

CONVERGENCE RATES FOR STEADY-STATE DERIVATIVE ESTIMATORS

Pierre L'ECUYER

*Département d'Informatique et de Recherche Opérationnelle, Université de Montréal,
C.P. 6128, Succ. A, Montréal, Canada H3C 3J7*

Abstract

We describe various derivative estimators for the case of steady-state performance measures and obtain the order of their convergence rates. These estimators *do not* use explicitly the regenerative structure of the system. Estimators based on infinitesimal perturbation analysis, likelihood ratios, and different kinds of finite-differences are examined. The theoretical results are illustrated via numerical examples.

Keywords: Discrete-event systems, gradient estimations, steady state.

1. Introduction

Estimating derivatives of expected performance measures with respect to some continuous parameters, in the context of stochastic discrete-event simulations, has received a great deal of attention lately [3,5,10,11,13,15,21–24,26]. Such derivative estimators are useful for sensitivity analysis, or can be used within stochastic optimization algorithms [3,10,17,24]. For finite-horizon simulations, Glynn [10] gives the convergence rates of different estimators, under given sets of assumptions. In that context, the infinitesimal perturbation analysis (IPA) and likelihood ratio (LR) (also called score function (SF)) estimators converge at the canonical rate of $n^{-1/2}$, where n is the number of replications (thanks to the central-limit theorem). For finite-difference (FD) schemes, things are not so easy, because the bias component must be taken into account. To make the bias go to zero, the FD interval must be reduced towards zero, but then the variance typically increases to infinity. Therefore, a compromise must be made and as a result, typically, one does not obtain the canonical convergence rate. Glynn [10] gives (subcanonical) convergence rates for forward and centered FD schemes, with and without common random numbers, under specific assumptions. On the other hand, L'Ecuyer and Perron [18] show that in most interesting cases where IPA applies, FD with common random numbers reaches the canonical rates.

The aim of this paper is to extend these results to derivative estimators of steady-state performance measures. The system is viewed as a discrete-time Markov chain with general state space. The model, with its assumptions, is stated in

section 2. Section 3 describes the derivative estimators that we consider and we derive bounds on their convergence rates. In this case, not only the number of replications but also the run length for the individual replications, should increase with the computer budget to get the initialization bias down to zero. Therefore, for a given budget, we have to compromise between run length and number of runs. It does not appear trivial, in this context, that the convergence rates will be the same as for the finite-horizon case. Indeed, it turns out that the straightforward LR estimators no longer reach the canonical rate. We derive their convergence rates with and without the control variate approach proposed in [16]. (Note, however, that there exist LR derivative estimators that converge at the canonical rate, but they use *explicitly* the regenerative structure [9,21].) For IPA and FD, we obtain the same rates as for the finite-horizon case.

2. Model and assumptions

As in [4], for any $f : \mathbb{N} \mapsto [0, \infty)$, we define $O(f(n))$ as the set of functions $g : \mathbb{N} \mapsto [0, \infty)$ such that for some constant $c > 0$, $g(n) \leq cf(n)$ for all n in \mathbb{N} . The set $\Omega(f(n))$ is defined in the same way, with \leq replaced by \geq , and $\Theta(f(n)) = O(f(n)) \cap \Omega(f(n))$.

The setting is similar as in [17]. We consider a Markov chain $\{X_j(\theta, \omega), j = 0, 1, \dots\}$ with general (Borel) state space S , defined over a probability space $(\Omega, \Sigma, P_\theta)$. Let $X_0(\theta, \omega) = s_0$ for some fixed initial state $s_0 \in S$. The sample point $\omega \in \Omega$ represents the “randomness” that drives the system. The probability measure P_θ depends (in general) on the parameter value θ . Here, $\theta \in (a, b)$, an open interval of \mathbb{R} .

A cost $g(\theta, x)$ is incurred whenever we visit state x (except for the initial state $X_0 = s_0$), where $g : (a, b) \times S \mapsto \mathbb{R}$ is assumed measurable. Let

$$h_t(\theta, \omega) = \frac{1}{t} \sum_{j=1}^t g(\theta, X_j) \tag{1}$$

be the average cost for the first t steps and let

$$\alpha_t(\theta) \stackrel{\text{def}}{=} \int_{\Omega} h_t(\theta, \omega) dP_\theta(\omega) = \frac{1}{t} \sum_{j=1}^t E_\theta[g(\theta, X_j)]. \tag{2}$$

Assume that for each $\theta \in (a, b)$,

$$|\alpha_t(\theta) - \alpha(\theta)| \in O(1/t), \tag{3}$$

where $\alpha(\theta)$ represents the steady-state average cost for running the system at parameter level θ . We suppose that the derivative $\alpha'(\theta)$ exists for all $\theta \in (a, b)$ and we are interested in estimating $\alpha'(\theta_0)$ for some $\theta_0 \in (a, b)$. We also suppose that

$$|\alpha'_i(\theta_0) - \alpha'(\theta_0)| \in O(1/t). \tag{4}$$

These assumptions hold for many systems of interest. For regenerative systems, for example, the main theorem in [19] implies that under very mild assumptions, (3) holds. However, since $\alpha'_i(\theta) = (1/t) \sum_{j=1}^t \partial E_\theta(g(\theta, X_j)) / \partial \theta$, the same result implies that (4) also holds under appropriate mild assumptions (just replace $E_\theta[g(\theta, X_j)]$ by its derivative).

Basically, our derivative estimators are based on simulations of the system for a finite number of transitions, each from initial state s_0 . We perform n replications of this. Some schemes (like IPA or LR) require only one simulation run per replication, while others may require more (like the usual finite-difference schemes, which require two). Replication i gives an estimate $\psi_{n,i}$ of the derivative and (except when stated otherwise) the overall derivative estimate is

$$Y_n = \frac{1}{n} \sum_{i=1}^n \psi_{n,i}. \tag{5}$$

Our “loss function” is the mean square error of Y_n and we are interested in how fast its square root converges to zero as a function of the total simulation time (computing cost) C_n required for the n replications. We assume that the simulation cost is directly proportional to the number of simulated transitions. We have

$$R_n = E[Y_n - \alpha'(\theta_0)]^2 = V_n + B_n^2, \tag{6}$$

where R_n , V_n , and B_n denote the mean square error, the variance, and the bias, respectively. The convergence rate is defined as $C^{-\beta^*/2}$ (in terms of the total CPU budget C), where β^* is the largest value of b for which $C_n^\beta R_n \in O(1)$ (as a function of n).

