**Les cahiers du GERAD** ISSN: 0711-2440

# Convergence Rates for Steady-state Derivative Estimators

Pierre L'Ecuyer

G–90–55

December 1990

# 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

November 1990

**Abstract**

We describe different derivative estimators for the case of steady-state performances 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.

**KEYWORDS:** Discrete event systems; gradient estimation; steady-state

**Résumé**

Nous introduisons différents estimateurs pour la dérivée d'une mesure de performance sur horizon infini et obtenons l'ordre de leur taux de convergence. Ces estimateurs n'utilisent *pas* explicitement la structure regénérative du système. Nous examinons des estimateurs basés sur l'analyse de perturbation infinitésimale, sur la méthode du rapport de vraisemblance, et sur différentes formes de différences finies.

# I. INTRODUCTION

Estimating derivatives of expected performance measures with respect to some continuous parameters, in the context of stochastic discrete-event simulations, has received a lot of attention lately [2, 6, 8, 9, 13, 14, 15, 17]. Such derivative estimators are useful for sensitivity analysis, or can be used within stochastic optimization algorithms [6, 11]. For finite-horizon simulations, Glynn [6] 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 get the canonical convergence rate. Glynn [6] gives (subcanonical) convergence rates for forward and centered FD schemes, with and without common random numbers, under specific assumptions. L'Ecuyer and Perron [12] 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 II. Section III describes the derivative estimators that we consider and derive 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 that context, that the convergence rates will be the same as for the finite-horizon case. Indeed, it turns out that the straightforward LR estimators do not reach the canonical rate anymore. We derive their convergence rates with and without the control variate approach proposed in [10]. (Note however that there exists derivative estimators that converge at the canonical rate, but they use *explicitly* the regenerative structure [5, 13].) For IPA and FD, we obtain the same rates as for the finite-horizon case.

## II. Model and assumptions

As in [1], for any $f : \mathbb{N} \to [0, \infty)$, we define $O(f(n))$ as the *set* of functions $g : \mathbb{N} \to [0, \infty)$ 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 [11]. We consider a Markov chain $\{X_j(\theta, \omega), j = 0, 1, \ldots\}$ 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 $P_\theta$ depends (in general) on the parameter value $\theta$. 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 $g : (a, b) \times S \to \mathbb{R}$ is assumed measurable. Let

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

be the average cost for the first $t$ steps and let

$$\alpha_t(\theta) = \int_\Omega h_t(\theta, \omega) dP_\theta(\omega). \tag{2}$$

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 $\theta$. 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

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

These assumptions hold for many systems of interest (like, e.g., typical regenerative systems).

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 that. 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 expended CPU time $C$), where $\beta^*$ is the largest value of $\beta$ 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 $t_n \to \infty$ as $n \to \infty$ to obtain $B_n \to 0$. 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 $E[Y_n] \neq \alpha'_{t_n}(\theta_0)$ (like for finite differences). A reasonable choice is

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

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

## III. Derivative estimators and their convergence rates

### A. *Finite differences*

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 $c_n > 0$. Let $\omega_1^-, \ldots, \omega_n^-, \omega_1^+, \ldots, \omega_n^+$ be $2n$ independent sample points, where each $\omega_i^-$ is generated under $P_{\theta - c_n}$ and each $\omega_i^+$ is generated under $P_{\theta + c_n}$. The *forward* FD estimator is

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

while the *central* FD estimator is

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

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 $h_{t_n}(\theta + c_n, \omega_i^+)$ (and similarly for $h_{t_n}(\theta - c_n, \omega_i^-)$ and $h_{t_n}(\theta + c_n, \omega_i^+)$ in (9)).

In typical stochastic simulations, the sample point $\omega$ (which represents the "randomness" that drives the system) can be viewed as a sequence of independent uniform variates between 0 and 1. $P_\theta \equiv P$ is then independent of $\theta$. In that context, one can generate only $n$ (independent) sample points $\omega_1, \ldots, \omega_n$ 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

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

while the *central* FDC estimator is

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

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 gives rise to large variability (Glynn 1989, Meketon 1987, Zazanis and Suri 1988). The variances of $\psi_{n,i}^{\mathrm{FDf}}$ and $\psi_{n,i}^{\mathrm{FDc}}$ are in $\Theta(c_n^{-2})$ 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 $\psi_{n,i}^{\mathrm{FDCf}}$ and $\psi_{n,i}^{\mathrm{FDCc}}$ are in $\Theta(c_n^{-1})$), Glynn (1989) 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 $t_n \in \Theta(n^p)$, $c_n \in \Theta(n^{-\gamma})$, and suppose that $\sup_{\theta\in(a,b)}[\text{ Var }(h_t(\theta,\omega))] \in O(1/t)$. For the forward case (FDf and FDCf), suppose that $\alpha''(\theta) \neq 0$ (and exists) in a neighborhood of $\theta_0$. For the central case (FDc and FDCc), suppose that $\alpha'''(\theta) \neq 0$ (and exists) in a neighborhood of $\theta_0$. For FDC, suppose (as in Glynn 1989) that the variance of $\psi_{n,i}$ is in $\Theta(c_n^{-1})$. Then, the "optimal" values of $p$, $\gamma$, $\beta$, 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 $\Theta(c_n)$ for forward differences and in $\Theta(c_n^2)$ for central differences. For (8), one has $V_n = \text{ Var }(Y_n^{\text{FDf}}) \in O(1/(nc_n^2 t_n)) = O(n^{-1-p-2\gamma})$, $B_n \in O(1/t_n + c_n) = O(n^{-p} + n^{-\gamma})$, 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 $\beta$ for this inequality to be satisfied is $\beta = 1/2$, with $p = \gamma = 1/3$. (In that 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

