I}. \end{cases}$$

By Lemma 1, $dl^i(t) > 0$ whenever $l^i(t) > 0$. By definition, $\mathbf{l}(A_k) = R^{-1}(\mathbf{y}(A_k) - \mathbf{x}(A_k)) = R^{-1}(\mathbf{Y}_{k+1} - \mathbf{Y}_k - \mathbf{W}_k + \mu A_k)$. Therefore, the set $\tilde{I} = I$ and (6) follows immediately.

Following a similar argument, we can have $B_k = A_k \cdot E + R\tilde{D}$ with

$$\tilde{D}_{ij} = \begin{cases} 0 & \text{if } i \notin I, \\ -A_k E_{ij} - \sum_{r \neq i} R_{ir} \tilde{D}_{rj} & \text{if } i \in I. \end{cases}$$

Therefore, we can conclude $\tilde{D} = A_k \cdot D$ and hence $B_k = A_k C_k$. □

Proposition 2 $B_{-n(\tau)-1} = C_{-n(\tau)-1} = 0$

Proof. Let $\mathbf{y}(t) = \Gamma(\mathbf{Y}_{-n(\tau)-1} + \mathbf{W}_{-n(\tau)-1} - \mu\cdot)(t)$ and then $\mathbf{Y}_{-n(\tau)} = \mathbf{y}(A_{-n(\tau)})$. By Assumption 2, $A_{-n(\tau)}$ is a continuous random variable, therefore,

$$P(\exists t < A_{-n(\tau)} \text{ such that } \mathbf{y}(s) = 0, \text{ for all } s \in [t, A_{-n(\tau)}]) = 1.$$

In other words, $d\mathbf{y}(A_{-n(\tau)}-) = 0$. As a result, if we let $\mathbf{l}_\tau = d\mathbf{l}(A_{-n(\tau)})$, we have

$$0 = -\mu + R\mathbf{l}_\tau \Rightarrow \mathbf{l}_\tau = R^{-1}\mu.$$

From Assumption 1, we have $R^{-1}(E[\mathbf{W}_{-n(\tau)-1}] - \mu E[A_{-n(\tau)-1}]) < 0$. Since R^{-1} has non-negative elements, $E[\mathbf{W}_{-n(\tau)-1}] \geq 0$ and $E[A_{-n(\tau)-1}] > 0$, we can conclude $\mathbf{l}_\tau = R^{-1}\mu > 0$. Then, plug in $I = \{1,2,...,d\}$ in (6) and we solve $D = -R^{-1}E$. Therefore,

$$C_{-n(\tau)-1} = E + RD = 0 \text{ and } B_{-n(\tau)-1} = A_{-n(\tau)-1}C_{-n(\tau)-1} = 0,$$

as required. □

4 NUMERICAL EXAMPLES

4.1 A System of Tandem Queues

We consider an example of a stochastic fluid network with 10 queues in tandem. In detail, $P_{i,i+1} = 1$ for $i = 1,2,...,9$ and $P_{10,j} = 0$ for all $j = 1,...,10$, and $W^i = 0$ for all $2 \leq i \leq 10$. We assume a Poisson arrival process with rate $\lambda = 1$ and the job sizes are exponentially distributed with unit mean. The service-rate vector $\mathbf{r} = \left(r^1,...,r^{10}\right)^T$ are given by $\mathbf{r} = (1.55, 1.5, 1.45, 1.4, 1.35, 1.3, 1.25, 1.2, 1.15, 1.1)$. We are interested in computing the steady-state gradient of the workload $E[Y^{*,i}]$ at server i with respect to the service rate at server j , i.e. $\partial E[Y^{*,i}]/\partial r^j$, for all $1 \leq i, j \leq 10$.