In most cases, all runs will have the same length t_n , which yields $C_n \in \Theta(nt_n)$. In general, it is necessary that $t_n \rightarrow \infty$ and $n \rightarrow \infty$ to obtain $B_n \rightarrow 0$. The bias component that is due to the fact that t_n is finite is in $O(1/t_n)$. There might also be other bias, that is, if $E[Y_n] \neq \alpha'_i(q_0)$ (like for finite differences). A reasonable choice for the general shape of t_n is

$$t_n = \lfloor T n^p \rfloor \tag{7}$$

for some constants $p \geq 0$ and $T > 0$. Then, $C_n = n \lfloor T n^p \rfloor \in \Theta(n^{p+1})$.

Of course, the best way of reducing the bias due to the finite run length would be to devote the entire budget to a single replication, that is, to take $n = 1$. However, for FD there will still remain the finite-difference bias, while for LR, the variance will then be so high (in general) to make the estimator totally useless. This has been pointed out in particular by Glynn [9] and Rubinstein [24], who suggested truncation approaches similar to the ones that we study here, but who did not derive convergence rates.

3. Derivative estimators and their convergence rates

3.1. FINITE DIFFERENCES

In typical stochastic simulations, the sample point ω (which represents the “randomness” that drives the system) can be viewed as a sequence of independent uniform variates between 0 and 1. Then, $P_\theta \equiv P$ is independent of θ . (This is a standard interpretation; for more details, see [15,18].)

Finite-difference (FD) schemes have been used for a long time to estimate derivatives. The most straightforward schemes use independent streams, as follows. Let n be the number of replications and $\varepsilon_n > 0$. Let $\omega_1^-, \dots, \omega_n^-, \omega_1^+, \dots, \omega_n^+$ be $2n$ independent sample points, where each ω_i^- and each ω_i^+ is generated under P . The *forward* FD estimator is

$$Y_n^{\text{FDf}} = \frac{1}{n} \sum_{i=1}^n \psi_{n,i}^{\text{FDf}} = \frac{1}{n} \sum_{i=1}^n \frac{h_{t_n}(\theta + \varepsilon_n, \omega_i^+) - h_{t_n}(\theta, \omega_i^-)}{\varepsilon_n}, \quad (8)$$

while the *central* FD estimator is

$$Y_n^{\text{FDc}} = \frac{1}{n} \sum_{i=1}^n \psi_{n,i}^{\text{FDc}} = \frac{1}{n} \sum_{i=1}^n \frac{h_{t_n}(\theta + \varepsilon_n, \omega_i^+) - h_{t_n}(\theta - \varepsilon_n, \omega_i^-)}{2\varepsilon_n}. \quad (9)$$

In practice, to compute each term of the sum in (8), one performs two different simulation runs of length t_n , with independent random numbers, to obtain $h_{t_n}(\theta, \omega_i^-)$ and $h_{t_n}(\theta + \varepsilon_n, \omega_i^+)$ (and similarly for $h_{t_n}(\theta - \varepsilon_n, \omega_i^-)$ and $h_{t_n}(\theta + \varepsilon_n, \omega_i^+)$ in (9)). (Technically, all ω_i^+ and ω_i^- should be viewed as part of a unique sample point, i.e. defined over the same probability space. However, this “abuse of notation” is no serious problem since one can always take the product space, for which the sample point effectively contains all the information.)

Alternatively, one can generate only n (independent) sample points $\omega_1, \dots, \omega_n$ from P and take $\omega_i^- = \omega_i^+ = \omega_i$ for each i . This gives finite-difference estimators with common random numbers (FDC). The *forward* FDC estimator is

$$Y_n^{\text{FDCf}} = \frac{1}{n} \sum_{i=1}^n \psi_{n,i}^{\text{FDCf}} = \frac{1}{n} \sum_{i=1}^n \frac{h_{t_n}(\theta + \varepsilon_n, \omega_i) - h_{t_n}(\theta, \omega_i)}{\varepsilon_n}, \quad (10)$$

while the *central* FDC estimator is

$$Y_n^{\text{FDCc}} = \frac{1}{n} \sum_{i=1}^n \psi_{n,i}^{\text{FDCc}} = \frac{1}{n} \sum_{i=1}^n \frac{h_{t_n}(\theta + \varepsilon_n, \omega_i) - h_{t_n}(\theta - \varepsilon_n, \omega_i)}{2\varepsilon_n}. \quad (11)$$

Again, in practice, two simulation runs are made to compute each term of the sum, but these runs are made with common random numbers.

In the context of *finite-horizon simulations* (where $C_n \in \Theta(n)$), it is well known that FD with independent streams give rise to large variability [10, 18, 20, 27]. The variances of $\psi_{n,i}^{\text{FDf}}$ and $\psi_{n,i}^{\text{FDc}}$ are in $\Theta(\varepsilon_n^{-2})$ in general. The best convergence rates for forward differences and central differences are in $O(n^{-1/4})$ and $O(n^{-1/3})$, respectively. For FDC and under a given set of assumptions (including the assumption that the variances of $\psi_{n,i}^{\text{FDCf}}$ and $\psi_{n,i}^{\text{FDCc}}$ are in $\Theta(\varepsilon_n^{-1})$), Glynn [10] obtains respective convergence rates in $O(n^{-1/3})$ and $O(n^{-2/5})$. In what follows, we show that the convergence rates for the *infinite-horizon* case are the same.