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

Now, instead of using the same $t_n$ and $c_n$ for all $n$ replications, one can use say $t_i \in \Theta(i^p)$ and $c_i \in \Theta(i^{-\gamma})$ for replication $i$, for $i = 1, \ldots, n$. $Y_n$ 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 $w_1, \ldots, w_n$. 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

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

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 remain 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 $\theta_0 + c_n$, then another one at $\theta_0$. 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

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

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

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

To get $(p+1)\beta - 1 - p + 2\gamma = (p+1)\beta - 2\gamma = 0$, 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 on 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.

### *B. Infinitesimal perturbation analysis*

As for FDC, let us view the sample point $\omega$ as a sequence of independent uniform variates between 0 and 1, and denote $P_\theta$ by $P$. Under some conditions (see [3, 9]), the random variable $h_t'(\theta,\omega) = dh_t(\theta,\omega)/d\theta$ can be used as an unbiased estimator of $\alpha_t'(\theta)$. This is the IPA derivative estimator.

One can get $n$ i.i.d. replicates of $h_{t_n}'(\theta,\omega)$ and use the estimator

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

Here, $\omega_1, \omega_2, \ldots, \omega_n$ are independent sample points generated under the common distribution $P$. (In practice, for complex simulations, $h'_{t_n}(\theta, \omega)$ is not always easy to compute. There are also many cases where it gives a biased or even totally meaningless derivative estimator. See [9].)

**Theorem 2.** *Assume that* $E[h'_t(\theta_0, \omega)] = \alpha'_t(\theta_0)$ *for each* $t > 0$ *and that the variance of* $h'_t(\theta_0, \omega)$ *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 $C_n^\beta R_n \in O(n^{(p+1)\beta}(n^{-p-1} + n^{-2p})) = O(1)$ 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 $E[h'_t(\theta_0, \omega)] = \alpha'_t(\theta_0)$ are given in [3, 9]. In most interesting cases where IPA applies, these conditions are satisfied. In the context of finite-horizon simulations, L'Ecuyer and Perron [12] show that whenever these conditions are satisfied, the variances of $\psi_{n,i}^{\mathrm{FDCf}}$ and $\psi_{n,i}^{\mathrm{FDCc}}$ are in $O(1)$ (instead of $\Theta(1/c_n)$) 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

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

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

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

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

### *C. Likelihood ratios*

The likelihood ratio (or score function) derivative estimation technique, aimed for the case where $P_\theta$ really depends on $\theta$, goes as follows [5, 9, 13, 14, 15]. To fix ideas, suppose that $\omega$ can be viewed as $\omega = (\zeta_1, \ldots, \zeta_t)$, where $\zeta_j$ is the value taken by a continuous random variable with density $f_{j,\theta}$, and the $\zeta_j$'s are independent. Given $X_{j-1}$, the value of $\zeta_j$ determines the next state $X_j$ of the Markov chain. (A similar development can be made for discrete or mixted random variables.) Let $G = P_{\theta_0}$ and suppose that $G$ dominates the $P_\theta$'s in a neighborhood of $\theta_0$. One can rewrite

$$\alpha_t(\theta) = \int_\Omega H_t(\theta, \omega) dG(\omega) \tag{14}$$

where

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

Now, generate $\omega$ according to $G$ and compute

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

as the LR derivative estimate, 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 one has, as in (13):

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

This can be viewed as a generalization of IPA [9]. On the other hand, as explained in [12], this is equivalent to applying IPA *after* replacing $h_t$ and $P_\theta$ by $H_t$ and $G$, 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 [9].

It is also well known that unfortunately, the variance of $H'_t(\theta,\omega)$ is typically in $\Theta(t)$, i.e. increases linearly with the simulation length. Basically, this is because the variance of the score function is in $\Theta(t)$. This means that the assumptions of Theorem 2 are usually not satisfied for that "special case" of IPA. L'Ecuyer and Glynn [10] 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). But this is no more the case here. Intuitively, we expect that the run lengths should be kept shorter to keep the variance down.

**Theorem 3.** *Assume that* $E[H'_t(\theta_0,\omega)] = \alpha'_t(\theta_0)$ *for each* $t > 0$. *Let* $t_n \in \Theta(n^p)$. *For LR, if the variance of* $H'_t(\theta_0,\omega)$ *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 of* $H'_t(\theta_0,\omega)$ *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 $i$, $i = 1,\ldots,n$, as explained in the context of FD. It can be seen easily that for LR and CLR, the orders of $V_n$ and $B_n$, and the convergence rates, remain the same as in the above Theorem.