Remark: In this example, whenever $Y_k^{*,1} > 0$, we have $Y_k^{*,i} = Y_{k-1}^{*,i} + 0.05A_{k-1}$ for $i \geq 2$ and therefore $(Y_k^{*,i} - Y_{k-1}^{*,i})$ are all equal for all $i \geq 2$. Since $P(Y_k^{*,1} > 0) > 0$, the event $\{Y_k^{*,1} > 0$ and $(Y_k^{*,i} - Y_{k-1}^{*,i})$ are all equal for all $i \geq 2\}$ happens infinitely often in a sample path of the workload process. Now, given any $\delta \neq 0$, if we change the service rate of Server 2 from μ^2 to $\mu^2 + \delta$, as $(0.05 - \delta)A_{k-1} \neq 0.05A_{k-1}$, the event $\{Y_k^{*,1} > 0$ and $(Y_k^{*,i} - Y_{k-1}^{*,i})$ are all equal for all $i \geq 2\}$ will never happen. As a result, the sample path space of the workload process corresponding to μ and $\mu + (0, \delta, 0, ..., 0)'$ have empty intersection. As a result, it is difficult to directly apply the benchmark regenerative algorithm as in Glynn (2006) .

For a network of tandem queues with exponentially distributed job size, it turns out, the true values of $E[Y^{*,i}]$ has an explicit expression in $\mathbf{r}$. See, for instance, Debicki, Dieker, and Rolski (2007) for reference. In particular,

$$E[Y^{*,1}] = 1/(r^1 - 1) \text{ and } E[Y^{*,i}] = 1/(r^i - 1) - 1/(r^{i-1} - 1). \tag{7}$$

Given the expression (7), we can derive an explicit expression for the true value of the gradient as well:

$$\frac{\partial E[Y^{*,i}]}{\partial r^j} = \begin{cases} -\frac{1}{(r^j-1)^2} & \text{if } i = j, \\ \frac{1}{(r^j-1)^2} & \text{if } i = j+1, \\ 0 & \text{otherwise.} \end{cases}$$

To estimate the gradient matrix $G = (\partial E[Y^{*,i}]/\partial r^j)$, we implement Algorithm 1 to compute an unbiased estimators $\hat{V}$ for the gradient matrix $V = (\partial E[Y^{*,i}]/\partial \mu^j)$ using 10000 stationary sample paths generated by the perfect sampling algorithm (Blanchet and Chen 2014). Recall that μ is the throughput rate of the network and by definition $\mu = R\mathbf{r}$. We then compute $\hat{G} = \hat{V}R$ as an unbiased estimator for G. The simulation results and the true values of the gradients are reported in Table 1. To illustrate the validity of Algorithm 1, we compare the absolute simulation error with the half-length of the estimated confidence interval in Table 2 which shows that the simulation estimator lies in the 95% confidence interval. To test the validity of simulated CI, we simulated 1000 independent 95% CI's for G_{11}. For each CI, we simulate 400 independent sample paths. Out of the 1000 simulated CI's, 950 cover the true value -3.3058, which indicates our estimator generated by Algorithm 1 is unbiased.

We do the implementation in Matlab. It took a few minutes in a laptop to generate 10000 independent sample paths and compute the gradient matrix.

Table 1: Unbiased estimates of gradient matrix $G_{ij} = \frac{\partial E[Y^{*,i}]}{\partial r^j}$.

Simulation	1	2	3	4	5	6	7	8	9	10
1	-3.2723	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	3.2723	-3.9618	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.0000	3.9618	-4.8906	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4	0.0000	0.0000	4.8906	-6.2615	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5	0.0000	- 0.0000	0.0000	6.2615	-8.0683	0.0000	0.0000	0.0000	0.0000	0.0000
6	0.0000	0.0000	0.0000	0.0000	8.0683	-11.1009	0.0000	0.0000	0.0000	0.0000
7	0.0000	-0.0000	0.0000	0.0000	0.0000	11.1009	-15.8137	0.0000	0.0000	0.0000
8	0.0000	0.0000	0.0000	0.0000	-0.0000	0.0000	15.8137	-24.7289	0.0000	0.0000
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	24.7289	-45.1027	0.0000
10	0.000	0.0000	-0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	45.1027	-99.8973
True Value	1	2	3	4	5	6	7	8	9	10
1	-3.3058	0	0	0	0	0	0	0	0	0
2	3.3058	-4.0000	0	0	0	0	0	0	0	0
3	0	4.0000	-4.9383	0	0	0	0	0	0	0
4	0	0	4.9383	-6.2500	0	0	0	0	0	0
5	0	0	0	6.2500	-8.1633	0	0	0	0	0
6	0	0	0	0	8.1633	-11.1111	0	0	0	0
7	0	0	0	0	0	11.1111	-16.0000	0	0	0
8	0	0	0	0	0	0	16.0000	-25.0000	0	0
9	0	0	0	0	0	0	0	25.0000	-44.4444	0
10	0	0	0	0	0	0	0	0	44.4444	-100.0000