THEOREM 1

Let $t_n \in \Theta(n^p)$, $\varepsilon_n \in \Theta(n^{-\gamma})$, and suppose that $\sup_{\theta \in (a,b)} [\text{Var}(h_i(\theta, \omega))] \in O(1/t)$. For the forward case (FDf and FDCf), suppose that $\alpha''(\theta) \neq 0$ (and exists) in a neighborhood of θ_0 . For the central case (FDc and FDCc), suppose that $\alpha'''(\theta) \neq 0$ (and exists) in a neighborhood of θ_0 . As in Glynn [10], we suppose that the variance of $\psi_{n,i}$ is in $\Theta(\varepsilon_n^{-1} t_n^{-1})$. Then, the “optimal” values of p, γ, β , and the corresponding convergence rates are, respectively,

- (i) $p = \gamma = 1/3, \beta = 1/2$, and $C^{-1/4}$ for (8);
- (ii) $p = 1/2, \gamma = 1/4, \beta = 2/3$, and $C^{-1/3}$ for (9);
- (iii) $p = \gamma = 1/2, \beta = 2/3$, and $C^{-1/3}$ for (10);
- (iv) $p = 2/3, \gamma = 1/3, \beta = 4/5$, and $C^{-2/5}$ for (11).

Proof

Using Taylor’s expansion, it is easily seen that the bias component due to the fact that we use finite differences is in $\Theta(\varepsilon_n)$ for forward differences and in $\Theta(\varepsilon_n^2)$ for central differences. For (8), one has $V_n = \text{Var}(Y_n^{\text{FDf}}) \in O(1/(n \varepsilon_n^2 t_n)) = O(n^{-1-p+2\gamma})$, $B_n \in O(1/t_n + \varepsilon_n) = O(n^{-p} + n^{-\gamma})$, $C_n \in \Theta(n^{p+1})$, and

$$\begin{aligned} C_n^\beta R_n &\in O(n^{(p+1)\beta} [V_n + B_n^2]) \\ &= O(n^{(p+1)\beta} [n^{-1-p+2\gamma} + (n^{-p} + n^{-\gamma})^2]) \\ &= O(n^{(p+1)\beta} [n^{-1-p+2\gamma} + n^{-2p} + n^{-2\gamma} + n^{-p-\gamma}]). \end{aligned}$$

Then, $C_n^\beta R_n$ will be in $O(1)$ if

$$(p + 1)\beta \leq \max(1 + p - 2\gamma, 2p, 2\gamma, p + \gamma).$$

The maximal value of β for this inequality to be satisfied is $\beta = 1/2$, with $p = \gamma = 1/3$. (In this case, the maximum is reached for all four values and equality holds.) Also, $C_n^\beta R_n \in \Omega(n^{(p+1)\beta} [n^{-1-p+2\gamma} - n^{-2\gamma}])$, so that $(p + 1)\beta$

$\leq \max(1 + p - 2\gamma, 2\gamma)$ is a necessary condition for $C_n^\beta R_n$ to be in $O(1)$. Again, $\beta = 1/2$ is the best possible value that satisfies this (with $4\gamma = p + 1$) and (i) follows. Results (ii) to (iv) are obtained in a similar way. The expressions of $C_n^\beta R_n$ for (ii) to (iv) are, respectively,

$$C_n^\beta R_n \in O(n^{(p+1)\beta} [n^{-1-p+2\gamma} + (n^{-p} + n^{-2\gamma})^2]);$$

$$C_n^\beta R_n \in O(n^{(p+1)\beta} [n^{-1-p+\gamma} + (n^{-p} + n^{-\gamma})^2]);$$

$$C_n^\beta R_n \in O(n^{(p+1)\beta} [n^{-1-p+\gamma} + (n^{-p} + n^{-2\gamma})^2]). \quad \square$$

Now, instead of using the same t_n and ε_n for all n replications, one can use, say, $t_i \in \Theta(i^p)$ and $\varepsilon_i \in \Theta(i^{-\gamma})$ for replication i , for $i = 1, \dots, n$. Y_n can then be defined as

$$Y_n = \frac{\sum_{i=1}^n w_i \Psi_{n,i}}{\sum_{i=1}^n w_i} \tag{12}$$

for some positive weights w_1, \dots, w_n . We have $C_n \in \Theta(\sum_{i=1}^n i^p) = \Theta(n^{p+1})$. For forward FD, the bias and variance expressions become

$$B_n \in \Theta\left(\frac{\sum_{i=1}^n w_i (i^{-p} + i^{-\gamma})}{\sum_{i=1}^n w_i}\right)$$

and

$$V_n \in \Theta\left(\frac{\sum_{i=1}^n w_i^2 (i^{-p+2\gamma})}{(\sum_{i=1}^n w_i)^2}\right).$$

One can take, for example, $w_i = 1$ for all i , or $w_i = t_i$ (weight proportional to work). In both cases, it is easy to see that $B_n \in \Theta(n^{-p} + n^{-\gamma})$ and $V_n \in \Theta(n^{-1-p+2\gamma})$, the same as for (8) above, so that theorem 1(i) still applies. It can be verified in the same way that (ii) to (iv) in theorem 1 also remains unmodified.

Instead of performing n replications from a given initial state, one can also just perform one very long replication, in order to diminish the bias component due to the initial state (transient). Perform one (long) simulation of length in $\Theta(nt_n) = \Theta(n^{p+1})$ at $\theta_0 + \varepsilon_n$, then another one at θ_0 . Here, n is just a parameter, not a number of replications. Equivalently, one can view the run length as being cut into n pieces of size in $\Theta(n^p)$, in a “batch-means” fashion. The (only) gain here is that the bias component due to the initial transient will now be in

$$\Theta\left(\frac{1}{n} \sum_{i=1}^n \frac{1}{it_n}\right) = \Theta\left(n^{-p-1} \sum_{i=1}^n \frac{1}{i}\right) = \Theta(n^{-p-1} \ln(n)).$$

Then, for forward FD, $B_n \in \Theta(n^{-p-1} \ln(n) + n^{-\gamma})$ and

$$C_n^\beta R_n \in O(n^{(p+1)\beta} [n^{-p-1+2\gamma} + (n^{-(p+1)} \ln n)^2 + n^{-2\gamma} + n^{-p-1-\gamma} \ln n]).$$

To obtain $(p + 1)\beta - 1 - p + 2\gamma = (p + 1)\beta - 2\gamma = 0$, we need $\beta = 1/2$ and $4\gamma = p + 1$. These conditions are also sufficient to obtain $C_n^\beta R_n \in O(1)$. So, there is no improvement of the convergence rate. It can be verified that the same also applies to the case of central differences and to the forward and central versions of FDC. This is no surprise. Recall that the convergence rates in theorem 1 are the same as for the finite-horizon case, for which there is no transient bias. Therefore, reducing the transient bias should not be expected to improve the convergence rate.