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 $O(1)$. But 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 $\ell_j \in \Theta(j^q)$ for $q \leq 1$. This gives the gradient estimator

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

The variance of $Z_j = g(\theta_0, X_j)\sum_{k=j-\ell_j+1}^{j} S_k(\theta_0,\zeta_k)$ is typically in $\Theta(\ell_j)$. 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

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

and the largest value of $\beta$ 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 $\ell_j = j$ $(q = 1)$ 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. But 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), but still in the same order.

Now, suppose that the process is regenerative. Let $0 = \tau_0 \leq \tau_1 \leq \tau_2, \ldots$ be the regeneration points. Conditional on $\tau_i$, what happens after $\tau_i$ is independent of what happened until $\tau_i$. Based on that, 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, $E[g(\theta_0, X_j)S_k(\theta_0,\zeta_k)|\tau_1,\tau_2,\ldots] = E[g(\theta_0, X_j)|\tau_1,\tau_2,\ldots]E[S_k(\theta_0,\zeta_k)|\tau_1,\tau_2,\ldots]$. But the problem is that in general, $E[S_k(\theta_0,\zeta_k)|\tau_1,\tau_2,\ldots] \neq E[S_k(\theta_0,\zeta_k)] = 0$, which means that one cannot really remove these terms. Correct (asymptotically unbiased) gradient estimators for the regenerative case are given in [5, 13].

## ACKNOWLEDGEMENTS

This work has been supported by NSERC-Canada grant # A5463 and FCAR-Québec grant # EQ2831. Gaétan Perron and Peter W. Glynn made helpful suggestions.

## REFERENCES

[1] Brassard, G. and Bratley, P., *Algorithmics, Theory and Practice*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.

[2] Cao, X. R., "Sensitivity Estimates Based on One Realization of a Stochastic System", *J. Stat. Comput. Simul.*, **27** (1987), 211–232.

[3] Glasserman, P., "Performance Continuity and Differentiability in Monte Carlo Optimization", *Proceedings of the Winter Simulation Conference 1988*, IEEE Press (1988), 518–524.

[4] Glasserman, P., "Derivative Estimates From Simulation of Continuous Time Markov Chains", Technical report, AT&T Bell Laboratories, Holmdel, New Jersey (1989).

[5] Glynn, P. W. "Likelihood Ratio Gradient Estimation: an Overview", *Proceedings of the Winter Simulation Conference 1987*, IEEE Press (1987), 366–375.

[6] Glynn, P. W. "Optimization of Stochastic Systems Via Simulation", *Proceedings of the Winter Simulation Conference 1989*, IEEE Press (1989), 90–105.

[7] Heidelberger, P., Cao, X.-R., Zazanis, M. A. and Suri, R., "Convergence Properties of Infinitesimal Perturbation Analysis Estimates", *Management Science*, **34**, 11 (1989), 1281–1302.

[8] Ho, Y.-C., "Performance Evaluation and Perturbation Analysis of Discrete Event Dynamic Systems", *IEEE Transactions of Automatic Control*, **AC-32**, 7 (1987), 563–572.

[9] L'Ecuyer, P., "A Unified View of the IPA, SF, and LR Gradient Estimation Techniques", to appear in *Management Science* (1990).

[10] L'Ecuyer, P. and Glynn, P. W., "A Control Variate Scheme for Likelihood Ratio Gradient Estimation", In preparation (1990).

[11] L'Ecuyer, P., Giroux, N., and Glynn, P. W., "Stochastic Optimization by Simulation: Convergence Proofs and Experimental Results for the GI/G/1 Queue in Steady-State", Submitted for publication (1990).

[12] L'Ecuyer, P. and Perron, G., "On the Convergence Rates of IPA and FDC Derivative Estimators for Finite-Horizon Stochastic Simulations", In preparation (1990).

[13] Reiman, M. I. and Weiss, A., "Sensitivity Analysis for Simulation via Likelihood Ratios", *Operations Research*, **37**, 5 (1989), 830–844.

[14] Rubinstein, R. Y., "The Score Function Approach for Sensitivity Analysis of Computer Simulation Models", *Math. and Computers in Simulation*, **28** (1986), 351–379.

[15] Rubinstein, R. Y., "Sensitivity Analysis and Performance Extrapolation for Computer Simulation Models", *Operations Research*, **37**, 1 (1989), 72–81.

[16] Suri, R., "Infinitesimal Perturbation Analysis of General Discrete Event Dynamic Systems", *J. of the ACM*, **34**, 3 (1987), 686–717.

[17] Suri, R., "Perturbation Analysis: The State of the Art and Research Issues Explained via the GI/G/1 Queue", *Proceedings of the IEEE*, **77**, 1 (1989), 114–137.

[18] Zazanis, M. A. and Suri, R., "Comparison of Perturbation Analysis with Conventional Sensitivity Estimates for Stochastic Systems", Submitted for publication (1988).