4.2 A Symmetric Reflected Brownian Motion (RBM)

Reflected Brownian motion (RBM) is an important class of stochastic process in queueing theory as it is the diffusion limit of a large family of queueing networks. For instance, in the Quality-and-Efficiency-Driven (QED) regime, the queuing length process of a Generalize Jackson network will converge to a reflected Brownian motion. Moreover, the stationary distribution of the pre-limit Generalized Jackson networks will also weakly converge to the stationary distribution of the limit RBM (see Budhiraja and Lee 2009). However, how to compute the stationary distribution of the RBM, either analytically or numerically, remains an open

Table 2: Simulation error and half length of 95% confidence interval for $G_{ii} = \frac{\partial E[Y^{*,i}]}{\partial r^i}$.

	1	2	3	4	5	6	7	8	9	10
error	0.0335	0.0382	0.0477	0.0115	0.0950	-0.0102	0.1863	0.2711	0.6582	-0.1027
95% CI	0.1275	0.1546	0.1883	0.2344	0.2977	0.4109	0.5783	0.8842	1.6334	3.5809

problem. In this section, we show that Algorithm 1 can be potentially applied to continuous stochastic process such as a reflected Brownian motion via numerical examples.

Mathematically, a reflected Brownian motion $\mathbf{Z}(t)$ is the solution to a Skorokhod problem with Brownian input process:

$$\mathbf{Z}(t) = \Gamma(\mathbf{X})(t),$$

where Γ is the Skorokhod map with a given reflection matrix R and $\mathbf{X}(t) = \mu t + \Sigma \mathbf{B}(t)$, where $\mathbf{B}(t)$ is a standard multi-dimensional Brownian motion. For the numerical example, we consider a 2-dimensional symmetric RBM with drift coefficient $\mu = [-1,-1]$, covariance rate matrix $\Sigma^2 = [1,0;0,1]$ and reflection matrix $R = [1,-0.2;-0.2,1]$. In such a special symmetric case, the stationary mean $E[Z^{*,1}] = E[Z^{*,2}] := m_1$ and has an explicit expression (see Dai and Harrison 1992):

$$m_1(\mu^1) = \frac{1}{-\mu^1} \cdot \frac{1}{2(1+r)},$$

where $r = -R_{12} = 0.2$. Note that the equation holds only in the symmetric case when $\mu^2 = \mu^1$, as a result we can derive that the gradient $\partial E[Z^{*,1}]/\partial \mu^1$ and $\partial E[Z^{*,1}]/\partial \mu^2$ must satisfy the following relation:

$$\frac{\partial E[Z^{*,1}]}{\partial \mu^1} + \frac{\partial E[Z^{*,1}]}{\partial \mu^2} = m_1'(\mu^1).$$

As a result, in this special 2-dimensional case, we can compute the true value $m_1 = 5/12 \approx 0.4167$ and $m_1'(-1) = m_1 = 0.4167$.

Now we use Algorithm 1 to estimate the sensitivity of $E[Z^{*,1}]$ with respect to the drift coefficient μ_1 by simple discretization. In detail, we replace the input Brownian motion $\mathbf{X}(t)$ by its discretized approximation $\tilde{\mathbf{X}}_n = \mathbf{X}(n\Delta)$ with step size $\Delta = 2^{-h}$. Let $\tilde{\mathbf{Z}}_n$ be the reflected process with input $\tilde{\mathbf{X}}_n$ and reflection matrix R. Then, we can apply Algorithm 1 to the discretized process $\tilde{\mathbf{Z}}_n$ to compute a gradient estimator $\hat{V}^h$ for $\partial E[\tilde{\mathbf{Z}}^*]/\partial \mu$. Table 3 reports the simulation results with different discretization step sizes. For each h, we