3.2. INFINITESIMAL PERTURBATION ANALYSIS

As for FDC, let us view the sample point ω as a sequence of independent uniform variates between 0 and 1, and denote P_θ by P . Under some conditions (see [6, 15]), the random variable $h'_i(\theta, \omega) = dh_i(\theta, \omega)/d\theta$ can be used as an unbiased estimator of $\alpha'_i(\theta)$. This is the IPA derivative estimator.

One can obtain n i.i.d. replicates of $h'_i(\theta, \omega)$ and use the estimator

$$Y_n^{\text{IPA}} = \frac{1}{n} \sum_{i=1}^n h'_i(\theta, \omega_i). \tag{13}$$

Here, $\omega_1, \omega_2, \dots, \omega_n$ are independent sequences of independent uniform variates. (In practice, for complex simulations, $h'_i(\theta, \omega)$ is not always easy to compute. There are also many cases where it gives a biased or even totally meaningless derivative estimator. See [15].)

THEOREM 2

Assume that $E[h'_i(\theta_0, \omega)] = \alpha'_i(\theta_0)$ for each $t > 0$ and that the variance of $h'_i(\theta_0, \omega)$ is in $\Theta(1/t)$. Let $t_n \in \Theta(n^p)$. Then, the convergence rate of $C^{-1/2}$ ($\beta = 1$) is obtained for any $p \geq 1$.

Proof

We have $C_n \in \Theta(n^{p+1})$, $V_n \in O(1/(nt_n)) = O(n^{-p-1})$, and $B_n^2 \in O(1/t_n^2) = O(n^{-2p})$. Therefore, $C_n^\beta R_n = C_n^\beta (V_n + B_n^2) \in O(n^{(p+1)\beta} (n^{-p-1} + n^{-2p})) = O(1)$ for $\beta = 1$ and $p \geq 1$. □

This is clearly the best convergence rate that one can expect. Again, as for FD, one can perform just one run of length in $\Theta(n^{p+1})$ to reduce the transient bias. This could reduce the mean square error, but will not improve the convergence rate.

Sufficient conditions under which $E[h'_i(\theta_0, \omega)] = \alpha'_i(\theta_0)$ are given in [6, 15]. In most interesting cases where IPA applies, these conditions are satisfied. In the context of finite-horizon simulations, L'Ecuyer and Perron [18] show that whenever these conditions are satisfied, the variances of $\psi_{n,i}^{\text{FDCf}}$ and $\psi_{n,i}^{\text{FDCc}}$ are in $O(1)$ (instead of $\Theta(1/\epsilon_n)$) and the convergence rates for both the forward and central versions of FDC are the same as for IPA. This is also true in the steady-state case. For forward FDC, one has in that case

$$C_n^\beta R_n \in O(n^{(\rho+1)\beta} [n^{-1-\rho} + (n^{-\rho} + n^{-\gamma})^2]).$$

A convergence rate of $C^{-1/2}$ is obtained when $\beta = 1, \rho \geq 1$, and $\gamma \geq \max(1, (\rho + 1)/2)$. For example, $\rho = 1$ and any $\gamma \geq 1$ will do. For central FDC, one has

$$C_n^\beta R_n \in O(n^{(\rho+1)\beta} [n^{-1-\rho} + (n^{-\rho} + n^{-2\gamma})^2]).$$

A convergence rate of $C^{-1/2}$ is obtained when $\beta = 1, \rho \geq 1$, and $\gamma \geq \max(1/2, (\rho + 1)/4)$. Here, with $\rho = 1$, any $\gamma \geq 1/2$ will do.

3.3. LIKELIHOOD RATIOS

The likelihood ratio (or score function) derivative estimation technique, aimed for the case where P_θ really depends on θ , goes as follows [3, 9, 11, 15, 21–24]. To fix ideas, suppose that ω can be “viewed” as $\omega = (\zeta_1, \dots, \zeta_t)$, where ζ_j is the value taken by a continuous random variable with density $f_{j,\theta}$, and the ζ_j 's are independent. Given X_{j-1} , the value of ζ_j determines the next state X_j of the Markov chain. (A similar development can be made for discrete or mixed random variables.) Let $G = P_{\theta_0}$ and suppose that G dominates the P_θ 's in a neighborhood of θ_0 . One can rewrite

$$\alpha_i(\theta) = \int_{\Omega} H_i(\theta, \omega) dG(\omega), \tag{14}$$

where

$$H_i(\theta, \omega) = h_i(\theta, \omega) \prod_{j=1}^t \frac{f_{j,\theta}(\zeta_j)}{f_{j,\theta_0}(\zeta_j)}. \tag{15}$$

Now, generate ω according to G and compute

$$H'_i(\theta, \omega) = h'_i(\theta, \omega) + h_i(\theta, \omega) \sum_{j=1}^t S_j(\theta, \omega) \tag{16}$$

as the LR derivative estimate at $\theta = \theta_0$, where

$$S_j(\theta, \omega) = \frac{d}{d\theta} (\ln f_{j,\theta}(\zeta_j)). \tag{17}$$

The sum in (16) is called the score function. For n replications of length t_n , one has, as in (13):

$$Y_n^{\text{LR}} = \frac{1}{n} \sum_{i=1}^n H'_i(\theta, \omega_i). \tag{18}$$

This can be viewed as a generalization of IPA [15]. On the other hand, as explained in [18], this is equivalent to applying IPA *after* replacing h_i and P_θ by H_i and G_i , i.e. on top of an importance sampling scheme. Consequently, known results on IPA can be applied to LR. In particular, the sufficient conditions for unbiasedness of IPA also apply to LR [15].

Since the ζ_j 's are independent, the variance of the score function is the sum of the variances of the S_j 's. When the latter variances are non-zero, the variance of the score function is typically in $\Theta(t)$. Under mild conditions [16], the variance of $H'_i(\theta, \omega)$ is also in $\Theta(t)$, i.e. increases linearly with the simulation length. This is a most unfortunate situation. This also means that the assumptions of theorem 2 are usually not satisfied when IPA is applied on top of an importance sampling scheme as in (16). On the other hand, in the "degenerate" special case where IPA is applied directly to (1), the score function is always zero and does not contribute to the variance.

L'Ecuyer and Glynn [16] have proposed a control variate scheme for LR under which that variance gets down to $\Theta(1)$. Let us denote it by CLR. When the variance of $\psi_{n,i}$ was in the order of $1/t_n$, as in theorems 1 and 2, V_n depended essentially only on the total simulation length C_n , and not on the way it was cut down into pieces (replications). However, this is no longer the case here. Intuitively, we expect that the run lengths should be kept shorter to keep the variance down. This is exactly what the next theorem tells us.

THEOREM 3

Assume that $E[H'_i(\theta_0, \omega)] = \alpha'_i(\theta_0)$ for each $t > 0$. Let $t_n \in \Theta(n^p)$. For LR, if the variance of $H'_i(\theta_0, \omega)$ is in $\Theta(t)$, then the best convergence rate is $C^{-1/4}$ ($\beta = 1/2$) and is obtained for $p = 1/3$. For CLR, if the variance $H'_i(\theta_0, \omega)$ is in $\Theta(1)$, then the best convergence rate is $C^{-1/3}$ ($\beta = 2/3$) and is obtained for $p = 1/2$.

Proof

For LR, we have $V_n \in \Theta(n^p/n)$, while for CLR, $V_n \in \Theta(1/n)$. Therefore, for LR, $C_n^\beta R_n \in O(n^{(p+1)\beta} (n^{p-1} + n^{-2p}))$, which is $O(1)$ for $\beta = 1/2$ and $p = 1/3$, while for CLR, $C_n^\beta R_n \in O(n^{(p+1)\beta} (n^{-1} + n^{-2p}))$, which is $O(1)$ for $\beta = 2/3$ and $p = 1/2$. □

One can also use different run lengths for different replications, like $t_i \in \Theta(i^p)$ for replication i , $i = 1, \dots, n$, as explained in the context of FD. It can easily be

seen that for LR and CLR, the orders of V_n and B_n , and the convergence rates, remain the same as in the theorem 3.

For CLR, one can also perform just one long simulation run, of length $t_n \in \Theta(n^{p+1})$, to diminish the initial transient bias (as for FD). (Now, n becomes just an index; it is no longer a number of replications.) If one does this straightforwardly, the variance of Y_n will be in $O(1)$. However, here $h_{t_n}(\theta, \omega)$ is an average of t_n terms (see eq. (1)). To diminish the variance of the score function, we can truncate it, that is, consider the derivative of each term $\alpha_j(\theta) - \alpha_{j-1}(\theta)$ individually and associate to this term a likelihood ratio based on a "window" of width, say, $l_j \in \Theta(j^q)$ for $q \leq 1$. This idea was proposed and explored in Arsham et al. [1] and Rubinstein [24]. This gives the gradient estimator

$$Y_n = h'_{t_n}(\theta_0, \omega) + \frac{1}{t_n} \sum_{j=1}^{t_n} g(\theta_0, X_j) \sum_{k=j-l_j+1}^j S_k(\theta_0, \zeta_k). \tag{19}$$

The variance of $Z_j = g(\theta_0, X_j) \sum_{k=j-l_j+1}^j S_k(\theta_0, \zeta_k)$ is typically in $\Theta(l_j)$. Assume (unrealistically, but optimistically) that these Z_j 's are independent. Then, the variance of Y_n is

$$V_n \in \Theta \left(n^{-2(p+1)} \sum_{j=1}^{n^{p+1}} j^q \right) = \Theta(n^{(p+1)(q-1)})$$

and the bias is

$$B_n \in \Theta \left(n^{-(p+1)} \sum_{j=1}^{n^{p+1}} j^{-q} \right) = \Theta(n^{-q(p+1)}).$$

This yields

$$C_n^\beta R_n \in \Theta(n^{(p+1)\beta} [n^{(p+1)(q-1)} + n^{-2q(p+1)}])$$

and the largest value of b for that to be in $O(1)$ is $\beta = 2/3$, with $q = 1/3$ and any $p > -1$. Therefore, there is no improvement on the convergence rate.

When $l_j = j$ ($q = 1$) in the above, (19) becomes

$$Y_n = h'_{t_n}(\theta_0, \omega) + \frac{1}{t_n} \sum_{j=1}^{t_n} g(\theta_0, X_j) \sum_{k=1}^j S_k(\theta_0, \zeta_k). \tag{20}$$

This is (16) with $t = t_n$, except that the terms $g(\theta_0, X_j) S_k(\theta_0, \zeta_k)$ for $j < k$ have been removed from the estimator. However, these terms have zero expectation because for $j < k$, $X_j(\theta_0, \omega)$ is independent of $S_k(\theta_0, \omega)$ and $E[S_k(\theta_0, \omega)] = 0$. Therefore, (20) is an unbiased estimator of $\alpha'_{t_n}(\theta_0)$. Typically, its variance is less than that of (16) (although there is no guarantee, but still in the same order).