Table 3: Simulation estimator and 95% confidence interval for $E[\tilde{Z}^{*,1}]$, $\partial E[\tilde{Z}^{*,1}]/\partial \mu^1$ and $\partial E[\tilde{Z}^{*,1}]/\partial \mu^2$ with step size equal to 2^{-h}.

h	$E[\tilde{Z}^{*,1}]$	$\partial E[\tilde{Z}^{*,1}]/\partial \mu^1$	$\partial E[\tilde{Z}^{*,1}]/\partial \mu^2$
8	0.3842 ± 0.0084	0.3540 ± 0.0121	0.0627 ± 0.0027
10	0.3986 ± 0.0081	0.3521 ± 0.0123	0.0621 ± 0.0027
12	0.4174 ± 0.0082	0.3599 ± 0.0124	0.0615 ± 0.0026

generate 10000 stationary sample paths to estimate the expectations and construct the 95% confidence interval. One can check that the true value of $E[Z^{*,1}] = 0.4167$ and $\partial E[\tilde{Z}^{*,1}]/\partial \mu^1(-1) + \partial E[\tilde{Z}^{*,1}]/\partial \mu^2 = 0.4167$ for the continuous reflected Brownian motion are both in the 95% intervals when $h = 12$. The numerical results indicate that Algorithm 1 is potentially applicable for continuous processes via simple discretization. Of course, rigorous proof is in need to validate the convergence and bound the convergence rate of the discretization scheme.

ACKONOWLEDGMENT

The author thanks Antonius Dieker and Guido Lagos for pointing out the possibility of the application of perfect sampling algorithm in gradient simulation. The author also thanks Jose Blanchet and Peter Glynn for all the helpful discussion, and the two referees for their useful comments.

REFERENCES

Blanchet, J., and X. Chen. 2014. "Steady-state simulation of reflected Brownian motion and related stochastic networks". *Submitted for publication*.

Budhiraja, A., and C. Lee. 2009. “Stationary distribution convergence for generalized Jackson networks in heavy traffic”. *Mathematics of Operations Research* 34 (1): 45–56.

Dai, J. G., and J. M. Harrison. 1992. “Reflected Brownian motion in an orthant: numerical methods for steady-state analysis”. *The Annals of Applied Probability* 2 (1): 65–86.

Debicki, K., A. B. Dieker, and T. Rolski. 2007. “Quasi-product forms for Lévy-driven fluid networks”. *Mathematics of Operations Research* 32 (3): 629–647.

Dieker, A. B., S. Ghosh, and M. S. Squillante. 2014. “Optimal resource capacity management for stochastic networks”. *Submitted for publication*.

Glasserman, P. 1992. “Stationary waiting time derivatives”. *Queueing Systems* 12 (3-4): 369–390.

Glynn, P. W. 2006. “Simulation Algorithms for Regenerative Processes”. In *Handbook in Operations Research and Management Science, vol. 13: Simulation*, edited by S. G. Henderson and B. L. Nelson, 477–500. North-Holland Publishing, Amsterdam.

Harrison, J. M., and M. I. Reiman. 1981. “Reflected Brownian motion on an orthant”. *The Annals of Probability* 9 (2): 302–308.

Kella, O., and W. Whitt. 1996. “Stability and Structural Properties of Stochastic Fluid Networks”. *Journal of Applied Probability* 33:1169–1180.

Mandelbaum, A., and K. Ramanan. 2010. “Directional Derivatives of Oblique Reflection Maps”. *Mathematics of Operations Research* 35 (3): 527–558.

AUTHOR BIOGRAPHIES

XINYUN CHEN is an Assistant Professor in the Department of Applied Mathematics and Statistics at Stony Brook University. She received a Ph.D in Operations Research from Columbia University. Her research deals with modeling and computation of stochastic systems with applications in financial and service engineering. Her email address is xinyun.chen@stonybrook.edu.