Now, suppose that the process is regenerative. Let $0 = \tau_0 \leq \tau_1 \leq \tau_2, \dots$ be the regeneration points. Conditional on τ_i , what happens after τ_i is independent of what happened until τ_i . Based on this, one would be tempted to remove from (20) the terms $g(\theta_0, x_j)S_k(\theta_0, \zeta_k)$ for which j and k belong to different regenerative cycles (say, $k \leq \tau_i < j$ for some i), since then $E[g(\theta_0, X_j)S_k(\theta_0, \zeta_k) | \tau_1, \tau_2, \dots] = E[g(\theta_0, X_j) | \tau_1, \tau_2, \dots]E[S_k(\theta_0, \zeta_k) | \tau_1, \tau_2, \dots]$. However, the problem is that in general $E[S_k(\theta_0, \zeta_k) | \tau_1, \tau_2, \dots] \neq E[S_k(\theta_0, \zeta_k)] = 0$, which means that one cannot really remove these terms. Correct (asymptotically unbiased) gradient estimators for the regenerative case are given in [9,21].

4. Two examples involving a simple queue

Consider a $GI/G/1$ queue [2] with service time distribution F_θ , which has a density f_θ and depends on a parameter θ . We assume that for all $\theta \in (a, b)$, the queue is stable. A $GI/G/1$ queue can be described in terms of a discrete-time Markov chain via Lindley's equation. Let W_j, Z_j , and $X_j = W_j + Z_j$ be the waiting time, service time, and sojourn time for the j th customer, and let A_j be the time between arrivals of the $(j - 1)$ th and j th customer. For our purposes, the sojourn time X_j will be the state of the chain at step j . The state space is $S = [0, \infty)$ and $X_0 = 0$, which corresponds to an initially empty system. For the first customer, one has $X_1 = Z_1$ and, for $j \geq 2$,

$$X_j = W_j + Z_j = (X_{j-1} - A_j)^+ + Z_j, \tag{21}$$

where x^+ means $\max(x, 0)$.

4.1. AVERAGE SOJOURN TIME PER CUSTOMER

Let $\alpha(\theta)$ be the average sojourn time in the system per customer, in steady state, at parameter level θ . The cost at step j is $g(\theta, X_j) = X_j$. It is proven in [17] that (3)–(4) hold in this case, under mild conditions.

For LR, one views ω as representing the sequence of interarrival and service times. The score function associated with customer j is $S_j(\theta, \omega) = \sum_{k=1}^j \partial \ln f_\theta(Z_k) / \partial \theta$. Sufficient conditions for the LR estimator to be unbiased are given in theorem 1 of L'Ecuyer [15].

Suppose that each service time Z_j is generated by inversion: $Z_j = F_\theta^{-1}(U_j)$, where the U_j 's are i.i.d. $U(0, 1)$ uniform variates. The IPA derivative estimator is $h'_j(\theta, \omega) = (1/t) \sum_{j=1}^t X'_j$, where for $j \geq 1$, $X'_j = X'_{j-1} + Z'_j$ if $X_{j-1} > A_j$, and $X'_j = Z'_j$ otherwise. Here, $Z'_j = \partial F_\theta^{-1}(U_j) / \partial \theta$.

For a numerical illustration, consider an $M/M/1$ queue with arrival rate λ and service rate μ . The traffic intensity is $\rho = \lambda/\mu$ and the parameter of interest is the mean service time $\theta = 1/\mu$. The true values of the average cost and of its derivative are given by $\alpha(\theta) = 1/(\mu(1 - \rho))$ and $\alpha'(\theta) = 1/(1 - \rho)^2$, respectively. In

this case, one has $\partial \ln f_{\theta}(Z_k)/\partial \theta = \partial \ln((1/\theta)e^{-Z_k/\theta})/\partial \theta = (Z_k - \theta)/\theta^2$ for LR and $Z'_j = \partial[\theta \ln(1 - U_j)]/\partial \theta = Z_j/\theta$ for IPA.

Tables 1 and 2 report numerical results for $\lambda = 1$ and $\rho = 1/4$ and $1/2$, respectively. In each case, we estimated the empirical mean square error (MSE) for $C_n = 10, 10^2, 10^3, 10^4,$ and 10^5 , where $C_n = nt_n$ denotes the total number of customers in n simulation runs of length t_n (the run length is measured in terms of the number of

Table 1

Sample MSE for derivative of mean sojourn time, *M/M/1* queue with $\rho = 1/4$.

	p	γ	$C_n = 10^1$	$C_n = 10^2$	$C_n = 10^3$	$C_n = 10^4$	$C_n = 10^5$
FDCc	1/2	1/4	2.10	7.07E-1	2.42E-1	5.58E-2	2.25E-2
FDCc	2/3	1/3	9.36E-1	3.10E-1	7.08E-2	1.63E-2	4.05E-3
FDCc	1	1	7.62E-1	1.57E-1	2.25E-2	2.47E-3	2.44E-4
FDCc	2	2	7.60E-1	1.48E-1	1.95E-2	2.08E-3	1.92E-4
IPA	1		5.86E-1	1.39E-1	1.97E-2	2.12E-3	1.47E-4
IPA	2		6.86E-1	1.45E-1	1.93E-2	2.12E-3	1.32E-4
LR	1/3		3.08	8.43E-1	2.25E-1	7.26E-2	1.99E-2
LRD	1/3		2.64	6.88E-1	1.69E-1	5.34E-2	1.59E-2
CLR	1/2		1.90	5.54E-1	1.23E-1	2.99E-2	4.42E-3
CLRD	1/2		1.77	4.56E-1	9.48E-2	1.80E-2	2.85E-3
CLRD	1		2.71	7.95E-1	1.61E-1	2.78E-2	1.19E-2

Table 2

Sample MSE for derivative of mean sojourn time, *M/M/1* queue with $\rho = 1/2$.

	p	γ	$C_n = 10^1$	$C_n = 10^2$	$C_n = 10^3$	$C_n = 10^4$	$C_n = 10^5$
FDCc	1/2	1/4	21.78	11.34	5.68	1.88	7.17E-1
FDCc	2/3	1/3	14.31	7.48	2.54	6.13E-1	1.23E-1
FDCc	1	1	9.80	3.26	6.48E-1	7.01E-2	7.38E-3
FDCc	2	2	7.27	1.99	4.08E-1	5.55E-2	4.84E-3
IPA	1		6.71	2.68	5.70E-1	7.20E-2	5.10E-3
IPA	2		5.58	1.85	4.15E-1	4.81E-2	6.88E-3
LR	1/3		10.58	7.05	4.15	2.46	1.13
LRD	1/3		10.10	6.86	4.09	2.44	1.11
CLR	1/2		8.84	5.19	2.53	9.34E-1	2.33E-1
CLRD	1/2		8.72	5.08	2.49	8.56E-1	2.14E-1
CLRD	1		9.02	4.41	1.62	3.28E-1	5.99E-2

customers). We took $t_n = \text{round}(n^p)$, with different values of p , depending on the derivative estimation methods. Here, $\text{round}(x)$ means x rounded to the nearest integer. For a given C_n , the number of runs was then $n = \text{round}((C_n)^{1/(p+1)})$. For example, if $p = 1/2$, then for $C_n = 10^3$, we have $t_n = 10$ and $n = 100$, while for $C_n = 10^5$, we

have $t_n = 46$ and $n = 2154$. On the other hand, if $p = 2$ and $C_n = 10^5$, then $t_n = 2154$ and $n = 46$, which means much less initialization bias. This illustration shows how LR has to sacrifice unbiasedness to avoid variance explosion. For the FD methods, we took $\epsilon_n = 0.1 n^{-\gamma}$, where γ is indicated in the tables, and tried only the centered version. Note that the FD methods used twice the computer budget than the other methods, since each of the n replications required two runs of length t_n .

To estimate the MSE, in each case, we repeated r times the computation of Y_n based on these n runs of length t_n . The value of r was larger for smaller C_n : we took $r = \min(20, 10^6/n)$. Thus, each sample MSE value in the tables is an average of r "square error" values. These sample MSE are quite noisy, despite the large amount of computer time involved in these experiments, so that the numerical results just give a rough idea of what is going on.

In this case, IPA applies, which implies that from our discussion at the end of section 3.2, FDC will have the same convergence rate if p and γ are large enough. This can be observed clearly in the results (for $p \geq 1$ and $\gamma \geq 1$). One can also see that IPA and FDC are much more effective than the other methods in this case. For LR, incorporating the control variate scheme of [16] (CLR) yields significant improvement. LRD means that we have used (20) instead of (16). This gives a small improvement. The last line of both tables is interesting: we took $p = 1$ instead of the "optimal" $p (= 1/2)$ prescribed by theorem 3. For $p = 1/4$, the results are then clearly worse. However, for $p = 1/2$, they look better! This can be explained by the fact that in the second case, the bias is more important. The optimal t_n is still in the order of n^p , but with a larger hidden constant. This is why taking $p = 1$ does better for (relatively) small values of C_n . However, if we continue increasing C_n , the results of the previous line ($p = 1/2$) will become better at some point.

4.2. PROPORTION OF CUSTOMERS WHOSE SOJOURN TIME IS MORE THAN L

Suppose now that the objective function $\alpha(\theta)$ is the steady-state fraction of customers whose sojourn time is more than some fixed constant L . If we define

$$\varphi_j = \begin{cases} 1, & \text{if } X_j > L; \\ 0, & \text{otherwise,} \end{cases}$$

then $h_t(\theta, \omega) = (1/t) \sum_{j=1}^t \varphi_j$. For the $M/M/1$ queue, the theoretical value of that fraction is given by $\alpha(\theta) = e^{-\mu(1-\rho)L}$ (see Kleinrock [14], p. 202). The derivative with respect to the mean service time θ is $\alpha'(\theta) = (L/\theta^2)e^{-(1-\rho)L/\theta}$.

In this case, straightforward IPA does not apply (i.e. it yields a biased estimate), since for a fixed sequence ω of underlying $U(0, 1)$ random variables, the sample performance measure $h_t(\theta, \omega)$ is discontinuous with respect to θ . See L'Ecuyer and Perron [18] for further details. Therefore, we will not try IPA for this example. Note, however, that it is possible to apply IPA after replacing each φ_j by an appropriate conditional expectation (see [18]). This is called *smoothed* perturbation

analysis [8]. We will not do it here. For LR, the score function is the same as for the previous example.

We performed the same kind of numerical experiments as for the previous example, with $\rho = 1/4$, $L = 1$, and $\varepsilon_n = 0.01 n^{-\gamma}$. The results appear in table 3. Now that IPA does not apply, FDC is not as good, i.e. its convergence rate is no more $C^{-1/2}$, but appears to correspond rather to the results of theorem 1(iv). In particular, large values of ρ and γ are bad in this case. The remainder of the results also agree with the theory.

Table 3

Sample MSE for derivative of $P(\text{sojourn time} > 1)$, $M/M/1$ queue with $\rho = 1/4$.

	ρ	γ	$C_n = 10^1$	$C_n = 10^2$	$C_n = 10^3$	$C_n = 10^4$	$C_n = 10^5$
FDc	1/2	1/4	33.97	12.95	4.59	1.03	2.33E-2
FDCc	2/3	1/3	3.45	8.97E-1	1.64E-1	2.93E-2	3.85E-3
FDCc	1	1	8.68	3.54	1.18	4.14E-1	2.21E-1
FDCc	2	2	14.17	7.02	3.63	1.37	7.93E-1
LR	1/3		9.44E-1	2.51E-1	6.08E-2	1.87E-2	5.56E-3
LRD	1/3		9.09E-1	2.36E-1	5.61E-2	1.79E-2	5.10E-3
CLR	1/2		9.81E-1	2.39E-1	5.00E-2	1.05E-2	1.90E-3
CLRD	1/2		9.63E-1	2.15E-1	4.17E-2	6.99E-3	1.58E-3
CLRD	1		1.39	3.30E-1	5.83E-2	1.15E-2	2.75E-3

5. Conclusion

In the context of a Markov chain model, we have described various derivative estimation methods that do not exploit the regenerative structure explicitly. In that context, longer simulations and (in the finite-difference case) shorter finite-difference intervals must be used to reduce the bias. However, that often implies a variance increase. A trade-off must then be made between bias and variance to minimize, say, the mean square error. For many different estimators, we have shown how such a trade-off can be optimized, under reasonable conditions, if the aim is to optimize the convergence rate of the mean square error as a function of the computational budget (i.e. total simulation length). Some methods, like LR or FD, have poor convergence rates because decreasing the bias is very expensive in terms of variance increase. In other cases, like for IPA, or for FDC when IPA applies, one can decrease the bias without increasing the variance and the convergence rate is much better: the canonical rate of $C^{-1/2}$ can be achieved.

We looked at a number of tricks that could be used to try to reduce the variance of the likelihood ratio method. Among these tricks, only the control variate scheme of [16] did improve the convergence rate. Other tricks, like truncating the score function by some "windowing" mechanism, or associating to each step of the chain its own score function based only on what has happened up to this step, could

yield small variance reductions but do not improve the convergence rate. We have numerical examples that illustrate these theoretical properties.

Acknowledgements

This work has been supported by NSERC-Canada Grant No. 110050 and FCAR-Québec Grant No. EQ2831. Gaétan Perron, Peter W. Glynn, and the Guest Editor, R. Rubinstein, made helpful suggestions.

References

- [1] H. Arsham, A. Feuerverger, D.L. McLeish, J. Kreimer and R.Y. Rubinstein, Sensitivity analysis and the "what if" problem in simulation analysis, *Math. Comput. Mod.* 12(1989)193–219.
- [2] S. Asmussen, *Applied Probability and Queues* (Wiley, 1987).
- [3] S. Asmussen, Performance evaluation for the score function method in sensitivity analysis and stochastic optimization, Manuscript, Chalmers University of Technology, Göteborg, Sweden (1991).
- [4] G. Brassard and P. Bratley, *Algorithmics, Theory and Practice* (Prentice–Hall, Englewood Cliffs, NJ, 1988).
- [5] X.R. Cao, Sensitivity estimates based on one realization of a stochastic system, *J. Stat. Comput. Simul.* 27(1987)211–232.
- [6] P. Glasserman, Performance continuity and differentiability in Monte Carlo optimization, *Proc. Winter Simulation Conf. 1988* (IEEE Press, 1988), pp. 518–524.
- [7] P. Glasserman, Derivative estimates from simulation of continuous-time Markov chains, *Oper. Res.* 40(1992)292–308.
- [8] P. Glasserman and W.-B. Gong, Smoothed perturbation analysis for a class of discrete event systems, *IEEE Trans. Auto. Control* AC-35(1990)1218–1230.
- [9] P.W. Glynn, Likelihood ratio gradient estimation: An overview, *Proc. Winter Simulation Conf. 1987* (IEEE Press, 1987), pp. 366–375.
- [10] P.W. Glynn, Optimization of stochastic systems via simulation, *Proc. Winter Simulation Conf. 1989* (IEEE Press, 1989), pp. 90–105.
- [11] P.W. Glynn, Likelihood ratio gradient estimation for stochastic systems, *Commun. ACM* 33 (10) (1990)75–84.
- [12] P. Heidelberger, X.-R. Cao, M.A. Zazanis and R. Suri, Convergence properties of infinitesimal perturbation analysis estimates, *Manag. Sci.* 34(1989)1281–1302.
- [13] Y.-C. Ho, Performance evaluation and perturbation analysis of discrete event dynamic systems, *IEEE Trans. Auto. Control* AC-32(1987)563–572.
- [14] L. Kleinrock, *Queueing Systems, Vol. 1: Theory* (Wiley, 1975).
- [15] P. L'Ecuyer, A unified view of the IPA, SF, and LR gradient estimation techniques, *Manag. Sci.* 36(1990)1364–1383.
- [16] P. L'Ecuyer and P.W. Glynn, A control variate scheme for likelihood ratio gradient estimation, in preparation.
- [17] P. L'Ecuyer and P.W. Glynn, Stochastic optimization by simulation: Convergence proofs for the GII/GI queue in steady-state (1992), submitted for publication.
- [18] P. L'Ecuyer and G. Perron, On the convergence rates of IPA and FDC derivative estimators (1990), submitted for publication.
- [19] M.S. Meketon and P. Heidelberger, A renewal theoretic approach to bias reduction in regenerative simulations, *Manag. Sci.* 28(1982)173–181.

- [20] M.S. Meketon, Optimization in simulation: A survey of recent results, *Proc. Winter Simulation Conf. 1987* (IEEE Press, 1987), pp. 58–67.
- [21] M.I. Reiman and A. Weiss, Sensitivity analysis for simulation via likelihood ratios, *Oper. Res.* 37(1989)830–844.
- [22] R.Y. Rubinstein, The score function approach for sensitivity analysis of computer simulation models, *Math. Comp. Simul.* 28(1986)351–379.
- [23] R.Y. Rubinstein, Sensitivity analysis and performance extrapolation for computer simulation models, *Oper. Res.* 37(1989)72–81.
- [24] R.Y. Rubinstein, How to optimize discrete-event systems from a single sample path by the score function method, *Ann. Oper. Res.* 27(1991)175–212.
- [25] R. Suri, Infinitesimal perturbation analysis of general discrete event dynamic systems, *J. ACM* 34(1987)686–717.
- [26] R. Suri, Perturbation analysis: The state of the art and research issues explained via the *GI/G/1* queue, *Proc. IEEE* 77(1989)114–137.*****
- [27] M.A. Zazanis and R. Suri, Comparison of perturbation analysis with conventional sensitivity estimates for stochastic systems (1988), manuscript.